UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACIÓN

**Proposal of a trust and reputation model
to provide recommendation in open
quality oriented environments**

TESIS DOCTORAL

Iñaki Martínez Sarriegui

Ingeniero de Telecomunicación

Madrid, 2016

Departamento de Tecnología Fotónica y Bioingeniería

Grupo de Bioingeniería y Telemedicina

E.T.S.I. de Telecomunicación

UNIVERSIDAD POLITÉCNICA DE MADRID

**Proposal of a trust and reputation model
to provide recommendation in open
quality oriented environments**

TESIS DOCTORAL

Autor:

Iñaki Martínez Sarriegui

Directora:

María Elena Hernando Pérez

Madrid, Febrero de 2016

Tribunal formado por el Magnífico y Excelentísimo Sr. Rector de la Universidad Politécnica de Madrid, el día 18 de enero de 2016.

Presidente:	D. Enrique J. Gómez Aguilera	Doctor Ingeniero de Telecomunicación Catedrático de Universidad Universidad Politécnica de Madrid
Vocales:	D ^a Isabel Román Martínez	Doctora Ingeniera de Telecomunicación Profesora Colaboradora a Tiempo Completo Universidad de Sevilla
	D. Pablo Lamata de la Orden	Doctor Ingeniero Biomédico Sir Henry Dale Fellow & Lecturer King's College of London
	Arturo González Ferrer	Doctor Ingeniero Informático Investigador Doctor Instituto de Investigación Sanitaria San Carlos (IdISSC)
Secretaria:	D ^a . Patricia Sánchez González	Doctora Ingeniera Biomédica Profesora Ayudante Doctor Universidad Politécnica de Madrid
Suplentes:	D ^a Paula de Toledo Heras	Doctora Ingeniera de Telecomunicación Profesora Titular de Universidad Universidad Carlos III de Madrid
	D. César Cáceres Taladriz	Doctor Ingeniero de Telecomunicación Profesor Contratado Doctor Universidad Rey Juan Carlos

Realizado el acto de defensa y lectura de la Tesis el día 4 de febrero de 2016 en Madrid, acuerda otorgarle la calificación de

El Presidente

Los Vocales

El Secretario

*"Asking a PhD student when he/she is going to finish his/her thesis
Is like asking a lady her age: disrespectful and tactless"*

Mith Buttermann

A mis padres y mi hermana,
por limitar la 'maldita' pregunta al máximo.

A Laura, por haber llegado en el peor,
o mejor momento de la tesis.

Acknowledgments

Por fin llega el día que parecía que no llegaría nunca. Y lo hace después de 10 años, una tesis que no fue, una que debió ser, otra que pudo ser, y la que al final ha sido. El hacer una tesis en un tema diferente de los proyectos en los que trabajas tiene inconvenientes, pero también tiene algunas ventajas importantes. Al dilatarse tanto mi pertenencia al GBT he aprendido muchas tecnologías diferentes y, sobre todo, he conocido a mucha gente. Llega el momento de recordarlos a todos y temo olvidarme de alguno o no estar a la altura de lo que se espera. Confío en hacerlo bien, si olvido a alguno por favor no me lo tengáis en cuenta.

En primer lugar me gustaría dar las gracias a mi directora de tesis, Elena, por la confianza depositada en mí todos estos años, sobre todo en esos momentos generalmente relacionados con la tesis cuando ni yo mismo habría confiado en mí. Y gracias por la ayuda de estos últimos meses, principalmente con la relacionada con los temas burocráticos y de normativa que me hubieran sepultado. Quería agradecer también a Enrique por su apoyo y por haberme tenido siempre en cuenta para cualquier cosa que hubiera que hacer en el grupo.

En segundo lugar me gustaría tener un reconocimiento especial para las dos personas que supervisaron mis estancias doctorales, pues marcaron lo que es esta tesis. *To the professor Hong Zhu, for his help with my first steps in the world of intelligent agents and for guiding me in my journey to a world full of abstractions and new concepts that shaped the way I've faced my doctorate. Special thanks also to the professor Michael Luck, for that terrifying question in our first meeting: Why? That question destroyed the foundations of my PhD up to that moment. But was the first stone upon this new thesis was built. For all your support, help and advice. Many thanks.* Y aprovechando que hablo de mis estancias... *To all my labmates and housemates in Oxford and London for all the stories and your friendship: Vikki, Lucia, José, Massimo, Chris, Lijun, Samhar.* A Pablo por su hospitalidad, ayuda, apoyo y amistad cuando estuve en Oxford, hiciste que me sintiera siempre como en casa. Y a Álvaro, por ese sofá que tan bien me vino las dos primeras semanas en Londres.

El tercer lugar de estos agradecimientos lo ocupan todos aquellos con los que he compartido trabajo todos estos años en forma de colaboraciones y proyectos de investigación. A Mercedes y a Belén, y al resto de gente del Hospital de Sabadell, por los econsulta, closed-loop y mobiguides; pero sobre todo por lo que va más allá del trabajo como esas comidas aceleradas, esas raciones de oreja, las visitas “guiadas” por Madrid o los grandes éxitos músico-comerciales. *And to all the mobiguiders, for the last four years and the great pleasure of working with you, but again specially for what it was not work. Special thanks to Mor for its willingness and to Val and Lucia for their time and their outstanding review. But my warmest acknowledgment is to all the baboos: Erez, Daniel, Enea, Boris, Arturo, Tom, Carlos, Ángel, Danielle, Guy and Denis. For those marathon Skype-based testing sessions, our dinners and our beers, and for all the good memories. Next time we meet Aperol is on me!!!*

Me gustaría también agradecer a todos mis proyectandos en el GBT. En esta tesis no hay trabajo directo vuestro, pero sí que hay mucho de vosotros en el tipo de investigador que soy. Espero haberos enseñado yo también algo. A Jimena, Jacob, Álvaro, Cristina, Álvaro, Mercedes y Paloma. Espero de veras que os vaya todo muy bien.

Me considero una persona con mucha suerte pues creo poder decir que tengo muchos amigos junto a mí. Todos vosotros formáis parte de mi vida, y unas torpes palabras en un apartado de agradecimientos difícilmente pueden expresar lo agradecido que estoy. Pero aun así lo haré de la mejor manera que se.

A los compañeros de toda una vida, los que siempre están ahí y estarán, a pesar de que nuestro contacto no sea tan habitual como lo fue. Si, a vosotros, a los de la plaza, a los de siempre. A David, a Javitu, a Pablo, a Álvaro y a Pey. Y a los que gracias a vosotros entraron en mi vida: Nacho, Edu, Manolo, Javi... Siempre nos quedará Hupe (como mayor fan que soy), y Pryca, que ya no es Pryca que es Carrefour.

A mis buenos tricantinos, a todos aquellos que me habéis aguantado estos últimos años, sobre todo en esta fase final. Por todos los buenos momentos pasados y por venir. Por las cenas, los cumpleaños, los viajes, las latitas místicas, las noches en el Venan. Por soportarme todas esas veces que soy seco y borde. Por crecer juntos, por madurar acompañado. Por lo que queda. A Candela, Floren, Ayensa, Dani, Herrerías, Alber, Rubén, Miguel y Fer. A Teresa, Lucía y Elena. Y a Elisa, Natalia, Miriam, Petra, Chus, Nuria, Jaime, Sebas y Paco. Y a las peques. Y a Marta por el empujón inicial para comenzar con el doctorado.

A esos ‘teleqos’ de moda. Porque lo que unió el grupo 14 no lo separe nadie. Por ese Barógrafo infame. Por esas cenas en casa de la Mary. Y por las casas rurales. Y porque sí, por una vez voy a ganar a Samu!! A Miguel y a Gema, por los dibujos y friqueces. A Yorch, eh? Por todos esos laboratorios. A Kike, por Road to Hell, la intensificación y todas nuestras charlas. A Javi, por esa perilla que sí que llevabas, y por ser fan número uno del zigzag. A Yisus, por la charca, por la maldita placa. Y porque por fin termino y dejarás de preguntarlo. A Bea, por sus grandes respuestas y decir lo que todos esperamos, y por asustarnos con el agua. A Mary, por tantas y tantas cenas compartiendo penas y glorias, por estar siempre ahí para hablar y para todo lo que se le necesita. Y por último a Samu. Parece mentira que al final sea yo doctor antes que tú, pero entiendo perfectamente por qué no lo eres todavía. A todos otra vez, porque la risa compensa.

A María, por ayudarme en el que probablemente haya sido mi momento más bajo. Por darme ánimos y sobre todo herramientas para rehacerme y poder llegar a este día de la mejor de las maneras. Gracias también por tus ayudas puntuales con la tesis. Y gracias por las pelis, cenas, cervezas y charlas. Y por las que quedan.

Y gracias a “la banda de la Remonta” y el resto de nuevos amigos que me ha traído la vida este año tan intenso.

Una persona sabia dijo: “poca gente puede presumir de que ir a trabajar sea como quedar con amigos, pero así es” (Oropesa 2012). Yo lo veo y envido más. Los gebetianos sois para mí como una segunda familia. Son tantos los buenos ratos que he pasado en estos más de 11 años, incluyendo el PFC, que se me hace muy duro imaginar el trabajo y la vida sin vosotros. Gracias a todos los que fuisteis, sois y seréis gebetianos. Por los cafeses, museSES, picnics, cenas de navidad, amigos invisibles y caseibs. Y las sidrerías y cenas del club. Y POR LOS MÁS GUAPOS!!!

A Nacho, por los viajes, cines, charlas, por más alegrías ya que tristezas o pupas ☺ Por todo y por lo que queda. Por ser como un hermano. A Paty, por muchos años aquí, por tu ayuda con la tesis, porque este tema ha llegado a buen puerto en gran parte por ti y por tu apoyo y asesoramiento. A Borja, porque conocerse desde hace 32 años y seguir siendo amigos, trabajando juntos y disfrutando de ratos inolvidables no es fácil de decir. A Paloma, a Dove, no me cambies nunca por favor, porque todos necesitamos tener a alguien positivo como tú cerca. Y porque así además Samu queda algo más cerca de mí, una de las personas más cariñosas que conozco. A Iniesta, por todo lo que hemos hablado, por intentar arreglar el mundo empezando por el GBT, y por todas esas ideas que algún día se realizarán. Y a Fer, por su inestimable ayuda con esta tesis, y sobre todo porque sin ti lo mismo no habría dos tesis en el GBT. A Marta, por tantas y tantas risas, charlas y por compartir juntos, y con ese Chemuco grande, estos últimos meses la carga de la tesis. A Ana B. por incluirme siempre en los mails de proyectos y por las visitas en el rectorado. A los gebetianos de adopción: Manu, David, Ana, Gustavo, y los babies.

Y no me podría olvidar de vosotros, figuras: a Rodri, Solana y Pedrito, sobre todo por esa noche en Tolosa con la furgo. Y al resto de miembros de tropeltxos y G.B.T. futbolero: Molina, Paco, Jaime, Luismi, Álvaro, Alexis. A todos los gebetianos que he conocido a lo largo de estos años sigáis por aquí o no: Vero, Silvia, Juan Luis, Gloria, Ale, Viveca, Paula, Jorge, Antonio, Jesús, Pablojo, Alicia, Bea, Omar, Eloy, Dirty, Brito, Javier, Laura, Carlos, Olga, Clara, Natalia, Rossella, Junior, Nachovi, Pacoga, Vinny, Fonta, Ruth, Alberto, Álvaro Serrano, Carmen, Agustín, mini Jose, Estefanía, Diego, Mailin, Jorge, etc. A César, por su eterna sonrisa y ofrecirme su ayuda y contactos. A todos y cada uno de los proyectandos y masterandos que habéis pasado por aquí y que he tenido el gusto de conocer. A Jipi, por ser el verdadero jefe.

A Gema, la G del GBT. Por ser mi “guefa”. Por todos los proyectos, proyectandos, las charlas, los viajes, smarthealth y demás proyectos diabéticos realizados y por realizar. Por ayudarme siempre que lo he necesitado sin pedírtelo. Por aguantar y soportar mis subidas y bajadas, los tacos en el laboratorio, cuando me he puesto a cantar, a reír o a llorar. Por ser tan buena persona, tan trabajadora y tan prudente siempre, ojalá aprendiera de ti. Por ser una gran amiga. Por haber tenido el enorme placer de haber trabajado contigo todos estos años, se me va a hacer muy difícil siquiera imaginar no seguir haciéndolo.

Y ya por último a mi familia, por estar siempre ahí, con el mejor de los apoyos, incondicional y constante. A mis padres y a mi hermana. A mi tía y mis primos. A mi abuela: Amona, que ya dejó de estudiar!! Bueno, eso nunca, pero termino la tesis. A Juani y las primas. Y a Laura. Os quiero!!

Madrid/Tres Cantos/Oxford/Londres, a 4 de febrero de 2016

P, S: Os tengo que hacer a vosotros un agradecimiento especial, pues fuisteis los que me recogisteis después de las primeras caídas, en el momento clave. Lo hicisteis de manera desinteresada, para corregir mis primeros pasos con la tesis. Esta no es aquella tesis, es otra, pero aun así, va por vosotros, pues no sé si lo habría conseguido. Por vuestro tiempo.

Table of Contents

SUMMARY.....	1
English	1
Spanish	3
 I. INTRODUCTION.....	 5
I.1. Motivation and overview.....	7
I.2. Document structure	8
 II. STATE OF THE ART.....	 9
II.1. Technology Enhanced Learning (TEL).....	11
II.2. Recommender Systems (RSs).....	11
II.2.1. General classification	12
II.2.2. Recommendation in TEL platforms.....	14
II.2.3. Summary classification	15
II.2.4. Evaluation of Recommender Systems.....	17
II.2.4.1. Evaluation of TEL-RSs	17
II.2.1. Commercial systems	18
II.3. Trust and Reputation (T&R) systems	19
II.3.1. Classification	20
II.3.1.1. Hard features	20
II.3.1.2. Soft features	21
II.3.2. Evaluation of T&R systems.....	26
II.4. Combining T&R with RS systems	28
II.5. T&R and RS systems in HealthCare.....	29
II.6. Objective assessment of resources	31
II.7. Agent-based simulation models	31
II.8. Discussion.....	32
 III. HYPOTHESES AND OBJECTIVES.....	 35
III.1. Hypotheses	37
III.2. Objectives	38

IV.	<i>MATERIAL AND METHODS</i>	39
IV.1.	<i>Chapter structure</i>	41
IV.2.	<i>Running example</i>	41
IV.3.	<i>TruQOs model</i>	42
IV.3.1.	Introduction	42
IV.3.2.	Resource Reputation	44
IV.3.3.	Provider Reputations	45
IV.3.3.1.	Reputation based on provider's past evaluations	45
IV.3.3.2.	Reputations based on past opinions about the provider	46
IV.3.3.2.1.	Reputation based on consumer's personal experience	46
IV.3.3.2.2.	Reputation based on social information	47
IV.3.4.	Reputation based on system's structure	49
IV.3.4.1.	Semi-dynamic architecture	49
IV.3.4.2.	Reputation based on Institutions' providers' evaluations	51
IV.3.4.3.	Relations based on structure	51
IV.3.4.4.	Structural Reputation	52
IV.3.5.	Building the puzzle	53
IV.3.6.	How TruQOs deals with newcomers	54
IV.3.7.	Reputation and Reliability aggregation	56
IV.4.	<i>Evaluation of the model</i>	57
IV.4.1.	TruQOs Simulation Tool (TST)	58
IV.4.1.1.	Presenting the simulation tool	58
IV.4.1.2.	User behaviour modelling	60
IV.4.1.2.1.	Consumer	60
IV.4.1.2.2.	Provider	62
IV.4.1.2.3.	Polyvalent	64
IV.4.1.3.	TruQOs recommendation process	64
IV.4.1.4.	TST-Agents	65
IV.4.1.4.1.	"Simulator", the SimulationAgent	65
IV.4.1.4.2.	"Reputer", the ReputationAgent	67
IV.4.1.4.3.	"Evaluator", the EvaluationAgent	68
IV.4.1.4.4.	"Hierarcher", the HierarchyAgent	70
IV.4.1.4.5.	"Socializer", the SocialAgent	71
IV.4.1.4.6.	"Databaser", the DatabaseAgent	73
IV.4.1.5.	Simulation process	74
IV.4.1.6.	Additional functionalities	76
IV.4.2.	Experimental design	77
IV.4.2.1.	Introduction	77
IV.4.2.2.	Description of evaluation scenarios	78
IV.4.2.3.	Selection of the Organisations and Hierarchies	78
IV.4.2.4.	Users' behavioural patterns	80
IV.4.2.5.	Evaluation parameters	80

IV.4.2.6.	Selection of the initial state	81
IV.4.2.1.	TruQOs functionality tests	83
IV.4.2.1.1.	S1: Effect of changes in the quality of resources	83
IV.4.2.1.2.	S2: Effect of changes in the consumers' opinions	84
IV.4.2.1.3.	S3: Changes in quality and opinions combined effect	85
IV.4.2.1.4.	S4: Number of attackers' effect	86
IV.4.2.1.5.	S5: Effect of Institution's position in the Ranking	87
IV.4.2.1.6.	S6: Hierarchy's effect	88
IV.4.2.1.7.	S7: Effect of changes in Reliability values	89
IV.4.2.1.8.	S8: Newcomers' effect	90
IV.4.2.1.9.	S9: Integrating resources and users recommendation	91
IV.4.2.1.10.	S10: Initial dispersion	91
IV.4.3.	Summary of evaluation scenarios	92
V.	RESULTS	93
V.1.	Introduction	95
V.2.	TruQOs functionality tests	96
V.2.1.	S1: Effect of changes in the quality of resources	96
V.2.2.	S2: Effect of changes in the consumers' opinions	98
V.2.3.	S3: Changes in quality and opinions combined effect	102
V.2.3.1.	S3a: Quality decrease	102
V.2.3.2.	S3b: Opinions-based attacks	104
V.2.3.3.	S3c: Compensating quality changes through opinions	105
V.2.4.	S4: Number of attackers' effect	108
V.2.5.	S5: Effect of Institution's position in the Ranking	109
V.2.5.1.	S5a: Attack from well reputed institutions (Competition)	110
V.2.5.2.	S5b: Attack from institutions with lower reputation (Discredit)	111
V.2.5.3.	S5c: Favours own institution's users (Promotion)	113
V.2.6.	S6: Hierarchy's effect	114
V.2.6.1.	S6a: Rivalry inside groups within an institution	114
V.2.6.2.	S6b: Competitiveness inside groups within an institution	116
V.2.6.3.	S6c: Nepotism within an institution	117
V.2.7.	S7: Effect of changes in Reliability values	119
V.2.7.1.	S7a: Resource's opinions and evaluations variability (q-Factor)	119
V.2.7.2.	S7b: Personal past experiences variability	121
V.2.7.3.	S7c: Quartiles Aggregation Method values variability	122
V.2.8.	S8: Newcomers' effect	124
V.2.8.1.	S8a: Resource Recommendation	124
V.2.8.2.	S8b: User Recommendation	126
V.2.9.	S9: Integrating resources and users recommendation	127
V.2.9.1.	S9a: Generation of Resources focus	127
V.2.9.2.	S9b: Personal service focus	129

V.2.10.	S10: Initial dispersion.....	130
V.3.	<i>Summary and discussion of results</i>	132
V.3.1.	Concerning TruQOS	132
V.3.2.	Concerning TruQOS Simulation Tool.....	135
VI.	<i>CONCLUSIONS AND FUTURE WORKS.....</i>	137
VI.1.	<i>Verification of research hypotheses.....</i>	139
VI.1.1.	TruQOs hypothesis	139
VI.1.2.	TST hypothesis	139
VI.2.	<i>Main contributions</i>	140
VI.3.	<i>Scientific publications</i>	142
VI.3.1.	International journals	142
VI.3.2.	Abstracts in international journals	143
VI.3.3.	Book chapters	143
VI.3.4.	International congresses.....	144
VI.3.5.	National congresses.....	145
VI.4.	<i>International collaborations.....</i>	146
VI.5.	<i>Future works.....</i>	147
VI.5.1.	Concerning the trust and reputation model.....	147
VI.5.2.	Concerning the simulation tool	148
VI.5.3.	Beyond TruQOs: portability and new research.....	149
VII.	<i>REFERENCES.....</i>	151
VIII.	<i>ANNEXES.....</i>	171
VIII.1.	<i>Annex A: Dictionary of TruQOs and TST terms</i>	173
VIII.2.	<i>Annex B: Summary of equations.....</i>	175
VIII.2.1.	TruQOs	175
VIII.2.1.1.	Resource Reputation.....	175
VIII.2.1.2.	Provider Reputations	175
VIII.2.1.2.1.	Reputation based on provider's past evaluations	175
VIII.2.1.2.2.	Reputations based on past opinions about the provider	175
VIII.2.1.2.2.1.	Reputation base on consumer's personal experience	175
VIII.2.1.2.2.2.	Reputation based on social information	176

VIII.2.1.3.	Reputation based on system's structure	176
VIII.2.1.3.1.	Reputation based on Institutions' providers' evaluations	176
VIII.2.1.3.2.	Structural Reputation	176
VIII.2.1.4.	Final Reputations	177
VIII.2.1.4.1.	Opinions-based provider Reputation	177
VIII.2.1.4.2.	Provider Reputation.....	177
VIII.2.1.4.3.	TruQOs Reputation	177
VIII.2.1.4.3.1.	Resource Recommendation	177
VIII.2.1.4.3.2.	User Recommendation.....	177
VIII.2.1.5.	Newcomers	177
VIII.2.1.5.1.	Resource Recommendation.....	177
VIII.2.1.5.2.	User Recommendation	177
VIII.2.1.5.3.	Weighting parameter K.....	178
VIII.2.2.	TST	178
VIII.2.2.1.	Probability distributions of Resource/Service quality.....	178
VIII.2.2.2.	OnlyOpinions Reputation.....	178
VIII.2.2.3.	UserAgent behaviours	178
VIII.2.2.3.1.	Consumer.....	178
VIII.2.2.3.2.	Provider	179
VIII.2.3.	Evaluation	179
VIII.2.3.1.	Ratios of recommendation	179
VIII.2.3.2.	Ratio of benefits.....	179
VIII.3.	<i>Annex C: Summary of user roles and parameters</i>	180
VIII.4.	<i>Annex D: Scenarios of evaluation.....</i>	181

Figure Index

Figure IV.3-1. Main goal of TruQOs: to provide an ordered recommendation of resources. _____	42
Figure IV.3-2. TruQOs' dimensions of reputation schema _____	43
Figure IV.3-3. Impact of parameter 'q' on reliability's assessment. X-axis: σ , Y-axis: reliability. _____	45
Figure IV.3-4. Witness triangulation: assessing provider's reputation using other users experiences _____	47
Figure IV.3-5. Impact of the socialTrust assessment in $f(x,y)$ (projection $y=0$ in $f(x,y)$). With a) $\varphi=1$ and b) $\varphi=1000$. _____	48
Figure IV.3-6. Impact of parameter 'c' on $f(x,y)$ with a) $\varphi=1$ and b) $\varphi=1000$ _____	49
Figure IV.3-7. Internal hierarchy of an Institution _____	50
Figure IV.3-8. Ranking of Institutions for Service N _____	50
Figure IV.3-9. Impact of HDS in Structural Reputation assessment. Axis x: levels' difference Δ _____	52
Figure IV.3-10. TruQOs reputation/reliability assessment diagram _____	53
Figure IV.3-11. Impact of promotion period on function K for newcomers' reputation weighting mechanism _____	56
Figure IV.3-12. Quartiles Aggregation Method: Resources are first ordered according to their Reliability and then their Reputation is modified by q_i factors to obtain Reputation'. _____	56
Figure IV.4-1. TruQOs Simulation Tool (TST) _____	58
Figure IV.4-2. TST structure and Agents communications: normal lines represent mainly interactions among users during the simulation but they may include also settings interaction, dotted lines are only setting interactions and the dash line is database communication _____	59
Figure IV.4-3. Consumer requests recommendation of a Resource to consume _____	61
Figure IV.4-4. Consumer requests recommendation of a Provider to interact with _____	62
Figure IV.4-5. Provider's Resource generation _____	63
Figure IV.4-6. TruQOs recommendation process _____	64
Figure IV.4-7. SimulationAgent's GUI: Management and Visualization of: a) Left.- Recommendation mode and Users' simulation parameters; b) Global simulation parameters and simulation status. _____	66
Figure IV.4-8. Global Users' activity during the simulation _____	66
Figure IV.4-9. Users' activity over time during the simulation _____	66
Figure IV.4-10. ReputationAgent's GUI: Reputation method and parameters selection _____	68
Figure IV.4-11. EvaluationAgent's GUI: Management and Visualization of: a) Left.- Simulation of quality of results offered by the Providers; b) Right.- Resources' Classification according to their quality and Service _____	69
Figure IV.4-12. HierarchyAgent's GUI: Management and Visualization of: a) Left.- System's structure and internal hierarchies; b) Right.- Institutions' Ranking for Services provided _____	71
Figure IV.4-13. SocialAgent' GUI: Users and their interactions 'sociogram'. Colours identify users from same Institution. Black triangle: dead user. _____	72
Figure IV.4-14. SocialAgent's GUI: Interactions' information panel. _____	73
Figure IV.4-15. SocialAgent's GUI: Resources generated by an User for each Service _____	73
Figure IV.4-16. Simulation process from Simulator's point of view with the different phases highlighted. _____	75
Figure IV.4-17. System's structure and internal hierarchies to be used in the evaluation. Orange=Provider, Blue=Consumer _____	79
Figure IV.4-18. Internal hierarchy of Institution 1* for scenarios S5 and S6. Orange=Provider, Blue=Consumer _____	79
Figure IV.4-19. Newcomers effect, scenario S7: Left.- Institution 4* new User in existing Institution; Right.- Institution 6 new User in new Institution. Orange=Provider, Blue=Consumer _____	79
Figure IV.4-20. S0 Accessed Resources over simulation time for each Provider. _____	82
Figure IV.4-21. Ratio of Resource Recommendation over simulation time for each S0 Provider. _____	83
Figure IV.4-22. S0 Results Summary _____	83

Figure IV.4-23. Quality of resources pattern by 'User 12' in Scenarios S1 and S3b; for Resources recommendation	84
Figure IV.4-24. Quality of resources and opinions received patterns by 'User 12' in Scenarios S2 and S3b; for Resources recommendation	85
Figure IV.4-25. Quality of resources and opinions received patterns by 'User 12' in Scenario S3b; for Resources recommendation	86
Figure IV.4-26. Quality of resources and opinions received patterns by 'User 12' in Scenario S4; for Resources recommendation	86
Figure IV.4-27. Quality of resources and opinions patterns by 'User 12' in a) Scenario S5a and b) Scenario S5b for Resources recommendation. In a) the quality of resources pattern for 'Institution 5' (Q (I5)) is also shown.	87
Figure IV.4-28. Quality of resources and opinions received patterns by 'User 27' in Scenario S5c; for Resources recommendation	88
Figure IV.4-29. Quality of resources and opinions received patterns by 'User 6' in a) Scenarios S6a and S6b and b) Scenario S6c; for Resources recommendation	89
Figure IV.4-30. Quality of resources and opinions received patterns by 'User 6' (equal values) in Scenario S7; for Resources recommendation	90
Figure IV.4-31. Quality of resources pattern by newcomer 'User 31' vs quality of resources of the rest of users (Q (rest)) pattern for a) Resources recommendation in Scenario S8a and b) Users recommendation in Scenario S8b	90
Figure IV.4-32. a) Quality of service during direct interactions (Qdir) pattern and qualities of resources patterns by 'User 12' in Scenario S9a; b) Quality of resources (Qres) pattern and qualities of service during direct interactions patterns by 'User 12' in Scenario S9b. Both for Resources and Users recommendation	91
Figure V.2-1. Ratio of Resources Recommendation over simulation time for 'User 12' and each of S1 tests, with unweighted evaluations. Comparison between considering all resources vs. only new resources. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) or decreased (red).	97
Figure V.2-2. Ratio of Resources Recommendation over simulation time for 'User 12' and each of S1 tests. Comparison between considering unweighted vs. weighted evaluations. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) or decreased (red).	98
Figure V.2-3. S1 Total ratios summary for 'User 12'.	98
Figure V.2-4. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S2 tests. Study of the effect of parameter ' ω ' in the calculation of Social Reputation. . Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple)	99
Figure V.2-5. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S2 tests. Comparison between considering all resources vs. only new resources. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple).	100
Figure V.2-6. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S2 tests. Comparison between considering unweighted vs. weighted opinions. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple)	101
Figure V.2-7. S2 Total ratios summary for 'User 12'.	102
Figure V.2-8. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S3a tests with weighted evaluations. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) or decreased (red).	103
Figure V.2-9. S3a Total ratios summary for 'User 12'.	104

Figure V.2-10. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S3b tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple)	104
Figure V.2-11. S3b Total ratios summary for 'User 12'.	105
Figure V.2-12. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S3c tests with weighted evaluations and opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) or decreased (red); and the periods where users of 'Institutions 1,2' criticize (purple) or praise (orange) 'User 12'.	107
Figure V.2-13. S3c Total ratios summary for 'User 12'.	107
Figure V.2-14. Ratio of Resources Recommendation over simulation time for 'User 12' for S4 tests with weighted evaluations and opinions. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) and the periods where other users criticize (purple) 'User 12'.	109
Figure V.2-15. S4 Total ratios summary for 'User 12'.	109
Figure V.2-16. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S5a tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 5' hate 'User 12' (purple)	111
Figure V.2-17. S5a Total ratios summary for 'User 12'.	111
Figure V.2-18. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S5b tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 2' hate 'User 12' (purple)	112
Figure V.2-19. S5b Total ratios summary for 'User 12'.	112
Figure V.2-20. Ratio of Resources Recommendation over simulation time for 'User 27' and a subset of S5c tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 27' resources is increased (blue) or decreased (red); and the periods where users of 'Institutions 5' love 'User 27' (orange)	113
Figure V.2-21. S5c Total ratios summary for 'User 27'.	113
Figure V.2-22. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S6a tests with weighted opinions. Marked with colour squares the period where 'Users 4,5' lye about 'User 6' (purple) giving bad opinions.	115
Figure V.2-23. S6a Total ratios summary for 'User 6'.	115
Figure V.2-24. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S6b tests with weighted opinions. Marked with colour squares the period where 'User 10' lye about 'User 6' (purple) giving bad opinions.	116
Figure V.2-25. S6b Total ratios summary for 'User 6'.	117
Figure V.2-26. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S6c tests with weighted opinions. Marked with colour squares the period where 'Users 2,8' lye about 'User 6' (purple) giving good opinions.	118
Figure V.2-27. S6c Total ratios summary for 'User 6'.	118
Figure V.2-28. Ratio of Resources Recommendation over simulation time for 'User 6' and each of S7a tests with weighted evaluations. Marked with colour squares the period where the quality of the resources created by 'User 6' is varied (pink).	120
Figure V.2-29. S7a Total ratios summary for 'User 6'.	120
Figure V.2-30. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S7b tests with weighted evaluations. Marked with colour squares the period where the quality of the resources created by 'User 6' is varied (pink).	122

Figure V.2-31. S7b Total ratios summary for 'User 6'.	122
Figure V.2-32. Ratio of Resources Recommendation over simulation time for 'User 6' and each of S7c tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of 'User 6' is varied (pink).	123
Figure V.2-33. S7c Total ratios summary for 'User 6'.	123
Figure V.2-34. Accessed Resources over simulation time for 'User 31' and each of S8a1 tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of 'User 31' is increased (blue); and when 'User 31' was still unborn (grey). Newcomer in 'Institution 3'.	124
Figure V.2-35. Accessed Resources over simulation time for 'User 31' and each of S8a2 tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of 'User 31' is increased (blue); and when 'User 31' was still unborn (grey). Newcomer created in a new institution.	125
Figure V.2-36. S8a1 Total ratios summary for 'User 31'. Newcomer in 'Institution 3'.	125
Figure V.2-37. S8a2 Total ratios summary for 'User 31'. Newcomer created in a new institution.	126
Figure V.2-38. Interactions Answered by 'User 31' over simulation time for each of S8b tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of 'User 31' is increased (blue); and when 'User 31' was still unborn (grey).	126
Figure V.2-39. S8b Total ratios summary for 'User 31'.	127
Figure V.2-40. Accessed resources created by 'User 12' over simulation time for S9a TruQOs tests. Marked with colour squares the period where the quality of the resources of 'User 12' is increased (blue).	128
Figure V.2-41. Interactions Answered by 'User 12' over simulation time for S9a TruQOs tests. Marked with colour squares the period where the quality of the resources of 'User 12' is increased (blue).	128
Figure V.2-42. S9a Total resource recommendation ratio (RR), user recommendation (UR) ratio, and ratio of benefits summary for 'User 12' and TruQOs' tests.	128
Figure V.2-43. Interactions Answered by 'User 12' over simulation time for S9b TruQOs tests. Marked with colour squares the period where the quality of the resources of 'User 12' is increased (light blue) and where the quality of the service provided by 'User 12' during direct interactions is increased (blue).	129
Figure V.2-44. Accessed resources created by 'User 12' over simulation time for S9a TruQOs tests. Marked with colour squares the period where the quality of the resources of 'User 12' is increased (light blue) and where the quality of the service provided by 'User 12' during direct interactions is increased (blue).	130
Figure V.2-45. S9b Total resource recommendation (RR) and user recommendation (UR) ratios summary in left axis, and Resource recommendation ratio difference with TruQOs_1 in right axis. All the results for 'User 12' and TruQOs' tests.	130
Figure V.2-46. S0 vs. S10 Results Comparison.	131

Table Index

<i>Table II.2-1. Summary Classification of Recommender Systems based on (Burke 2007; Park et al. 2012)</i>	16
<i>Table II.3-1. Trust and Reputation systems' classification, modified from (Noorian & Ulieru 2010). In grey the six T&R systems of Noorian et al. study.</i>	23
<i>Table IV.3-1. Classification of reputation dimensions into source of information families.</i>	43
<i>Table IV.4-1. TST's database tables' structure</i>	73
<i>Table IV.4-2. Evaluation Framework's Scenarios</i>	92
<i>Table V.1-1. Evaluation Tests Summary</i>	95
<i>Table V.2-1. Scenario S1: TruQOs' tests description</i>	96
<i>Table V.2-2. Scenario S2: TruQOs' tests description</i>	99
<i>Table V.2-3. Scenarios S3a-S3b: TruQOs' tests description</i>	102
<i>Table V.2-4. Scenario S3c: TruQOs' tests description</i>	105
<i>Table V.2-5. Scenario S4: TruQOs' tests description</i>	108
<i>Table V.2-6. Scenario S5: TruQOs' tests description</i>	110
<i>Table V.2-7. Scenario S6a: TruQOs' tests description</i>	114
<i>Table V.2-8. Scenario S6b: TruQOs' tests description</i>	116
<i>Table V.2-9. Scenario S6c: TruQOs' tests Tests description</i>	117
<i>Table V.2-10. Scenario S7a: TruQOs' tests description</i>	119
<i>Table V.2-11. Scenario S7b: TruQOs' tests description</i>	121
<i>Table V.2-12. Scenario S7c: TruQOs' tests description</i>	123
<i>Table V.2-13. Scenario S8: TruQOs' tests description</i>	124
<i>Table V.2-14. Scenario S9a: TruQOs' tests description</i>	127
<i>Table V.2-15. Scenario S9b: TruQOs' tests description</i>	129
<i>Table V.2-16. Scenario S10: TruQOs' tests description</i>	131
<i>Table V.2-17. Dispersion effectiveness analysis results</i>	131

Au village, sans prétention,
J'ai mauvaise réputation.
Qu'je m'démène ou qu'je reste coi
Je pass' pour un je-ne-sais-quoi!
Je ne fais pourtant de tort à personne
En suivant mon chemin de petit bonhomme.
Mais les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Non les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Tout le monde médit de moi,
Sauf les muets, ça va de soi.

Le jour du Quatorze Juillet
Je reste dans mon lit douillet.
La musique qui marche au pas,
Cela ne me regarde pas.
Je ne fais pourtant de tort à personne,
En n'écoutant pas le clairon qui sonne.
Mais les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Non les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Tout le monde me montre du doigt
Sauf les manchots, ça va de soi.

Quand j'croise un voleur malchanceux,
Poursuivi par un cul-terreux;
J'lance la patte et pourquoi le taire,
Le cul-terreux s'retrouv' par terre
Je ne fais pourtant de tort à personne,
En laissant courir les voleurs de pommes.
Mais les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Non les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Tout le monde se rue sur moi,
Sauf les culs-de-jatte, ça va de soi.

Pas besoin d'être Jérémie,
Pour d'viner l'sort qui m'est promis,
S'ils trouv'nt une corde à leur goût,
Ils me la passeront au cou,
Je ne fais pourtant de tort à personne,
En suivant les ch'mins qui n'mènent pas à Rome,
Mais les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Non les brav's gens n'aiment pas que
L'on suive une autre route qu'eux,
Tout l'mond' viendra me voir pendu,
Sauf les aveugles, bien entendu.

Georges Brassens (1921-1981), La Mauvaise Réputation

SUMMARY

English

Internet is changing everything, and this revolution is especially present in traditionally offline spaces such as medicine. In recent years health consumers and health service providers are actively creating and consuming Web contents stimulated by the emergence of the Social Web. Reliability stands out as the main concern when accessing the overwhelming amount of information available online. Along with this new way of accessing the medicine, new concepts like *ubiquitous or pervasive healthcare* are appearing. Trustworthiness assessment is gaining relevance: open health provisioning systems require mechanisms that help evaluating individuals' reputation in pursuit of introducing safety to these open and dynamic environments.

Technical Enhanced Learning (TEL) -commonly known as eLearning- platforms arise as a paradigm of this Medicine 2.0. They provide an open while controlled/supervised access to resources generated and shared by users, enhancing what it is being called informal learning. TEL systems also facilitate direct interactions amongst users for consultation, resulting in a good approach to ubiquitous healthcare. The aforementioned reliability and trustworthiness problems can be faced by the implementation of mechanisms for the trusted recommendation of both resources and healthcare services providers.

Traditionally, eLearning platforms already integrate recommendation mechanisms, although this recommendations are basically focused on providing an ordered classifications of resources. For users' recommendation, the implementation of trust and reputation systems appears as the best solution. Nevertheless, both approaches base the recommendation on the information from the subjective opinions of other users of the platform regarding the resources or the users.

In this PhD work a novel approach is presented for the recommendation of both resources and users within open environments focused on knowledge exchange, as it is the case of TEL systems for ubiquitous healthcare. The proposed solution adds the objective evaluation of the resources to the traditional subjective personal opinions to estimate the reputation of the resources and of the users of the system. This combined measure, along with the reliability of that calculation, is used to provide trusted recommendations. The integration of opinions and evaluations, subjective and objective, allows the model to defend itself against misbehaviours. Furthermore, it also allows 'colouring' cold evaluation values by providing additional quality information such as the educational capacities of a digital resource in an eLearning system. As a result, the recommendations are always adapted to user requirements, and of the maximum technical and educational quality.

To our knowledge, the combination of objective assessments and subjective opinions to provide recommendation has not been considered before in the literature. Therefore, for the evaluation of the trust and reputation model defined in this PhD thesis, a new simulation tool will be developed following the agent-oriented programming paradigm. The multi-agent approach allows an easy modelling of independent and proactive behaviours for the simulation of users of the system, conforming a faithful resemblance of real users of TEL platforms.

For the evaluation of the proposed work, an iterative approach have been followed, testing the performance of the trust and reputation model while providing recommendation in a varied range of scenarios. A comparison with two traditional recommendation mechanisms was performed: a) using only users' past opinions about a resource and/or other users; and b) not using any reputation assessment and providing the recommendation considering directly the objective quality of the resources.

The results show that the developed model improves traditional approaches at providing recommendations in Technology Enhanced Learning (TEL) platforms, presenting a higher adaptability to different situations, whereas traditional approaches only have good results under favourable conditions. Furthermore the promotion period mechanism implemented successfully helps new users in the system to be recommended for direct interactions as well as the resources created by them. On the contrary OnlyOpinions fails completely and new users are never recommended, while traditional approaches only work partially.

Finally, the agent-oriented programming (AOP) paradigm has proven its validity at modelling users' behaviours in TEL platforms. Intelligent software agents' characteristics matched the main requirements of the simulation tool. The proactivity, sociability and adaptability of the developed agents allowed reproducing real users' actions and attitudes through the diverse situations defined in the evaluation framework. The result were independent users, accessing to different resources and communicating amongst them to fulfil their needs, basing these interactions on the recommendations provided by the reputation engine.

Spanish

El auge y penetración de las nuevas tecnologías junto con la llamada Web Social están cambiando la forma en la que accedemos a la medicina. Cada vez más pacientes y profesionales de la medicina están creando y consumiendo recursos digitales de contenido clínico a través de Internet, surgiendo el problema de cómo asegurar la fiabilidad de estos recursos. Además, un nuevo concepto está apareciendo, el de *pervasive healthcare* o sanidad ubicua, motivado por pacientes que demandan un acceso a los servicios sanitarios en todo momento y en todo lugar. Este nuevo escenario lleva aparejado un problema de confianza en los proveedores de servicios sanitarios.

Las plataformas de eLearning se están erigiendo como paradigma de esta nueva Medicina 2.0 ya que proveen un servicio abierto a la vez que controlado/supervisado a recursos digitales, y facilitan las interacciones y consultas entre usuarios, suponiendo una buena aproximación para esta sanidad ubicua. En estos entornos los problemas de fiabilidad y confianza pueden ser solventados mediante la implementación de mecanismos de recomendación de recursos y personas de manera confiable.

Tradicionalmente las plataformas de eLearning ya cuentan con mecanismos de recomendación, si bien están más enfocados a la recomendación de recursos. Para la recomendación de usuarios es necesario acudir a mecanismos más elaborados como son los sistemas de confianza y reputación (*trust and reputation*) En ambos casos, tanto la recomendación de recursos como el cálculo de la reputación de los usuarios se realiza teniendo en cuenta criterios principalmente subjetivos como son las opiniones de los usuarios.

En esta tesis doctoral proponemos un nuevo modelo de confianza y reputación que combina evaluaciones automáticas de los recursos digitales en una plataforma de eLearning, con las opiniones vertidas por los usuarios como resultado de las interacciones con otros usuarios o después de consumir un recurso. El enfoque seguido presenta la novedad de la combinación de una parte objetiva con otra subjetiva, persiguiendo mitigar el efecto de posibles castigos subjetivos por parte de usuarios malintencionados, a la vez que enriquecer las evaluaciones objetivas con información adicional acerca de la capacidad pedagógica del recurso o de la persona. El resultado son recomendaciones siempre adaptadas a los requisitos de los usuarios, y de la máxima calidad tanto técnica como educativa.

Esta nueva aproximación requiere una nueva herramienta para su validación in-silico, al no existir ninguna aplicación que permita la simulación de plataformas de eLearning con mecanismos de recomendación de recursos y personas, donde además los recursos sean evaluados objetivamente. Este trabajo de investigación propone pues una nueva herramienta, basada en el paradigma de programación orientada a agentes inteligentes para el modelado de comportamientos complejos de usuarios en plataformas de eLearning. Además, la herramienta permite también la simulación del funcionamiento de este tipo de entornos dedicados al intercambio de conocimiento.

La evaluación del trabajo propuesto en este documento de tesis se ha realizado de manera iterativa a lo largo de diferentes escenarios en los que se ha situado al sistema frente a una amplia gama de comportamientos de usuarios. Se ha comparado el rendimiento del modelo de confianza y reputación propuesto frente a dos modos de recomendación tradicionales: a) utilizando sólo las opiniones subjetivas de los usuarios para el cálculo de la reputación y por extensión la recomendación; y b) teniendo en cuenta sólo la calidad objetiva del recurso sin hacer ningún cálculo de reputación.

Los resultados obtenidos nos permiten afirmar que el modelo desarrollado mejora la recomendación ofrecida por las aproximaciones tradicionales, mostrando una mayor flexibilidad y capacidad de adaptación a diferentes situaciones. Además, el modelo propuesto es capaz de asegurar la recomendación de nuevos usuarios entrando al sistema frente a la nula recomendación para estos usuarios presentada por el modo de recomendación predominante en otras plataformas que basan la recomendación sólo en las opiniones de otros usuarios.

Por último, el paradigma de agentes inteligentes ha probado su valía a la hora de modelar plataformas virtuales complejas orientadas al intercambio de conocimiento, especialmente a la hora de modelar y simular el comportamiento de los usuarios de estos entornos. La herramienta de simulación desarrollada ha permitido la evaluación del modelo de confianza y reputación propuesto en esta tesis en una amplia gama de situaciones diferentes.

I. INTRODUCTION

The first chapter of this PhD dissertation introduces the reader into the research work described in the rest of the document. First the motivations of the work are presented, followed by the description of the structure or outline of the PhD thesis. The ultimate goal is to provide a general panorama of the work, facilitating the reading of the document as well as the search of relevant information within the document.

I.1. Motivation and overview

Internet is changing everything, even ourselves. It has changed our means to communicate and to find life partners, how we access the information or how the businesses operate, and the way we shop and consume (Veer 2011). And this transformation not only has not finished but it will be boosted by the irruption of new technologies as the Internet of Things (IoT) (Tan 2010), Big Data (Cukier & Mayer-Schönberger 2013) or the Semantic Web (Berners-Lee et al. 2001).

This Internet revolution is especially present in traditionally offline spaces such as medicine: healthcare combines communication between users, access to information, knowledge sharing, business and services provision/consumption. In recent years health consumers and health service providers are actively creating and consuming Web contents stimulated by the emergence of the Social Web and creating the concept of Medicine 2.0 (Eysenbach 2008). Moreover, as the Web matures, consumers are gaining access to personalized applications adapted to their health needs and interests (Fernandez-Luque et al. 2011).

Reliability stands out as the main concern when accessing the overwhelming amount of information available online. The creation of open or semi-open virtual medical information environments oriented to health education (to both patients and professionals), where information and resources are medically controlled, supervised and/or verified; will provide the required security when accessing health-related information. Recommender systems appear as the perfect technological solution to filter the information and to personalize it to users' tastes and needs.

Along with this new way of accessing the medicine, new concepts like *ubiquitous or pervasive healthcare* are appearing. Trustworthiness assessment is gaining relevance: open health provisioning systems require mechanisms that help evaluating individuals' reputation in pursuit of introducing safety to these open and dynamic environments. Trust and reputation systems, extensively introduced to e-business applications or peer-to-peer networks, can be applied to healthcare platforms to provide 'soft security' (Resnick & Zeckhauser 2002).

Technology Enhanced Learning (TEL) -commonly known as eLearning- platforms arise as a paradigm of both concepts. They provide an open while controlled/supervised access to resources generated and shared by users, enhancing what it is being called informal learning (Collis & Moonen 2001; Colley et al. 2002) TEL systems also facilitate direct interactions amongst users for consultation, resulting in a good approach to ubiquitous healthcare.

In this PhD dissertation a novel approach is presented for the recommendation of resources and users within open environments focused on knowledge exchange, as it is the case of TEL systems for ubiquitous healthcare. This approach is based on the following assumption: "*An objective assessment of the resources generated within an environment can be obtained, at the moment of creation of the resource, by automatic methods or impartial reviewers, resulting in an undeniable evaluation rating*". Section II.6 lays the foundations for the legitimacy of this affirmation, although its validity might be discussed. However, demonstrating the above affirmation is out of the scope of this thesis, and it will be taken as ground assumption.

The proposed solution adds the objective evaluation of the resources with the traditional personal opinions to estimate the reputation of the resources and users of the system. This measure, along with the reliability of that calculation, is used to provide trusted recommendations. The resulting union is TruQOs (**Trust in Quality-Oriented systems**): a trust and reputation model capable of filtering the finest and most relevant resources for each user, as well as recommending always the best expert for consultation while providing ‘soft-security’ in user-to-user interactions by means of reputation assessment.

The work includes an agent-based simulation tool, dubbed TruQOs Simulation Tool (TST), to model complex TEL systems with consultation capabilities amongst users, incorporating TruQOs as the recommendation engine. The multi-agent approach allows an easy modelling of independent and proactive behaviours for the users of the system, conforming a faithful resemblance of real users. The tool is used to perform an extensive evaluation of the model, testing TruQOs in a wide range of scenarios.

I.2. Document structure

This PhD thesis document is structured in 6 chapters, with the addition of 4 annexes and the list of all the references to other works from the literature:

- In this **Chapter I**, the motivation of this research work has been presented, briefly describing the problem to solve, its challenges and introducing the solution.
- **Chapter II**, covers the systematic state of the art review of the most relevant facets of this PhD thesis. Understanding the problem and knowing the existing options for its solution explains the approach followed on this research work.
- In **Chapter III** the Hypotheses on which this research work is based are presented, followed by the objectives that this thesis aims at fulfilling.
- **Chapter IV** presents TruQOs, the trust model for reputation assessment in quality oriented societies and the proposed material and methods for its evaluation, including TST the simulation tool implemented for this validation task.
- The main results of TruQOs’ evaluation are summarized in **Chapter V**: an extensive and thorough evaluation was carried out, forcing worst case scenarios in the simulation tool and testing the robustness of the proposed model under extreme conditions.
- To conclude, **Chapter VI** summarizes the main conclusions and contributions extracted from this PhD dissertation research. It also proposes a discussion about the validity of TruQOs as trust and reputation model for resources’ and users’ recommendation. The document concludes with a description of possible future works to continue the research started in this PhD thesis.

II.STATE OF THE ART

In this second chapter the relevant state of the art for this research work is presented. The exposition of precedent work aims at providing the reader with the appropriate information to understand the problem this PhD thesis faces and justifying the selected approach to solve it.

II.1. Technology Enhanced Learning (TEL)

Technology Enhanced Learning (TEL) aims at designing, developing and testing socio-technical innovations that will support and enhance learning practices of both individuals and organizations (Manouselis et al. 2011). TEL can be considered as a macro-context (Vuorikari & Berendt 2009) with different dimensions such as educational level, formality, delivery settings and different user roles. Two different types of learning can be distinguished: formal learning and informal learning. **Formal learning** includes learning offers from educational institutions within a curriculum and is characterised as teacher-directed, synchronous, accreditation-oriented, involving domain experts and in environments with person-to-person interactions. Research on formal learning has often followed a *top-down* approach, greatly depending on expert knowledge and a large amount of user-related information has to be encoded in the content creation phase (Cristea 2005).

On the other hand, **informal learning** is described as a phase of so-called lifelong learners responsible of their own learning process (Collis & Moonen 2001; Colley et al. 2002) and depends to large extent on individual preferences or choices, being self-directed (Brockett & Hiemstra 1991). It is characterised as learner-centred, asynchronous, without curriculum and flexible. The resources might come from a variety of sources: from expert communities to friends. Learning resources emerge from the *bottom-up* through the communities, without experts control and at any moment during the life cycle of the systems. There is an overwhelming variety of learning resources and users need tools that help them to find and to identify suitable learning resources, learning paths and relevant peer-learners, thusly, TEL systems can benefit from services provided by recommender systems.

II.2. Recommender Systems (RSs)

Generally speaking, recommender systems (or recommendation systems) are algorithms, SW tools, models or techniques used to provide final users with suggestions about contents, products or services (such as books, films, music, travels, etc.); filtering information to help users making decisions whether or not to use or to consume the proposed items or services (Resnick & Varian 1997; Mahmood & Ricci 2009). These recommendations are personalized, different users or groups receive different suggestions (Ricci et al. 2011).

The idea behind these systems is that it is in human nature to rely on recommendations provided by others in making daily decisions (Mahmood & Ricci 2009; McSherry & Mironov 2009). RSs appeared to help users to cope with the overwhelming amount and variety of information available on the Web and the rapid introduction of new e-business services (Ricci et al. 2011). The study of recommender systems and recommendation techniques started in mid-1990s (Goldberg et al. 1992; Anand & Mobasher 2005; Mahmood & Ricci 2009; McSherry & Mironov 2009) and in recent years it has become an important research area with thousands of publications; dedicated conferences and workshops (such as ACM Recommender Systems, RecSys); RSs courses at institutions of higher education; books and special issues at academic journals (such as AI Communications, IEEE Intelligent Systems or ACM Transactions on Computer-Human Interaction). Moreover, many authors have published extensive reviews on

the field (Frias-Martinez et al. 2006; Burke 2007; Frias-Martinez et al. 2009; Kim et al. 2010; Park et al. 2012; Sezgin & Ozkan 2013; Lu et al. 2015).

II.2.1. General classification

One of the best classifications of RSs is the taxonomy defined by Burke (Burke 2007). Burke's classification distinguishes between six types of recommendation approaches, in regard to the recommendation algorithm used. Burke's classification is shown here with the addition of some other recommender systems from literature. Table II.2-1 gathers some RSs examples from literature.

- Content-based (CB). The recommendations are generated from two sources: the product/service features and the ratings generated by the users. The system classifies items for the user's likes and dislikes and the classification depends on the ratings of the users with the same tastes. This way of sorting may cause overspecialized classifications that only include items very similar to choose (Blanco-Fernandez et al. 2008).
- Collaborative filtering (CF). The system collects ratings from members in a community and uses that information to recommend items to other users, "people-to-people correlation" (Schafer et al. 2001). The underlying idea is that different people have different tastes and rate things differently according to those tastes (Jøsang et al. 2007). If two users rate a set of items similarly we can deduce they have similar ideas, considering them as neighbours. CF is the most widely implemented technique in RSs.
- Demographic (DM). Recommends items based on the demographic profile of the user and the recommendations could be targeted to specific demographic niches (for example the language) combining ratings of the users of that niche.
- Knowledge-based (KB). These systems suggest products based on domain knowledge inferences about users' preferences. A similarity function estimates how much the user needs match the recommendations. KB systems tend to work better at the beginning of systems' life but they need to be provided with learning techniques to ensure their efficiency (Ricci et al. 2011).
- Community-based (CoB). These techniques base the recommendations on the maxim: "tell me who your friends are, and I will tell you who you are". People rely more on friends' advice than on anonymous individuals with similar tastes (even with more similarities than with friends) (Sinha & Swearingen 1999). These systems are also known as social recommender systems and, thusly, the recommendations are based on ratings provided by user's social environment. CoB systems are similar to CF systems, and in many cases they do not provide better recommendations. However, other authors have demonstrated that adding social network data to traditional CF systems improves considerably recommendations results.
- Hybrid recommender systems (Hyb). Hybrid systems are constructed combining multiple recommendation techniques, usually to deal with the cold-start problem (handling new items or new users). Burke (Burke 2007) identify seven types of hybridization techniques:
 - **Weighted (W)**. The score of different recommendation components are combined numerically.

- **Mixed (M).** Recommendations from different recommenders are presented together.
- **Switching (S).** The system chooses among recommendation components and applies the selected one. The choice of the switching criterion is important, some authors use internal criteria and others use external criteria.
- **Feature combination (FC).** Features derived from different knowledge sources are combined together and given to a single recommendation algorithm.
- **Feature Augmentation (FA).** One recommendation technique is used to compute a feature or set of features and the result is the input of the next technique.
- **Cascade (Ca).** Recommenders are given strict priority, with the lower priority ones breaking ties in the scoring of the higher ones.
- **Meta-level (ML).** One recommendation technique is applied and produces some sort of model, which is then the input used by the next technique.

Other authors, like Hee Park et al. (Park et al. 2012), only consider two categories: collaborative filtering and content-based filtering. They performed an extensive classification into application fields and data mining techniques used as recommendation algorithms. Hee Park et al. distinguished eight *application fields* (extending the taxonomy by Montaner et al. (Montaner 2003)): book, document, image, movie, music, shopping, TV program and others; and also eight *data mining techniques*:

- **Association rule (AR).** Refers to the discovery of all association rules that are above user-specified minimum support and minimum confidence levels.
- **Clustering (CI).** These methods identify a finite set of categories or clusters to describe data. The most popular are *K-means* (kM) and *self-organizing map* (SOM).
- **Decision tree (DT).** These are the most popular methods and are based on the construction of a decision tree to label or categorize cases.
- **k-Nearest Neighbour (kNN).** The k-nearest neighbour model is a typical traditional CF recommender system that makes recommendations through phases. The k-NN phases are threefold: (1) the system constructs a user profile from explicit ratings of items or indirectly from usage information; (2) the system applies statistical or machine learning techniques to discover k users, the recommenders, with similar behaviours, forming a neighbourhood of equals; (3) once the neighbourhood is formed, RSs make a top-n item set analysing tastes and interests of neighbours.
- **Neural network (NN).** These systems use artificial neural networks (ANN) to predict future behaviours of users based on different characteristics such as tastes, interests or rating values.
- **Link analysis (LA).** This technique discovers relations between domains in large databases. One type of link analysis is social network analysis, and studies patterns relationships and interactions to find a fundamental social structure.
- **Regression (R).** Analyses associative relationships between variables and one or more independent variables, and have been used for curve fitting and prediction.
- **Other heuristic methods (OHM).** Such as ontology methods.
- **Semantic Web (SeW).** Semantic web technologies mainly for resource filtering.

One of the latest and most complete efforts to summarize and classify recommender systems was published last year (Lu et al. 2015). In this work, the authors review more than 150 up-to-date real-world application developments of recommender systems, clusters their applications into eight main categories and summarizes the related recommendation techniques used in each category. This work provides also some useful guidelines about how to develop and apply recommender systems in different domains to support users in various decision activities.

II.2.2. Recommendation in TEL platforms

The development of recommender systems (RSs) has been related to a number of relevant user tasks supported by the RSs within some particular application content. Herlocker et al. (Herlocker et al. 2004) related user's tasks with specific recommendation goals and Manouselis et al. (Manouselis et al. 2011) modified this classification with the addition of Technology Enhanced Learning (TEL) users' specific tasks and how these tasks can be supported by RSs.

TEL environments have several particularities and transferring recommender system from an existing platform to TEL may not accurately meet the needs of the users. The main particularity of TEL platforms, as pointed by Drachsler et al. (Drachsler et al. 2008), is the existence of two different types of learning. The highly structured environment of **formal learning** platforms provides a great variety of information about the users that can be used by recommender systems to suggest courses, learning paths or resources to the final users (Baldoni et al. 2007). Although recommender systems oriented to formal learning environments can be found in literature; the characteristics of **informal learning**, as described in section II.1, have made this area a major focus for the research on TEL recommender systems over the last 15 years.

One of the first attempts has been the Altered Vista system (Walker et al. 2004) a collaborative filtering (CF) system based on word-of-mouth recommendation. Another example of CF is CYCLADES system (Avancini & Straccia 2005), based on ratings of the digital resources. Other attempts rely on knowledge-based methods such as QSIA (Question Sharing and Interactive Assignments) systems (Rafaeli et al. 2005) which relies on a mostly user-controlled process. Or systems using rule-based methods which, as stated by Abel et al. (Abel et al. 2008) have proved to be more useful than others systems as rules evolve in parallel to users' knowledge evolution or new ones are included.

Content-based methods were also used by many systems like the CoFind prototype (Dron et al. 2000) that also used digital resources from the Web but apply for the first time folksonomies (tags) for recommendation. Huang et al. (Huang et al. 2009) also propose a sequencing system which uses Markov chain model to calculate transition probabilities of possible learning objects in a sequenced course of study.

In the last years, the hybrid recommendation approach has been followed by many authors like Manouselis et al. (Manouselis & Costopoulou 2007). They proposed a multi-attribute recommendation system to consider multi-dimensional ratings from users. The solution proposed by Tang and McCalla (Tang & McCalla 2005) consists on an evolving e-learning system, open into new learning resources that may be found online, including a hybrid recommendation algorithm combining a clustering module technique and a collaborative filtering module.

Another example is the RACOFI (Rule-Aplying Collaborative Filtering) system (Lemire et al. 2005) that combined two recommendation approaches: a collaborative filtering systems and an inference rule engine. Shen and Shen followed a different approach (Shen & Shen 2004) and proposed a system based on sequencing rules that guide users through the concepts of an ontology of topics.

Hummel et al. (Hummel et al. 2007) proposed ISIS system which was successfully evaluated with real learners (Drachsler, Hummel, et al. 2009). The same group developed ReMashed (Drachsler et al. 2010), a mash-up environment that combines using hybrid techniques information coming from different Web2.0 services like Flickr¹, Delicious² or Slideshare³. Other proposals include the integration of case-based reasoning recommenders (Gomez-Albarran & Jimenez-Diaz 2009) or contextual recommendations (Santos 2008). Romero et al. go a little further, introducing a new abstraction level, meta-rules, empowering the rule adaptation process (Romero Zaldivar & Burgos 2010).

In recent years a considerable number of TEL platforms explored the application of Semantic Web (SeW) technologies which aim at adding semantic information to web contents to create an environment where software agents will be capable of doing tasks efficiently (Berners-Lee et al. 2001; Breslin et al. 2011). Semantic Web techniques have been applied to RSs (Ziegler 2004) and in the last years several efforts have been done in the application of these technologies to TEL (Stojanovic et al. 2001; Devedzic 2006; Bittencourt et al. 2009; Sánchez-Vera et al. 2012; Montuschi et al. 2015) primarily oriented to the provision of feedback in online assessment environments for improving the teaching-learning process. Recommender Systems based on Semantic Web technologies can be considered as a special type of knowledge-based recommender systems.

According to (Lu et al. 2015) concept drift techniques (Gama et al. 2014) should be introduced into recommender systems to model users' preference drift and improve the recommendation performance in a fast changing environment (Koren 2009) like eLearning platforms. They argue that the distribution of data, such as the users' behaviours towards their interests/preferences and the functionality of some items keeps changing, and using the outdated data to predict users' current preferences will result in poor performance.

II.2.3. Summary classification

Table II.2-1 provides a summary classification of all RSs presented in previous sections with the addition of a selected set of examples from the overwhelming number of recommender systems available in the state of the art. It includes both TEL specific recommender systems and the systems focused on other application domains. The classification followed the taxonomy proposed by Burke (Burke 2007) combined with the data-mining techniques' classification provided by Hee Park et al. (Park et al. 2012).

¹ <https://www.flickr.com/>

² <https://delicious.com/>

³ <http://www.slideshare.net/>

Table II.2-1. Summary Classification of Recommender Systems based on (Burke 2007; Park et al. 2012)

Characteristics		Application Domain	
		Other applications	Informal learning TEL platforms
CB	Unsorted	(Jennings & Higuchi 1993; Lang 1995; Pazzani et al. 1996; Chen & Sycara 1998)	CoFind (Dron et al. 2000), (Huang et al. 2009)
	AR	(Cho et al. 2002)	(Zaiane 2002; Biletskiy et al. 2009)
	CI	Uns.	(Xue et al. 2005; Shinde & Kulkarni 2012; Ghazanfar & Prügel-Bennett 2014)
		kM	(Berry & Linoff 2004; KIM & AHN 2008)
		SOM	(Lihua et al. 2005)
	DT	(Kim et al. 2002)	
	NN	(Ibnkahla 2000), AIMED (Hsu et al. 2007)	
CF	Unsorted	(Goldberg et al. 1992; Resnick et al. 1994; Hill et al. 1995; Shardanand & Maes 1995)	Altered Vista (Walker et al. 2004), CYCLADES (Avancini & Straccia 2005)
	kNN	(H. K. Kim et al. 2009; Pappas & Popescu-Belis 2015)	
	LA	PageRank (Cai et al. 2004)	
DM		(Krulwich 1997; Pazzani 1999)	
KB	Unsorted	(Burke 1999; Burke 2000; Bridge et al. 2005)	PLRS (Lu 2004), QSIA (Rafaeli et al. 2005), AHA! (Romero et al. 2009), FIRT (Chen et al. 2004; Chen & Duh 2008), (Santos et al. 2014)
	SeW	(Ziegler 2004; Rosa et al. 2015)	(Stojanovic et al. 2001; Devedzic 2006; Bittencourt et al. 2009; Sánchez-Vera et al. 2012; Montuschi et al. 2015)
CoB		Improve CF	CourseAgent (Farzan & Brusilovsky 2006)
		Not Improve CF	
Hyb	W	(Claypool et al. 1999; Mobasher et al. 2004; Chen & Chen 2015; Zhang & Min 2016)	
	M	(Smyth & Cotter 2000)	(Gomez-Albarran & Jimenez-Diaz 2009)
	S	Internal Criteria	(Santos 2008)
		External Criteria	
	FC	(Basu et al. 1998; Mooney & Roy 1999; Christakou & Stafylopatis 2005; Shambour & Lu 2011; Nilashi et al. 2014)	(Manouselis & Costopoulou 2007; Wu et al. 2015), IWT (Capuano et al. 2014)
	FA	(Melville et al. 2002; Wilson et al. 2003)	(Tang & McCalla 2005; Porcel et al. 2009), RACOFI (Lemire et al. 2005), ISIS (Drachsler, Hummel, et al. 2009), ReMashed (Drachsler et al. 2010)
	Ca	(McSherry 2002; Kai Yu, Volker Tresp 2004)	
	ML	(Balabanović 1997; Pazzani 1999)	(Romero Zaldivar & Burgos 2010)
	R	(Malhotra 1999)	
	OHM	(Cantador et al. 2008)	(Shen & Shen 2004; Cobos et al. 2013)

II.2.4. Evaluation of Recommender Systems

The design of recommender systems (RSs) is highly dependent on applications' domain, and the algorithm approach should be selected considering the specific characteristics of the domain in which it is going to be used (Ricci et al. 2011). Consequently, evaluation is also a major challenge for RSs, presenting a lack of standardization or definition of accepted testbeds.

Evaluation is required at different stages of the systems' life cycle for various purposes (Adomavicius et al. 2005; Burke 2007). Firstly, at design phase, an off-line evaluation should be conducted comparing different algorithms on existing public benchmark data (if available) or collected data, and following known experiment to ensure reliable results (Bailey 2008). Later, when the system has been launched, an on-line evaluation is needed. Burke (Burke 2007) proposes an evaluation based on the accuracy of an algorithm's prediction and related measures, while Shani and Gunawardana (Shani & Gunawardana 2011) go further in their discussion and study of evaluation for RSs and focus their approach on property-directed evaluation, providing a large set of properties that are relevant to the system success, along with the experiments and measures that best suit the evaluation of these properties. Nevertheless, , to our knowledge, there's no testbed for the evaluation of RSs.

II.2.4.1. Evaluation of TEL-RSs

Until today, evaluation of recommender systems (RSs) gives emphasis to rather 'technical' measures coming from information retrieval research, although the importance of including user-related evaluation has been highlighted (Herlocker et al. 2004; McNee 2006). In TEL-RSs evaluation becomes an even more demanding task, considering the particularities of the educational contexts (Manouselis et al. 2011). In the world of consumer RSs, several datasets are available used as benchmark to test and compare the performance of different recommender systems and algorithms. However, these datasets are domain specific and their application to other domains is not usually feasible, although some authors (Verbert et al. 2011) explored some available datasets that capture learners' interactions with the system and used them to evaluate and to compare the performance of different recommendation algorithms.

On the other hand, to evaluate pedagogy-driven RSs, no standardised data are available. Manouselis et al. (Manouselis et al. 2011) described some guidelines to perform the evaluation of TEL recommender systems but, to our knowledge, there is no testbed, standard or majority-accepted benchmark that can be used for the evaluation of TEL systems. This lack of dataset results in a considerable number of recommender systems in literature with no evaluation, as pointed by (Erdt et al. 2015) in one of the latest and biggest surveys concerning evaluation in recommender systems for Technology Enhanced Learning (42% of the publications surveyed by them had no evaluation whatsoever).

Erdt et al. studied a total of 235 articles from major conferences, workshops, journals, and books where relevant work have been published between 2000 and 2014. In the survey the diverse evaluation methods that have been applied to evaluate TEL-RSs are investigated and classified. They conclude that, although there is a growing awareness in the research community of the

necessity for more elaborate evaluations, there is still substantial potential for further improvements. Additionally, they highlight some of the possible future trends.

II.2.1. Commercial systems

Recommender system research is being conducted with a strong emphasis on practice and commercial applications. The recommender algorithm used is not always published or reported but the computation is mostly based on the summation or average of ratings. One of the most famous systems is **eBay's feedback forum**⁴. Both buyers and sellers rate each other as positive, negative or neutral after transactions. The ratings are computed summing the positive ones and deducting the negatives over three different time windows. Empirical studies (Resnick & Zeckhauser 2002) observed that ratings are surprisingly positive and there's a high correlation between buyer and seller ratings (Resnick et al. 2006). This is problematic if the goal of the system was obtaining honest and fair ratings but it seems to have a strong positive impact on eBay as a marketplace, as it can be seen by its success.

Another well-known commercial system is the one implemented by **Amazon**⁵. Amazon's users rate items from 1 to 5 and add prose text to their review. The average of all ratings is used as rating score for the items and text reviews from users is shown as additional information. Users can also vote on reviews as being helpful or not helpful and the number of helpful votes versus total votes is displayed with each review. As a function of the number of helpful votes received (as well as other parameters not publicly revealed), reviewers are given a rank and a status. Amazon has a system of *Favourite People* where members can choose other members as favourite reviewers.

The web page ranking system of **Google**⁶ could be also considered as a recommender system. Google's search engine is based on the PageRank algorithm (Page et al. 1998) with some not public modifications. Basically, if there are many hyperlinks to a web page its ranking is increased whereas many hyperlinks from a web page contribute to decrease the ranking for that web page.

As a conclusion, the importance of recommender systems can be demonstrated by the increasing number of sites that use any of the techniques mentioned in section II.2.1: **YouTube**⁷, **Yahoo**⁸, **TripAdvisor**⁹, **IMDb**¹⁰, etc. Moreover, many media companies are developing and deploying RSs as part of their services, such as **Netflix**¹¹, that created a million dollar prize for the team that first succeeded in improving substantially the performance of its recommender system (Koren et al. 2009).

⁴ <http://www.ebay.com>

⁵ <http://www.amazon.com>

⁶ <http://www.google.com>

⁷ <http://www.youtube.com>

⁸ <http://www.yahoo.com>

⁹ <http://www.tripadvisor.com>

¹⁰ <http://www.imdb.com>

¹¹ <http://www.netflix.com>

II.3. Trust and Reputation (T&R) systems

Trust and Reputation (T&R) systems represent a significant trend in decision support for Internet mediated service provision (Jøsang et al. 2007). Generally stated, the underlying goal of all T&R systems is to predict the trustworthiness and proficiency of peers in future actions. This prediction is based on information gathered from their past behaviour in the environment, their peers' view towards their history (Ruohomaa et al. 2007), and from the system. The idea behind is that service consumers can be confident that they will get the service they expect as long as they trust the provider, without having the opportunity of trying the service in advance.

People have clues to assess trustworthiness through traditional styles of interaction and physical encounters that are not present in computer mediated communication. T&R systems aim at finding online substitutes for the traditional clues used to assess trustworthiness, and to take advantage of information technologies (IT) and the Internet to collect efficiently additional information on the system from which reputation measures can be derived to support decision making. To establish a reputed street presence takes time and investment, in sharp contrast to the relative simplicity and low cost of establishing a good looking Internet presence which gives little evidence about the solidity of the organization behind it (Jøsang et al. 2007).

There is a rapidly growing literature on the theory and applications of T&R systems, nonetheless there also seems to be a lack of coherence with different definitions. According to Gambetta's definition (Gambetta 2000) **trust** is the subjective probability by which an individual expects that another individual performs a given action on which its welfare depends. This definition is unsatisfactory, as pointed by Falcone and Castelfranchi (Castelfranchi & Falcone 2001), who recognise that having high trust in a person could not be enough to decide to depend on that person. MacKnight and Chervany (Mcknight & Chervany 1996) extend the definition of trust by adding the concept of risk: trust gives a feeling of relative security, even though negative consequences are possible. Thusly, T&R systems can be considered *soft security* mechanisms, in contrast to *hard security* mechanisms like authentication and access control (Resnick & Zeckhauser 2002), which provide social control mechanisms reducing risk likelihood. However, only a few computational trust models explicitly take risk into account (Manchala 1998; Jøsang & Lo Presti 2004).

The concept of reputation is closely linked to the concept of trust, but there is a clear and important difference. Trust is a personal and subjective process based on various factors due to direct interactions among partners. On the other hand, **reputation** can be considered as a collective measure of trustworthiness based on the referrals from members in a community and derive from the underlying social network which is globally visible to all members of the network (Marsden & Lin 1982). Therefore, reputation can relate to a group or to an individual: individuals belonging a given group will inherit reputation based on group's reputation, as showed in Tadelis' study (Tadelis 2003).

Usually trust and reputation are gathered together to create systems capable of providing partners with additional information that can be used by them to decide whether or not to interact with other elements of the system, combining both individual and social perspectives (Huynh et al. 2006).

II.3.1. Classification

In the last years many authors have published surveys on this area (Grandison & Sloman 2000; Ramchurn et al. 2005; Artz & Gil 2007; Jøsang et al. 2007; Noorian & Ulieru 2010; Wahab et al. 2015) that tried to classify T&R systems, attempting to reduce confusion around the terminology used to describe trust and reputation systems. Classifying T&R systems is not an easy task. As pointed by Jøsang et al. (Jøsang et al. 2007) there is a lack of coherence in T&R systems' literature and terminology confusion with authors proposing new systems from scratch, without trying to extend and enhance previous proposals. Furthermore, T&R systems are context sensitive: their design has been dependent on the target domain and the related specific requirements, hindering the effort of creating a global and/or standard classification.

One of the most generic efforts on classifying T&R systems is the work of Noorian and Ulieru (Noorian & Ulieru 2010), complete not in the number of systems in the study but in the proposed feature topology. They propose a comparison framework that targets the distinctive aspects of T&R systems mapping them against a variety of features along several dimensions. They apply the proposed framework for the comparison of six renowned and reputed, it bears repeating, T&R systems. The dimensions are classified in two main categories: *hard features* and *soft features*. This section follows their classification including additional concepts, dimensions and examples of other systems in literature to complement the classification.

Table II.3-1 presents a set of selected examples of some of the most representative systems from the overwhelming number available in the literature. Table II.3-1 gathers systems, from different application domains, classified according to the categories defined by Noorian et al. In the table, 'Y' means that the system successfully fulfils the specific feature, and 'P' expresses that the system partially fulfils the feature. The field is left empty if the system does not present the feature or if no information was found to support or discard the inclusion in the category.

II.3.1.1. Hard features

Hard features represent the underlying characteristics which are fundamental to establish any T&R system, and includes the following dimensions:

- Rating Approaches (HF1). Two different rating systems. **Single-criterion/binary (HF1a)** where participants rate the overall interaction with two levels (satisfactory or unsatisfactory). And **multiple-criterion (HF1b)** where participants rate different aspects of the interaction and not just the overall impression.
- Witness Locating Approaches (HF2). This dimension refers to the architecture of the T&R system and three different structures can be distinguished. **Centralised (HF2a)**, where ratings are stored in a central authority/element where reputation is also calculated. **Distributed (HF2b)**, where each participant records its own ratings and generates the reputation score. And **hybrid (HF2c)** structure combines both the benefits of centralized and distributed approaches.
- Reputation Computation Engines (HF3). There is a wide variety of mechanisms to calculate reputation scores:

- **Deterministic approach (HF3a)**, trust values are calculated from handcrafted formulae to yield the desired results obtaining systems with great flexibility but sometimes very conditioned by the field of application (the formulae are usually designed for a specific domain and its portability depends on the similarities between domains).
- **Simple summation/average (HF3b)**, sum of positive ratings minus negative ratings or computing the reputation score as the average (weighted or not) of all ratings like many recommender systems.
- **Belief models (HF3c)**, Dempster-shafter theory (DST) (Shafer 1976) is an extension to probability theory with the advantage of being able to model uncertainty and it has been used to represent reputation scores by some authors where different opinions of several outcomes are evaluated and combined to assess reputation. Other authors present a simpler model where users evaluate a binary outcome, agent A is trustworthy or it is not trustworthy.
- **Fuzzy models (HF3d)**, trust and reputation can be represented as linguistically fuzzy concepts, where membership functions describe to what degree an agent can be trusted and Fuzzy Logic is used to determine reputation scores.
- **Bayesian approach (HF3e)**, reputation scores are calculated by statistical updating of beta probability density functions (PDF), the *a posteriori* reputation score is computed by combining the *a priori* reputation score with the new rating, and the shape parameters of the PDF are usually the number of positive and negative interactions.
- Information sources (HF4). T&R systems gather information from different sources to guarantee the provision of reputation scores in any condition. Some of the possible sources are **direct experiences (HF4a)** or **witness observations (HF4b)**, used by the vast majority of T&R systems.
Furthermore, when the above sources are not available some systems consider other sources of information. The first source is the **certified reputation (HF4c)** where target agents release information indicating their former interaction results. Other systems use **role-base trust (HF4d)**, where agents have predefined reputation scores associated to their role and relationships in the system, normally assessed through *sociograms*. Finally, only a few group of systems consider **quality indicators (HF4e)**, usually basic metrics related to the level of tasks' completion.

II.3.1.2. Soft features

Soft features embrace the attributes and traits which help to enhance the performance of the system and quality of outcomes, and includes the following dimensions:

- Context Diversity Checking (SF1). Reputation scores are always associated to the specific context in which the transaction among entities takes place; thusly, entities should **check the similarities between the context (SF1a)** of the present transaction and previous one to assess the reliability of reputation assessment received. On the other hand, some T&R systems are applied to fields where the **context is time invariable (SF1b)** and, consequently, there's no need to use context information.

- Adaptability (SF2). T&R systems should implement techniques to deal with the dynamism of the environment in which they operate. Especially delicate is **bootstrapping (SF2a)** problem or how newcomers' reputation is assessed. Bootstrapping helps hindering malicious entities from changing their identity and re-enter to the system but it may lead to the exclusion of newcomers with no acquaintance with the environment and low initial reputation values (Huynh 2006). Furthermore, the **dynamism of environments (SF2b)** may also force T&R systems to deal with unanticipated events and they must be prepared to anticipate those situations.
- Reliability and Honesty Measurement (SF3). Under this category, Noorian and Ulieru (Noorian & Ulieru 2010) describe a set of features implemented by T&R systems to alleviate the effect of dishonest information providers and spurious ratings.
 - **Behaviours' trend and volatility detection (SF3a)**, allows predicting future behaviours or detecting fraudulent activity.
 - **Proficiency in giving correct ratings (SF3b)**: ratings could be honest but, however, they could be inefficient due to inexperience or insufficient interactions. This feature is not easy to address and it is not typically implemented by trust and reputation systems.
 - **Ballot box stuffing (SF3c)**: also known as double counting or correlated evidence problem, where coalition of entities cast negative votes against the outsiders and positive votes in favour of their allies (Jøsang et al. 2007) deriving in the exclusion of entities. Most T&R systems implement techniques to avoid it by defining maximum number of eligible votes like or restricting participants to release only their personal experience.
 - **Reliability in giving honest ratings (SF3d)** refers to the fact that entities could give misleading ratings due to many factors like jealousy, competitive, etc., even though they could be considered trustworthy entities in the way they always fulfil their commitments. Therefore, most T&R systems implement methods to measure the reliability of entities when they assess reputation, that is, entities' credibility. The discount of unfair ratings could be done in two ways. The first method is the *endogenous discounting (SF3d1)*, where the exclusion criteria is based on analysing and comparing the rating values themselves assuming that unfair ratings can be recognised by their statistical properties. The second method is *exogenous discounting (SF3d2)*, and the externally determined reputation is used to determine the weight given to ratings, under the assumption that raters with low reputation are likely to give unfair ratings and vice versa.
 - **Detection of discrimination (SF3e)**: determining if an entity or a group of them are discriminating other entities is a tough question and, as far as we know, there's no T&R system that successfully addresses this situation, although some systems partially solve it.
- Low incentive for providing rating. Lack of incentives for providing ratings is a general problem in T&R systems and might require specific incentive mechanisms. Some systems propose **incentives for providing rates (SF4a)**, but the majority of systems rely on **willingness of entities on providing rates (SF4b)**.

Table II.3-1. Trust and Reputation systems' classification, modified from (Noorian & Ulieru 2010). In grey the six T&R systems of Noorian et al. study.

T&R system	T&R Feature																										
	Hard Features															Soft Features											
	HF1		HF2			HF3					HF4					SF1		SF2		SF3					SF4		
	a	b	a	b	c	a	b	c	d	e	a	b	c	d	e	a	b	a	b	a	b	c	d		e	a	b
																						1	2				
BRS (Jøsang & Ismail 2002)	Y		Y	Y	Y					Y	Y	Y			P		Y	Y	Y	P		P		Y	P		Y
ReGreT (Sabater & Sierra 2002a)		Y	Y	Y	Y				Y		Y	Y		Y		Y		Y				Y		Y			Y
(Yu & Singh 2003)			Y					Y			Y	Y			P	P	Y							Y			Y
PeerTrust (Li & Ling 2004)	Y		Y			Y					Y	Y			P		Y	Y	Y	P		Y		Y	P		Y
FIRE (Huynh et al. 2004; Huynh et al. 2006)		Y	Y			Y					Y	Y	Y	Y				Y	Y	P	P	Y		Y	P		Y
TRAVOS (Teacy et al. 2006)	Y		Y	Y	Y					Y	Y	Y			P		Y					Y		Y	P		Y
(Manchala 1998)									Y		Y	Y															Y
(Zacharia et al. 1999)											Y	Y						Y									Y
(Dellarocas 2000)											Y	Y										Y					Y
(Aberer et al. 2001)				Y							Y	Y															Y
Subjective (Jøsang 2001)								Y			Y	Y															Y
(Chen & Singh 2001)											Y	Y										Y					Y
(Mui et al. 2002)	Y									Y	Y	Y			P												Y
(Cornelli et al. 2002)				Y							Y	Y											Y				Y

T&R system	T&R Feature																										
	Hard Features														Soft Features												
	HF1		HF2			HF3					HF4					SF1		SF2		SF3					SF4		
	a	b	a	b	c	a	b	c	d	e	a	b	c	d	e	a	b	a	b	a	b	c	d		e	a	b
(Damiani et al. 2002)				Y							Y	Y											1	2			Y
(Fahrenheit et al. 2002)				Y							Y	Y															Y
(Resnick & Zeckhauser 2002)							Y				Y	Y															Y
(Ekström M A 2002)											Y	Y												Y			Y
(Miller et al. 2002)											Y	Y														Y ¹²	
(Gupta et al. 2003)				Y							Y	Y															Y
EigenTrust (Kamvar et al. 2003)	Y			Y							Y	Y															Y
(Jurca & Faltings 2003)											Y	Y														Y ¹²	
(Whitby et al. 2004)									Y		Y	Y											Y				Y
(Buegger & Le Boudec 2004)											Y	Y												Y			Y
(Maximilien & Singh 2005)											Y	Y			P	Y			Y					Y			Y
(Chang et al. 2006)						Y					Y	Y				Y											Y
(Sherchan et al. 2006)								Y			Y	Y			P	Y							Y	Y			Y

¹² Scheme based on financial rewards or payments

T&R system	T&R Feature																										
	Hard Features															Soft Features											
	HF1		HF2			HF3					HF4					SF1		SF2		SF3						SF4	
	a	b	a	b	c	a	b	c	d	e	a	b	c	d	e	a	b	a	b	a	b	c	d		e	a	b
																						1	2				
(Malik & Bouguettaya 2008)											Y	Y			P	Y			Y			Y					Y
(Malik & Bouguettaya 2009)						Y									P	Y			Y			Y					Y
(Paradesi et al. 2009)		Y				Y					Y	Y			P	Y											Y
(Yahyaoui 2010)		Y									Y	Y			P												Y
(Thurrow & Delano 2010)								Y							P	Y								Y			Y
(Nguyen et al. 2010)						Y					Y	Y			P	Y						Y	Y				Y
(Hang et al. 2012)		Y						Y							P	Y								Y			Y
(Yahyaoui 2012)		Y								Y	Y	Y			P									Y			Y
(Mehdi et al. 2013)		Y								Y					P	Y											Y
(Vavilis et al. 2014)											Y	Y			P									Y			Y
(Chae et al. 2015)	Y					Y				Y	Y					Y			Y	Y	Y			Y			Y
(Gupta & Singh 2015)	Y			Y		Y					Y	Y			P		Y		Y			Y			P		Y

II.3.2. Evaluation of T&R systems

Evaluation is one of the main challenges in the research on T&R systems. As Kerr et al. pointed (Kerr & Cohen 2010): along with the multitude of T&R systems proposed on literature have come a similar large number of methods to evaluate the proposals. This diversity on methods with different features and performances make very difficult to design an evaluation process that cover all the possibilities. Nevertheless, in the last years a significant effort on this area has been invested. In the year 2000, Dingleline et al. (Dingleline et al. 2000) proposed a set of criteria for judging the quality and soundness of T&R systems:

- 1) Accuracy for long-term performance. The system must reflect the confidence of a given score.
- 2) Weighting toward current behaviour. The system must recognise and reflect recent trends in entity performance.
- 3) Robustness against attacks. The system should resist attempts of entities to manipulate reputation scores.
- 4) Smoothness. Adding any single rating should not influence the score significantly.

Taking these criteria into consideration some approximations to an evaluation testbed were used to evaluate T&R systems such as the *Prisoner's Dilemma* of game theory, that was used with some changes by Marsh (Marsh 1994) and Schillo et al. (Schillo et al. 2000); and the *Sporas test model*, created by Zacharia et al. (Zacharia et al. 1999) to test Sporas model and later was used by Sabater and Sierra (Sabater & Sierra 2001a) to compare ReGreT with Sporas and Amazon Auctions' mechanism. Carbo et al. (Carbo et al. 2002; Carbo et al. 2003) defined an extended version of these experiments for the comparison of their model AFRAS with Yu and Singh's model (Yu et al. 2002), two online reputation mechanisms (eBay⁴ and Bizrate¹³) and the two aforementioned reputation models, Sporas and ReGreT. Another notable approximation is the freely configurable simulation framework proposed by Schlosser et al. (Fullam, Klos, et al. 2005). The framework provides some objective metrics to evaluate trust problems; however it is not easily extended to compare multiple algorithms.

These approximations cannot be considered real testbeds and Sabater (Jordi Sabater 2004b) goes a little further and proposes a modular and open testbed, defining a set of parameters that can be considered to build a general testbed for T&R systems. The Agent Reputation and Trust (ART) Testbed initiative¹⁴ picks up the baton of Sabater (Sabater 2004b) and Fullam and Barber (Fullam, K.; Barber 2004) to create the first full operative and accepted testbed for the evaluation of T&R systems, which full architecture can be found in (Fullam, Voss, et al. 2005).

ART Testbed can operate in two modes: competition mode and experimentation mode. The framework defines the art appraisal domain in which both modes run (Fullam, Klos, et al. 2005). In **competition mode** each participant/researcher controls an agent which competes against

¹³ <http://www.bizrate.com>

¹⁴ ART Testbed Team. <http://www.lips.utexas.edu/~kfullam/competition/>

other agents of the systems in different game sessions. The results of the games are used to compare different researchers' strategies. In each session 'dummy agents', not controlled by other researchers, may be included. In **experimentation mode**, researchers download the testbed and use it out of a competition. ART environment allows easily-repeatable experimentation and, thereby, results may be compared among researchers.

Since its creation, ART Testbed has been extensively used to evaluate T&R systems. However, some authors state that ART Testbed has not suited the evaluation of their work (Harbers et al. 2007; Hang et al. 2008). Kerr et al. (Kerr & Cohen 2010) state that, although ART is well-designed for evaluating agents in a competitive manner using a small social trust scenario, it is not appropriate for broad use in the experimental evaluation of T&R systems due to a number of design choices. They propose a new Testbed, TREET: Trust and Reputation Experimentation and Evaluation Testbed, designed specifically to support generally-purpose experimentation and evaluation, based on the platform used previously by them to defeat a number of noteworthy T&R system of the literature (Kerr & Cohen 2009).

TREET models a centralized 'advertised-price' marketplace scenario where sellers offer goods for sale and buyers choose whether to buy or not and from whom, deeply described in (Kerr & Cohen 2009) along with the full architecture of TREET. A simulation is run with rounds of interactions among buyers and sellers in the marketplace and the simulation's parameters can be customized. TREET platform is released as open source software and can be used by researchers to try defeating existing T&R technologies. TREET allows a great degree of flexibility of experimentation while generating meaningful benchmarks and comparisons with a set of useful tool to generate graphical representation of the performance of T&R systems in comparison.

Also oriented to online marketplaces is Euphemus (Adamopoulou & Symeonidis 2014) an adaptable, customizable, multivariate agent testbed for the simulation of open online markets and the study of various factors affecting the quality of the service consumed. Euphemus was evaluated by testing the performance of four known trust and reputation models: ReGreT (Sabater & Sierra 2002a), BRS (Jøsang & Ismail 2002), FIRE (Huynh et al. 2006) and (Yu & Singh 2003) showing good results. However promising, both Euphemus and the aforementioned TREET testbeds are not suitable for all T&R systems but only systems that can function in marketplace environments.

Another approach is the one presented by Hermoso et al. (Hermoso & Centeno 2007). TOAST emerges with the intention of providing a tool to evaluate trust and reputation mechanisms that take advantage of Electronic Institutions and Virtual Organizations concepts, used in open agent societies to impose a certain structure. In these organizational environments (Hermoso et al. 2006) the evolution of the organization has also to fulfil its goals and not only the agents. TOAST provides users with different views to monitor simulation experiments at a macro level (organization) and micro level (agents). Although supposing a clear advance to solve a problem not well addressed before: the analysis of trust models and behaviour patterns at the level of group of agents, this specific virtual organization approach makes also this testbed domain dependant.

One of the latest efforts is Alpha Testbed¹⁵ (ATB) (Jelenc et al. 2013). This testbed proposes a novel approach over the hypothesis that the choice of decision making mechanism used for the evaluation of trust and reputation models is relevant. They evaluated five models either with or without decision making mechanism: BRS (Jøsang & Ismail 2002), TRAVOS (Teacy et al. 2006), EigenTrust (Kamvar et al. 2003), (Singh & Sycara 2004) and (Abdul-Rahman & Hailes 2000); confirming the initial hypothesis. The influence of the decision making process is not addressed by any other available test and the possibility of considering it or not in the evaluation of the trust and reputation model makes ATB a powerful testbed. Nevertheless, it focuses on peer-to-peer systems, where users are agents with simple behaviours. Thus, the validity of this testbed is limited to certain trust and reputation models of specific application domains.

II.4. Combining T&R with RS systems

In opposition to recommender systems (RSs), ratings on trust and reputation (T&R) systems are not typically used to filter information but to judge the performance of a user, a transaction or the quality of service (QoS) consistently. These ratings are produced and consumed by all the members of the system and, thusly, are not addressed only to a subset of members sharing similar tastes. The ultimate purpose of the ratings is to sanction poor service providers, with the aim of giving an incentive for them and the term “collaborative sanctioning” (CS) has been used to refer to these systems (Mui et al. 2001).

RSs and T&R systems can be used together to combine the advantages of both approximations: RSs will be used to filter resources according to members’ tastes and create recommendations and T&R systems will provide soft security (Resnick & Zeckhauser 2002) filtering the recommendations generated by the RSs system according to the quality of the resources, the source of these resources and other system parameters; protecting against information providers that can act deceitfully or by providing poor, false or misleading information. Damiani et al. (Damiani et al. 2002) proposed a combined approach to separate between provider reputation and resource reputation in P2P networks.

In contrast, Massa et al. (Massa & Avesani 2004; Massa & Bhattacharjee 2004; Massa & Avesani 2007) propose to extend RSs with trust-awareness, allowing users to express their ‘*web of trust*’ (Guha 2003) and propagating trust throughout the global trust network, thusly, creating Trust-aware Recommender Systems which allow coping with the problem of new users (cold start), data sparseness and detection of malicious or unreliable users. Massa et al., as well as other authors (Levien 2003; Golbeck et al. 2003; Ziegler & Lausen 2004) proposed using Semantic Web technologies to calculate and propagate trust, on contrast to the application of specific T&R methodologies as the ones mentioned section II.2.4.

Anwar et al. (Anwar & Greer 2011) discussed the application of trust-aware recommender systems to eLearning environments. The creation of trust-based enhanced recommender systems, as a sort of combination of RSs with trust and reputation models, seems to be the most

¹⁵ <http://atb.fri.uni-lj.si/>

accepted solution, according to the increasing number of research work in this area in the last years (Kim & Phalak 2012; Lee & Zhu 2012; Dwivedi & Bharadwaj 2013; Eirinaki et al. 2014; Shripad & Vaidya 2014; Oh et al. 2015; Alahmadi & Zeng 2015)

II.5. T&R and RS systems in HealthCare

The application of recommender systems (RSs) to healthcare has been limited, compared to other fields like e-commerce, and relevant review studies are rare in the field (Sezgin & Ozkan 2013). RSs have already been employed in health services for dietary and activity assistance purposes to patients (Sami et al. 2008; J.-H. Kim et al. 2009), or to help physicians in the medical imaging diagnostic decision (Monteiro et al. 2015). But their major focus has been the recommendation to patients of the best doctor to treat them. Katzenbeisser and Petkovic (Katzenbeisser & Petkovic 2008) proposed a medical recommendation system where relevant patient's information (symptoms, diseases, medical history, etc.) is compared with a list of doctors and the best matching expertise is given as recommendation to the patient. Patients' have the security that the recommended doctor has the skills to treat their problem but they have no guarantee of being the best option.

The recommendation of optimal doctors has been also addressed by commercial systems. Online ratings for medical doctors are already in use in some countries like USA: Healthcare Reviews¹⁶, Vimo¹⁷, BetterDoctor¹⁸ and RateMDs¹⁹; or Spain: Doctoralia²⁰. These websites typically allow people to rate doctors and to provide written review anonymously, and provided ratings are aggregated to create a doctors' ranking.

Patients using these systems have to confront to the problem of unknown healthcare providers and other security issues (Valarmathi et al. 2011). Over the last decade, as Internet, wireless technologies and mobile devices are becoming a reality in our lives; concepts like pervasive or ubiquitous healthcare are gaining presence in the research areas of technology applied to health. Pervasive healthcare enables patients to access healthcare services anywhere, anytime ubiquitously (Dearden et al. 2010).

Yuan et al. (Yuan et al. 2007) discussed the challenges and problems derived from the use of pervasive systems concluding the need of trust and privacy assessments in such environments. This need relates to the information asymmetries arising from the specialist nature of medical knowledge, the uncertainty regarding the competence of the medical service provider that assists the patient (Rowe & Calnan 2006). Furthermore, this trust and privacy need has an indirect influence on health outcomes (Safran et al. 1998).

¹⁶ <http://www.healcarereviews.com>

¹⁷ <http://vimo.com/reports>

¹⁸ <https://betterdoctor.com/>

¹⁹ <http://www.ratemds.com>

²⁰ <http://www.doctoralia.es>

The framework presented by Hoens et al. (Hoens et al. 2010) attempts to match patients with the optimal doctor for their relevant conditions while preserving the privacy of patient data: the reliability of data is maintained mitigating system's abuse from dishonest users. Their recommendation system consists in a simple average aggregation of ratings but the novelty of their proposal is the design of two different architectures for the platform as well as the communication protocols, required to maintain the privacy. A similar approach was followed by Valarmathi et al. (Valarmathi et al. 2011). Instead of using recommender systems, they combine reputation and credential models to give secure and reliable information about practitioners to the patients using the system.

Trust and reputation models have focused traditionally in the provision of soft security in ubiquitous healthcare systems by assigning rights to the system (Yuan et al. 2008); using cloud theory to create trust mechanisms (Athanasiou et al. 2015); or creating a whole trust assessment architecture (Bahtiyar & Çağlayan 2014). In recent years the efforts seem to concentrate in the provision of trusted identity management (IdM) solutions for consumer-centric approaches (Au & Croll 2008; Dolera Tormo et al. 2013), to provide security in the interactions between users of ubiquitous healthcare systems: patients and doctors.

In contrast with the solutions presented above where the focus is in the *interaction plane*, understood as the plane where patients select the best or the most trustworthy doctor to interact with, other authors' research is oriented to filtering medical information and selection of the best data or resources. In this area, Sánchez-González et al. (Sánchez-González et al. 2011) proposed a TEL environment for the training of surgeons in Minimally Invasive Surgery (MIS) where a recommender system (Romero Zaldivar & Burgos 2010) is used to suggest video resources to users according to their interests. Recent works in the recommendation of relevant and trustworthy healthcare-related information focus on the application of social media and social data enhanced with semantic web technologies (Li & Zaman 2014; Zaman & Li 2014).

Also designed for the recommendation of resources but different objectives, the Hedaquin system (van Deursen et al. 2008) provides healthcare professionals with an indication of the quality of health data in Personal Health Records (PHR) using an extension of the Beta reputation system (Jøsang & Ismail 2002) to cover specific PHR requirements. To determine the quality of the data they consider three factors: the credentials of the health data supplier, the ratings for the health data supplier and metadata supplied by measuring devices.

Recommender systems (RSs) have been also applied to assess the quality of this Personal or Electronic Health Record (PHR/EHR) data, renamed as Health Recommender Systems (HRS) (Choon-oh Lee et al. 2008; Wiesner & Pfeifer 2010) Other authors (Alhaqbani & Fidge 2009) suggest using a time-variant trustworthiness model metadata attached to incoming medical data to assess the variation over time of the reputation of both the practitioner and the healthcare organization from which the data was acquired.

And finally some authors (Krishna et al. 2014) combine trust-based approaches with recommender systems to solve traditional recommender systems such as cold start and scalability by accommodating a trust factor in a collaborative filtering RS, reaping the efficiencies of the k-means++ algorithm, which provides the threshold rating for the cold start users.

II.6. Objective assessment of resources

The process of objective assessment can be modelled as a three-sided problem (Oropesa et al. 2011): 1) the application side²¹; 2) the technological side; and 3) the analytical side. The application side is domain-dependent and different metrics would be used for each domain and/or even for the different roles of individuals of a system. For instance clinical metrics might not be applied to an email agent. The technological side is related to the analysis of parameters that could be associated to quality of data (QoD) (Larburu Rubio et al. 2014) and/or quality of service (QoS) assessment like probability of response, delay, accuracy, data compression ratio, etc. Finally, the analytical side studies the use of statistical analysis, in order to ascertain whether or not automatic classification systems to perform objective assessment are viable. High level statistical analysis and machine learning techniques can be employed to infer knowledge and correlate metrics to specific evaluation values.

The work proposed in this PhD dissertation is built upon the strong conviction that an automatic objective value can be obtained in the majority of domains by at least one of the dimensions exposed by Oropesa et al. Some examples of specific domains are presented in section VI.5.3, when presenting the possible future works beyond this thesis. The reliability of those values will determine their applicability in the reputation assessment. However, the combination of objective assessments with users' opinions in TruQOs reduces the impact of poor evaluations.

Finally, to complement the automatic value, or in domains where no metric can be found, the objective assessment can be produced by a panel of experts. Providing that the panel is composed by 3rd party individuals, not involved in the resources' generation process, the evaluation values created by them can be considered as objective²². One example of this panel of experts are film, hotels or restaurant critics.

II.7. Agent-based simulation models

The main characteristics of software agents (Wooldridge 2002) are:

- **Autonomy:** Agents can operate without the direct intervention of any user or software.
- **Sociability or cooperation:** Agents interact with other agents to solve complex problems that cannot be solved by only one agent.
- **Proactivity:** Agents do not only react in response to external events (i.e. a remote method call) but they also exhibit a goal-directed behaviour and, where appropriate, are able to take initiative.
- **Reaction/Adaptability:** Agents should not be limited to answer only to direct interactions but they must also react and adapt themselves to changes in the environment.

²¹ This first dimension is a generalization of the dimension used by Oropesa et al. where it was named clinical side.

²² Or at least pseudo-objective, considering that all the resources will be evaluated with the same subjectivity.

- **Mobility:** Agents can change their location, moving from one environment to another one in the same or different machines.
- **Temporal continuity:** After their creation, Agents remain in an idle state until an event, a direct interaction or a change in their environment activates them and motivates them to perform any action. Once that action is finished they go back to their waiting state until another event comes.
- **Customizable:** To adapt themselves to the tasks, actions, behaviours, etc., they will carry out during their life cycle.
- **Intelligence:** This characteristic refers to the capacity of Agents to learn how to solve new problems using the experience gained in the past.

The users of open TEL systems share many of the characteristics listed before. Furthermore, some of them such as the proactivity, the adaptability or the temporal continuity are also of great utility to develop a simulator. The use of agent-based models (ABMs) for research and management is growing rapidly in a number of fields, especially for simulation. Some of the most commonly used ABM platforms follow the “framework and library” paradigm, providing a framework —a set of standard concepts for designing and describing ABMs— along with a library of software implementing the framework and providing simulation tools (Railsback et al. 2006).

Examples of this category are **Swarm**²³ (Nelson Minar, Roger Burkhart Y 1996), **Repast**²⁴ or **MASON**²⁵ (Luke et al. 2005). Railsback et al. (Railsback et al. 2006) performed a deep analysis of these ABMs platforms along with **NetLogo**²⁶. NetLogo is a high-level platform whose primary objective was to allow students down to the elementary level to build and learn from simple ABMs. Their main conclusion was that software development remains a significant barrier to the use of ABMs, especially for the types of modelling and analysis needed to address substantial, real-world scientific problems.

II.8. Discussion

The main conclusion obtained from the study of the state of the art is that, although in the last years the application of recommender systems and trust and reputation models to healthcare has been increased, there is a lot of work to conduct in this field. Additionally, it seems that, to our knowledge, proposed solutions focus either on the recommendation of resources or the recommendation of users (doctors). Thus, any solution combining both types of recommendation would be well received, especially by patients, the main end-users of ubiquitous healthcare systems.

²³ <http://www.swarm.org>

²⁴ <http://repast.sourceforge.net>

²⁵ <http://cs.gmu.edu/~eclab/projects/mason/>

²⁶ <http://ccl.northwestern.edu/netlogo/>

In this PhD dissertation a new trust and reputation model is proposed, exploring a new source of information to assess the final reputation value: the objective evaluation of digital resources. To our knowledge, no other model has effectively included this reputation dimension in the assessment process. The model combines these objective values with the traditional subjective personal opinions from past experiences to provide trusted recommendations in open environments focused on knowledge exchange.

The selected application domain for the model is Technology Enhanced Learning (TEL) platforms for ubiquitous healthcare. TEL systems are very rich scenarios, with different types of interactions between users depending on the type of learning to which the platform is oriented: formal or informal learning. The type of learning also conditions user behaviours at sharing digital resources. Traditionally, the recommendation on these systems is performed by means of recommender systems, usually consisting in a mere filtering the system. These mechanisms do not provide the required trustworthy atmosphere. Furthermore, they are not designed for the recommendation of users.

The trust and reputation model presented in this thesis aims at providing the needed security, trustworthiness and reliability required in the envisioned TEL platforms for ubiquitous healthcare through the recommendation of both resources and users using the assessment of reputation as trust measure.

III. HYPOTHESES AND OBJECTIVES

The hypotheses and objectives of the present PhD research work are detailed in this chapter.

III.1. Hypotheses

The hypotheses on which this PhD thesis is built are:

H1.- A trust and reputation model combining evaluations of resources with opinions from system's users helps mitigating subjective behaviour punishments while enriching evaluation numbers with additional pedagogical information. This combination improves traditional approaches that use only personal appraisals or quality ratings.

Traditionally, recommendation systems are built upon reputations assessed using users' opinions about a resource or about another users. Users' opinions are usually affected to some extent by variables other than the intrinsic quality of the resource that it is being consumed or the excellence of the service they are receiving. Consumers' mood at the moment of consuming the resource or service or the existing relationship between them and the provider conditions the process of rating the interactions; punishing a good quality resource due to users' rivalry, for example. In quality-oriented societies, such as medical or academic, this situation is clearly unfair and the consequences of these misbehaviours are not acceptable.

To face these misbehaviours, objective assessments of resources may be used to provide the recommendations. However, unbiased evaluations are cold and have a lack of peripheral information. For example, a tutorial video clip could have a good technical quality and/or describe all the steps to perform a task, but it can have limited explanations, making another resource more suitable to really understand the goal of the tutorial.

The integration of opinions and evaluations, subjective and objective, allows the model to defend itself against misbehaviours. Furthermore, it also allows 'colouring' cold evaluation values by providing additional quality information such as the educational capacities of a digital resource in an eLearning system. As a result, the recommendations are always adapted to user requirements, and of the maximum technical and educational quality.

H2.- Intelligent agents' paradigm allows modelling user behaviours in Technology Enhanced Learning platforms, optimizing multi-scenario testing and personalization.

Evaluating the effectiveness of a recommendation model may take a long time. And the outcomes of that evaluation can differ from the expected ones, forcing system administrators to switch to another model. The consequences of this change might be worse than the mere loss of time: in a competitive world like the digital world it can be the difference between success and failure, between having a system with users and watching the users' exodus to other platforms. Testing the recommendation model in a simulation tool appears as an efficient way of reducing, if not eliminating, the aforementioned risks.

The trust and reputation model defined in this PhD thesis proposes the combination of objective assessments and subjective opinions to provide recommendation, a novel approach not considered before, to our knowledge, in the literature. Hence, there are no available testbeds to evaluate a model as the one presented in this work. Therefore, a new simulation tool will be developed following the agent-oriented programming paradigm.

III.2. Objectives

The main objective of this PhD thesis is the **definition of a trust and reputation model for the recommendation of resources and users** (to interact with) within an open system for knowledge exchange.

To attain this main goal, the following specific objectives were defined:

- Analysis of the characteristics of medical Technology Enhanced Learning (TEL) platforms and other open or semi-open quality-oriented environments based on knowledge exchange, focusing on their trustworthiness and 'soft security' needs.
- Study of the methods and techniques used to provide resources' and users' recommendation; with special interest on trust and reputation models, recommender systems, and their combinations.
- Definition of a new trust and reputation model built not only on personal opinions from system's users but also on objective assessments of resources.
- Design and implementation of a simulation tool to allow modelling complex systems and their users, to ease and to speed up the evaluation of the model.
- Evaluation of the proposed trust and reputation system in a wide range of different situations of 'real-life' open healthcare and TEL platforms.

IV. MATERIAL AND METHODS

This chapter presents the material and methods defined, implemented and/or employed in this PhD to fulfil the objectives and, consequently, to support the hypotheses defined in Chapter III. It includes the trust and reputation model TruQOs, the testbed implemented for the simulation of virtual Technology Enhanced Learning (TEL) platforms and, finally, the framework proposed to carry out the evaluation of the model.

IV.1. Chapter structure

The chapter is divided in three parts. After an introductory section, the second section presents TruQOs, the trust and reputation model for the recommendation of resources and users proposed as the main contribution of this PhD thesis. The chapter continues with a third section where the evaluation of the model is introduced, which is in turn divided in two parts: the simulation tool designed and developed as testbed for the evaluation of TruQOs; and the scenarios and tests created to assess the effectiveness of the proposed trust and reputation model and the validity of the simulation platform.

Throughout this chapter a set of relevant terms is used, with close meanings but an exact and different purpose in the model's definition. To avoid confusions and misleading explanations, a section introducing these parameters in the context of a system for the exchange of clinical knowledge between users opens the chapter. Additionally, Annex A (section VIII.1) presents a dictionary of all these terms, and Annex B (section VIII.2) a summary of all the equations used in this chapter.

IV.2. Running example

The target domain selected in this PhD dissertation for the application of the proposed trust and reputation model is a digital Technology Enhanced Learning (TEL) platform where physicians of different clinical institutions can share their knowledge through the creation and exchange of medical resources such as therapies, scientific papers, photos, videos or voice recordings. All the generated **resources are automatically evaluated at the moment of their creation**, obtaining an objective value $ev_i \in [0,1]$, which is stored in the evaluations' database.

Patients willing to know more about their health status can also access this TEL platform. The trust and reputation model helps them to filter the relevant information from all the available resources, recommending them the best ones for their own interests.

Additionally, users of this platform (both patients and physicians) can communicate amongst them for clinical consultation: patient-patient, patient-physician, and physician-physician. These direct interactions can be made through different channels such as chat, email, videoconference, or even on face-to-face meetings. The result of these interactions is not objectively evaluated.

In this TEL platform most of the users (physicians) have a **polyvalent** role, and they can behave as both **providers** of resources or services through direct consultation, or they can act as resources' or services' **consumers**. All the interactions within the system, both direct and indirect by accessing a resource, are the result of a recommendation. The system supports two types of recommendation:

- **Resources Recommendation:** A consumer requests the recommendation of a resource and TruQOs assess the reputation of the all the resources generated by each provider to provide a classification of all the available resources in the system.

- **Users Recommendation:** A consumer requests the recommendation of another user and TruQOs assesses the reputation of each one of the available providers in the system to provide a classification.

In clinical platforms like the one described, the physicians of the system have different fields of expertise and create resources (or answer to requests) of varied **services**. A topology of services can be built with all the available services and the relation amongst them. The proposed trust and reputation model filters the resources (or providers) of the service required by the consumer and offers only those matching the desired service.

After any resource's consumption or direct consultation consumers decide whether or not to give an opinion or a review, depending on their *laziness* behaviour. These opinions are rating values $op_i \in [0,1]$ and are stored in the opinions' database. They will be retrieved, along with the evaluations, to calculate the reputation of the candidates (resources or providers) during the recommendation process. The proposed trust and reputation model considers all the opinions about the candidates, including both the past opinions gave by the consumer requesting the recommendation, as well as opinions offered by other consumers within the system after interactions with the specific candidate. Those other consumers are called **witnesses** and their opinions are used to calculate social reputation.

IV.3. TruQOs model

IV.3.1. Introduction

TruQOs stands for *Trust in Quality-Oriented Systems* and it is an evolution of THRIOS (Martínez-Sarriegui et al. 2012) where the main concepts of the model were introduced. The main goal of TruQOs, Figure IV.3-1, is to provide an **ordered classification of resources** based on their reputation in the platform, as shown in Figure IV.3-1. Besides, TruQOs provides in addition **users recommendation** by using only the information about the author of a resource, the provider, but not about the resource itself.

Figure IV.3-1. Main goal of TruQOs: to provide an ordered recommendation of resources.

Traditionally, recommendations in online platforms are based on consumers' opinions about the different available **resources** created by the providers within the platform. TruQOs includes information about **providers' history** in the system into the reputation assessment process.

Furthermore, providers' reputations are calculated combining: 1) **past experiences** of consumers with other resources created by the provider, and 2) an unbiased measure of **provider's proficiency**, assessed with information coming from the objective evaluations of his/her resources. To our knowledge, TruQOs is the first model that uses this latter dimension as one of the main sources of information to provide recommendations. Additionally, TruQOs incorporates information obtained from **system's structure** to weight consumers' opinions.

Summarizing, TruQOs follows a modular approach built upon the aggregation of information coming from different unitary reputation dimensions, as represented in the schema of Figure IV.3-2.

Figure IV.3-2. TruQOs' dimensions of reputation schema

These unitary dimensions can be grouped in three different families. Table IV.3-1 shows dimensions' classification into families along with a brief description of each dimension.

Table IV.3-1. Classification of reputation dimensions into source of information families.

Family	Reputations		Description
Resource	Resource Reputation		Calculated using all the opinions expressed by system's users about the specific resource whose reputation is being assessed
Provider	Provider's Evaluations Reputation		Calculated considering the quality of the resources generated by a provider. Gives information on how good is a user generating resources.
	Provider's Opinions Reputation	Personal Reputation	Calculated with all the past opinions of the consumer about other resources created by the provider.
		Social Reputation	Calculated using all the past personal experiences of other consumers when interacting with the provider.

Family	Reputations	Description
System	Structural Reputation	Calculated using the position of users inside system's structure and their dependency relations. It is used as a measure of social trust to weight opinions in Social Reputation.
	Institution Reputation	Calculated the global reputation of each Institution of the system as the average Provider's Reputation of its providers. It is used to determine the Institutions' Ranking and update system's structure

TruQOs complements the assessment of the reputation with the reliability of the calculated reputation. This concept is taken from ReGreT (Sabater & Sierra 2001b; Sabater & Sierra 2002a; Sabater 2004a), although it was first introduced in Sporas (Zacharia et al. 1999). As defined by Sabater (Sabater 2004a): *"This measure tells the agent how confident the system is on that value according to how it has been calculated. Thanks to this measure, the agent can decide, for example, if it is sensible or not to use the trust and reputation values as part of the decision making mechanism"*. This PhD dissertation proposes aggregating these two values, the **Reputation** R_{TruQOs} , and its **Reliability** RL_{TruQOs} , in a unique value, called **Reputation'** R'_{TruQOs} and to use this final value to perform the recommendation of both resources and users.

Once introduced all TruQOs dimensions, the following sections describe in detail how all those concepts are translated into a numeric reputation and reliability.

IV.3.2. Resource Reputation

The reputation of the 'Resource X' is assessed through a weighted mean of all the opinions expressed by all the users of the system that have consumed that specific resource. TruQOs gives more importance to the newest opinions through the function $\rho(t, t_j)$. The reliability of the reputation is calculated considering the variability of the ratings given by all the users, using the Uncorrected Sample Standard Deviation ' σ '.

According to this, the formulae for resource reputation and reliability are:

$$R_{Resource} = \sum_N \rho(t, t_j) \cdot op_j^{ResX} \in [0, 1] \quad (IV.3-1)$$

$$RL_{Resource} = (1 - 2 \cdot \sigma)^q \in [0, 1] \quad (IV.3-2)$$

Where ' q ' is a variable that allows either to establish a linear relationship between variations in ratings and Reliability or, on the contrary, high or low variations are punished. This impact of the ' q ' parameter is illustrated in Figure IV.3-3 with different variations of the parameter.

Figure IV.3-3. Impact of parameter ‘q’ on reliability’s assessment. X-axis: σ , Y-axis: reliability.

Finally, N is the number of opinions about ‘Resource X’ in the system while $\rho(t, t_j)$ and ‘ σ ’ are calculated using the following equations:

$$\rho(t, t_j) = \frac{t_j}{t} \quad (\text{IV.3-3})$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (op_i^{ResX} - \overline{op^{ResX}})^2} \quad (\text{IV.3-4})$$

IV.3.3. Provider Reputations

IV.3.3.1. Reputation based on provider’s past evaluations

TruQOs stands over the assumption that the quality of the resources offered by the individuals of the system can be objectively assessed or estimated by means of any automatic process or metric, or by impartial experts. The outcome of this objective assessment is a value $ev_i \in [0, 1]$. If no quality indicator can be found, this reputation’s dimension is zero.

To calculate Provider’s Evaluations reputation TruQOs gives more importance to the latest evaluations under the supposition that they will be closer to the evaluation value of future resources. With this time weighting TruQOs aims at encouraging the continuous generation of new resources of the highest quality.

$$R_{ProviderEv} = \sum_N \rho(t, t_j) \cdot ev_j \in [0, 1] \quad (\text{IV.3-5})$$

$$RL_{ProviderEv} = (1 - 2 \cdot \sigma)^q \in [0, 1] \quad (\text{IV.3-6})$$

Where now the reliability of the reputation indicates the average quality of the resources created by a provider and it is calculated considering the variability of the evaluations of the resources. The effect of the 'q' parameter is the same as the one presented in Figure IV.3-3. Additionally, $\rho(t, t_j)$ is the same as in equation (IV.3-3) and ' σ ' is again the Uncorrected Sample Standard Deviation as in (IV.3-4), calculated now with the following formula:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (ev_i - \bar{ev})^2} \quad (IV.3-7)$$

IV.3.3.2. Reputations based on past opinions about the provider

Reputations based on past opinions about the provider are related to the consumers' subjective perception about the outcome of their direct interactions with the provider. These reputations are inspired by ReGreT (Sabater & Sierra 2001b; Sabater & Sierra 2002a; Sabater 2004a), sharing with it some characteristics such as their personal and social approach or the 'shape' of some of the equations.

IV.3.3.2.1. Reputation based on consumer's personal experience

For the assessment of Personal Reputation and Reliability the model uses the formulae proposed for Individual Reputation in ReGreT (Sabater & Sierra 2001b)²⁷:

$$R_{Personal} = R_{personal}(C \rightarrow P) = \sum_N \rho(t, t_j) \cdot op_j \in [0, 1] \quad (IV.3-8)$$

$$RL_{Personal} = RL_{personal}(C \rightarrow P) = (1 - \mu) \cdot N_i(IDB) + \mu \cdot D_t(IDB) \in [0, 1] \quad (IV.3-9)$$

Where N is, for **user recommendation**, the number of all the opinions given to the provider, including both opinions on resources and opinions on direct interactions. For **resource recommendation** the system uses the opinions of direct interactions and all the resources' opinions except the opinions related to 'Resource X', the resource whose reputation is being assessed (previously used for $R_{Resource}$).

$\rho(t, t_j)$ is the same as in equation (IV.3-3). Finally, $N_i(IDB)$ is the number of impressions used to calculate the reputation value with 'itm' as the intimacy level; and $D_t(IDB)$ is the impressions' rating deviation, both concepts directly taken from ReGreT (Sabater & Sierra 2001a):

$$N_i(IDB) = \begin{cases} \sin\left(\frac{\pi}{2 \cdot itm} |IDB|\right) & |IDB| \in [0, itm] \\ 1 & otherwise \end{cases} \quad (IV.3-10)$$

²⁷ In ReGreT opinions are in the range [-1,1] while in TruQOs they are in the range [0,1]

$$Dt(IDB) = 1 - \sum_{impressions} \rho(t, t_j) \cdot |op_i - R_{Personal}| \quad (IV.3-11)$$

IV.3.3.2.2. Reputation based on social information

In some situations, a consumer might not be able to assign a trustworthy reputation value to a provider based on direct interactions: either because the interactions between them had been scarce, or because the reliability of the personal reputation is low. Or maybe it just wants to have opinions from other consumers. In those situations, to fill in the information gap, the model uses the witness reputation R_w , calculated with the information coming from other individuals regarding their past direct interactions with the provider.

Following the running example, to assess the reputation of 'User P' (the provider) requested by 'User C' (the consumer) using third party information, the model asks 'User w_1 ' and 'User w_2 ' (the witnesses) about their opinions on 'User P' and calculates the reputation through a triangulation. The model balances the reputation of 'User w_1 ' and 'User w_2 ' from the point of view of 'User C' with the reputation of 'User P' from the point of view of 'User w_1 ' and 'User w_2 '. Figure IV.3-4 illustrates this process.

Figure IV.3-4. Witness triangulation: assessing provider's reputation using other users experiences

The Witnesses Reputation is based on the homonymous reputation defined for ReGreT system (Sabater & Sierra 2002b). TruQOs adopts the concept of social trust and the method used to identify the witnesses. The model doesn't collect the opinions of all the users that had had interactions with 'User B', but only the most representative ones using a domain dependant

‘sociogram’ and a heuristic metric based on the work by Hage and Harary (Hage & Harary 1984). The selection of only representative witness aims at avoiding the “correlated evidence problem” (Pearl 1988): the opinions of different witnesses trend to unify due to the high amount of information being shared. However, while Sabater and Sierra use fuzzy rules to relate and assign a value to the social relation between the consumer and each witness, the **socialTrust**, and between witnesses and provider; TruQOs uses the different hierarchical relationships existing amongst them.

$$R_W = \frac{1}{N_W} \cdot \left\{ \omega \cdot \left[\sum_{j \in |W|} G(C \rightarrow w_j) \right] + (1 - \omega) \cdot \left[\sum_{j \in |W|} G(w_j \rightarrow P) \right] \right\} \in [0,1] \quad (IV.3-12)$$

With: $G(a \rightarrow b) \equiv f(\text{socialTrust}(a \rightarrow b), R_{\text{Personal}}(a \rightarrow b))$ (IV.3-13)

Being: $f(x, y) = 1 - e^{-c(x+y)}$ (IV.3-14)

$$\text{socialTrust}(a \rightarrow b) = \frac{R_{\text{Structural}}(a \rightarrow b)}{\varphi} \quad (IV.3-15)$$

Where N_W is the number of witnesses and parameter ‘ φ ’ allows the modulation of socialTrust’s impact on the calculation of witnesses’ reputation. Figure IV.3-5 shows the impact of socialTrust in the calculation of $f(x,y)$ when the personal reputation is zero (projection $y=0$) with variations of the parameter ‘ c ’. Figure IV.3-5 a) is obtained with $\varphi = 1$ and a high impact of system’s structure while Figure IV.3-5 b) is obtained with $\varphi = 1000$, thus reducing the structural effect. Additionally, Figure IV.3-6 shows the combined effect in witnesses’ reputation of socialTrust and Personal Reputations variations. Again, a) is obtained with $\varphi = 1$ and a high impact of system’s structure while b) is obtained with $\varphi = 1000$

Figure IV.3-5. Impact of the socialTrust assessment in $f(x,y)$ (projection $y=0$ in $f(x,y)$). With a) $\varphi=1$ and b) $\varphi=1000$.

Figure IV.3-6. Impact of parameter 'c' on $f(x,y)$ with a) $\phi=1$ and b) $\phi=1000$

Ultimately, the Witnesses Reputation is again weighted with the existing social trust relationship between the consumer and the provider. Therefore, the formula for social reputation assessment is:

$$R_{Social} = f(\text{socialTrust}(C \rightarrow P), R_W) \in [0, 1] \quad (\text{IV.3-16})$$

Similar effects as the one illustrated by Figure IV.3-5 and Figure IV.3-6 are presented in the calculation of R_{Social} .

Finally, the social reliability is assessed using only as information the different reliabilities of the personal reputations without any hierarchical consideration.

$$RL_{Social} = \frac{1}{N_W} \cdot \left\{ \omega \cdot \left[\sum_{j \in |W|} RL_{Personal}(C \rightarrow w_j) \right] + (1 - \omega) \cdot \left[\sum_{j \in |W|} RL_{Personal}(w_j \rightarrow P) \right] \right\} \in [0, 1] \quad (\text{IV.3-17})$$

IV.3.4. Reputation based on system's structure

IV.3.4.1. Semi-dynamic architecture

For TruQOs each user of the system belongs to one, and only one, Institution and a user cannot be in the system without affiliation. Furthermore, inside the Institution the users are organized in an internal hierarchy, as shown in Figure IV.3-7. TruQOs considers a fixed hierarchy with a typical tree shape. Each node of the tree has a level "l" associated: the highest level is numbered with number 1 and, by consensus increasing natural numbers are applied for the rest of the

levels. As an example, in Figure IV.3-7's hierarchy black circles have level $L=1$; striped circles $L=2$, white circles $L=3$ and squared circles $L=4$.

Figure IV.3-7. Internal hierarchy of an Institution

In addition to the internal hierarchy, the institutions are also organized in a tree, in this case a linear tree or ranking, as shown in Figure IV.3-8. This dynamic ranking, updated periodically, determines which institution provides the best resources within the system and TruQOs calculates one different ranking for each of the services available in the system: an institution could provide high quality resources for 'Service 1' but low quality ones for 'Service 2'. Each Institution has a level ' L ' associated, and the highest level is numbered with number 1. As an example in Figure IV.3-8's ranking Institution 2 has level $L=1$, Institution 3 $L=2$ and Institution 1 $L=3$.

Figure IV.3-8. Ranking of Institutions for Service N

The position of the institutions within the ranking is determined by their Institution Reputation, and ordered from higher to lower reputation value. The following section introduces Institution Reputation, defining the equations for its calculation.

IV.3.4.2. Reputation based on Institutions' providers' evaluations

The underlying concept of Institution reputation refers to the prestige or fame of an Institution within the system. If 'User A' belongs to a 'Institution I_A' renowned for the high quality of the resources created by its members, then the perception of the quality of the resources generated by 'User A' would be conditioned by the assumption that 'User A' would also give good results. Thus, the reputation and reliability of 'Institution I_A' take into account the Provider's Evaluations Reputation and Reliability of all of its members.

$$R_{Institution} = \frac{1}{N} \sum_N R_{ProviderEv}(j) \in [0, 1] \quad (IV.3-18)$$

$$RL_{Institution} = \frac{1}{N} \sum_N RL_{ProviderEv}(j) \in [0, 1] \quad (IV.3-19)$$

Where N is the number of providers within the institution.

IV.3.4.3. Relations based on structure

The perception that one user has of another user in the system, as well as the interactions between them, are conditioned by three factors:

1. The position of both users within their own institution.
2. The existing hierarchical relation between them.
3. The position of each user's Institution in the ranking tree.

The ideas behind this affirmation are threefold. The *first idea* is related with the criterion for the selection of users' position, usually based on their authority, experience, professional standing or ability. The *second idea* is that, although the main relationship between users of the same institution is cooperation, too often partners are caught in infightings motivated by competition between them (Hamel et al. 1989). A balance between these two forces, competition and cooperation, should be achieved (Teece 1992). Finally, the *third idea* is that groups or institutions influence the point of view of their members (Karlins & Abelson 1970). Thus, and if the provider is not in consumer's institution, consumer's perception of provider can be estimated by using the fame of provider's institution, its position in the ranking tree.

TruQOs defines the *Hierarchical Dependency Strength (HDS)* between two users of the system as the type and intensity of dependency relation between the two users, in four different situations:

- a) Both users are in the same institution ($I_A=I_B$) and there is a direct hierarchical dependency between them. **HDS = hds_high**. U1-U7 or U2-U6 in Figure IV.3-7.
- b) Both users are in the same institution ($I_A=I_B$) with no direct hierarchical dependency between them but both have a common ancestor in the tree. **HDS = hds_medium**. U9-U10 or U8-U4 in Figure IV.3-7.
- c) Both users are in the same institutions ($I_A=I_B$) but they belong to different branches in the hierarchy tree. **HDS = hds_low**. U3-U5 in Figure IV.3-7.

- d) The agents belong to different institutions ($I_A \neq I_B$). **HDS = hds_none**. For example one user of 'Institution 2' and another one from 'Institution 3' in Figure IV.3-8.

IV.3.4.4. Structural Reputation

Structural Reputation is always relative to the relation between two users and it is used in the calculation of Social Reputation as a measure of trust between users, to weight their opinions, as it is described in section IV.3.3.2.2. TruQOs considers this dimension of reputation as undeniable, giving to it the highest reliability.

$$R_{Structural} = R_{Structural}(a \rightarrow b) = \frac{1}{1 + e^{-(HDS+\Delta)}} \in [0, 1] \quad (IV.3-20)$$

$$RL_{Structural} = RL_{Structural}(a \rightarrow b) = 1 \quad (IV.3-21)$$

Where HDS is the numeric value assigned to each of the Hierarchical Dependency Strength cases as defined before, and Δ is the difference between users' levels, and it is calculated with the following equation:

$$\Delta = \begin{cases} \frac{l_B - l_A}{\max(l) - 1} & I_A = I_B \\ \frac{L(I_B) - L(I_A)}{\max(L) - 1} & I_A \neq I_B \end{cases} \quad (IV.3-22)$$

' $L(I_j)$ ' are the levels of users' Institution in the Institutions' Ranking, and ' l_j ' are the users' levels inside the internal hierarchy of their Institution. Note that high numbers indicate low levels. Figure IV.3-9 shows the impact of different HDS values in the final Structural reputation (axis y) for different values of levels' difference, Δ .

Figure IV.3-9. Impact of HDS in Structural Reputation assessment. Axis x: levels' difference Δ

IV.3.5. Building the puzzle

Once calculated all the dimensions of reputation, the final step is to combine them. The solution proposed in this PhD research work is described by the diagram of Figure IV.3-10.

Figure IV.3-10. TruQOs reputation/reliability assessment diagram

First the model calculates Social and Personal Reputations and combines them in a new value called **Provider's Opinions Reputation**:

$$R_{ProviderOp} = k_{Personal} \cdot R_{Personal} + k_{Social} \cdot R_{Social} \in [0,1] \quad (IV.3-23)$$

$$RL_{ProviderOp} = k_{Personal} \cdot RL_{Personal} + k_{Social} \cdot RL_{Social} \in [0,1] \quad (IV.3-24)$$

Where:

$$k_{Personal} + k_{Social} = 1 \quad (IV.3-25)$$

The next step is to combine provider's opinions with the evaluations of the resources created by the provider, in what it is known in TruQOs as **Provider Reputation**:

$$R_{Provider} = k_{ProviderEv} \cdot R_{ProviderEv} + k_{ProviderOp} \cdot R_{ProviderOp} \in [0,1] \quad (IV.3-26)$$

$$RL_{Provider} = k_{ProviderEv} \cdot RL_{ProviderEv} + k_{ProviderOp} \cdot RL_{ProviderOp} \in [0,1] \quad (IV.3-27)$$

Where:

$$k_{ProviderEv} + k_{ProviderOp} = 1 \quad (IV.3-28)$$

Finally, the last step varies depending on the type of recommendation:

- **Resource recommendation:** the model combines the reputation related to the provider of the resource with the opinions expressed by the users about the specific resource.

$$R_{TruQOs} = k_{Resource} \cdot R_{Resource} + k_{Provider} \cdot R_{Provider} \in [0,1] \quad (IV.3-29)$$

$$RL_{TruQOs} = k_{Resource} \cdot RL_{Resource} + k_{Provider} \cdot RL_{Provider} \in [0,1] \quad (IV.3-30)$$

Where:

$$k_{Resource} + k_{Provider} = 1 \quad (IV.3-31)$$

- **User recommendation:** the provider's reputation is directly used.

$$R_{TruQOs} = R_{Provider} \in [0,1] \quad (IV.3-32)$$

$$RL_{TruQOs} = RL_{Provider} \in [0,1] \quad (IV.3-33)$$

Through ' k_j ' coefficients the model modulates the use of the information coming from all the reputation dimensions to calculate the final TruQOs reputation and reliability by weighting their values. These coefficients are selected by the designer of the system and allow adjusting TruQOs to the characteristics of the target domain.

IV.3.6. How TruQOs deals with newcomers

As stated in Chapter II, trust and reputation models should implement techniques to deal with the dynamism of the environment in which they operate, especially to face the **bootstrapping** problem, or how newcomers' reputation is assessed. Bootstrapping may lead to the exclusion of newcomers with no acquaintance with the environment and low initial reputation values (Huynh 2006).

TruQOs deals with this issue defining a **promotion period** ('**prom**'). During this period, the initial reputation for the new providers is calculated providing a temporal value according to the type of recommendation being assessed:

- **Resources Recommendation:** If a 'User N' enters the system and does not generate any resource there is no possibility to be recommended as a resource generator. On the contrary, if 'User N' creates resources during the promotion period, all those resources created are promoted by using the Provider's Evaluations reputation as the final TruQOs reputation value.

- **Users Recommendation:** As 'User N' has no opinions or evaluations, his/her Author's Reputation is 0. Hence, to give newcomers the chance to be selected the model gives a fixed value of 0.75 as final reputation. This value has been selected to help 'User N' to get opinions during the promotion period to position 'User N' in relation to other users with more experience in the system.

To encourage the new providers into the generation of new resources, during the first half of the promotion period the final reputation is the calculated temporal value. But for the rest of the period the temporal reputation is weighted considering the moment of calculation of the reputation within the promotion period and the moment when the newcomer entered the system, t_{Debut} . Therefore, the equations for the reputation and reliability are:

$$R_{TruQOs}(ResRec) = K \cdot R_{ProviderEv} \in [0,1] \quad (IV.3-34)$$

$$R_{TruQOs}(UserRec) = K \in [0,1] \quad (IV.3-35)$$

$$RL_{TruQOs}(ResRec) = RL_{ProviderEv} \quad (IV.3-36)$$

$$RL_{TruQOs}(UserRec) = 1 \quad (IV.3-37)$$

Where K is calculated with the following formula:

$$K = \begin{cases} 1 & 0 \leq \Delta_t \leq \left\lceil \frac{prom}{2} \right\rceil \\ 1 - \frac{1}{prom - \left\lceil \frac{prom}{2} \right\rceil + 1} \left(\Delta_t - \left\lceil \frac{prom}{2} \right\rceil \right) & \left\lceil \frac{prom}{2} \right\rceil < \Delta_t \leq prom \end{cases} \quad (IV.3-38)$$

$$\Delta_t = t - t_{Debut} \quad (IV.3-39)$$

The effect of selecting different promotion periods in the final weight, 'K', applied to the assessment of newcomers' reputation is better understood with Figure IV.3-11, where the x axis is Δ_t , the difference between the actual simulation time and the moment when the new provider entered the system.

The promotion period is additionally used in the assessment of Institution Reputation (see section IV.3.4.2). If a provider of an Institution has not created any resource yet, the model considers whether or not to take this user into consideration for $R_{Institution}$, based on the promotion period. During their promotion period providers with no results are not applied to this reputation's calculation and the cardinality N in (IV.3-18) and (IV.3-19) is not increased. When the provider finally creates a resource or the promotion period is over, the reputation of the provider is considered, even if the contribution is a zero.

Summarizing, the inclusion of this **promotion period** parameter fulfils three needs of the system: 1) to help newcomers to the system to be recommended; 2) not penalising Institutions' trustworthiness when a new user enters the institution and has not had time to create

resources; and 3) encouraging lazy users to generate good resources by means of group pressure: if they do not create resources or create bad ones their Institution will be punished.

Figure IV.3-11. Impact of promotion period on function K for newcomers' reputation weighting mechanism

IV.3.7. Reputation and Reliability aggregation

Once calculated the reputation and reliability of all the possible providers and resources, the model needs to combine these related information values. This combination can be performed in many ways, including the possibility of letting the consumer to make it himself/herself. It is on system's designer's hand to choose the aggregation method that best fits system's requirements. For the purposes of the simulations and evaluations carried out on TruQOs, the recommendation of individuals and resources is made through a three-step process, "**Quartiles Aggregation Method**" first proposed in this PhD work, and illustrated in Figure IV.3-12.

Figure IV.3-12. Quartiles Aggregation Method: Resources are first ordered according to their Reliability and then their Reputation is modified by q_i factors to obtain Reputation'.

First, individuals or resources are ordered according to their calculated reliability, from higher to lower, and divided into four groups or quartiles. Next, their reputation is modified according to the quartile they belong. A different factor multiply the reputation of the individuals: 'q1' for the first group, 'q2' for the second group, 'q3' for the third, and 'q4' for the last group then, where $q1 \geq q2 \geq q3 \geq q4$.

Finally another order of the providers or resources is performed, according to their new processed reputation, *Reputation'*. This ordered list with no reference to their reputation and reliability is then sent to the consumer as the recommendation.

IV.4. Evaluation of the model

The essential conclusion obtained after the analysis of the State of the Art regarding the evaluation of trust and reputation (T&R) models is that it is still one of the main challenges in the research on these systems. There are nearly as many T&R systems as methods to evaluate them (Kerr & Cohen 2010). Since its creation, ART Tesbed (Fullam, Klos, et al. 2005) has been extensively used to evaluate T&R systems. Nevertheless, although it is well-designed for evaluating agents in a competitive manner it is not appropriate in scenarios with complex social relationships (Harbers et al. 2007; Hang et al. 2008). Another testbed candidate is TREET (Kerr & Cohen 2010) but it is primarily focused on trust and reputation models in marketplace environments where users are not grouped in social structures as institutions.

TOAST testbed (Hermoso & Centeno 2007) is probably the one that better fits the requirements, albeit it does not support the main features of TruQOs. TOAST emerges with the intention of providing a tool to evaluate trust and reputation mechanisms that take advantage of Electronic Institutions and Virtual Organizations concepts, used in open agent societies to impose a certain structure. In TruQOs this organizational structure is also presented although the structure of the institutions is more rigid and the influence of hierarchies on users' interactions is not addressed in Virtual Organizations. However, the essential drawback is related to TruQOs' Quality of Results reputations family. TOAST generates a *Utility* value as the result of each interaction but it is more a measure used to simulate individuals' subjective perception than a real objective value of outcomes' quality.

The model proposed in this PhD thesis presents a novel approach to the assessment of trust and reputation process for the recommendation of resources and/or users. To our knowledge, TruQOs is the only model that specifically takes into account objective evaluation of resources to provide reputation values. Thus none of the existing testbeds or evaluation mechanisms can be fully used in the validation. Therefore, a new tool adapted to the features of TruQOs has been defined.

The evaluation of the model is divided in two sections. In section IV.4.1 the application developed to evaluate TruQOs is presented. It models and simulates virtual open quality-oriented systems for knowledge exchange based on a recommendation engine. Lastly, section IV.4.2 is devoted to the evaluation framework proposed to test the efficiency and robustness of the proposed model in a wide range of scenarios.

IV.4.1. TruQOs Simulation Tool (TST)

IV.4.1.1. Presenting the simulation tool

The simulation tool developed in this PhD research work is called TruQOs Simulation Tool and its short name, used all along this section, is TST. Figure IV.4-1 shows the initial screen of the application.

Figure IV.4-1. TruQOs Simulation Tool (TST)

From the different available languages and platforms for agent-based programming (Kravari & Bassiliades 2015), the selection to implement TST is **JADE**²⁸ (Java Agent DEvelopment framework), in its **4.2.0 version**, for two main reasons: a) it has been widely used in the last years and has a life support community; and b) it is a *freeware* framework fully implemented in Java, a language in which the author of this PhD thesis has an expertise of more than 10 years. Eclipse Kepler (version 4.3) was used as the IDE for the programming.

JADE simplifies the implementation of multi-agent systems through a middleware that claims to comply with the FIPA²⁹ (Foundation for Intelligent Physical Agents) specifications (Page et al. 2006). In addition to some already implemented Agents to help in the management of the platform such as the *Agent Management System (AMS)* or the *Directory Facilitator (DF)*, JADE provides the developer with a complete API to create user-defined agents and to allow their interaction with FIPA system agents. New agents are created using two main classes (Bellifemine et al. 2007):

1. **Class Agent:** An agent is simply an instance of a user defined Java class that extends the base Agent class. This implies the inheritance of features to accomplish basic interactions with the agent platform such as registration, configuration or remote management; and a basic set of methods that can be called to implement the custom

²⁸ <http://jade.cselt.it/>

²⁹ De facto standard for agent-oriented programming since 2005, <http://www.fipa.org>

behaviour of the agent, for example to send/receive messages. The computational model of an agent is multitask, where tasks (or behaviours) are executed concurrently. Agents interact with each other through *asynchronous* Agent Communication Language (ACL) Messages, as defined by FIPA.

2. **Class Behaviour:** Each functionality/service provided by an agent should be implemented as one or more behaviours. Behaviours work like co-operative threads. A scheduler, internal to the base Agent class and hidden to the programmer, carries out a round-robin non-preemptive scheduling policy among all behaviours available in the ready queue. In order to avoid an active wait for messages, every single Behaviour is allowed to block its computation. All blocked behaviours are rescheduled as soon as a new message arrives.

Figure IV.4-2 shows the multi-agent system (MAS) that constitutes the structure of the simulation tool (the figure excludes jade-agents –see Annex A in section VIII.1 for definitions-). The lines represent the interactions amongst the agents: normal lines represent interactions among users, mainly during the simulation process but they may include also settings' interaction. Dotted lines are only settings' interactions and the dash line represents the communication of the "Databaser" with the physical database of the system.

Figure IV.4-2. TST structure and Agents communications: normal lines represent mainly interactions among users during the simulation but they may include also settings interaction, dotted lines are only setting interactions and the dash line is database communication

There are different types of agents in TST. The orange circles in Figure IV.4-2 are internal system agents, from now on *tst-agents*, in charge of different tasks in the design, recommendation and simulation process and only one agent of each type can be presented in the platform: *SimulationAgent* ("Simulator"), *ReputationAgent* ("Reputer"), *EvaluationAgent* ("Evaluator"), *HierarchyAgent* ("Hierarcher"), *SocialAgent* ("Socializer"), and *DatabaseAgent* ("Databaser"). The **Provider** and the **Consumer** are *UserAgents* (or *user-agents*) and represent the final users of

an open quality-oriented system using TruQOs as the recommendation engine. In TST the number of agents of type UserAgent depends on the system that the designer wants to simulate.

IV.4.1.2. User behaviour modelling

UserAgents represent the final users of the open system being modelled with TST and in which the designer wants to test and to integrate TruQOs as the recommendation engine. Users are created in the system by the HierarchyAgent and they are configured with the parameters and behaviours selected by the designer using the graphical interface of the “Simulator. In TST three different types of UserAgents can be defined, with different parameters and roles: **Consumer**, **Provider** and **Polyvalent**. Annex C (section VIII.3) summarizes these roles and their parameters.

IV.4.1.2.1. Consumer

Consumer users in the system request recommendation to access resources created by other users and/or recommendations of users to whom they can directly interact to request a service.

When a consumer agent receives the first simulation event, it decides whether or not to start an interaction process according to a *Geometric* probability distribution with a certain parameter called **Talkativeness**. This parameter is in the range of [0,1] and models the average number of simulation events that the consumer will wait to communicate: higher values are translated into a lower delay in consumers’ activation. On each round this number is decreased and when it equals 0 the consumer requests a recommendation, calculating again a new number of events to wait for the next interaction. The geometric probability distribution is calculated using the JSC³⁰ library, obtaining the average number of steps until the user becomes active, $nSteps_{RequestRec}$:

$$nSteps_{RequestRec} = \text{round}\left(\frac{1 - \text{talkativeness}}{\text{talkativeness}}\right) \quad (\text{IV.4-1})$$

Once a consumer has decided to request an interaction there are two options: requesting a Resource recommendation, where its sequence of actions is shown in Figure IV.4-3, or requesting a User recommendation to interact to for the provision of a service, with a sequence of actions represented in Figure IV.4-4.

In both cases the consumer looks for resources or providers of the **Service** they have as configuration. The decision of which type of recommendation should be applied depends on the simulation mode. If only one type of recommendation is enabled then there is no decision to make. On the contrary, if the recommendation mode in the SimulationAgent is set to ‘*Resources and Users*’ the decision is determined by the parameter **Resources:Users**. This integer parameter defines the periodicity of users’ recommendation over resources’: a value of ‘n’ indicates that after ‘n’ resource recommendations the next time the consumer is active a user recommendation will be requested. After this user recommendation, the decision process starts over again with a resource recommendation.

³⁰ <http://www.jsc.nildram.co.uk/>

Figure IV.4-3. Consumer requests recommendation of a Resource to consume

After receiving a recommendation from the “Reputer” the consumer uses this recommendation to either consume the resource (if that resource was not previously consumed in the past), Figure IV.4-3, or to communicate with the first of the recommended users, Figure IV.4-4. If the first user in the recommendation list does not answer, the consumer keeps trying with the next user in the list until one satisfies the request or the list is finished.

Lastly, if the interaction was successful, the consumer decides whether or not to rate subjectively the outcome of the interaction, based on a probability calculated through a Bernoulli distribution with a parameter called **Laziness**, with the probability mass function calculated through the following equation:

$$f(\text{opining}; \text{laziness}) = \begin{cases} 1 - \text{laziness} & \text{giving opinion} \\ \text{laziness} & \text{not opining} \end{cases} \quad (\text{IV.4-2})$$

There are seven different **Opinions’ Behaviours** for consumers while providing subjective ratings of the outcomes:

- Fair: the subjective opinion equals the objective evaluation of the resource or the simulated quality of the interaction’s result.
- Praiser: the subjective opinion equals to the square root of the objective value³¹, overrating the quality of the resource or interaction’s result.
- Critic: the subjective value is the square of the objective evaluation³¹, underrating the quality.

³¹ Note that objective evaluations $ev \in [0,1]$ with $\sqrt{ev} > ev$ and $ev^2 < ev$

- Fickle: consumer's behaviour changes randomly from fair to critic or flatterer.
- Swindler: the subjective opinion is the opposite of the objective value.
- Lover: consumer always loves provider's results giving unrealistic randomly high values in the range [0.7,1].
- Hater: consumer always hates provider's results giving unrealistic randomly low values in the range [0,0.3].

These behaviours when giving opinions can be applied to the resources and interactions of all the users or only to one specific user or Institution through the **Target** parameter. In the case a specific target has been selected the opinions' behaviour is applied only to the target(s) while 'fair' behaviour is used to rate the rest of the users.

Figure IV.4-4. Consumer requests recommendation of a Provider to interact with

The final action for both types of recommendation, as it is depicted at the bottom of the sequence diagrams of Figure IV.4-3 and Figure IV.4-4, is to inform the "Socializer" about the result of the interaction, in case it was successful. If the consumer has decided to give an opinion about the resource or the interaction with the provider, the SocialAgent in turn sends this opinion to the "Databaser" for its storage.

IV.4.1.2.2. Provider

Providers in TST are users whose main goal is the generation of new digital resources. Additionally, they offer their services for direct interaction or consultation to other users. For providers, the reception of a simulation event supposes making the decision of generating or not a new resource. The approach is similar as the one taken for consumers, modelling the average number of simulation events that the provider will wait to generate a resource with a Geometric

probability distribution with a parameter called **Productivity**, in the range of [0,1] and with higher values translated into a lower delay in provider's activation. Again, on each round this number is decreased and when it equals 0 the provider generates a resource and calculates a new number of events to wait to create new resources. JSC³⁰ library is once more used to obtain the average number of steps until the user becomes active, $nSteps_{CreateRes}$:

$$nSteps_{CreateRes} = \text{round}\left(\frac{1 - productivity}{productivity}\right) \quad (IV.4-3)$$

The sequence of actions after the provider decides to generate a resource is described in the diagram of Figure IV.4-5. The resource is then sent³² to the "Evaluator" for its evaluation and, once an objective evaluation has been calculated, the EvaluationAgent informs the "Socializer" and the "Databaser" to store the evaluation value.

Additionally, regardless a resource was created or not during the simulation round, providers wait for possible direct requests from consumers, as in the diagram of Figure IV.4-4, and produce a result in response. However, not all the providers fulfil all the requests. Their **availability** is modelled with the homonymous parameter and a Bernoulli distribution using again the JSC³⁰ library, being its probability mass function:

$$f(\text{answering}; \text{availability}) = \begin{cases} \text{availability} & \text{answering} \\ 1 - \text{availability} & \text{not answering} \end{cases} \quad (IV.4-4)$$

Figure IV.4-5. Provider's Resource generation

³² Note that resources creation is simulated and no real resources are generated in TST

IV.4.1.2.3. Polyvalent

Polyvalent users can behave as consumers or providers but not at the same time. They can generate resources and answer to direct consultations, as well as request recommendations to access resources or interact with other users.

When the first simulation event is received these users calculate the number of steps to create a new resource ($nSteps_{CreateRes}$) with equation (IV.4-3) and the number of steps to consume a resource, and so to request a recommendation ($nSteps_{RequestRec}$) with equation (IV.4-1). On each round these numbers are decreased and when any of them equals 0, polyvalent user takes the corresponding role: provider if $nSteps_{CreateRes}$ equals 0 or consumer if $nSteps_{RequestRec}$ equals 0; calculating then a new $nSteps_{CreateRes}$ or $nSteps_{RequestRec}$. In the case both numbers of steps equal 0 in the same simulation step, the role of the polyvalent user is randomly determined with a Binomial distribution with equal probability for each role.

IV.4.1.3. TruQOs recommendation process

The recommendation process is shown in detail in Figure IV.4-6.

Figure IV.4-6. TruQOs recommendation process

The processes for both recommendation types (resources and users) are equivalent. The only difference appears in the calculation of resources' authors, which is only executed for resources recommendation. The rest of actions are identical except for the nature of the data to be collected from the database, and the equations to calculate reputation.

The first step of the process, after the identification of the candidates and the determination of resources' authors (in case of resource recommendation), is to identify the witnesses, that is, finding the users that interacted with the provider in the past and whose opinions will be considered by TruQOs to assess the Social reputation. The "Reputer" delegates this task to the "Socializer". The second step is to obtain from the "Databaser" all the past evaluations of resources created by the provider. Once the witnesses have been discovered and the provider's evaluations retrieved, the "Reputer" requires to the "Hierarcher" to calculate the Hierarchical Dependence Strength (HDS) values of the relations: consumer-provider (cp), consumer-witnesses (cw), and witnesses-provider (wp).

The final step is to get from the database all the opinions in the system about the provider: the opinions about the resource whose reputation is being assessed, and the opinions of the rest of provider's resources and/or interactions of other users with him/her. With all the information retrieved in all the steps, the "Reputer" calculates the reputation of the candidate and the reliability of that reputation, aggregates both values and offers and ordered classification of resources or users to the consumer.

IV.4.1.4. TST-Agents

IV.4.1.4.1. "Simulator", the *SimulationAgent*

The "**Simulator**" is in charge of orchestrating all the simulation process in TST, from the personalization of simulation mode and parameters, to the visualization of the results of the simulation. This agent coordinates the execution of all the involved agents in the simulation: tst-agents and user-agents. Using the *SimulationAgent*'s GUI, left part of Figure IV.4-7, the designer of the system can select the recommendation mode to be used in the simulation: *Only Resources*, *Only Users* or *both*. Additionally, the GUI allows selecting the number of Services in the system and to personalize the role and behaviour of each one of the users of the system.

The right part of the GUI is devoted to the control and visualization of the results of the simulation process. TST enables two modes of simulation: **Step-by-Step**, where only one simulation step is executed when pressing the 'Simulate' button; or **Free-Run** in which the designer selects the number of steps to be executed before the simulation stops.

A summary of the simulation after each round is shown in a *TextArea* occupying most of the right part of the screen. For each user the simulation status is represented through seven flags: 1) number of resources created (**ResCreated**); 2) times that a resource created by that user has been accessed (**ResAccessed**); 3) number of resources consumed (**ResConsumed**); 4) number of simulation rounds a user has been active (**Started**); 5) number of active rounds that lead to a successful result (**Successful**); 6) number of unsuccessful active rounds (**Failed**); and 7) the number of direct interactions answered (**Answered**).

The global activity of users is represented in a histogram type graph as shown in Figure IV.4-8. Additionally, the *TextArea* shows also the result of the latest simulation round for each of the consumers, pointing out either if the user has consumed a resource or the user has interacted with another user. A temporal representation of the results, Figure IV.4-9, can be also accessed. The Java library **JFreeChart**³³ in its **version 1.0.14** was used to create both types of graphs.

Figure IV.4-7. SimulationAgent's GUI: Management and Visualization of: a) Left.- Recommendation mode and Users' simulation parameters; b) Global simulation parameters and simulation status.

Figure IV.4-8. Global Users' activity during the simulation

Figure IV.4-9. Users' activity over time during the simulation

³³ <http://www.jfree.org/jfreechart/>

IV.4.1.4.2. “Reputer”, the ReputationAgent

The “Reputer” performs reputation assessment within the simulation platform. Through its graphical interface, Figure IV.4-10, the designer of the system can select the ‘*Reputation mode*’ and its parameters.

Three different modes are enabled for calculating reputation:

- 1) **NoReputation:** In this mode the recommendation is not based on reputation values but it is performed according to the objectively assessed quality of the resources. For users recommendation the ReputationAgent offers a randomized classification of all the available users.
- 2) **OnlyOpinions:** The reputation is assessed using only resource’s opinions or, in the case of user recommendation, using all the opinions about past provider’s interactions, according to the following formulae:

$$R_{OnlyOpinions}(ResRec) = \frac{1}{N} \sum_N op_j^{ResX} \in [0, 1] \quad (IV.4-5)$$

$$R_{OnlyOpinions}(UserRec) = \frac{1}{N} \sum_N op_j \in [0, 1] \quad (IV.4-6)$$

For this mode of reputation all the opinions are equally considered not giving more weight to the most recent ones. No reliability calculation is performed in this mode.

- 3) **TruQOs:** The recommendation is offered using the trust and reputation model presented in this PhD dissertation and defined in section IV.3. The designer can select from this interface all the parameters of the model, including the possibility of not considering $p(t, t_i)$, as in OnlyOpinions reputation mode.

During the simulation process, the “Reputer” receives several recommendation’s requests from different users. For each one of these requests it retrieves all the reputation’s candidates (resources or users) and calculates in parallel all the reputations of each one of the candidates. Once all the reputations and reliabilities for one consumer are assessed, the ReputationAgent calculates the **Reputation’** (see IV.2.6) and uses it to return an ordered recommendation.

“Reputer” makes use of a combination of *parallelBehaviours* and *sequentialBehaviours*, both provided by JADE. Using *parallelBehaviours* allows assessing the reputation for each one of the possible candidates at the same time, speeding up the recommendation process. With the *sequentialBehaviour* “Reputer” assures that all the reputation process is executed in the correct order: 1) get candidates, 2) get witnesses, 3) get evaluations and opinions, 4) get hierarchical levels and HDS values, and 5) calculate reputation value.

The screenshot shows the 'Reputer' application window. It features a 'Reputation Mode' section at the top with radio buttons for 'NO REPUTATION', 'TruQOs RECOMMENDATION' (selected), and 'ONLY OPINIONS'. Below this is a 'Temporal effect opinions/evaluations?' section with 'No' and 'Yes' (selected) options. The main area is divided into several parameter groups: 'HIERARCHY REPUTATION PARAMETERS' (HDS_HIGH, MEDIUM, LOW, NONE factors), 'RESOURCES RELIABILITY PARAMETERS' (q factor), 'NEWCOMERS PARAMETERS' (Promotion period), 'TRUQOS PARAMETERS' (Evaluations weight, Opinions weight), 'R:RL AGGREGATION PARAMETERS' (q1 R:KL, q2 R:RL, q3 R:KL, q4 R:RL), 'OPINIONS REPUTATION PARAMETERS' (Personal weight, Social weight), 'Personal Parameters' (Intimacy level, Mu parameter), 'Social Parameters' (c factor, Phi parameter, Omega parameter), and 'Resource weight' and 'Author weight'. At the bottom, there is a 'DEBUG Mode' section with radio buttons for 'Print User/Resource RECOMMENDATION' (selected), 'Print RANKING', 'Print FINAL reputation/reliability', and 'Print ALL dimensions' reputation/reliability'. Below the debug mode are four buttons: 'Set params', 'Reset Params', 'Save Params', and 'Load Params'.

Figure IV.4-10. ReputationAgent's GUI: Reputation method and parameters selection

IV.4.1.4.3. "Evaluator", the EvaluationAgent

Quality of resources and/or results of direct interactions for a service in TruQOs Simulation Tool (TST) is simulated by the **"Evaluator"**. When a provider generates a new resource the EvaluationAgent receives a request to evaluate its quality and the "Evaluator" simulates resources' objective evaluation. The "Evaluator" is also requested to make an evaluation when a direct interaction between two users has just occurred. Usually, users' opinions about the outcome of an interaction are related to the personal perception of quality of that outcome. TST models this by requesting the "Evaluator" to provide a quality number every time an interaction occurs and returning this value to the consumer³⁴. In turn, the consumer uses it to generate an opinion rating from the evaluation value received according to different opinion's rating behaviours, as described in section IV.4.1.2.1.

Through EvaluationAgent's GUI, left part of Figure IV.4-11, the designer can select for each provider the evaluation mode and the probabilistic distribution used to simulate the quality and its parameters. The evaluation mode allows the designer of the system to promote either indirect (access to resources) or direct interactions, or simulating users whose resources and direct interactions' results have similar quality.

³⁴ This evaluation value is not stored in the database.

Figure IV.4-11. EvaluationAgent's GUI: Management and Visualization of: a) Left.- Simulation of quality of results offered by the Providers; b) Right.- Resources' Classification according to their quality and Service

The simulated objective values are calculated using the **Java Statistical Classes³⁵ (JSC)** library, allowing three probability distributions: **Uniform**, **Normal** (Gaussian) and **Beta**:

$$\text{Uniform}(a, b) \equiv F(x; a, b) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1} \quad a, b \in (0, 1) \quad (\text{IV.4-7})$$

$$\text{Normal}(\mu, \sigma) \equiv F(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \mu, \sigma \in (0, 1) \quad (\text{IV.4-8})$$

$$\text{Beta}(\alpha, \beta) \equiv F(x; \alpha, \beta) = I_x(\alpha, \beta) = \frac{B(x; \alpha, \beta)}{B(\alpha, \beta)} \quad \alpha, \beta > 0 \quad (\text{IV.4-9})$$

Where $I_x(\alpha, \beta)$ is the regularized incomplete beta function, $B(x; \alpha, \beta)$ is the incomplete beta function and $B(\alpha, \beta)$ is the beta function:

$$B(x; \alpha, \beta) = \int_0^x u^{\alpha-1} \cdot (1-u)^{\beta-1} du \quad (\text{IV.4-10})$$

³⁵ <http://www.jsc.nildram.co.uk/>

$$B(\alpha, \beta) = \frac{(\alpha - 1)! \cdot (\beta - 1)!}{(\alpha + \beta - 1)!} \quad (\text{IV.4-11})$$

Additionally, a constant value in the range [0,1] can also be used as objective evaluation, despite not being very realistic.

Apart from its main functionality of creating objective evaluations, the “Evaluator” also keeps record of all the resources generated within the system and classifies them into a ranking from highest to lowest quality, as it can be seen in the right side of EvaluationAgent’s interface in Figure IV.4-11. This resources’ tree can be filtered using a selector to show all the available resources or only the resources of a specific service.

IV.4.1.4.4. “Hierarcher”, the HierarchyAgent

Internal hierarchies and Institutions’ Ranking are two important concepts in TruQOs and the “Hierarcher” is the agent responsible for their management. From its graphical interface (GUI) the designer of the system can add or remove Institutions and Users, creating also the internal hierarchy of each Institution, as seen in the left part of Figure IV.4-12. Internally, the HierarchyAgent creates (or kills) the Institutions as JADE containers (Bellifemine et al. 2002) and the Users as *UserAgents* in the container of the institution they belong to, according to the hierarchy tree.

Furthermore, the “Hierarcher” acts as the sentinel of TST controlling that only the agents allowed in the system are created and killing the agents detected as strangers and potentially threats to the security. Allowed agents include jade-agents, tst-agents and user-agents manually created by the designer. Once the structure of the system is set, the “Hierarcher” sends it to all the tst-agents for their knowledge.

In addition to supervising the creation of users in the system, the “Hierarcher” is also in charge of managing the Ranking of Institutions, calculated through the Institution Reputation as seen previously in the section IV.3.4.1 of this PhD document. Using the GUI the designer can decide if TruQOs is going to use this information in the reputation assessment process or, on the contrary, all the Institutions are going to be considered as having the same ranking position. In the case of considering a dynamic Institutions’ Ranking, the designer must select the periodicity of this ranking renovation, shown in the top right side of the GUI in Figure IV.4-12. The rest of the right part of the GUI is devoted to the visualization of this ranking for each of the available Services in the system.

Besides the management functionalities, during each reputation assessment process the *ReputationAgent* consults the “Hierarcher” regarding the levels of the consumer, the provider and the witnesses (if existing), in both the internal hierarchies and the Institutions’ Ranking. In addition, it also calculates the different Hierarchical Dependency Strength (HDS) parameters: $HDS(C \rightarrow P)$, $HDS(C \rightarrow w_i)$ and $HDS(w_i \rightarrow P)$.

Figure IV.4-12. HierarchyAgent's GUI: Management and Visualization of: a) Left.- System's structure and internal hierarchies; b) Right.- Institutions' Ranking for Services provided

IV.4.1.4.5. "Socializer", the SocialAgent

Social relations amongst users of the system play an important role in TruQOs. All direct interactions and accesses to resources created by a user are annotated in TST and conform the 'sociogram' of the system being modelled in the simulation tool. This graphic representation of social links is used to determine the witnesses and then to calculate the social reputation as shown in Figure IV.4-13.

The "Socializer" is the tst-agent in charge of creating and managing the 'sociogram', and also performs the witnesses' discovery process. An example of social graph can be found in Figure IV.4-13. Users of the system are represented with a coloured circle, with all the users belonging to the same Institution sharing the same colour.

There is one exception to this: deleted users, as 'User 13' in Figure IV.4-13, are depicted with a black triangle. TST keeps record of removed users for one fundamental reason: although they may not be available for direct interaction and will not be recommended, the resources they created could still be accessed. The system needs to know who the author of those resources was and who interacted with him in the past.

Figure IV.4-13 shows the SocialAgent GUI, where the curved black lines in represent the existing interactions between users. Every time a user reports an interaction with another user or the access to a resource, the SocialAgent draws a line linking the two involved users. For the sake of clarity this line is only draw once when the first communication between them occurs. The following interactions are annotated in an internal **interactions map**. Every edge of the graph has then bidirectional information of the communications between two agents stored in the interactions map. This map is consulted during the witnesses' discovery process.

Figure IV.4-13. SocialAgent' GUI: Users and their interactions 'sociogram'. Colours identify users from same Institution. Black triangle: dead user.

To build up and display the graph the **Java Universal Network/Graph framework³⁶ (JUNG)** in its **2.0.1 version** was used. This framework provides a common and extendible language for the modelling, analysis, and visualization of data that can be represented as a graph or a network.

The "Socializer" does not consider all the users that have interacted with the provider as potential witnesses. A parameter determines the minimum number of interactions a user has to have with another user to be considered as a potential witness. This parameter can be configured from the SocialAgent's GUI as shown at the top of Figure IV.4-13.

Even though the usability is not a priority in TST, a user-friendly approach has been followed in the design of the graphical interfaces when possible. In the case of the "Socializer", two "mouse methods" have been implemented. The first one is executed by pressing and holding the left button over a user. The designer (final user of the simulation tool) can grab that user and move it to a different position, with the edges of the graph moving accordingly. The second method is triggered by pressing the right button of the mouse over an interaction or a user:

- **Interaction:** A frame with all the information about that bidirectional interaction is shown (Figure IV.4-14). It specifies for each service the total interactions and the

³⁶ <http://jung.sourceforge.net/>

number of those interactions that were rated (the user gave an opinion). From all the interactions, the frame also shows which were related to the consumption of a resource and which of those accesses were rated. All this information is shown for both directions of the edge, from 'User A' to 'User B' and from 'User B' to 'User A'.

- **User:** In this case a frame with the resources created by that user for each Service is shown (Figure IV.4-15).

Figure IV.4-14. SocialAgent's GUI: Interactions' information panel.

Figure IV.4-15. SocialAgent's GUI: Resources generated by an User for each Service

Finally, the SocialAgent's GUI has in the top part a choice selector (see Figure IV.4-13, Figure IV.4-14 or Figure IV.4-15) that allows filtering the social graph to show: a) all the users and interactions; b) only users with interactions; and c) only the users with interaction of a specific Service (with as many options as available Services in the system).

IV.4.1.4.6. "Databaser", the DatabaseAgent

All the resources' evaluations and opinions expressed by the simulated users of TST are stored in two tables of a database implemented using MySQL server 5.5: **tbResources** and **tbOpinions** with the structure shown in Table IV.4-1. These tables store all the information used to calculate all the reputation dimensions of TruQOs. There is a logical (not physical) relationship between the two tables, the '*resourceId*'. Through this parameter the system can link the resources, and their objective evaluation, with each of the opinions given by the users that have consumed that resource. This is valid for the recommendation of resources but, if the rating comes as a result of a direct interaction between users, the '*resourceId*' in **tbOpinions** is set to 0, leaving empty that logical union.

Table IV.4-1. TST's database tables' structure

tbResources						
resourceId	provider	institution	evaluation	simTime	service	resourceName
tbOpinions						
opinionId	consumer	provider	rating	simTime	resourceId	

The “**Databaser**” centralizes all the communications with the physical database. It is in charge of receiving query requests from the *ReputationAgent* and returning all the resources and opinions needed for the reputation assessment. The *DatabaseAgent* also receives the results of the objective assessment of the generated resources from the *EvaluationAgent*, and the outcomes of the interactions user-resource or user-user from the *SocialAgent*, storing everything in the corresponding tables of the database.

IV.4.1.5. Simulation process

Keeping the synchronism between simulation rounds is the main issue to be addressed when designing and implementing simulation tools. In linear processes achieving a good synchronization is easy but in multi-agent systems like the one presented in this PhD dissertation it arises as a big challenge. In TruQOs Simulation Tool (TST) **user-agents** are proactive, with independent behaviours that move them to autonomously decide whether or not they are going to be active during a simulation round. Moreover, giving opinions and/or satisfying requests is also optional and behaviour-dependant. Hence, the time spent by user-agents during a simulation round varies among users and between simulation steps.

Although being mainly reactive to users’ and Simulator’s actions or requests several **tst-agents**, are configured with a set of behaviours which execution’s duration is also variable. This variability is especially present in two agents: *DatabaseAgent* and *ReputationAgent*. In the case of the “**Databaser**”, parallel communications with the database condition its execution delaying the completion of agent’s actions during a simulation round, and being the main *bottleneck* of the system. For “**Reputer**” the source of execution’s time fluctuations is the reputation assessment engine: for each agent a new execution’s thread is created where the different reputations of all the possible candidates (resources or users) are calculated in parallel, again in different threads. Thus, an undetermined number of threads are executed in parallel with different durations, making impossible to plan in advance the time that the “**Reputer**” needs to fulfil all its actions.

Finally, the agents’ communication mechanism stands in the way of achieving simulation’s synchronism. FIPA²⁹, and by extension JADE²⁸, define the exchange of FIPA-ACL compliant messages as the standard communication protocol amongst agents of the MAS platforms. The messages are sent from one agent to another and when received they are stored in agents’ ‘mailbox’. Whenever a message is received the agent is notified. However, when, or if, the agent picks up the message from the queue for processing is a design choice (Bellifemine et al. 2007). Message exchange is by definition an asynchronous mechanism thus, to achieve a synchronized simulation, additional mechanisms need to be implemented.

TST implements a signalling mechanism to let the Simulator know the completion status of agents’ behaviours during simulation. The simulation process in TST is described in Figure IV.4-16 using an activity diagram.

Figure IV.4-16. Simulation process from Simulator's point of view with the different phases highlighted.

The simulation activity during one simulation round can be decomposed in different phases. The figure shows, from the point of view of the Simulator, the different phases of the process with coloured squares and numbers:

- 1) **Simulation start.** (Green square) The Simulator sends a simulation event to all the agents in the system: to the *tst-agents* to let them know that they should be prepared to receive requests from other agents; and to the *user-agents* to activate them.
- 2) **Users' activation.** (Blue square) When user-agents receive the simulation event they decide to be active or not, generating a resource or requesting a recommendation. The Simulator creates one message receiver per existing user and waits for the activation signal coming from the users. Each type of users indicates the action they are performing during the present simulation cycle. A timeout is set for every receiver to avoid waiting for non-activating users. At the end of this phase the Simulator knows the exact number of users being active in the system and uses this information to estimate the time it may wait for the completion of the following phases. If no users are active the simulation continues with phase 6.

- 3) **Recommendation's requests.** (Orange square) From the previous phase the Simulator knows exactly how many consumers are requesting a recommendation. When the Reputer finishes calculating the recommendation for every consumer it sends to the Simulator a message. Once the Simulator has received all the expected recommendation's messages it knows that this phase has been completed and it will be receiving interactions' results soon.
- 4) **Interactions' reception.** (Brown square) Continuing with the signalling policy to assure simulation's synchronism, consumers send a message to the Simulator when they consume a resource or interact with a provider. If the interaction was unsuccessful because there was no recommendation available at that moment or the consumer did not find a reachable provider to interact with, a message is also sent. The result of these interactions is annotated and later shown in SimulationAgent's GUI (see Figure IV.4-7). Again, the Simulator knows the exact number of interaction results to be received and when to continue with the simulation.
Additionally, the results need to be also annotated by the Socializer. Another waiting loop is implemented until the Socializer notifies that all the expected interactions have been stored.
- 5) **Database's storage completion.** (Red square) The Databaser receives a considerable number of requests in parallel to retrieve information from past simulation rounds and to insert the resources and the opinions generated in the present cycle. The insertions are especially time-consuming: the Simulator could be ready to end the simulation while the Databaser is still busy storing interactions' results or even storing the resources generated during phase 2. The Simulator is aware of all the insertions that have to be done and waits until their completion.
- 6) **Ranking renovation.** (Purple square) Just before finishing the simulation round the Simulator sends a message to the Hierarcher to renovate the Institutions' Ranking. At its reception, the Hierarcher checks if it is time to renew it according the renewal periodicity (see section IV.4.1.4.2) and informs back the Simulator with information about whether or not the renewal was performed.
- 7) **Simulation resolution.** (Light Green Square) Finally, the last phase is the decision of continuing or not with the simulation. If the simulation mode is set to Step-by-Step or if it is set to Free-Run and this was the last of the expected simulation cycles, the simulation finishes. On the contrary, if more cycles are expected then the simulation continues by the execution again of phase 1.

IV.4.1.6. Additional functionalities

Apart from the obvious testing and simulation purposes, a testbed should offer some other functionalities, being the ability of importing and exporting simulations the principal one. The designer of the system has the possibility of exporting the results of a simulation to perform a deeper analysis of the results. Later, that or another simulation may be imported into the system to run different simulations with different parameters in the certainty that the start point of all the simulations is the same.

In TST this functionality can be accessed from the bottom right side of the SimulationAgent's GUI (Figure IV.4-7). By pressing the Export button the Simulator offers the designer to select a location and a name to store a file with proprietary extension '.tst' with all the needed information to restore later the state of the simulation. Internally, the Simulator sends a message to all the agents, both tst-agents and user-agents, with the order to export their data. In turn, each agent stores in files all the information they may need to restore their status, including the content of the database. Lastly, pressing the Import button the designer is allowed to select a valid .tst file to restore the state of all the agents and continue with simulation at the same point it was stored.

Additionally, and aiming at easing the parameters' selection process, an individual Export/Import method in some tst-agents has been implemented. It allows a) to save or load the system's structure in the Hierarcher (bottom left part of Figure IV.4-12); to save or load reputation mode and parameters (bottom of Figure IV.4-10); and to export or import evaluation and simulation parameters (bottom left of Figure IV.4-11 and Figure IV.4-7 respectively).

IV.4.2. Experimental design

IV.4.2.1. Introduction

The functionality of TruQOs is evaluated through different scenarios where iteratively the **effect of the different reputation dimensions and parameters** is tested. To achieve this main goal, the following specific objectives were defined:

- Modelling and simulation of a varied range of user roles and associated behaviours.
- Definition and simulation of a set scenarios focused at analysing the impact of all TruQOs parameters under different contexts.
- Comparison of TruQOs with other traditional approaches to provide recommendation: not using any reputation's assessment, **NoReputation** mode; and using only users' past opinions about a resource and/or other users, **OnlyOpinions** mode.

In the coming sections the scenarios defined to evaluate TruQOs' performance are introduced, briefly describing their main goal. The chapter continues with the description of the different user topologies and hierarchical structures used to model the clinical TEL platform defined at the beginning of the chapter (section IV.2). The different users' behaviours patterns implemented in the scenarios are presented next; followed by the definition of the main parameters to be used in the analysis of the results.

Subsequently, an initial state of the system for all the upcoming simulations is defined. The outcomes of this initial state are used to select the target users on which the recommendation performance is analysed. Finally, the chapter finishes with an extended description of the evaluation scenarios.

IV.4.2.2. Description of evaluation scenarios

For the design of the scenarios an iterative approach have been followed. Each new scenario introduces a new concept related with the assessment of TruQOs reputation, and the tests focus on the evaluation of each one of the different dimensions of the reputation. Scenarios **S1-S7** aim at setting all the main parameters of TruQOs and at confirming or discarding hypothesis H1 of this PhD work. Being the recommendation of resources the main goal of TruQOs, the selected recommendation mode for these scenarios is **Resources Recommendation**. Scenario **S8** introduces **Users Recommendation** and adds it to the resources recommendation to test Newcomers' effect, or how TruQOs deals with new users in the system.

Scenario **S9** analyses the existing relation between creating good resources and being recommended as a user to interact with. And vice versa, the relation between providing good results in direct interactions and improving the recommendation of own resources. The recommendation mode is set to **Resources and Users Recommendation**. Finally, in **S10** the optimal TruQOs parameters fixed in previous scenarios are tested under the same conditions of S0, comparing the results of both scenarios. This scenario is also tested using the OnlyOpinions mode and again a comparison between the three reputation modes is analysed.

IV.4.2.3. Selection of the Organisations and Hierarchies

Figure IV.4-17 shows the systems' organization of users and internal hierarchies of the institutions used for the simulation in all the scenarios except S6 and S7. The figure represent different hierarchy topologies to help evaluating system's structure impact in TruQOs while modelling possible real structures of clinical institutions: e.g. department director, staff physicians and resident physicians or nurses. This structure presents a balanced number and distribution of consumers (in blue) and providers (in orange) inside the Institutions.

To better study the effect of internal hierarchy scenario S6 is designed to use a smaller system's structure that the one used for the previous tests (Figure IV.4-17). The result is a structure where 'Institution 5' from the global structure is used. The selected Institution's internal hierarchy is shown in Figure IV.4-18 with all the possible HDS situations represented (with the exception of *hds_none*). The reduced hierarchy presents also a balanced number and distribution of consumers (in blue) and providers (in orange). Selecting a reduced hierarchy aims at focusing only in the effects of the institutions' internal hierarchy and to obtain results free of other side effects like the position in the ranking. Furthermore, a reduced hierarchy will also speed up the tests. The structure of Figure IV.4-18 is used again in scenario S7 aiming at reducing complexity in the system to better study the effect of reliability's assessment on the *Reputation*' value.

Finally, scenario S8 tests the performance of the proposed trust and reputation model when dealing with new users entering the system. To evaluate the so-called newcomers' impact two subscenarios were design where the new user, 'User 31' is created inside an existing institution ('Institution 3') from the general system's structure of Figure IV.4-17, see Figure IV.4-19 left; or in a new institution ('Institution 6'), see Figure IV.4-19 right.

Figure IV.4-17. System's structure and internal hierarchies to be used in the evaluation. Orange=Provider, Blue=Consumer

Figure IV.4-18. Internal hierarchy of Institution 1* for scenarios S5 and S6. Orange=Provider, Blue=Consumer

Figure IV.4-19. Newcomers effect, scenario S7: Left.- Institution 4* new User in existing Institution; Right.- Institution 6 new User in new Institution. Orange=Provider, Blue=Consumer

IV.4.2.4. Users' behavioural patterns

For the sake of repeatability of tests, the random aspects of the simulations have been limited to the maximum, simplifying the behaviours of the users. 18 basic user behaviours were modelled and variations of them were used in the simulations:

- 2 behaviours related to the role of the user: provider and consumer.
- 3 behaviours for providers related to the resources' generation pattern or availability to receive direct inquiries: generate a resource 3 out of 5 cycles, generate a resource every cycle, and answer all the incoming direct requests.
- 2 behaviours for consumers related to recommendations' request pattern or laziness at giving opinions: request a recommendation every cycle, and always give an opinion after an interaction.
- 3 behaviours related to the type of recommendation requested by consumers: only resources, only users, and resources and users alternatively.
- 3 evaluations behaviours: increasing or decreasing the quality over a period of time, and increasing/decreasing the quality alternatively.
- 5 opinions' behaviours: fair, critic, praiser, hater and lover.

Reduction of randomness is achieved by 'forcing' these behaviours and creating condensed patterns of resources' generation and recommendations' requests. It was decided not to use polyvalent roles due to their random behaviour at determining which role they will represent each round because it makes difficult the testing. Additionally, all the providers in the system were set to offer the same service ('Service 1') during direct consultation and to create only resources of this service. Consequently, consumers only request resources' or providers' recommendation of 'Service 1'.

In all the scenarios consumers are active all the time both requesting and rating direct and indirect interactions. As for consumers, in S0 and S10 they generate a resource 3 out of 5 simulation rounds but starting from S1 all the providers generate resources every simulation round. From S8 providers always answer to direct requests.

Despite being forced patterns, they model the operation of a real digital platform for sharing resources and consultation between users. A simulation round represents a fixed time window in the lifetime of the platform, allowing modulating the time to fit application domain's needs. In the scenarios of the evaluation framework the time span of each simulation step is set to represent a week. Generating a resource or requesting a recommendation within a week accurately portrays the platform being modelled.

IV.4.2.5. Evaluation parameters

Fixed patterns allow the definition of two **ratios of recommendation**, the main parameters used as the outcome of the tests and as the evaluation's success indicator. These parameters, applied only to providers, give a trustworthy measure of the recommendations' effectiveness for both resources and users, and they are calculated with the following formulae:

$$Ratio_{ResRec} = \frac{Resources_{Accessed}}{Resources_{Generated}} \quad (IV.4-12)$$

$$Ratio_{UserRec} = \frac{Interactions_{Answered}}{n_{Steps}} \quad (IV.4-13)$$

If one provider creates better resources this provider will be more recommended and accessed, increasing the ratio in comparison to other consumers' $Ratio_{ResRec}$. A similar explanation is applied for $Ratio_{UserRec}$, if one provider offers better services during direct interactions, he/she will be more recommended, answering more interactions and increasing the ratio. Moreover, these ratios provide more information than the mere number of accessed resources or interactions answered. They reflect the consumption and service provision patterns over time, facilitating the analysis of relation between dynamic changes in quality and/or opinions and changes in the number of times a provider is being recommended.

If providers receive some sort of retribution, for example money, every time a resource is accessed or they satisfy a direct request, then these ratios give a clear measure of profit and efficiency. During the complete set of scenarios the evolution of the ratio as a consequence of adding new TruQOs parameters and reputation's dimensions is analysed for one (or sometimes two) users.

Additionally, another evaluation parameter is defined, the **ratio of benefits**. This provider's parameter is only of application for *Resources and Users* recommendation mode. It measures the existing feedback between creating resources of a high quality and being more recommended for direct interactions. And vice versa the relation between providing good service in direct requests and an increase in the number of own resources being recommended.

$$Ratio_{Benefits} = \frac{Resources_{Accessed}}{Interactions_{Answered}} \quad (IV.4-14)$$

IV.4.2.6. Selection of the initial state

This scenario models a clinical platform for sharing resources and consultation between users **where no recommendation mechanism is installed** but in which the introduction of one of these mechanisms is being considered. It was designed to serve as initial state for all the scenarios except S6 and S7 where a reduced system's structure is selected to ease and speed up the simulation. The ultimate goal is to eliminate the warm-up phase, making the evaluation of TruQOs more effective. S0 provides a starting point of the system with an important set of resources already created and the opinions given by the consumers to that resources.

In S0 all providers have been modelled with the same global behaviour, generating resources approximately 3 out of 5 simulation rounds (or cycles), all of them with 0.5 as their quality. The consumers request a recommendation every round and always give opinions about the resources they consume with a fair behaviour.

The system is left working until the system has enough resources to guarantee the success of the upcoming simulations and an initial dispersion of resources and opinions amongst the providers is achieved, with no user soaring above the rest. This state is considered to have been achieved after 20 cycles and the results are used later in other scenarios. Furthermore, from the analysis of the results in S0 the target user(s), on which the evaluation will focus, are selected.

The objective is to select a user belonging to the providers' Institution (Institution 3) to study most of the effects. In situations where the effect of own institution's consumers wants to be studied (like scenarios S3c, S6 or S7), the target user is selected from a mixed institution (Institutions 1, 4 or 5). In both cases, the selected user is positioned in the middle of the hierarchy, to have more flexibility at studying all the effects and not to be conditioned in extreme by the Structural Reputation in socialTrust assessment. All the studied effects will be applied on these selected users, and some initial actions are applied to assure that the target user is the most recommended and in that way to facilitate the visualization of the desired effects.

Figure IV.4-20 shows the number of times a resource generated by each provider has been accessed over simulation time. A value of 'n' resources accessed means that a) one resource has been accessed 'n' times, b) 'n' resources have been accessed once, or c) any possible combination between a) and b).

Figure IV.4-20. S0 Accessed Resources over simulation time for each Provider.

Being the quality of the generated resources by all the users the same, $Q=0.5$, the differences in the number of accesses could be motivated by a difference on the number of generated resources: more resources in the system, more accesses. To be able of making a reliable comparison amongst users, the ratio of resource recommendation (defined in previous section and calculated with equation (IV.4-12)) was used, as it is depicted in Figure IV.4-21.

Figure IV.4-21. Ratio of Resource Recommendation over simulation time for each S0 Provider.

This ratio shows a deeper concentration of results between users, proving that the test attained the expected outcomes. The ultimate expected result of S0 is a system with most of its user being equally recommended. Figure IV.4-22 shows a summary of the scenario. In addition to the created and accessed resources and the ratio of recommendation, the mean and the standard deviation of this ratio through the 20 simulation cycles are also represented.

At the light of the results, the selected users are 'User 12' from provider's institution and 'User 27' from mixed institution. Both users were chosen due to the similarity between the ratio and its mean value.

Figure IV.4-22. S0 Results Summary

IV.4.2.1. TruQOs functionality tests

IV.4.2.1.1. S1: Effect of changes in the quality of resources

The scenario S1 evaluates the effect that changes in quality of resources' have in the recommendation process. A two-stage temporary fall in the quality of the results provided by 'User 12', from 0.5 to 0.49, is modelled. The analysis compares TruQOs reputation with using only quality information (not using reputation) or considering only opinions for providing the recommendation.

To facilitate analysis the impact of decreasing the quality, 'User 12' is set to produce the best results at the beginning of the test, making 'User 12' the most recommended user by increasing the quality of 'User 12' results from 0.5 to 0.51 for 5 cycles and then is set back to $Q=0.5$. The resulting resources' quality pattern is shown in Figure IV.4-23.

TruQOs' recommendation tests consider only **Provider's Evaluations Reputation** and are carried out twice: one with temporal weighting of the evaluations through $p(t,t_j)$ as defined in (IV.3-2) and (IV.3-3), and another one giving to all the evaluations the same weight.

Figure IV.4-23. Quality of resources pattern by 'User 12' in Scenarios S1 and S3b; for Resources recommendation

IV.4.2.1.2. S2: Effect of changes in the consumers' opinions

Scenario S2 evaluates the effect of other users' opinions in the recommendation of 'User 12'. Now a temporary attack from consumers of two Institutions towards 'User 12' is modelled. The attack is carried out by giving lower opinions compared to the real evaluation values of the resources. As for S1, 'User 12' is set to be the most recommended user by increasing the quality of 'User 12' results from 0.5 to 0.51 for 5 cycles and then is set back to $Q=0.5$. Next, two temporary attacks³⁷ are modelled by configuring attacking users with the *critic*³⁸ behaviour and lying about 'User 12' with $op=0.25$. Both resources' quality and opinions' patterns are shown in Figure IV.4-24.

TruQOs' recommendation tests consider only **Provider's Opinions Reputations** and are carried out twice: one with temporal weighting of the personal opinions through $p(t,t_j)$ as defined in (IV.3-8) and (IV.3-3), and another one giving to all the opinions the same weight.

³⁷ The word attack is being used here because it is easier to understand the effect, but the critics in the opinions could be motivated by a decrease in the pedagogical quality of the resources, a decrease that it is not reflected in the objective assessment.

³⁸ Opinions for critic behaviour are calculated through the square value of the evaluation: $0.5^2=0.25$

Figure IV.4-24. Quality of resources and opinions received patterns by ‘User 12’ in Scenarios S2 and S3b; for Resources recommendation

IV.4.2.1.3. S3: Changes in quality and opinions combined effect

The third scenario has as goal the determination of the combined effect that variations in evaluations and opinions have in the recommendation process. In addition a preliminary set of optimal TruQOs parameters will be obtained, ‘ k_j ’ values in equations from (IV.3-23) to (IV.3-31), as a result of selecting the best ‘ k_j ’ combination from the tests. Scenarios S1 and S2 are repeated adding to the reputation assessment process the opinions’ dimensions to S1, resulting in scenario S3a; and evaluations’ dimension to S2, to conform scenario S3b. The objective is to set the provider reputation weights. The quality of resources’ and opinions’ patterns are the same as the ones used in S1, Figure IV.4-23, and in S2, Figure IV.4-24.

Finally, a new sub scenario, S3c, is designed, where the evaluations’ and opinions’ effects are combined. **Resource Reputation** is added to the best results from S3a and S3b to set all TruQOs weights. To study this effect a new behaviour when giving opinions is introduced: *praiser*, used to promote users by giving better opinions than the objective evaluation. Thus, the selected target user must belong to an institution with consumers that can give these good opinions to promote him/her. Thus, ‘User 12’ is not valid and ‘User 27’ was selected.

Again 30 cycles are simulated and the quality of resources’ and opinions’ patterns are shown in Figure IV.4-25. When ‘User 27’ increases the quality of the resources to $Q=0.51$ some users (from ‘Institutions 1 and 2’) are configured with the *critic*³⁹ behaviour to discredit ‘User 27’ by lying with $op=0.2601$; and vice versa, when the quality of resources drops to $Q=0.49$ then ‘User 27’ is artificially promoted by users of ‘Institution 5’ which are configured with the *praiser*⁴⁰ behaviour to give opinions about ‘User 27’ with $op=0.7$.

³⁹ Opinions for critic behaviour are calculated through the square value of the evaluation: $0.51^2=0.2601$

⁴⁰ Opinions for praiser behaviour are calculated through the square root of the evaluation: $\sqrt{0.49}=0.7$

Figure IV.4-25. Quality of resources and opinions received patterns by ‘User 12’ in Scenario S3b; for Resources recommendation

IV.4.2.1.4. S4: Number of attackers’ effect

Scenario S4 studies the effect, on ‘User 12’ recommendation, of increasing the number of users attacking him. Starting from scenario S0 situation, four different tests of 30 simulation cycles will be performed. In all of them both the quality of resources’ and opinions’ patterns will be the same, as shown in Figure IV.4-26. The simulation starts with ‘User 12’ offering a slightly higher quality on its resources than the rest of users and, at some point, users of other institutions start attacking ‘User 12’ by configuring them with the *critic*³⁹ behaviour. On each new test an institution is added to the attackers set. First only 2 users, from ‘Institution 1’, criticize ‘User 12’. Later, the 5 consumers of ‘Institution 2’ are added, to sum up a total of 7 users criticizing. In the last two test, users from ‘Institution 4’ and finally users from ‘Institution 5’ are added to the attackers set, resulting in 10 and 15 attackers respectively.

The tests are run using TruQOs as recommendation engine and compared only to OnlyOpinions recommendation mode. The recommendation with NoReputation mode is excluded from this test, as this mode is never affected by the opinions.

Figure IV.4-26. Quality of resources and opinions received patterns by ‘User 12’ in Scenario S4; for Resources recommendation

IV.4.2.1.5. S5: Effect of Institution's position in the Ranking

Scenario S5 analyses the effect that the position in the Institutions' Ranking has in the assessment of reputation. Three different experiments have been designed. In all of them the 'hds_none' is varied in different situations where consumers of several institutions attack the target user. The effect of these attacks changes depending on the institution to which those users belong.

S5a: Attack from well reputed institutions (Competition)

In S5a a 30 cycles' simulation is performed, divided in two phases. In the first phase, the quality of the resources offered by 'Institution 5' is increased during 10 cycles. Quality is set to $Q=0.51^{41}$. With this increment in the quality of the resources 'Institution 5' progresses to the top position within the Institutions' Ranking and the effect of an attack coming from the most famous institution is tested. After these 10 cycles the quality of the resources of 'User 12' is changed by increasing it to $Q=0.55$ for 20 cycles. Additionally, users of 'Institution 5' are configured with *hater* behaviour to hate 'User 12' providing opinions with a rating value $op=0.1^{42}$. All these quality' and opinions' patterns are better understood by looking at Figure IV.4-27 a).

S5b: Attack from institutions with lower reputation (Discredit)

This sub scenario is devoted to study the impact of gratuitous attacks from users that are not competing with the target user, aiming at discrediting it with no reasonable intention. Users of 'Institution 2' (an institution with only consumers) start hating 'User 12' with $op=0.1^{42}$, and maintain this behaviour until the end of the simulation. 'Institution 2' is in the lowest position of the ranking and the discredit effect tests the impact of an attack coming from an institution with less fame. Additionally, 10 cycles after the start of the attack 'User 12' increases its quality to $Q=0.6$ during the remaining 20 cycles in an attempt of counteracting the attack. These patterns are shown in Figure IV.4-27 b).

Figure IV.4-27 Quality of resources and opinions patterns by 'User 12' in a) Scenario S5a and b) Scenario S5b for Resources recommendation. In a) the quality of resources pattern for 'Institution 5' (Q (I5)) is also shown.

⁴¹ Note that S0 is considered as starting point, where the quality of all the resources was $Q=0.5$.

⁴² Hating behaviour has been forced to eliminate its random nature and to offer always the same value

S5c: Favours own institution's users (Promotion)

Finally the effect of blind promotion is studied. To study this effect the selected target user must belong to an institution with consumers that can give good opinions to promote him/her and 'User 27' was selected. Users of 'Institution 5' promote 'User 27' to combat a drop in the quality of its resources. 'User 27' is set to be the most recommended user by increasing the quality of 'User 12' results from 0.5 to 0.51 for 5 cycles and then is set back to $Q=0.5$. Later 'User 27' decreases temporarily its quality to $Q=0.49$. Users of 'Institution 5' react to this decrease in the quality by loving 'User 27' with $op=0.9^{43}$ during 10 cycles. These patterns for quality of resources and opinions are shown in Figure IV.4-28.

Figure IV.4-28. Quality of resources and opinions received patterns by 'User 27' in Scenario 5c; for Resources recommendation

IV.4.2.1.6. S6: Hierarchy's effect

As mentioned before in section IV.4.2.3, scenario S6 uses a reduced hierarchy (Figure IV.4-18). A new system structure requires a new starting point for the simulation, as S0 was executed with the global structure and thus is no longer valid for this scenario and new structure. A separate simulation of 40 cycles is executed in two phases of 20 cycles each one, being the first phase the equivalent of S0, a warm-up period to create an initial dispersion in the system. In the second phase the effect of the variation of HDS parameters on reputation's assessment is applied.

Three different effects are studied, **Rivalry inside groups within an institution** (S6a), **Competitiveness inside groups within an institution** (S6b) and **Nepotism within an institution** (S6c). In the first two situations, S6a and S6b, users of different tree in the hierarchy (Rivalry) – 'Users 4,5'- and users of same tree and different branches (Competitiveness) – 'Users 10'- are configured with the *critic*⁴⁴ behaviour to attack 'User 6' by lying with $op=0.25$, resulting in the quality of resources and opinions patterns shown in Figure IV.4-29 a).

⁴³ Loving behaviour has been forced to eliminate its random nature and to offer always the same value

⁴⁴ Opinions for critic behaviour are calculated through the square value of the evaluation: $0.5^2=0.25$

In the latter, *S6c*, the opposite effect is simulated, users of the same tree in the hierarchy and same branches –‘Users 2,8’- are configured with the *praiser*⁴⁵ behaviour to promote ‘User 6’ with $op=0.707$, resulting in the patterns of Figure IV.4-29 b).

Figure IV.4-29. Quality of resources and opinions received patterns by ‘User 6’ in a) Scenarios S6a and S6b and b) Scenario S6c; for Resources recommendation

IV.4.2.1.7. S7: Effect of changes in Reliability values

As seen in section IV.3.7, *Reputation*’ is calculated by aggregating R_{TruQOs} and RL_{TruQOs} to create the resources’ classification offered to the consumer as recommendation. To analyse the reliability, variations in the behaviour of a user need to be forced. This effect is achieved by defining a resources’ quality variation pattern. In this scenario again the reduced hierarchy of Figure IV.4-18. This reduced hierarchy will speed up the simulation without influencing the expected results.

All users’ quality is set to $Q=0.7$ and the simulation will run as many cycles as required to observe the effect of changing the reliability. In the same way as in scenario S6, the first 20 cycles is the warm-up period and NoReputation mode is used. After this 20 cycles, the pattern illustrated by Figure IV.4-30 is forced in ‘User 6’: 1 cycle $Q=0.71$, 1 cycle $Q=0.69$, 1 cycle $Q=0.71$, and so on. This pattern is maintained for 10 cycles and then ‘User6’ is set back to have the same quality as the rest of the users, $Q=0.7$, for the remaining 30 cycles.

The impact of the reliability is analysed by varying the ‘q’ factor used for the calculation of $RL_{Resource}$ and $RL_{ProviderEv}$; the $RL_{Personal}$ parameters and by playing with the quartiles’ weights (q_i) in the aggregation method.

⁴⁵ Opinions for praiser behaviour are calculated through the square root of the evaluation: $\sqrt{0.5}=0.707$

Figure IV.4-30. Quality of resources and opinions received patterns by 'User 6' (equal values) in Scenario S7; for Resources recommendation

IV.4.2.1.8. S8: Newcomers' effect

Scenario S8 evaluates how TruQOs faces the problem of **bootstrapping** and if the model is able to deal with newcomers entering the system. The impact of new users entering the system has been studied for both types of recommendations: considering only *Resources Recommendation* (S8a) and only *Users Recommendation* (S8b). In the first subscenario, *S8a*, 'User 31' generates resources with high quality, $Q=0.9$, following the pattern of Figure IV.4-31 a). In the second, *S8b*, 'User 31' is created offering services with a slightly higher quality than the rest of the users, $Q=0.51$, following the pattern of Figure IV.4-31 b).

For both situations scenario S0 is used again as starting point and an analysis of the ratio of recommendation for 'User 31' will be performed. The simulation is run for 30 cycles. After 10 cycles 'User 31' is created within the hierarchy of Figure IV.4-17, modifying the global system's structure. For each type of recommendation two different situations have been studied, as represented in Figure IV.4-19, *S8a1* and *S8b1* with 'User 31' being created within an existing institution, 'Institution 3'; or being created in a new institution, *S8a2* and *S8b2*.

Figure IV.4-31. Quality of resources pattern by newcomer 'User 31' vs quality of resources of the rest of users (Q (rest)) pattern for a) Resources recommendation in Scenario S8a and b) Users recommendation in Scenario S8b

IV.4.2.1.9. S9: Integrating resources and users recommendation

Previous scenarios worked with only *Resources Recommendation* mode (S0-S8a) or only *Users Recommendation* mode (S8b). In this scenario both recommendation are combined and *Resources and Users Recommendation* mode is applied in two sub scenarios. In both scenarios consumers are configured with the parameter **Resources:Users** to request one cycle a resource recommendation and next cycle a user recommendation.

In the first subscenario, *S9a*, the quality of the resources created by 'User 12' is increased with different values of $Q_{Resource}$, as it can be seen in Figure IV.4-32 a). The effect on the number of direct interactions recommended is studied through an analysis of both the ratio of user recommendation and the ratio of benefits. In *S9b* the best test of the previous subscenario is selected and a new simulation is run. $Q_{Resource}$ is set back to $Q=0.5$ and different values of Q_{Direct} are tested, as shown in Figure IV.4-32 b). The ratio of resources recommendation is studied and compared to the situation $Q_{Resource}=Q_{Direct}$. 'User 12' is used as target user.

Figure IV.4-32. a) Quality of service during direct interactions (Q_{dir}) pattern and qualities of resources patterns by 'User 12' in Scenario S9a; b) Quality of resources (Q_{res}) pattern and qualities of service during direct interactions patterns by 'User 12' in Scenario S9b. Both for Resources and Users recommendation

IV.4.2.1.10. S10: Initial dispersion

In this scenario the initial scenario S0 is repeated with the same resources' generation patterns but using TruQOs and OnlyOpinions as recommendation engines, instead of using a recommendation based on quality assessment (NoReputation). The set of optimal parameters obtained as result of all the tests will be used and the results will be compared with the ones obtained in scenario S0 for each one of the providers in the system.

IV.4.3. Summary of evaluation scenarios

Table IV.4-2 shows a summary of all the scenarios proposed in the evaluation framework, and a more detailed table can be found in the Annex D (section VIII.4) of this document.

Table IV.4-2. Evaluation Framework's Scenarios⁴⁶

TruQOs FUNCTIONALITY TESTS						
Scenario		Name	Recommendation type	Reputation Mode	Cycles	
S0		Initial state	Only Resources	No reputation	20	
S0	S1		Evaluations' effect	Only Resources	No Rep, Only ops, TruQOs	30
	S1a	No Ev Temporal weight	Only Resources	Param. Variation	5+5+10+5+5	
	S1b	Ev Temporal weight	Only Resources	Param. Variation	5+5+10+5+5	
S0	S2		Opinions' effect	Only Resources	No Rep, Only ops, TruQOs	30
	S2a	No Op Temporal weight	Only Resources	Param. Variation	5+5+10+5+5	
	S2b	Op Temporal weight	Only Resources	Param. Variation	5+5+10+5+5	
S0	S3		Combined effect	Only Resources	No Rep, Only ops, TruQOs	30
	S3a	Quality decrease	Only Resources	Param. Variation	5+5+10+5+5	
	S3b	Opinons attack	Only Resources	Param. Variation	5+5+10+5+5	
	S3c	Q +Op variation	Only Resources	Param. Variation	5+5+10+5+5	
S0	S4		Attackers size Effect	Only Resources	Only ops, TruQOs	10+20
S0	S5		Ranking's effect	Only Resources	No Rep, Only ops, TruQOs	30
	S5a	Competition	Only Resources	Param. Variation	10+20	
	S5b	Discredit	Only Resources	Param. Variation	5+5+20	
	S5c	Promotion	Only Resources	Param. Variation	5+5+10+10	
S6		Hierarchy's effect	Only Resources	No Rep, Only ops, TruQOs	40	
	S6a	Rivalry	Only Resources	Param. Variation	20+20	
	S6b	Competitiveness	Only Resources	Param. Variation	20+20	
	S6c	Nepotism	Only Resources	Param. Variation	20+20	
S7		Reliability's effect	Only Resources	No Rep, Only ops, TruQOs	30	
	S7a	q-Factor effect	Only Resources	Param. Variation	10+10+10	
	S7b	RLPersonal effect	Only Resources	Param. Variation	10+10+10	
	S7c	Rep-Rel effect	Only Resources	Param. Variation	10+10+10	
S0	S8		Newcomers' effect	Resources and Users	No Rep, Only ops, TruQOs	30
		S8a	ResRecommendation	Only Resources	Param. Variation	10+10
		S8a1	Existing Institution	Only Resources	Param. Variation	10+10
		S8a2	New Institution	Only Resources	Param. Variation	10+10
		S8b	UserRecommendation	Only Users	Param. Variation	10+10
		S8b1	Existing Institution	Only Users	Param. Variation	10+10
	S8b2	New Institution	Only Users	Param. Variation	10+10	
S0	S9		Direct requests' effect	Resources and Users	No Rep, Only ops, TruQOs	32/48
	S9a	Producer service	Resources and Users	Param. Variation	16+16	
	S9a S9b	Personal service	Resources and Users	Param. Variation	32+16	
S10		Initial dispersion	Only Resources	Only ops, TruQOs	20	

⁴⁶ The table shows a short version of scenarios' names. Longer and more self-descriptive names have been used in the text.

V. RESULTS

This chapter summarizes the main outcomes of the evaluation of TruQOs and the simulation tool. The results were obtained following the framework presented in the previous chapter and presented in a visual way through sets of graphics.

V.1. Introduction

A total of 275 effective tests were performed to evaluate both TruQOs reputation model and the simulation tool presented in this PhD dissertation. The distribution of these tests in the different scenarios of the evaluation framework defined in section IV.4.3 is shown in Table V.1-1. For each of the scenarios the number of new tests and repeated or reused from other scenarios is specified. A considerable amount of tests were discarded due to a bad definition of the tests leading to unexpected or inconclusive results.

Table V.1-1. Evaluation Tests Summary⁴⁷

TruQOs FUNCTIONALITY TESTS					
Scenario		Name	N of Tests		
			Total	New	Repeated
S0		Initial state	1	1	0
S0	S1	Evaluations' effect	6	4	2
	S1a	No Ev Temporal weight	3	3	0
	S1b	Ev Temporal weight	3	1	2
S0	S2	Opinions' effect	18	16	2
	S2a	No Op Temporal weight	9	9	0
	S2b	Op Temporal weight	9	7	2
S0	S3	Combined effect	78	68	10
	S3a	Quality decrease	23	20	3
	S3b	Opinions attack	23	16	7
	S3c	Q +Op variation	32	32	0
S0	S4	Attackers size Effect	12	12	0
S0	S5	Ranking's effect	51	51	0
	S5a	Competition	17	17	0
	S5b	Discredit	17	17	0
	S5c	Promotion	17	17	0
S6		Hierarchy's effect	33	33	0
	S6a	Rivalry	11	11	0
	S6b	Competitiveness	11	11	0
	S6c	Nepotism	11	11	0
S7		Reliability's effect	25	20	5
	S7a	q-Factor effect	7	7	0
	S7b	RLPersonal effect	11	9	2
	S7c	Rep-Rel effect	7	4	3
S0	S8	Newcomers' effect	24	20	4
	S8a	ResRecommendation	12	10	2
	S8a1	Existing Institution	6	6	0
	S8a2	New Institution	6	4	2
	S8b	UserRecommendation	12	10	2
	S8b1	Existing Institution	6	6	0
	S8b2	New Institution	6	4	2
S0	S9	Direct requests' effect	24	24	0
	S9a	Producer service	12	12	0
	S9a S9b	Personal service	12	12	0
S10		Initial dispersion	3	3	0
			275	252	23

⁴⁷ The table shows a short version of scenarios' names. Longer and more self-descriptive names have been used in the text.

The success of a test is measured through the analysis of the ratio of recommendation, as defined in section IV.4.2.5, in two ways:

- a) **Temporal Evolution:** The temporal evolution of the ratio of recommendation is studied to determine the effectiveness of the recommendation mode at combating attacks and/or reflecting changes in the quality of the resources.
- b) **Final Value:** The final value of the ratio of recommendation is taken as the final indicator of success of the recommendation engine. Higher ratios indicate that the target user was more accessed, as the number of created resources is the same for all the tests within the same subscenario.

Thus, to determine the test with the best⁴⁸ results, both the temporal evolution and the final values are analysed; selecting the test(s) with the temporal performance that better matches the desired or expected behaviour and, at the same time, obtains the highest final ratio.

The outcomes of the tests are introduced throughout the following sections, using sets of graphics to provide visual information regarding the effectiveness of TruQOs as a recommendation engine; easier to understand than only numbers.

V.2. TruQOs functionality tests

Each subsection is devoted to a specific scenario. An iterative approach have been followed, with each new scenario introducing a new concept related with the TruQOs assessment of reputation. Tests focus on studying the effect that the variation of the different TruQOs parameters has in the recommendation of resources and/or users. Ratios are multiplied by 10 to facilitate their interpretation in the global graphs, but not in the temporal evolution graphs.

V.2.1. S1: Effect of changes in the quality of resources

A variation on the quality of the results created by a user has a direct impact on the recommendation. This effect is studied in scenario S1. The scenario considers the *Provider's Evaluations Reputation* to calculate the final TruQOs reputation value, as shown in Table V.2-1 and compares the results with NoReputation and OnlyOpinions reputation modes. The quality of the resources created by 'User 12' is varied. The periods where the quality is increased to Q=0.51 are marked in the results figures with a blue square and when the quality is decreased to Q=0.49 are marked with red. During the rest of the simulation the quality is maintained in Q=0.5.

Table V.2-1. Scenario S1: TruQOs' tests description

Test	Description	Parameters
TruQOs	Only Evaluations	k _{Resource} =0; k _{Provider} =1; k _{ProviderEv} =1; k _{ProviderOp} =0

⁴⁸ Best is used with a relative meaning, always compared to the rest of the tests run under the same circumstances, and not with an absolute meaning as the best possible results.

Figure V.2-1 gathers the results of the tests and compares the accessed resources considering only the new resources generated in S1 (series marked with *), or considering all the available resources (including the resources created in S0).

Figure V.2-1. Ratio of Resources Recommendation over simulation time for ‘User 12’ and each of S1 tests, with unweighted evaluations. Comparison between considering all resources vs. only new resources. Marked with colour squares the periods where the quality of ‘User 12’ resources is increased (blue) or decreased (red).

The ratio of recommendation is always lower for new resources and this difference is greater for TruQOs mode than for the other reputation modes. The figure also shows that evaluations-based reputation modes are more sensible to quality variations than opinions. In this sense, TruQOs reputation based on evaluations shows a higher delay in the detection of quality variations in both senses, increasing and decreasing variations. NoReputation detects faster these quality fluctuations and reacts accordingly. Furthermore, TruQOs reputation presents a higher inertia in the recommendation which is clearly presented if all the resources are considered.

The main objective of scenario S1 was to evaluate the impact of temporal weighting to give more weight to recent evaluations of the resources’ quality in the reputation assessment process, as illustrated by Figure V.2-2. For this graph only new resources generated within S1 were considered. For TruQOs reputation, calculated only with the Provider’s Evaluations dimension, the temporal weighting enhances the sensibility to quality increments, provoking an earlier detection of quality changes: both the beginning and ending of the increased quality period is detected 1 cycle earlier if the temporal weighting is considered. The effect of decreasing the quality is similar in both cases, with the aforementioned one cycle difference.

To summarize, temporal weighting achieves higher recommendation ratio during the quality increase periods than no-weighting, with slightly lower ratio after the two periods of quality decrease. Moreover, it also softens the drops in the quality presented in NoReputation mode obtaining a higher final ratio. The OnlyOpinions mode experiments a big delay in the detection of the quality changes, needing an excessive number of cycles and opinions.

Figure V.2-2. Ratio of Resources Recommendation over simulation time for 'User 12' and each of S1 tests. Comparison between considering unweighted vs. weighted evaluations. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) or decreased (red).

A summary of the final ratios obtained in S1 is shown in Figure V.2-3.

Figure V.2-3. S1 Total ratios summary for 'User 12'.

V.2.2. S2: Effect of changes in the consumers' opinions

Scenario S2 studies the effect in the recommendation ratio of changes in the value of the opinions given by the consumers, with special analysis on the impact of temporal weighting of these opinions in the final TruQOs reputation. This scenario considers only the *Provider's Opinions Reputation* to calculate the final TruQOs reputation value, varying the weight given to the different dimensions of reputation based on opinions: *Personal Reputation* and *Social Reputation*, as shown in Table V.2-2.

The scenario also compares the results obtained with TruQOs based tests with NoReputation and OnlyOpinions reputation modes.

Table V.2-2. Scenario S2: TruQOs' tests description

Test	Description	Parameters
TruQOs_1	Only Personal Opinions	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=1; k_{Social}=0;$
TruQOs_2	Social Opinions, $\omega=0.3$	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=0; k_{Social}=1$ - $\omega=0.3$
TruQOs_3	Social Opinions, $\omega=0.5$	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=0; k_{Social}=1$ - $\omega=0.5$
TruQOs_4	Social Opinions, $\omega=0.7$	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=0; k_{Social}=1$ - $\omega=0.7$
TruQOs_5	Personal-Social(50-50), $\omega=0.7$	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=0.5;$ $k_{Social}=0.5$ - $\omega=0.7$
TruQOs_6	Personal-Social(70-30), $\omega=0.7$	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=0.7;$ $k_{Social}=0.3$ - $\omega=0.7$
TruQOs_7	Personal-Social(30-70), $\omega=0.7$	$k_{Resource}=0; k_{Provider}=1; k_{ProviderEv}=0; k_{ProviderOp}=1; k_{Personal}=0.3;$ $k_{Social}=0.7$ - $\omega=0.7$

The first analysis focus in the study of the impact of the parameter ' ω ' in the determination of Social Reputation. Theoretically, as described in section IV.3.3.2.2, this parameter reflects the weight given to the opinion that the consumer has about the witnesses versus the opinion of the witnesses about the provider. As it can be seen in Figure V.2-4, from all the three tests considering only the social reputation (TruQOs_2, TruQOs_3 and TruQOs_4), the one that obtains the highest final ratio of recommendation and detects faster the period of quality increase is equally balancing the opinions with $\omega=0.5$ (TruQOs_3). However, the results are very similar for values giving more importance to one or another of the aforementioned opinions.

Figure V.2-4. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S2 tests. Study of the effect of parameter ' ω ' in the calculation of Social Reputation. . Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple)

A value of $\omega=0.7$ is used in the rest of the tests and in future scenarios. This was decided to protect the consumer from 3rd party opinions making him or her more suspicious and relying more in his or her opinions about the witnesses, but maintaining the information coming from other users as social reputation.

In order to minimize the number of graphs and to ease the reading, Figure V.2-4, shows also the evolution of the ratio of recommendation for the test considering only the personal opinions to assess the reputation (TruQOs_1) and a test where personal and social opinions were used together in the recommendation (TruQOs_7). The combination of personal and social opinions appears as a better strategy than using them individually, obtaining a final recommendation ratio a 23.8% higher than using only personal opinions, and a 132.6% higher than using only social information. Furthermore, TruQOs_7 overcomes in a 9% the results obtained considering only quality indicators to provide the recommendation (NoReputation)

After selecting the value for ' ω ' parameter, the performed analysis mirrors the one done in the previous section for the study of the effect of changes in the quality of resources, analysing first the difference between the ratio considering only new resources or considering all the available resources, as shown in Figure V.2-5, marking again with * the tests where only the resources created in scenario S2 are considered.

Figure V.2-5. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S2 tests. Comparison between considering all resources vs. only new resources. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple).

A similar effect to section V.2.1 was observed, the ratio of recommendation is always lower for new resources and this difference is greater for TruQOs mode than for the other reputation modes. Figure V.2-5 only reflects TruQOs reputation using Personal Reputation to simplify the graph but the aforementioned effect is present in all the seven TruQOs tests (see Table V.2-2).

Again NoReputation is very sensible to a quality increase (blue square period), and it is immune to the periods where users of 'Institutions 1 and 2' criticize⁴⁹ 'User 12', (purple squares over x axis). On the contrary, OnlyOpinions mode is totally permeable to 3rd party opinions and its ratio decreases during the whole simulation. TruQOs reputation based on personal opinions presents a delay detecting the period where the quality was increased to $Q=0.51$ until the simulation cycle 29, but the periods of attack are detected faster, producing flat periods in the ratio, as it can be seen from simulation cycle 33. After this first analysis, the rest of the tests consider only the new resources created within S2.

The last analysis carried out in scenario S2 focuses on the study of the effect of temporal weighting (Figure V.2-6). The figure shows some of the performed tests to simplify the graph and ease the visualization. In those tests, the temporal weighting enhances the ratio of recommendation, fastening the detection of higher quality periods and reducing the effect of malicious opinions. These behaviours are clearly present in TruQOs_1 and TruQOs_5, where the weighted test obtained a final ratio a 46% and 44% higher, respectively. For TruQOs_7, the increment is only of a 4.6%, although it is the test with the highest difference in the detection of the higher quality period, being the weighted version 4 simulation cycles faster.

Figure V.2-6. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S2 tests. Comparison between considering unweighted vs. weighted opinions. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple)

⁴⁹ The critics in the opinions could be motivated by a decrease in the pedagogical quality of the resources, a decrease that it is not reflected in the objective assessment.

A summary of the final resources recommendation ratios obtained for all the tests in scenario S2 is shown in Figure V.2-7. As it can be seen, the benefits of applying a temporal weighting are more visible for the opinions than for the evaluations (Figure V.2-3).

Figure V.2-7. S2 Total ratios summary for 'User 12'.

V.2.3. S3: Changes in quality and opinions combined effect

Scenario S3 combines the effects studied in the first two scenarios. Temporal weighting is used for both resources' quality evaluations and consumers' opinions in the calculation of reputation values.

V.2.3.1. S3a: Quality decrease

Scenario S3a is run under the same conditions of scenario S1 through the tests described in Table V.2-3. Unlike scenario S1 where the final TruQOs reputation was assessed considering only *Provider's Evaluations Reputation*, scenario S3a adds the *Provider's Opinions Reputation* in the calculation. Moreover, the different tests evaluate the different unitary dimensions of reputation separately, as well as their combination.

Table V.2-3. Scenarios S3a-S3b: TruQOs' tests description

Test	Description	Parameters
TruQOs_1	Evaluations	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=1$; $k_{ProviderOp}=0$
TruQOs_2	Personal Opinions	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=1$; $k_{Social}=0$;
TruQOs_3	Social Opinions, $\omega=0.7$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0$; $k_{Social}=1 - \omega=0.7$
TruQOs_4	Personal-Social(50-50), $\omega=0.7$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.5$; $k_{Social}=0.5 - \omega=0.7$
TruQOs_5	Personal-Social(70-30), $\omega=0.7$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.7$; $k_{Social}=0.3 - \omega=0.7$
TruQOs_6	Personal-Social(30-70), $\omega=0.7$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$
TruQOs_7	Evaluations-Personal(50-50)	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=1$; $k_{Social}=0$;
TruQOs_8	Evaluations-Personal(70-30)	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.7$; $k_{ProviderOp}=0.3$; $k_{Personal}=1$; $k_{Social}=0$;
TruQOs_9	Evaluations-Personal(30-70)	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=1$; $k_{Social}=0$;

Test	Description	Parameters
TruQOs_10	Evaluations-Social(50-50), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0$; $k_{Social}=1 - \omega=0.7$
TruQOs_11	Evaluations-Social(70-30), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.7$; $k_{ProviderOp}=0.3$; $k_{Personal}=0$; $k_{Social}=1 - \omega=0.7$
TruQOs_12	Evaluations-Social(30-70), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0$; $k_{Social}=1 - \omega=0.7$
TruQOs_13	Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.5$; $k_{Social}=0.5 - \omega=0.7$
TruQOs_14	Evaluations-Opinions(50-50) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.7$; $k_{Social}=0.3 - \omega=0.7$
TruQOs_15	Evaluations-Opinions(50-50) _Personal-Social(30-70), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$
TruQOs_16	Evaluations-Opinions(70-30) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.7$; $k_{ProviderOp}=0.3$; $k_{Personal}=0.5$; $k_{Social}=0.5 - \omega=0.7$
TruQOs_17	Evaluations-Opinions(70-30) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.7$; $k_{ProviderOp}=0.3$; $k_{Personal}=0.7$; $k_{Social}=0.3 - \omega=0.7$
TruQOs_18	Evaluations-Opinions(70-30) _Personal-Social(30-70), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.7$; $k_{ProviderOp}=0.3$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$
TruQOs_19	Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.5$; $k_{Social}=0.5 - \omega=0.7$
TruQOs_20	Evaluations-Opinions(30-70) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.7$; $k_{Social}=0.3 - \omega=0.7$
TruQOs_21	Evaluations-Opinions(30-70) _Personal-Social(30-70), $\omega(0.7)$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$

Figure V.2-8 shows the temporal evolution of the resources recommendation ratio for a selected set of tests, presenting only the ones with the results that better help with the explanation and obtain the best results, understood as the tests with the temporal performance that better matches the desired or expected behaviour and, at the same time, obtains the highest final ratio, as mentioned in section V.1.

Figure V.2-8. Ratio of Resources Recommendation over simulation time for ‘User 12’ and a subset of S3a tests with weighted evaluations. Only new resources considered. Marked with colour squares the periods where the quality of ‘User 12’ resources is increased (blue) or decreased (red).

As it can be seen, the combination of opinions and evaluations (TruQOs_19, TruQOs_20 and TruQOs_21) is beneficial for the recommendation of 'User 12', obtaining a final ratio 14% higher than using only the evaluations (TruQOs_1). Furthermore, these combinations also overcome the results obtained for NoReputation mode in more than a 9 %.

Figure V.2-9 illustrates the advantages of combining evaluations and opinions showing only the final ratios for considering only new resources or considering all the available resources. For both situations, the ratio is higher for tests on the right side of the graph (TruQOs_7 and higher), corresponding to test where evaluations and opinions were combined, especially when compared to using only opinions (TruQOs_2 to TruQOs_6).

Figure V.2-9. S3a Total ratios summary for 'User 12'.

V.2.3.2. S3b: Opinions-based attacks

Scenario S3b repeats the conditions of scenario S2 but, as in S3a, considering combinations of *Provider's Evaluations Reputation* and *Provider's Opinions Reputation*, for the assessment of the final TruQOs reputation. The main results are shown in Figure V.2-10 for Table V.2-3's tests.

Figure V.2-10. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S3b tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 1,2' criticize 'User 12' (purple)

In this scenario the combination of evaluations and opinions is also beneficial but in some tests the impact of evaluations exceed the desired. That is the case of TruQOs_10 where evaluations are combined with social information coming from other consumers. The obtained inertia from the evaluations is so high that the effect of the increment of the quality of the resources is maintained for all the simulation. Adding personal opinions information to the aforementioned information (evaluations and social information like in TruQOs_10) helps compensating this inertia and reducing the over recommendation, while not being too conditioned by the attacking periods. The figure only shows TruQOs_19 to TruQOs_21 but this compensation mechanism is present in all TruQOs tests that consider all the available dimensions of reputation (from TruQOs_13).

In general, and as illustrated by Figure V.2-9, the combination of evaluations and opinions obtains better results in this scenario than in S2 where only the opinions were used, showing a lower sensibility to the attacks. However, the evaluations have a bigger impact than expected.

Figure V.2-11. S3b Total ratios summary for 'User 12'.

V.2.3.3. S3c: Compensating quality changes through opinions

The final subscenario analyses the combined effects, but in “opposite” directions, of changes in the quality of the resources and changes in consumers’ opinions. In this subscenario the opinions always compensate the changes in the quality, with better opinions to promote a user when the quality of the resources decreases, and vice versa. To study this promotion effect, and as indicated in section IV.4.2.6, the selected user is now ‘User 27’.

All TruQOs reputation dimensions are finally used, with the addition for the first scenario of the *Resource Reputation*. The resulting set of tests is described in Table V.2-4.

Table V.2-4. Scenario S3c: TruQOs’ tests description

Test	Description	Parameters
TruQOs_1	Evaluations	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=1$; $k_{ProviderOp}=0$
TruQOs_2	Personal-Social(30-70), $\omega=0.7$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$
TruQOs_3	Resource	$k_{Resource}=1$; $k_{Provider}=0$
TruQOs_4	Resource-Evaluations(50-50)	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=1$; $k_{ProviderOp}=0$
TruQOs_5	Resource-Evaluations(70-30)	$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=1$; $k_{ProviderOp}=0$
TruQOs_6	Resource-Evaluations(30-70)	$k_{Resource}=0.3$; $k_{Provider}=0.7$; $k_{ProviderEv}=1$; $k_{ProviderOp}=0$
TruQOs_7	Resource-Opinions(50-50) _Personal-Social(30-70)_ $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$

Test	Description	Parameters
TruQOs_8	Resource-Opinions(70-30) _Personal-Social(30-70), $\omega(0.7)$	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0; k_{ProviderOp}=1;$ $k_{Personal}=0.3; k_{Social}=0.7 - \omega=0.7$
TruQOs_9	Resource-Opinions(30-70) _Personal-Social(30-70), $\omega(0.7)$	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0; k_{ProviderOp}=1;$ $k_{Personal}=0.3; k_{Social}=0.7 - \omega=0.7$
TruQOs_10	Resource-Evaluation(50-50) _Evaluations-Personal(30-70)	$k_{Resource}=0.5; k_{Provider}=0.5; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=1; k_{Social}=0$
TruQOs_11	Resource-Evaluation(70-30) _Evaluations-Personal(30-70)	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.7;$ $k_{ProviderOp}=0.3; k_{Personal}=1; k_{Social}=0$
TruQOs_12	Resource-Evaluation(30-70) _Evaluations-Personal(30-70)	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0.3;$ $k_{ProviderOp}=0.7; k_{Personal}=1; k_{Social}=0$
TruQOs_13	Resource-Evaluations(50-50) _Evaluations-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5; k_{Provider}=0.5; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0; k_{Social}=1 - \omega=0.7$
TruQOs_14	Resource-Evaluations(70-30) _Evaluations-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0; k_{Social}=1 - \omega=0.7$
TruQOs_15	Resource-Evaluations(30-70) _Evaluations-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0; k_{Social}=1 - \omega=0.7$
TruQOs_16	Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5; k_{Provider}=0.5; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$
TruQOs_17	Resource-Evaluations(70-30) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$
TruQOs_18	Resource-Evaluations(30-70) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$
TruQOs_19	Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.5; k_{Provider}=0.5; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0.7; k_{Social}=0.3 - \omega=0.7$
TruQOs_20	Resource-Evaluations(70-30) _Evaluations-Opinions(50-50) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0.7; k_{Social}=0.3 - \omega=0.7$
TruQOs_21	Resource-Evaluations(30-70) _Evaluations-Opinions(50-50) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0.5;$ $k_{ProviderOp}=0.5; k_{Personal}=0.7; k_{Social}=0.3 - \omega=0.7$
TruQOs_22	Resource-Evaluations(50-50) _Evaluations-Opinions(70-30) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.5; k_{Provider}=0.5; k_{ProviderEv}=0.7;$ $k_{ProviderOp}=0.3; k_{Personal}=0.7; k_{Social}=0.3 - \omega=0.7$
TruQOs_23	Resource-Evaluations(70-30) _Evaluations-Opinions(70-30) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.7;$ $k_{ProviderOp}=0.3; k_{Personal}=0.7; k_{Social}=0.3 - \omega=0.7$
TruQOs_24	Resource-Evaluations(30-70) _Evaluations-Opinions(70-30) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0.7;$ $k_{ProviderOp}=0.3; k_{Personal}=0.7; k_{Social}=0.3 - \omega=0.7$
TruQOs_25	Resource-Evaluations(50-50) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5; k_{Provider}=0.5; k_{ProviderEv}=0.3;$ $k_{ProviderOp}=0.7; k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$
TruQOs_26	Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.3;$ $k_{ProviderOp}=0.7; k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$
TruQOs_27	Resource-Evaluations(30-70) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.3; k_{Provider}=0.7; k_{ProviderEv}=0.3;$ $k_{ProviderOp}=0.7; k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$

Test	Description	Parameters
TruQOs_28	Resource-Evaluations(50-50) _Evaluations-Opinions(30-70) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.7$; $k_{Social}=0.3$ - $\omega=0.7$
TruQOs_29	Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.7$; $k_{Social}=0.3$ - $\omega=0.7$
TruQOs_30	Resource-Evaluations(30-70) _Evaluations-Opinions(30-70) _Personal-Social(70-30), $\omega(0.7)$	$k_{Resource}=0.3$; $k_{Provider}=0.7$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.7$; $k_{Social}=0.3$ - $\omega=0.7$

The main results are shown in Figure V.2-12 and a summary in Figure V.2-13. Figure V.2-12 shows that considering only the Resource reputation (TruQOs_3) is unable to detect any change in the quality of the resources, probably because needs more simulation cycles to have enough information. Adding evaluation information (TruQOs_5) or opinions' information about the provider (TruQOs_9) solves this lack of information. The evaluations offer a faster detection of the periods where a change in resources' quality is modelled: TruQOs_5 is a 19.4 % faster than TruQOs_9 detecting the high quality period, and this delay makes TruQOs_9 unable to detect the quality decrease period.

Figure V.2-12. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S3c tests with weighted evaluations and opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue) or decreased (red); and the periods where users of 'Institutions 1,2' criticize (purple) or praise (orange) 'User 12'.

Figure V.2-13. S3c Total ratios summary for 'User 12'.

The combination of all types of information sources: providers' evaluations, opinions about the provider and opinions about the resource is, as expected, the solution offering the best results, as it can be seen for TruQOs_16 and, especially, for TruQOs_27, the only tests that detect the quality increment almost immediately, and also the promotion period, being even too sensible to third party opinions.

V.2.4. S4: Number of attackers' effect

The first three scenarios evaluated the influence of the main reputation sources, testing their impact on the final recommendation through the variation of the main parameters used to calculate a final TruQOs reputation value. These parameters are related to the weight that the different reputation dimensions have in the recommendation, and determine the global behaviour of the trust and reputation model, configuring it to be more or less sensible to variations in the quality of the resources or, on the contrary, to be more or less susceptible to other users' opinions. Or even to trust more in consumer's own past experiences.

Before starting with the analysis of the secondary parameters, scenario S4 analyses the impact or effect of varying the number of users attacking the target user, 'User 12' for this scenario. The main goal of the scenario is determining the robustness of TruQOs at combating attacks, compared to traditional approaches based on opinions, OnlyOpinions reputation mode.

This scenario aims as well at confirming the global behaviours observed in the previous scenarios regarding the effects of giving more importance to each reputation dimension against the others, and how some reputation dimensions are more effective against attacks than others. As shown in Table V.2-5, two different TruQOs configurations were tested in four situations, with an increasing number of attackers on each test.

Table V.2-5. Scenario S4: TruQOs' tests description

Test	Institutions attacking	Number of attackers
OnlyOp_1	TruQOs_1	Institution 1
OnlyOp_2	TruQOs_2	Institutions 1 and 2
OnlyOp_3	TruQOs_3	Institutions 1, 2 and 4
OnlyOp_4	TruQOs_4	Institutions 1, 2, 4 and 5
Subtest	Description	Parameters
A	Resource-Evaluations(30-70) _Evaluations-Opinions(70-30) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.3$; $k_{Provider}=0.7$; $k_{ProviderEv}=0.7$; $k_{ProviderOp}=0.3$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$
B	Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$

Figure V.2-14 shows the results for the best and worst situations (only 2 attackers versus all consumers attacking, 15 attackers) for TruQOs and OnlyOpinions tests. The tests using TruQOs as recommendation engine obtain a better protection to attacks than the traditional approach, although, the difference between the configuration TruQOs_4B and OnlyOpinions_4, with all the consumers attacking 'User 12', is insignificant: a ratio of recommendation of 0.72 for TruQOs_4B against a ratio of 0.70 for OnlyOpinions_4. For the same situation, TruQOs_4A overcomes their final recommendation ratio in a 22% and 25% respectively.

As for the best situation, where only two consumers attack ‘User 12’, TruQOs_1A clearly overcomes the rest, being the only test that successfully obtains a global protection against the attacks and detects the increase in the quality of the resources created by ‘User 12’ and starts recommending this user by the end of scenario, from simulation cycle 45.

Figure V.2-14. Ratio of Resources Recommendation over simulation time for ‘User 12’ for S4 tests with weighted evaluations and opinions. Marked with colour squares the periods where the quality of ‘User 12’ resources is increased (blue) and the periods where other users criticize (purple) ‘User 12’.

A summary of the final recommendation ratio for all S4 tests is shown in Figure V.2-15, where the impact of increasing the number of attackers is clearly visible, for all the modes.

Figure V.2-15. S4 Total ratios summary for ‘User 12’.

V.2.5. S5: Effect of Institution’s position in the Ranking

The following scenarios, starting from this one, focus on the study of the effect of secondary parameters: those responsible for obtaining a fine tuning of the model and better adapting the global behaviours obtained with the selection of the main parameters to the desired behaviour expected for the whole recommendation engine.

The first parameter to study is one of the *Hierarchy Dependence Strength (HDS)* parameters, **hds_NONE** as defined in section IV.3.4.3, and the effect of the position of institutions in the Institutions' Ranking. This effect is studied in three different subscenarios where users of different institutions attack or promote the target user, 'User 12' for *S5a* and *S5b* and 'User 27' for *S5c*. In all the situations, the number of attackers or promoters is the same, 5, a 33.33% of all the consumers in the system, and the different TruQOs configurations used in the tests are described in Table V.2-6.

Table V.2-6. Scenario S5: TruQOs' tests description

Test	Description	Parameters
TruQOs_1	Personal-Social(30-70), $\omega=0.7$	$k_{Resource}=0$; $k_{Provider}=1$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$
TruQOs_2	Resource-Opinions(70-30) _Personal-Social(30-70)_ $\omega(0.7)$	$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0$; $k_{ProviderOp}=1$; $k_{Personal}=0.3$; $k_{Social}=0.7 - \omega=0.7$
TruQOs_3	Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.5$; $k_{Social}=0.5 - \omega=0.7$
Subtest	Description	Parameters
A	No Ranking computation	$c=5$, $\varphi=1000$
B	Ranking computation	$c=2.5$, $\varphi=1 - hds_NONE=-1$
C		$c=2.5$, $\varphi=1 - hds_NONE=-0.5$
D		$c=2.5$, $\varphi=1 - hds_NONE=0.5$
E		$c=2.5$, $\varphi=1 - hds_NONE=1$

V.2.5.1. S5a: Attack from well reputed institutions (Competition)

The first subscenario, *S5a*, evaluates the impact of attacks between competitor institutions. Users of an Institution with a higher position in the ranking start attacking to users of another Institution that offer similar resources and to whom a competence relationship is established. Users of 'Institution 5' were set to offer higher quality resources ($Q=0.51$) during 10 cycles, positioning their institution as the best one in the Ranking. After those 10 cycles the quality of the users is set back to $Q=0.5$ and they start attacking 'User 12' by hating him with fixed low opinions, $op=0.1$. 'User 12' increases the quality of the generated resources to $Q=0.55$.

The effect of **hds_NONE** parameter is studied through variations in its value while dimensions' weights were left unvaried. A comparison with whether or not considering the ranking or hierarchy relationships on social reputation was performed. The results for a selected set of tests can be seen in Figure V.2-16, while the summary of the final result of all the tests is illustrated with Figure V.2-17.

Variations in **hds_NONE** parameter allow configuring TruQOs with a higher sensibility to quality increments, while enhancing the protection against the attack from other users, reaching a higher final recommendation ratio than OnlyOpinions mode. NoReputation mode, considering only the evaluation of the resources overcomes any other test.

Figure V.2-16. Ratio of Resources Recommendation over simulation time for ‘User 12’ and a subset of S5a tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of ‘User 12’ resources is increased (blue); and the periods where users of ‘Institutions 5’ hate ‘User 12’ (purple)

Figure V.2-17. S5a Total ratios summary for ‘User 12’.

V.2.5.2. S5b: Attack from institutions with lower reputation (Discredit)

This second subscenario continues with the study of the impact of institutions’ position in the Ranking and the effect of the *Hierarchy Dependence Strength (HDS)* parameter **hds_NONE**. The attacker now is a consumers’ Institution with no other intention than discrediting ‘User12’. The behaviour of the users of ‘Institution 2’ is not motivated by the will of promoting one of the users in their Institution as resource provider and can only be understood as a malicious attack. ‘Institution 2’ is in the lowest position of the ranking while in S5a the attackers occupied the highest position.

The results for a selected set of tests are represented in the graph of Figure V.2-18.

Figure V.2-18. Ratio of Resources Recommendation over simulation time for 'User 12' and a subset of S5b tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 12' resources is increased (blue); and the periods where users of 'Institutions 2' hate 'User 12' (purple)

Here again variations in hds_NONE parameter allow a faster defence against attacks, although it is not as effective as in scenario S5a. The increase in the quality of the resources is indeed detected by TruQOs_3A and TruQOs_3B in simulation cycles 33 and 40. However, they succumb to the attack and the ratio of recommendation decreases until the end of the simulation. Nevertheless, the final recommendation ratio is still higher for TruQOs than for OnlyOpinions mode. As for the increase in the recommendation ratio observed for TruQOs_3C in simulation cycle 26 no satisfactory explanation was found.

Figure V.2-19 shows a summary of the final ratio of recommendations for all the tests in the scenario. As it can be seen, and as it was expected, the highest ratio is achieved for the NoReputation mode. This mode considers only quality indicators to provide recommendation and, thus, it is not affected by the attacks, offering the best results when the quality of the resources of a user is higher than the quality of the resources created by the rest.

Figure V.2-19. S5b Total ratios summary for 'User 12'.

V.2.5.3. S5c: Favouring own institution's users (Promotion)

This subscenario evaluates the self-promotion carried out by users of an institution when one of its users decreases the quality of the results. For this scenario the user selected for the study was 'User 27' ('User 12' belongs to a providers' institution and no promotion can be simulated). The first 5 cycles were used to position 'User 27' as the most recommended, by increasing temporally the quality of his/her resources to $Q=0.51$. Later, the quality is decreased to $Q=0.49$ for 10 cycles while users of 'Institution 5' promote him/her giving abnormally high opinions, $op=0.9$. The tests' results are gathered in Figure V.2-20.

Figure V.2-20. Ratio of Resources Recommendation over simulation time for 'User 27' and a subset of S5c tests with weighted opinions. Only new resources considered. Marked with colour squares the periods where the quality of 'User 27' resources is increased (blue) or decreased (red); and the periods where users of 'Institutions 5' love 'User 27' (orange)

As in scenarios S5a and S5b, variations in `hds_NONE` help setting TruQOs to detect and to protect the system against these misbehaviours related to the subjective nature of the opinions. The global summary of the recommendation ratios is shown in Figure V.2-21 confirming what it was discovered in the other subscenarios: an adequate selection of `hds_NONE` parameter allows overcoming OnlyOpinions mode.

Figure V.2-21. S5c Total ratios summary for 'User 27'.

V.2.6. S6: Hierarchy's effect

Scenario S6 focus on the study of the effect of the position of consumers and providers within the internal hierarchy of the institutions. This effect is evaluated in three different subscenarios, each one devoted to testing one of the *Hierarchy Dependence Strength (HDS)* parameters: **hds_LOW**, **hds_MEDIUM** and **hds_HIGH**. These parameters, introduced in section IV.3.4.3, determine the existing structural relation between two users of the system, usually the consumer and the provider, and it is used to assess the *Structural Reputation*.

The design of the three subscenarios continues with the iterative approach followed for the whole evaluation framework, and the result of each subscenario is used in the following ones. To simplify the execution and facilitate the analysis only one Institution is used (see section IV.4.2.3). The selection of a new reduced system structure forces the selection of a new warm-up period in substitution of scenario S0, as the results obtained in S0 consider the global structure and can't be extrapolated. After that warm-up period (comprising 20 cycles), a 20 cycles' simulation is run where different behaviours when giving opinions are applied and their effects are analysed. 'User 6' was selected as the subject of evaluation for his/her position in the middle of the hierarchy tree, with both ascendant and descendant users.

V.2.6.1. S6a: Rivalry inside groups within an institution

Rivalry is defined within the TruQOs' evaluation framework as the competition relation between users of the same institution but belonging to different trees in the internal hierarchy. To test the effect of this relation the consumers of the second tree of the hierarchy: 'Users 4,5' were configured to start criticizing 'User 6' by giving low opinions during the last 10 cycles of the simulation. The simulation was performed using a fixed configuration of the main parameters of TruQOs and a variation of HDS parameters, with special focus on the parameter **hds_LOW**, as described in Table V.2-7.

Table V.2-7. Scenario S6a: TruQOs' tests description

Test	Description	Parameters	
Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$		$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$	
TruQOs_1	No Hierarchy computation	$c=5$, $\varphi=1000$	
TruQOs_2	Hierarchy computation	$c=2.5$, $\varphi=1$	$hds_LOW=-0.5$; $hds_MEDIUM=0$; $hds_HIGH=0$
TruQOs_3			$hds_LOW=0$; $hds_MEDIUM=0$; $hds_HIGH=0$
TruQOs_4			$hds_LOW=1$; $hds_MEDIUM=0$; $hds_HIGH=0$
TruQOs_5			$hds_LOW=1.5$; $hds_MEDIUM=0$; $hds_HIGH=0$
TruQOs_6			$hds_LOW=-0.5$; $hds_MEDIUM=1$; $hds_HIGH=1$
TruQOs_7			$hds_LOW=0$; $hds_MEDIUM=1$; $hds_HIGH=1$
TruQOs_8			$hds_LOW=1$; $hds_MEDIUM=1$; $hds_HIGH=1$
TruQOs_9			$hds_LOW=1.5$; $hds_MEDIUM=1$; $hds_HIGH=1$

Figure V.2-22 gathers the results of a selected set of the tests. A comparison between not considering hierarchy information and different values of the **hds_LOW** parameter is shown in the graph. Most of the **hds_LOW** parameters obtain worse results than not using hierarchy

information. But one of the tests success at combating the attack and obtains the best results among all the tests, TruQOs_3 with hds_LOW=0. The ratio of recommendation for this test does not show a decrease motivated by the attack. Furthermore, it even presents an increase from simulation cycle 37. Additionally, two of the tests with varied hds_LOW parameter (TruQOs_3 and TruQOs_4), along with not using hierarchy information (TruQOs_1), overcome the results of NoReputation and OnlyOpinions modes.

Figure V.2-22. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S6a tests with weighted opinions. Marked with colour squares the period where 'Users 4,5' lye about 'User 6' (purple) giving bad opinions.

All the results of the graph were obtained considering only new resources created after the warm-up period. With all the resources of S5a the results are similar with an offset, as it can be appreciated in Figure V.2-23.

Figure V.2-23. S6a Total ratios summary for 'User 6'.

V.2.6.2. S6b: Competitiveness inside groups within an institution

Scenario S6b analysed the competitiveness between users, understood as rivalry between users of the same institution and the same tree in the internal hierarchy, but belonging to different branches with a common ascendant. One example of this type of relationship is the possible rivalry between siblings within a family. Competitiveness effect is obtained by configuring ‘User 10’ to criticize ‘User 6’, simulated during the last 10 cycles. The performed analysis is similar to the one in S6a with the tests of Table V.2-8.

Table V.2-8. Scenario S6b: TruQOs’ tests description

Test	Description	Parameters
Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$		$k_{Resource}=0.7; k_{Provider}=0.3; k_{ProviderEv}=0.3; k_{ProviderOp}=0.7;$ $k_{Personal}=0.5; k_{Social}=0.5 - \omega=0.7$
TruQOs_1	No Hierarchy computation	$c=5, \varphi=1000$
TruQOs_2	Hierarchy computation	$c=2.5, \varphi=1$
TruQOs_3		
TruQOs_4		
TruQOs_5		
TruQOs_6		
TruQOs_7		
TruQOs_8		
TruQOs_9		

The results for this scenario are depicted in Figure V.2-24 for a selected set of tests with fixed $hds_LOW=0$ from S6a and varying hds_MEDIUM .

Figure V.2-24. Ratio of Resources Recommendation over simulation time for ‘User 6’ and a subset of S6b tests with weighted opinions. Marked with colour squares the period where ‘User 10’ lye about ‘User 6’ (purple) giving bad opinions.

One of the hds_MEDIUM values (TruQOs_9) clearly overcomes the rest, being the only test that succeeded at combating the critics attack. NoReputation mode is immune to opinions attack but incapable of increasing the ratio of recommendation in a scenario with all the users offering resources with the same quality. Three of the hds_MEDIUM values (TruQOs_2, TruQOs_7 and TruQOs_8), and also the test not considering hierarchy information (TruQOs_1) offer very poor results: the final ratio of recommendation is low and they do not offer any kind of defence against attacks. Figure V.2-25 supports for all the tests the conclusions obtained in the analysis of the selected set of tests.

Figure V.2-25. S6b Total ratios summary for 'User 6'.

V.2.6.3. S6c: Nepotism within an institution

The scenario S6c models the opposite behaviour of the one in scenario S6b, nepotism, defined as the practice of favouring relatives or friends. To model this behaviour, 'Users 2,8' in the same tree and branch of 'User 6', promote him/her giving abnormally high opinions.

Parameter hds_LOW was set to 0 and hds_MEDIUM was varied between 1.5 and 2 to test possible differences with the results obtained during S6b. As for the previous S6 subscenarios different test not considering hierarchy information and varying hds_HIGH were conducted, as described in Table V.2-9.

Table V.2-9. Scenario S6c: TruQOs' tests description

Test	Description	Parameters	
Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$		$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$	
TruQOs_1	No Hierarchy computation	$c=5$, $\varphi=1000$	
TruQOs_2	Hierarchy computation	$c=2.5$, $\varphi=1$	$hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=0.5$
TruQOs_3			$hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=1$
TruQOs_4			$hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2$
TruQOs_5			$hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$
TruQOs_6			$hds_LOW=0$; $hds_MEDIUM=2$; $hds_HIGH=0.5$
TruQOs_7			$hds_LOW=0$; $hds_MEDIUM=2$; $hds_HIGH=1$
TruQOs_8			$hds_LOW=0$; $hds_MEDIUM=2$; $hds_HIGH=2$
TruQOs_9			$hds_LOW=0$; $hds_MEDIUM=2$; $hds_HIGH=2.5$

The results of the tests are shown in Figure V.2-26 with better results for tests varying the target parameter than previous scenarios, obtaining better final recommendation ratios. Three tests (TruQOs_3, TruQOs_4 and TruQOs_5) give better results than not considering hierarchy information (TruQOs_1), although two of them (TruQOs_4 and TruQOs_5) present a high sensibility to the influence of promotion opinions. This sensibility may not be the desired effect, and the selection of the adequate hds_HIGH parameter must be done according to the desired behaviour for the system. A summary of all the tests is shown in Figure V.2-27.

Figure V.2-26. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S6c tests with weighted opinions. Marked with colour squares the period where 'Users 2,8' lye about 'User 6' (purple) giving good opinions.

Figure V.2-27. S6c Total ratios summary for 'User 6'.

V.2.7. S7: Effect of changes in Reliability values

All the previous scenario focus on the study of the impact of the different parameters and dimensions in the final TruQOs reputation value. However, as it was defined in Chapter IV, the recommendation considers the reputation and the reliability to perform an ordered classification of resources. Scenario S7 evaluates the impact of the different parameters used by TruQOs to assess the reliability of the reputations, as well as the parameters for the aggregation of reputation and reliability (section IV.3.7). For execution time and complexity reasons, this scenario uses again the same reduced structure of S6 with only one Institution and with 'User 6' as the subject of evaluation. This structure reduces the execution time but requires more simulation cycles to visualize the desired effects. Despite increasing the number of cycles, the global time needed for every test is still lower than for the same test with the whole structure.

Reliability changes are obtained by making quality to fluctuate temporally between two values. 60 simulation cycles were executed with all users' quality set to $Q=0.7$. The first 20 cycles the system was configured in NoReputation mode, as warm-up period. After the warm-up period, the variation pattern is forced in 'User 6': 1 cycle $Q=0.71$, 1 cycle $Q=0.69$, 1 cycle $Q=0.71$, and so on. This pattern is maintained for 10 cycles (marked in pink in the figures) and then 'User 6' is set back to have the same quality as the rest of the users, $Q=0.7$, for the remaining 30 cycles.

For all the different subscenarios, TruQOs-based tests were compared with NoReputation and OnlyOpinions modes. Both modes don't consider reliability of the calculated reputation for the assessment of the final recommendation. Thus, the same results were used for the three situations described in the subscenarios of S7. As it will be seen in the following sections, the best results are always obtained for NoReputation mode, being the mode with highest sensibility to quality changes, as it is seen in the pink area through the ramp shape. Results with OnlyOpinions mode are poor in comparison with NoReputation but similar to the ones achieved with TruQOs.

V.2.7.1. S7a: Resource's opinions and evaluations variability (q-Factor)

The first subscenario analyses the effect of the parameter dubbed **q-Factor** in the calculation of $RL_{Resource}$ and $RL_{ProviderEv}$ with equations (IV.3-2) and (IV.3-6). This 'q' parameter allows modulating the impact in the reliability of deviations in the values of the opinions given to a resource and deviations in the quality of the resource; and by extension in the evaluation values. Five different values of this parameter were tested (Table V.2-10).

Table V.2-10. Scenario S7a: TruQOs' tests description

Test	Parameters
Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $hds_NONE=-1$; $hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$
TruQOs_1	$q=1$
TruQOs_2	$q=0.25$
TruQOs_3	$q=0.5$
TruQOs_4	$q=2$
TruQOs_5	$q=4$

The results are shown in Figure V.2-28 for the temporal evolution of the ratio of recommendation, and in Figure V.2-29 as a summary of final ratios of recommendation. The effects of variations in the reliability show a great delay in the opinions: 22 simulation cycles for the fastest test (TruQOs_1) and 32 for the slowest (TruQOs_5). This lag is present in all the tests except the NoReputation, due to the TruQOs' parameters combination selected, where opinions through Resource and Opinions dimensions have a prevailing role. The lack of immediacy of opinions-based reputations was previously detected and conditioned the selection of a longer experimental period. Despite this delay, the effect of variations in the target parameter can be seen in the graph. The best outcomes are obtained for a value of $q=4$ (TruQOs_5), whose ratio of recommendation overcomes the results for OnlyOpinions mode, but far from NoReputation mode: the final ratio of recommendation is a 30% lower.

Figure V.2-28. Ratio of Resources Recommendation over simulation time for 'User 6' and each of S7a tests with weighted evaluations. Marked with colour squares the period where the quality of the resources created by 'User 6' is varied (pink).

Figure V.2-29. S7a Total ratios summary for 'User 6'.

To summarize, all the tests based on TruQOs except TruQOs_2, obtain better final results than the traditional approaches based on opinions (OnlyOpinions mode), although at the expense of experiencing a higher delay. This delay is not present in NoReputation mode, offering the best results among all the tests performed in this scenario.

V.2.7.2. S7b: Personal past experiences variability

The second effect to be studied was the impact of the **intimacy** ('itm') and ' **μ** ' ('mu') parameters in the calculation of the reliability of the reputation based on past personal experiences, $RL_{Personal}$ (as presented in section IV.3.3.2.1). Personal reputation and reliability, as well as the parameters for their assessment, were taken from ReGreT (Sabater & Sierra 2001a; Sabater & Sierra 2002b) which was properly evaluated (Sabater 2004a).

In this subscenario, 'itm' and ' μ ' are evaluated again through nine combinations of the two parameters, as shown in Table V.2-11. The objective of these tests is to measure the impact of the aforementioned parameters in the final TruQOs reputation value.

Table V.2-11. Scenario S7b: TruQOs' tests description

Test	Parameters
Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $hds_NONE=-1$; $hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$ - $q=4$
TruQOs_1	itm=3, $\mu=0.3$
TruQOs_2	itm=3, $\mu=0.5$
TruQOs_3	itm=3, $\mu=0.7$
TruQOs_4	itm=5, $\mu=0.3$
TruQOs_5	itm=5, $\mu=0.5$
TruQOs_6	itm=5, $\mu=0.7$
TruQOs_7	itm=7, $\mu=0.3$
TruQOs_8	itm=7, $\mu=0.5$
TruQOs_9	itm=7, $\mu=0.7$

The temporal evolution of the ratio of resources recommendation is illustrated by the graph in Figure V.2-30.

The best results, understood as the tests that obtain a higher final recommendation ratio, are obtained with $\mu=0.3$ (TruQOs_1, TruQOs_4 and TruQOs_7). A value of $\mu=0.3$ gives more weight to the number of impressions (N_i), that is to the familiarity between the consumer and the provider, than to the deviation in the values (D_i).

The intimacy level indicates a close relation between consumer and provider and more experiences will not increase the reliability of consumer's opinion from then on. Higher values mean that close relation state is harder to achieve in the system. In the performed tests, for the same value of $\mu=0.3$, higher values, itm=7, create earlier detection of the quality variations: simulation cycle 34 for TruQOs_7 vs simulation cycle 37 for TruQOs_4; although the final results for itm=5 are slightly better: TruQOs_4 final ratio overcomes in a 3% the ratio for TruQOs_7. These final ratios are easier observed in Figure V.2-31.

Figure V.2-30. Ratio of Resources Recommendation over simulation time for 'User 6' and a subset of S7b tests with weighted evaluations. Marked with colour squares the period where the quality of the resources created by 'User 6' is varied (pink).

Figure V.2-31. S7b Total ratios summary for 'User 6'.

V.2.7.3. S7c: Quartiles Aggregation Method values variability

This final reliability scenario presents the results related to the **reputation-reliability aggregation** method to calculate the final *Reputation'* and to provide the recommendation, as explained in section IV.3.7. Four different combinations of ' q_i ' parameters were used with all other TruQOs parameters set from previous tests, as shown in Table V.2-12. Each ' q_i ' value is the weight given to each quartile in the calculation of the final *Reputation'* value.

Figure V.2-32 collects the temporal results for all the tests while the summary of the final ratios is shown in Figure V.2-33. The best results amongst all TruQOs tests, the ones with the higher ratio of recommendation, are obtained with the combination $q_1=1$, $q_2=0.9$, $q_3=0.7$ and $q_4=0.5$ (TruQOs_2), which was used in all the previous scenarios. The results for this combination are followed by the results obtained with $q_1=1$, $q_2=0.75$, $q_3=0.5$ and $q_4=0.25$ (TruQOs_4), with a ratio of recommendation a 13.35% lower. Both combinations overcome the results for OnlyOpinions

test as well as the other two q_i combinations (TruQOs_1 and TruQOs_3); but are far from NoReputation mode (47% lower for TruQOs_2).

Table V.2-12. Scenario S7c: TruQOs' tests description

Test	Parameters
Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $hds_NONE=-1$; $hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$ - $q=4$ - $itm=5$, $\mu=0.3$
TruQOs_1	$q1(1)-q2(1)-q3(1)-q4(1)$
TruQOs_2	$q1(1)-q2(0.9)-q3(0.7)-q4(0.5)$
TruQOs_3	$q1(1)-q2(0.8)-q3(0.6)-q4(0.4)$
TruQOs_4	$q1(1)-q2(0.75)-q3(0.5)-q4(0.25)$

Figure V.2-32. Ratio of Resources Recommendation over simulation time for 'User 6' and each of S7c tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of 'User 6' is varied (pink).

Figure V.2-33. S7c Total ratios summary for 'User 6'.

V.2.8. S8: Newcomers' effect

How to deal with newcomers in the system is one of the main issues that every trust and reputation model must face. TruQOs implements two techniques, one for each type of recommendation (see section IV.3.6). Scenario S8 evaluates these two techniques and the impact of variations in the **promotion period parameter**, '**prom**'. Two situations were studied: a) 'User 31' is created within an existing institution, 'Institution 3'; and b) the new user, 'User 31', is created in a new institution. In both situations 'User 31' enters the system generating resources with a slightly higher quality ($Q=0.51$) and four tests were run (Table V.2-13).

Table V.2-13. Scenario S8: TruQOs' tests description

Test	Parameters
Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{\text{Resource}}=0.5$; $k_{\text{Provider}}=0.5$; $k_{\text{ProviderEv}}=0.5$; $k_{\text{ProviderOp}}=0.5$; $k_{\text{Personal}}=0.5$; $k_{\text{Social}}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $\text{hds_NONE}=-1$; $\text{hds_LOW}=0$; $\text{hds_MEDIUM}=1.5$; $\text{hds_HIGH}=2.5$ - $q=0.25$ - $\text{itm}=5$, $\mu=0.3$ - $q1(1)$ - $q2(0.9)$ - $q3(0.7)$ - $q4(0.5)$
TruQOs_1	prom=0
TruQOs_2	prom=1
TruQOs_3	prom=3
TruQOs_4	prom=5

V.2.8.1. S8a: Resource Recommendation

To measure the effectiveness of the promotion period technique, the number of accessed resources is studied. The results are shown in Figure V.2-34 for 'User 31' created in an existing institution and in Figure V.2-35 for 'User 31' created in a new institution.

Figure V.2-34. Accessed Resources over simulation time for 'User 31' and each of S8a1 tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of 'User 31' is increased (blue); and when 'User 31' was still unborn (grey). Newcomer in 'Institution 3'.

Figure V.2-35. Accessed Resources over simulation time for ‘User 31’ and each of S8a2 tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of ‘User 31’ is increased (blue); and when ‘User 31’ was still unborn (grey). Newcomer created in a new institution.

As expected NoReputation offers the best results for Resources Recommendation, motivated by its sensibility to quality variations. As for TruQOs, the promotion period allows ‘User 31’ to be recommended at the beginning and obtain good opinions. Moreover, these opinions help in a later recommendation. Increasing promotion period increases the number of resources accessed and the final ratio, as observed in Figure V.2-36 and Figure V.2-37.

Figure V.2-36. S8a1 Total ratios summary for ‘User 31’. Newcomer in ‘Institution 3’.

Figure V.2-37. S8a2 Total ratios summary for ‘User 31’. Newcomer created in a new institution.

Not considering promotion period, as for $p=0$ (TruQOs_1) or for OnlyOpinions mode results in a total absence of recommendation. Finally, the promotion technique show better results for ‘User 31’ created in a new institution than within an existing institution where competence with other users is present.

V.2.8.2. S8b: User Recommendation

As it can be seen in Figure V.2-38 for all promotion period values an initial effect is observed, but only for $p=5$ (TruQOs_4) the initial effect continues in the time. For other values of ‘p’ the number of interactions answered by ‘User 31’ remains in the number achieved after the promotion period. Figure V.2-39 gathers the final results for all the tests. Not using promotion period guarantees no recommendation.

Figure V.2-38. Interactions Answered by ‘User 31’ over simulation time for each of S8b tests with weighted evaluations. Marked with colour squares the period where the quality of the resources of ‘User 31’ is increased (blue); and when ‘User 31’ was still unborn (grey).

The results for the recommendation of 'User 31' for direct interactions does not make any difference in relation to where 'User 31' is created, and thus only one of the two situations is shown in the graphs.

Figure V.2-39. S8b Total ratios summary for 'User 31'.

V.2.9. S9: Integrating resources and users recommendation

Scenario S9 applies Resources and Users recommendation mode to evaluate the existing correlation between providing good resources and being recommended as user and vice versa. S8 is divided in two subscenarios where consumers were configured to request one resources recommendation and later on the recommendation of a user.

V.2.9.1. S9a: Generation of Resources focus

This first subscenario analysed the potential relation between providing good quality resources and being more recommended as a user for direct interactions, through the tests described in Table V.2-14.

Table V.2-14. Scenario S9a: TruQOs' tests description

Test	Parameters
Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.5$; $k_{Provider}=0.5$; $k_{ProviderEv}=0.5$; $k_{ProviderOp}=0.5$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $hds_NONE=-1$; $hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$ - $q=0.25$ - $itm=5$, $\mu=0.3$ - $q1(1)-q2(0.9)-q3(0.7)-q4(0.5)$ - $prom=5$
TruQOs_1	$Q_{Resource}=0.5$, $Q_{Direct}=0.5$
TruQOs_2	$Q_{Resource}=0.55$, $Q_{Direct}=0.5$
TruQOs_3	$Q_{Resource}=0.6$, $Q_{Direct}=0.5$
TruQOs_4	$Q_{Resource}=0.65$, $Q_{Direct}=0.5$

Figure V.2-40 shows the accessed resources over simulation time and Figure V.2-41 the temporal evolution of the interactions answered by 'User 12'. There is a clear relation for $Q_{Resource}=0.6$ (TruQOs_3) and $Q_{Resource}=0.65$ (TruQOs_4) tests between the increment in the number of accessed resources and the number of interactions answered by 'User 12'. This relation was measured with the ratio of benefits (see section IV.4.2.5 and equation (IV.4-14)), and easily observed in Figure V.2-42.

Figure V.2-40. Accessed resources created by ‘User 12’ over simulation time for S9a TruQOs tests. Marked with colour squares the period where the quality of the resources of ‘User 12’ is increased (blue).

Figure V.2-41. Interactions Answered by ‘User 12’ over simulation time for S9a TruQOs tests. Marked with colour squares the period where the quality of the resources of ‘User 12’ is increased (blue).

Figure V.2-42. S9a Total resource recommendation ratio (RR), user recommendation (UR) ratio, and ratio of benefits summary for ‘User 12’ and TruQOs’ tests.

V.2.9.2. S9b: Personal service focus

In the second scenario the opposite effect has been evaluated, which is the potential relation between providing good quality service during direct interactions and seeing own resources being more recommended.

The best test of the previous subscenario is selected, $Q_{\text{Resource}}=0.65$, and a new simulation of 16 cycles is run. Q_{Resource} is set back to $Q_{\text{Resource}}=0.5$ and different values of Q_{Direct} are tested, as defined in Table V.2-15. The ratio of resource recommendation is studied and compared to the situation $Q_{\text{Resource}}=Q_{\text{Direct}}$, the first of the test performed in this scenario

Table V.2-15. Scenario S9b: TruQOs' tests description

Test	Parameters
Resource-Evaluations(50-50) _Evaluations-Opinions(50-50) _Personal-Social(50-50), $\omega(0.7)$	$k_{\text{Resource}}=0.5$; $k_{\text{Provider}}=0.5$; $k_{\text{ProviderEv}}=0.5$; $k_{\text{ProviderOp}}=0.5$; $k_{\text{Personal}}=0.5$; $k_{\text{Social}}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $\text{hds_NONE}=-1$; $\text{hds_LOW}=0$; $\text{hds_MEDIUM}=1.5$; $\text{hds_HIGH}=2.5$ - $q=0.25$ - $\text{itm}=5$, $\mu=0.3$ - $q1(1)$ - $q2(0.9)$ - $q3(0.7)$ - $q4(0.5)$ - $\text{prom}=5$
TruQOs_1	$Q_{\text{Resource}}=0.5$, $Q_{\text{Direct}}=0.5$
TruQOs_2	$Q_{\text{Resource}}=0.5$, $Q_{\text{Direct}}=0.6$
TruQOs_3	$Q_{\text{Resource}}=0.5$, $Q_{\text{Direct}}=0.7$
TruQOs_4	$Q_{\text{Resource}}=0.5$, $Q_{\text{Direct}}=0.8$

Figure V.2-43 shows the interactions answered by 'User 12' and Figure V.2-44 the accessed resources. There is again a clear relation between the increment in the quality of the direct interactions answered by 'User 12' and the number of accessed resources increasing. This relation is not linear though, as it was expected, and the best results are obtained for $Q_{\text{Direct}}=0.6$ (TruQOs_2) and not $Q_{\text{Direct}}=0.8$ (TruQOs_4), being the number of accessed resources a 6.25% higher for TruQOs_2 than for TruQOs_4.

All the results were compared with the $Q_{\text{Direct}}=0.5$ situation (TruQOs_1), as it is illustrated by Figure V.2-45.

Figure V.2-43. Interactions Answered by 'User 12' over simulation time for S9b TruQOs tests. Marked with colour squares the period where the quality of the resources of 'User 12' is increased (light blue) and where the quality of the service provided by 'User 12' during direct interactions is increased (blue).

Figure V.2-44. Accessed resources created by ‘User 12’ over simulation time for S9a TruQOs tests. Marked with colour squares the period where the quality of the resources of ‘User 12’ is increased (light blue) and where the quality of the service provided by ‘User 12’ during direct interactions is increased (blue).

Figure V.2-45. S9b Total resource recommendation (RR) and user recommendation (UR) ratios summary in left axis, and Resource recommendation ratio difference with TruQOs_1 in right axis. All the results for ‘User 12’ and TruQOs’ tests.

V.2.10. S10: Initial dispersion

The last TruQOs evaluation scenario is a reedition of the first one, scenario S0, where the NoReputation strategy is substituted by the OnlyOpinions mode and the TruQOs reputation mode using the parameters set during the S1-S9 scenarios. The objective of this scenario is to evaluate if TruQOs and/or OnlyOpinions modes can be used during the initialization phase of a system and in situations where all the providers generate resources with the same quality or, on the contrary, a warm-up period where no recommendation engine is used will be needed before integrating any other recommendation technique.

Two different TruQOs configurations were used, as described in Table V.2-16

Table V.2-16. Scenario S10: TruQOs' tests description

Test	Description	Parameters
TruQOs_1	Resource-Evaluations(70-30) _Evaluations-Opinions(30-70) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0.3$; $k_{ProviderOp}=0.7$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $hds_NONE=-1$; $hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$ - $q=0.25$ - $itm=5$, $\mu=0.3$ - $q1(1)$ - $q2(0.9)$ - $q3(0.7)$ - $q4(0.5)$ - $prom=5$
TruQOs_2	Resource-Evaluations(70-30) _Evaluations-Opinions(20-80) _Personal-Social(50-50), $\omega(0.7)$	$k_{Resource}=0.7$; $k_{Provider}=0.3$; $k_{ProviderEv}=0.2$; $k_{ProviderOp}=0.8$; $k_{Personal}=0.5$; $k_{Social}=0.5$ - $\omega=0.7$ - $c=2.5$, $\varphi=1$ - $hds_NONE=-1$; $hds_LOW=0$; $hds_MEDIUM=1.5$; $hds_HIGH=2.5$ - $q=0.25$ - $itm=5$, $\mu=0.3$ - $q1(1)$ - $q2(0.9)$ - $q3(0.7)$ - $q4(0.5)$ - $prom=5$

As seen in Figure V.2-46 the results obtained in this scenario for TruQOs_1 clearly differ from the results obtained in scenario S0. The initial dispersion for TruQOs_1 is lower than in S0 and the recommendations are concentrated in a reduced set of providers (5 out of 15) with other providers not being recommended whatsoever. A similar situation is observed for OnlyOpinions mode, although this methods achieves a better dispersion (7 out of 15), but being also insufficient. The only test that achieves a successful dispersion is TruQOs_2, where all the providers are recommended at least once.

Figure V.2-46. S0 vs. S10 Results Comparison.

To determine the effectiveness of the dispersion, the root mean square error (RMSE) was calculated from the optimal dispersion, understood as the situation where all the providers are recommended the same number of times. The results are gathered in Table V.2-17, showing that the best dispersion is achieved by NoReputation but closely followed by TruQOs_2.

Table V.2-17. Dispersion effectiveness analysis results

Test	Optimal Accessed Resources per provider	RMSE
NoReputation	19	12.51862021
OnlyOpinions	18.86666667	25.5573951
TruQOs_1	19	40.28947478
TruQOs_2	19.13333333	17.65411714

V.3. Summary and discussion of results

V.3.1. Concerning TruQOS

The trust and reputation model proposed in this thesis has been tested throughout a varied range of scenarios, S1 to S9, in an iterative and progressive path. Each scenario inherits the results of the previous ones and includes a new effect to study and a new parameter to analyse its impact on the reputation assessment process. Most of the scenarios start from an initial state where the system was executed using no reputation calculation to provide the recommendation. Scenario **S0** created an initial dispersion of resources and opinions amongst the providers where no user soared above the rest. The results of this scenario allowed selecting the target users whose behaviours were varied and analysed to test the outcomes of the recommendation, as explained in detail when introducing the evaluation framework in the previous chapter.

The first three scenarios, S1, S2 and S3, aimed at studying the effect that changes in the quality of the resources (and thus the evaluations) and/or changes in the opinions have in the final reputation value. These effects were studied first separately, *evaluations* in **S1** and *opinions* in **S2**, where the impact of performing a temporal weighting of evaluations/opinions was also studied. For both scenarios a similar conclusion can be achieved: temporal weighting of evaluations and opinions enhances both the final resource recommendation ratio as well as its temporal evolution.

Scenario **S1** demonstrated that temporal weighting helps fastening the detection of higher quality periods and reducing the impact of transitory low quality periods. A bad providers' personal state or mood could cause these low quality periods and the aforementioned reduction of the impact of low quality periods allows not punishing in excess quality drops, in the hope that the quality of the new resources will be increased again. If low quality periods continue in time, TruQOs will detect this extension and the reduction of the impact will not be applied, thus decreasing the recommendation of the provider.

In a similar way, in scenario **S2** temporal weighting of opinions allows reducing the effect of malicious behaviours of consumers when rating their interactions with providers' resources. The combination of personal and social opinions appears as a better strategy than using them individually, obtaining higher final recommendation ratios (23.8% over considering only personal opinions and 132.6% over social opinions), and even overcoming the results obtained for NoReputation mode (9% higher recommendation ratio).

Scenario **S3**, studied the combining effect of varying evaluations and opinions. The combination of all TruQOs dimensions obtained higher resource recommendation ratios than using the dimensions separately (14% over in *S3a*). Moreover, TruQOs overcomes NoReputation and OnlyOpinions modes for most of the parameters combinations in *S3a* and *S3b*. However, some effects were discovered in some tests with evaluations-based reputation prevailing over opinions-based reputations. In those tests, the obtained inertia from the evaluations is so high that the effect of the initial increment of the quality of the resources is maintained for all the simulation, exceeding the desired impact (for example for TruQOs_1 and TruQOs_10 in *S3b*).

In general, recommendations based on opinions present a higher delay than those only based on evaluations in detecting changes in quality or opinions periods. Combining Provider's opinions with Resource's opinions increments this delay, as observed in scenario *S3c*, generating ramp-shaped temporal series. This delay can be partially balanced out by combining all types of information sources: providers' evaluations, opinions about the provider and opinions about the resource, as it is the case of *TruQOs_27*. However, for most of tests, the aforementioned delay make difficult to determine the real impact of compensating changes in resources' quality with changes in opinions. The scenario *S3c* was designed to last only 50 simulation cycles and, to study those effects for some *TruQOs* parameters combinations, this duration was too short. In following scenarios the simulation was run until the desired or expected results were achieved. Despite this issue regarding the delay, results show that for all the scenarios it is possible to set up a combination of *TruQOs* parameters to obtain the desired behaviour.

Once evaluated the main parameters of the model, and before studying the effects of secondary parameters, scenario **S4** analysed the impact or effect of varying the number of users attacking the target user. For the worst possible situation, where all the consumers in the system attack one provider, *TruQOs* obtains a better protection to attacks than the traditional approach, *OnlyOpinions* mode, achieving a final ratio of resources recommendation a 25% higher.

Scenarios *S5* to *S8* studied the parameters responsible for obtaining a fine tuning of the model, adapting the global behaviours achieved with the selection of the main parameters to the desired recommendation performance. Scenarios **S5** and **S6** evaluated the impact of *Structural Reputation* in the social trust used to asses *Social Reputation*. The main conclusion is that considering information about the position of consumers and providers within the structure of the system improves the results obtained with the *OnlyOpinions* reputation mode, as well as the results not considering structural information. Furthermore, if no sustained quality increment is simulated to test the desired effects, as in scenario **S6**, *TruQOs* also overcomes the results of *NoReputation* mode. In both scenarios some *TruQOs* tests showed a higher sensibility to 3rd party opinions than the desired or expected

The results obtained for scenario **S7**, where variations in reliability were tested, showed that here again *TruQOs*-based tests get better results, in terms of temporal evolution and final values of the resources recommendation ratio. All the subscenarios were run using a reduced system structure that speeded up the testing, although at the expense of experiencing a higher delay in the detection of the effects of changes in the reliability in subscenario *S7a*: 22 simulation cycles for the fastest test (*TruQOs_1*). This delay is reduced to 13 cycles in *S7b* and to 8 in *S7c*. This reduction in the delay along with the extended duration of the whole experiment⁵⁰ validate the decision of using a reduced system structure. The desired impact of reliability variations is always observed and, despite increasing the number of simulation cycles, the global time needed for every test is still lower than for the same test with the whole system structure.

⁵⁰ 60 simulation cycles versus 50 simulation cycles used in scenarios *S1-S5*

Scenario **S8** demonstrated that TruQOs is the only recommendation mechanism that successfully deals with the *bootstrapping* issue. The ‘promotion period’ technique allows the recommendation of new users entering the system for both types of recommendations: resources and users. NoReputation mode only works as long as the new user enters the system creating resources or providing services during direct interactions of a higher quality than the rest of the existing users in the system (or equal quality due to its random behaviour when there is equality). As for OnlyOpinions mode, it fails completely at recommending new users due to the lack of opinions.

Scenario **S9** was design to test the existing relation between providing good results in direct interactions and improving the recommendation of own resources, and vice versa. This bidirectionality can serve as incentive, encouraging system’s user to generate resources and to offer direct consultation to increase their reputation. At the light of the results this bidirectionality has been proved, although the encouraging effect can only be tested in platforms with real users.

The final scenario is **S10**. Scenario S0 is repeated under the same circumstances but substituting NoReputation with OnlyOpinions and TruQOs as recommendation engines. As expected, OnlyOpinions mode did not achieve a satisfactory dispersion in terms of number of providers being recommended. On the contrary, TruQOs can be personalized to achieve this desired dispersion, demonstrating the capacity of adaptation to different scenarios of the trust and reputation model proposed in this PhD research work.

Summarizing, the outcomes obtained with TruQOs are very satisfactory. The versatility and personalization ability offered by TruQOs allow adapting the model to the situation in which it is applied, offering in general better results than NoReputation and OnlyOpinions modes. Using provider’s past history of both evaluations and opinions grants an easier detection of periods where the quality of resources is increased, as well as reducing the effect of a decrease in the quality. The provider’s history is also effective against users’ misbehaviours when rating an interaction through opinions.

The evaluation of TruQOs was performed with a reduced set of users (and even smaller for scenarios S6 and S7), far from the overwhelming number of users of environments like the one envisioned in this thesis. Simulating a system with thousands or even hundreds of thousands is unfeasible in terms of memory and time consumption. Furthermore, it is even unnecessary. Mathematically, none of the equations presented in this PhD dissertation for the calculation of the different reputation values have size limitation regarding the number of opinions or evaluations, numbers after all, used as input. A higher number of users will be translated only in a higher number of resources created in the system (evaluations) and a higher number of opinions about the resources. The number of consumers attacking a provider is always translated to a percentage of all the consumers in the system, and scenario S4 already evaluated successfully that influence in terms of percentage.

Furthermore, the different system structures used in the evaluation scenarios reflect all the possible Hierarchy Dependence Strength (HDS) relations considered for the assessment of *Structural Reputation* and bigger structures will not add any useful information. In conclusion, we firmly believe that the results obtained in the evaluation of TruQOs and the conclusions obtained from their analysis are valid and they can be extrapolated to systems of any size and structure.

To end this section and the discussion of evaluation results, when the quality of the resources is increased and maintained for a long period, the best recommendation mode is NoReputation, regardless the presence or not of variations in the behaviour of users at giving opinions. In the description of the scenarios, these behaviours were considered as attacks, thus with misleading intentions. In that cases, neglecting the opinions is with no doubt the best solution and NoReputation has no rival. Nevertheless, sometimes higher evaluations mean higher technical quality but not higher overall quality: the educational capacities of the resource could be worse and bad opinions are not attacks. The recommendation mechanism should reflect such situations and NoReputation mode is no longer the best. From the recommendation modes being compared in this research work, TruQOs is the only solution that can adapt effectively to all the situations.

As final remark, although only one service was used during the evaluation scenarios to facilitate and speed up the testing, through a taxonomy of different services TruQOs can filter among all the possible results of the desired service and provide the best recommendation. Therefore, TruQOs can be considered as a sort of a feature augmentation hybrid recommender system conformed by first step of filtering followed by the trust and reputation model.

V.3.2. Concerning TruQOS Simulation Tool

The evaluation of TruQOs Simulation Tool (TST) was performed *indirectly*. The validity of TST was tested studying its suitability for modelling complex user behaviours. These modelling methods allowed easily mimicking users' actions and attitudes in all the situations being tested. Additionally, TST supported efficiently the simulations for all the tests, with very few errors, mostly related to database communication errors. Furthermore, no significant errors were found at exporting and/or importing the state of any simulation.

TST constitutes a powerful tool for modelling and simulating complex platforms oriented at knowledge sharing through the creation and consumption of digital resources. Moreover, the versatility and generic design of TruQOs Simulation Tool allows its use for other application domains, especially the module for user behaviours modelling. To our knowledge, the level of personalization of simulated users offered by TST is hardly achieved by the different testbeds available in the state of the art. This feature, along with the simulation of an objective evaluation of both resources and service provision, makes TST unique and its release to the scientific community could be of great use in the development and evaluation of new recommendation mechanisms in a wide range of applications.

VI. CONCLUSIONS AND FUTURE WORKS

This chapter establishes a discussion focused on the verification of the research hypotheses, and leading to the formulation of the main conclusion and contributions of this PhD work. Additionally, possible future works to this thesis are postulated.

VI.1. Verification of research hypotheses

Chapter III defined the main hypotheses of the research work, and the evaluation framework introduced in Chapter IV aimed at confirming them. This section confronts the outcomes of the validation tests presented in Chapter V with the research hypotheses and analyses their degree of validity.

VI.1.1. TruQOs hypothesis

H1.- A trust and reputation model combining evaluations of resources with opinions from system's users reduces subjective behaviour punishments while enhancing evaluation numbers with additional pedagogical information. This combination improves traditional approaches that use only personal appraisals or quality ratings.

Hypothesis confirmed

The results presented in Chapter V showed that TruQOs improves traditional approaches at providing recommendations in Technology Enhanced Learning (TEL) platforms, thus confirming hypothesis H1. Throughout the different scenarios of the evaluation framework a varied range of situations were tested. The iterative design allowed introducing new parameters and effects covering the main specifications of TEL platforms. TruQOs overcomes the outcomes obtained considering OnlyOpinions and NoReputation recommendation modes in most of the scenarios. Moreover, TruQOs shows a higher adaptability to different situations, whereas traditional approaches only have good results under favourable conditions.

A particular focus should be given to the bootstrapping problem, as presented in section IV.3.6 and tested in scenario S8 (V.2.8). The promotion period mechanism implemented in TruQOs successfully helps new users in the system to be recommended for direct interactions as well as the resources created by them. On the contrary OnlyOpinions fails completely and new users are never recommended, while NoReputation only works partially.

VI.1.2. TST hypothesis

H2.- Intelligent agents' paradigm allows modelling user behaviours in Technology Enhanced Learning platforms, optimizing multi-scenario testing and personalization.

Hypothesis confirmed

The agent-oriented programming (AOP) paradigm has proven its validity at modelling user behaviours in TEL platforms. Intelligent software agents' characteristics matched the main requirements of the simulation tool. The proactivity, sociability and adaptability of the developed agents allowed reproducing real users' actions and attitudes through the diverse situations defined in the evaluation framework. The result were independent users, accessing to different resources and communicating amongst them to fulfil their needs, basing these interactions on the recommendations provided by the reputation engine.

22 subscenarios were simulated with variations of 18 basic user behaviours:

- 2 behaviours related to the role of the user: provider and consumer.
- 3 behaviours for providers related to the resources' generation pattern or availability to receive direct inquiries: generate a resource 3 out of 5 cycles, generate a resource every cycle, and answer all the incoming direct requests.
- 2 behaviours for consumers related to recommendations' request pattern or laziness at giving opinions: request a recommendation every cycle, and always give an opinion after an interaction.
- 3 behaviours related to the type of recommendation requested by consumers: only resources, only users, and resources and users alternatively.
- 3 evaluations behaviours: increasing or decreasing the quality over a period of time, and increasing/decreasing the quality alternatively.
- 5 opinions' behaviours: fair, critic, praiser, hater and lover.

The agent paradigm has little to be blamed for regarding the performance issues showed by the simulation tool that are caused by a design that can be improved with some of the solutions exposed in the future works at the end of this chapter. Moreover, as discussed in the last section of Chapter V (V.3.2), AOP could help at reducing their impact.

VI.2. Main contributions

Once the validity of the research hypotheses has been analysed, the chapter continues with a summary of the main contributions of this PhD.

Contribution 1

MC1.- "The design and validation of a new trust and reputation model for resources and users recommendation, based on the combination of subjective opinions and objective evaluations".

The main contribution of this PhD research work is TruQOs, the trust and reputation model, presented in Chapter IV, for the recommendation of users and resources in quality oriented environments. Its main application domain is Technology Enhanced Learning (TEL) platforms. The model proposes a novel approach, generating the recommendation over the basis of adding objective evaluations of digital resources to the traditional subjective opinions given by the users after consuming a resource. Through the different parameters, the effect of which were presented in Chapter V, the model can be customized and adapted to different scenarios, enabling an effective recommendation in all contexts.

The goals to be attained by a recommendation engine are conditioned by the application domain. The modular design proposed in this thesis allows combining the different reputation dimensions in multiple ways. TruQOs proposes one possible combination to fulfil all the requirements detected for TEL platforms, but different solutions are easily achieved by selecting different combinations of pieces of the puzzle described in this PhD dissertation.

Contribution 2

MC2.- “The definition of a method for the aggregation of two related variables to provide an ordered classification”.

The “**Quartiles Aggregation Method**” presented in section IV.3.7 allows performing an ordered classification of elements according to two associated variables. Candidates are ordered according to one of the variables and the result is divided in four groups or quartiles. Each quartile has a different weight, in decreasing order, which is then used to modify the other variable of every element. Finally, the elements are ordered again according to this new modified variable. TruQOs uses this method to order providers or resources through the aggregation of their Reputation and the associated Reliability of those calculated reputation values. But this method is generic and can be used in many other fields to generate classifications. As an example, the Quartiles Aggregation Method could be used to create an ordered classification of measurements according to their mean value and standard deviation.

Contribution 3

MC3.- “The development of an agent-based generic tool for the modelling of complex user behaviours in Technology Enhanced Learning platforms”.

The ultimate success of ‘in silico’ evaluations of digital platforms lies in the credibility of the model created to mimic real users’ actions and attitudes within the platform. A generic tool was presented in Chapter IV that allows modelling complex users’ behaviours in TEL systems. Through this tool different user roles can be selected and configured. Each role enables different patterns for: a) the creation of resources; b) the availability for receiving direct inquiries; c) the willingness for consuming a resource or requesting a service; and d) the laziness at giving opinions. The tool also allows selecting users’ behaviour when giving opinions and modelling the quality over time of the resources created by the users. The tool uses the agent programming paradigm to apply the customized behaviour to the different digital users, creating individualized and independent user profiles.

Contribution 4

MC4.- “The design, implementation and evaluation of a simulation framework for multiple-scenario testing of trust and reputation models to provide recommendation in environments devoted to knowledge exchange”.

The tool for user behaviours modelling mentioned in MC3 was integrated with the recommendation mechanism based on TruQOs reputation assessment, and a simulation engine. TruQOs Simulation Tool (TST) supports multiple different configurations of users’ behaviours and TruQOs’ parameters, empowering the evaluation capabilities of the tool in a wide range of scenarios, and allowing the testing of the performance of trust and reputation models as recommendation mechanisms. In addition, TST has different modules for the graphic visualization of results, facilitating the analysis and interpretation of simulation outcomes.

VI.3. Scientific publications

VI.3.1. International journals

Iñaki Martínez-Sarriegui, Gema García-Sáez, Fernando García-García, J. M. Iniesta and Elena Hernando, "A Physical Activity Detector based on Intensity Levels for Patients with Exercise Constraints", *Methods of Information in Medicine* (Submitted 2015)

Gema García-Sáez, **I. Martínez-Sarriegui**, M. Rigla, Pons-Pons B., M. Villaplana, E.J. Gómez and M. E. Hernando, "Design evaluation of a prototype user interface to support a guideline-based decision support system in gestational diabetes". (2014), *Diabetes Technology and Therapeutics*

Gema García-Sáez, Mercedes Rigla, **Iñaki Martínez-Sarriegui**, Erez Shalom, Mor Peleg, Tom Broens, B. Pons, Estefania Caballero-Ruiz, Enrique J. Gómez and E. Hernando, "Patient-oriented Computerized Clinical Guidelines for Mobile Decision Support in Gestational Diabetes". (2014), *Journal of Diabetes Science and Technology*, 8(2):238-246

C. Gonzalez, P. Herrero, J. M. Cubero, J. M. Iniesta, M. E. Hernando, G. García-Sáez, A. J. Serrano, **I. Martínez-Sarriegui**, C. Pérez-Gandía, E.J. Gómez, E. Rubinat, V. Alcantara, E. Brugués, A. Chico, E. Mato, O. Bell, R. Corcoy and A. de Leiva, "PREDIRCAM eHealth platform for individualized telemedical assistance for lifestyle modification in the treatment of obesity, diabetes, and cardiometabolic risk prevention: a pilot study (PREDIRCAM 1) " (2013), *Journal of Diabetes Science and Technology*, 7:1932-2968 (Electronic), 1932-2968 (Linking), 4:888--897

B. Pons, **I. Martínez-Sarriegui**, M. E. Hernando, A. Montero, G. García-Sáez, J. M. Iniesta and M. Rigla, "ECONSULTA: Videoconsulting for Endocrinology and Primary Care Communication Improvement", (2011), *Diabetes Technology & Therapeutics*

I. Martínez-Sarriegui, G. García-Sáez, M. Rigla, E. Brugués, A. de Leiva, E.J. Gómez and M. E. Hernando, "How Continuous Monitoring Changes the Interaction of Patients with a Mobile Telemedicine System". (2011), *Journal of Diabetes Science and Technology*, 1(5):5-12

Gema García-Sáez, M. Elena Hernando, **Iñaki Martínez-Sarriegui**, Mercedes Rigla, Veronica Torralba, Eulalia Brugués, Alberto de Leiva and Enrique J. Gómez, "Architecture of a Wireless Personal Assistant for Telemedical Diabetes Care". (2009), *International Journal of Medical Informatics*, 78(6):391-403

M. E. Hernando, Gema García-Sáez, **Iñaki Martínez-Sarriegui**, A. Rodríguez-Herrero, Carmen Pérez-Gandía, Mercedes Rigla, Alberto de Leiva, Ismael Capel and E. Gómez, "Automatic data processing to achieve a safe telemedical artificial pancreas". (2009), *Journal of Diabetes Science and Technology*, 3(5):1039-1046

Alberto De Leiva, M. Elena Hernando, Mercedes Rigla, Ismael Capel, Eulalia Brugués, Belén Pons, Luisa Erdozain, Agustina Prados, R. Corcoy, Enrique J. Gómez, Gema García-Sáez, **Iñaki Martínez-Sarriegui**, Agustín Rodríguez-Herrero and C. Pérez-Gandía, “Telemedical Artificial Pancreas: PARIS (Páncreas Artificial Telemédico Inteligente) Research Project” (2009), *Diabetes Care*, 32(2):S211-S216

M. Elena Hernando, Mario Pascual, Carlos H. Salvador, Gema García-Sáez, A. Rodríguez-Herrero, **Iñaki Martínez-Sarriegui** and Enrique J. Gómez, “Definition of Information Technology Architectures for Continuous Data Management and Medical Device Integration in Diabetes”. (2008), *Journal of Diabetes Science and Technology*, 2(5):899-905

VI.3.2. Abstracts in international journals

E. Caballero-Ruiz, Gema García-Sáez, M. Rigla, B. Pons, **I. Martínez-Sarriegui**, Enrique Gomez and M. E. Hernando, “Personalized decision support for guiding women with gestational diabetes: automatic blood glucose classification”. (2013), *Diabetes Technology & Therapeutics*, pages A104

F. García-García, A. Marcano-Cedeño, **Iñaki Martínez-Sarriegui**, Pedro J. Benito, Enrique J. Gómez and M. Elena Hernando, “An Artificial Intelligence Algorithm for Automatic Assessment of Physical Activity Intensity and Metabolic Type Using Multiaxial Accelerometry and Heart Rate”. (2012), *Journal of Diabetes Science and Technology*, pages 462–465, A50

B. Pons, **Iñaki Martínez-Sarriegui**, M. Elena Hernando, Gema García-Sáez, Lydia Padró and Mercedes Rigla, “eCONSULTA: An Integrated Diabetes Care Model (Specialist–Primary) Based On Video”. (2012), *Journal of Diabetes Science and Technology*, A152

P. Herrero, M. E. Hernando, R. Garrigos, G. García-Sáez, J. M. Cubero, **I. Martínez-Sarriegui**, R. Corcoy, E. Rubinat, E.J. Gómez and A. de Leiva, “Feasibility Study of a Technological Platform for the Prevention of Diabetes Mellitus and Cardio Metabolic Risk”. (2010), *Journal of Diabetes Science and Technology*, 4(2), A57

VI.3.3. Book chapters

F. García-García, G. García-Sáez, P. Chausa, **Iñaki Martínez-Sarriegui**, Pedro J. Benito, E.J. Gómez and M. E. Hernando, “Statistical Machine Learning for Automatic Assessment of Physical Activity Intensity Using Multi-axial Accelerometry and Heart Rate”. (2011), *Artificial Intelligence in Medicine -Lecture Notes in Artificial Intelligence*, pages 70-79, Springer Berlin Heidelberg

G. García-Sáez, José M. Alonso, Javier Molero, Mercedes Rigla, **Iñaki Martínez-Sarriegui**, Alberto de Leiva, Enrique J. Gómez and M. Elena Hernando, “Mealtime Blood Glucose Classifier Based on Fuzzy Logic for the DIABTel Telemedicine System”. (2009), *Lecture Notes in Artificial Intelligence*, 5651(295–304)

Iñaki Martínez-Sarriegui, Gema García-Sáez, M. Elena Hernando, Mercedes Rigla, Eulalia Brugués, Alberto De Leiva and Enrique J. Gómez, “Mobile Telemedicine: A Computing and Networking Perspective”. (2008), Mobile Telemedicine for Diabetes Care, pages 143-159, Auerbach Publications

VI.3.4. International congresses

M. E. Hernando, **I. Martínez-Sarriegui**, Gema García-Sáez, Silvana Quaglini, Mercedes Rigla and Carlo Napolitano, “Mobile patient feedback through continued monitoring and guideline-based decision support”. (2015), 7th International Joint Workshop KR4HC 2015 - ProHealth 2015, pages 134-139, 2015

Arturo González-Ferrer, Mor Peleg, Enea Parimbelli, Erez Shalom, Carlos Marcos, Guy Klebanov, **Iñaki Martínez-Sarriegui**, Nick Lik San Fung and Tom Broens, “Use of the Virtual Medical Record Data Model for Communication among Components of a Distributed Decision-support System”. (2014), 2nd IEEE Biomedical and Health Informatics International Conference

I. Martínez-Sarriegui, F. García-García, Gema García-Sáez, M. E. Hernando Pérez and Michael Luck, “TRHIOS: Trust and reputation in hierarchical and quality-oriented societies”. (2012), 7th Iberian Conference on Information Systems and Technologies, CISTI 2012, pages 1-4

C. Pérez-Gandía, M. Rigla, B. Pons, D. Subías, G. García-Sáez, **I. Martínez-Sarriegui**, J. M. Iniesta, A. Rodríguez-Herrero, E.J. Gómez and M. E. Hernando, “Accuracy Evaluation of an On-line Neural Network Glucose Prediction Algorithm in a Cross-correlation Study” (2011), Proceedings of the 4th International Conference on Advanced Technologies and Treatments for Diabetes, London (UK). Journal of Diabetes Technology & Therapeutics 13(2) pp. 261

F. García-García, **Iñaki Martínez-Sarriegui**, E.J. Gómez, M. Rigla and M. E. Hernando, “Automatic Assessment of Physical Activity Using Multi-Axial Accelerometry and Heart Rate”. (2011), 4th International Conference on Advanced Technologies and Treatments for Diabetes (ATTD), pages 182, Journal of Diabetes Technology & Therapeutics

D. Subías, C. Pérez-Gandía, M. E. Hernando, B. Pons, G. García-Sáez, **I. Martínez-Sarriegui**, J. M. Iniesta, M. Gallach and M. Rigla, “First Clinical Experience Using an On-line Glucose Prediction Algorithm for Interprandial Optimisation in Type 1 Diabetes (DM 1)”. (2011), Proceedings of the 4th International Conference on Advanced Technologies and Treatments for Diabetes, London (UK). Diabetes Technology & Therapeutics, pages 276

C. Pérez-Gandía, G. García-Sáez, **I. Martínez-Sarriegui**, J. M. Iniesta, A. Rodríguez-Herrero, M. Rigla, E.J. Gómez and M. E. Hernando, “PREDEVAL: A New Tool for the Evaluation of Any Glucose Prediction Algorithm”. (2011), Proceedings of the 4th International Conference on Advanced Technologies and Treatments for Diabetes, London (UK). Journal of Diabetes Technology & Therapeutics 13(2) pp. 261

Iñaki Martínez-Sarriegui, Hong Zhu, Lijun Shan, Gema García-Sáez, Enrique J. Gómez and M. Elena Hernando, "Modelling and Formal Specification of a MultiAgent Telemedicine System for Diabetes Care". (2009), 1st International Conference on Agents and Artificial Intelligence, ICAART 09

Gema García-Sáez, M. Elena Hernando, Javier Molero, Mercedes Rigla, **Iñaki Martínez-Sarriegui**, Alberto De Leiva and Enrique J. Gómez, "Automatic Blood Glucose Classification Based on Fuzzy Logic Integrated in a Telemedicine System", 2nd Conference on Advanced Technologies & Treatments for Diabetes". (2009). Athens 25-28 Feb.

P. Herrero, M. E. Hernando, R. Garrigos, G. García-Sáez, R. Corcoy, **I. Martínez-Sarriegui**, C. Pérez-Gandía, A. Chico, J. M. Cubero, E. Rubinat, E. Mato, O. Bell, L. Roa, A. de Leiva, E.J. Gómez and E. Brugués, "PREDIRCAM: Technological platform for the prevention of Diabetes Mellitus and cardiometabolic risk". (2009), 2nd Conference on Advanced Technologies & Treatments for Diabetes (ATTD), Athens, pages 174

Gema García-Sáez, María Elena Hernando, Mercedes Rigla, Javier Molero, Ismael Capel, **Iñaki Martínez-Sarriegui**, Alberto de Leiva and Enrique Javier Gómez, "Automatic Alarms to detect anomalous patterns in continuous glucose measurements with a Telemedicine System for Diabetes Management". (2008), 1st Conference on Advanced Technologies & Treatments for Diabetes

Iñaki Martínez-Sarriegui, M. Elena Hernando, F. J. Brito Hernández, Gema García-Sáez, Javier Molero, Mercedes Rigla, Eulalia Brugués, Alberto De Leiva and Enrique J. Gómez, "Electronic Report Generation Web Service evaluated within a Telemedicine System". (2008), 4th European Congress for Medical and Biomedical Engineering, MBEC08, Ambers, pages 994-997

VI.3.5. National congresses

I. Martínez-Sarriegui, Gema García-Sáez, B. Pons, M. Villaplana, M. Rigla, A. Serrano Rubio, E. J. Gómez Aguilera and M. E. Hernando Pérez, "Tour Virtual: Una forma eficiente de realizar evaluaciones tempranas en Aplicaciones de Paciente". (2014), XXXII Congreso Anual de la Sociedad Española de Ingeniería Biomédica

Alberto Lago, **I. Martínez-Sarriegui**, Gema García-Sáez, Mercedes Rigla, E. J. Gómez Aguilera and M. E. Hernando, "Aplicación prototipo para comunicarse con glucómetros según el estándar ISO/IEEE11073-10417". (2012), XXIII Congreso Nacional de la Sociedad Española de Diabetes, pages 18, Sociedad Española de Diabetes

Iñaki Martínez-Sarriegui, A. Montero Rodríguez, B. Pons, Mercedes Rigla, Gema García-Sáez, J. M. Iniesta and M. Elena Hernando, "eCONSULTA: Integración de un Sistema de Videoconsultoría Web entre Asistencia Primaria y Atención Especializada". (2011), Libro de Actas del XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica

- P. Chausa Fernández, C. Cáceres Taladriz, F. J. García Peces, **I. Martínez-Sarriegui**, A. Montero Rodríguez, D. Varela Pérez, J. García Gutiérrez, P. Álvarez Díaz, I. Pérez Medina, M. E. Hernando Pérez, J. M. Gatell Artigas and E. J. Gómez Aguilera, "Hospital Virtual: Sistema de información clínica y telecuidado de pacientes VIH/SIDA basado en tecnologías Web 2.0". (2010), XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica
- F. J. Brito Hernández, **Iñaki Martínez-Sarriegui**, M. Elena Hernando, Gema García-Sáez, Javier Molero, Mercedes Rigla, Eulalia Brugués, Alberto De Leiva and Enrique J. Gómez, "Desarrollo de un Servicio Web para generación de informes electrónicos en un sistema de Telemedicina". (2007), XXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica, pages 342-345
- Gema García-Sáez, M. Elena Hernando, Mercedes Rigla, V. Torralba Piqueras, **Iñaki Martínez-Sarriegui**, B. Ares Alonso, Francisco del Pozo, Eulalia Brugués, Alberto De Leiva and Enrique J. Gómez, "El sistema de telemedicina DIABTel y la medida de glucosa en tiempo real optimizan el autocontrol del paciente". (2006), XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica, pages 125-128
- Iñaki Martínez-Sarriegui**, M. Elena Hernando, Gema García-Sáez, B. Ares Alonso, V. Torralba Piqueras, Francisco del Pozo, Mercedes Rigla, Eulalia Brugués and Enrique Javier Gómez, "Evaluación de un Asistente Inteligente para pacientes diabéticos tipo 1". (2006), XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica, pages 129-132
- B. Ares Alonso, M. Elena Hernando, Gema García-Sáez, **Iñaki Martínez-Sarriegui**, V. Torralba Piqueras, Mercedes Rigla, Eulalia Brugués, Alberto De Leiva, Francisco del Pozo and Enrique J. Gómez, "Evaluación del uso del sistema de telemedicina DIABTel combinado con un sensor continuo de glucosa". (2006), XXIV Congreso Anual de la Sociedad de Ingeniería Biomédica

VI.4. International collaborations

Department of Computing and Communication Technologies at Oxford Brookes University⁵¹.

The collaboration was started during a short stay under Dr. Hong Zhu's⁵² coordination. This short stay focused on the study of the agent-oriented programming paradigm and agent-base system formalisation, resulting in a publication in an international conference:

Iñaki Martínez-Sarriegui, Hong Zhu, Lijun Shan, Gema García-Sáez, Enrique J. Gómez and M. Elena Hernando, Modelling and Formal Specification of a MultiAgent Telemedicine System for Diabetes Care, in: 1st International Conference on Agents and Artificial Intelligence, ICAART 09 2009

⁵¹ Dep. of Computing and Communication Technologies home page: <http://cct.brookes.ac.uk/index.html>

⁵² Dr. Hong Zhu's personal home page: <http://cms.brookes.ac.uk/staff/HongZhu/>

Department of Informatics at King's College London⁵³. The collaboration was started during a short stay under Dr. Michael Luck's⁵⁴ coordination. This short stay focused on the analysis of existing trust and reputation models and their applications, resulting in the proposal of a new model published in a publication in an international conference:

I. Martínez-Sarriegui, F. García-García, Gema García-Sáez, M. E. Hernando Pérez and Michael Luck, TRHIOS: Trust and reputation in hierarchical and quality-oriented societies, in: 7th Iberian Conference on Information Systems and Technologies, CISTI 2012, pages 1-4, 2012

VI.5. Future works

The research presented in this PhD dissertation is not but the first step in the path. The road goes ever on and on, down from the thesis where it began; with many treads still to be taken. This section presents some future works, with the objective of enhancing both the trust and reputation model and the simulation tool.

Furthermore, the road forks in several different paths from the conclusions of the work. Some of them lead to the proposal of exploitation plans for the dissemination of the obtained results. Finally, the road continues until it joins some larger way where many paths and errands meet. And wither then?

The section concludes with the exploration of some of these new paths, describing new research lines and application domains that can benefit from the results obtained in this PhD.

VI.5.1. Concerning the trust and reputation model

Extension of results:

Although the evaluation of TruQOs offered good results, they were obtained in a simulated environment and these promising outcomes must be endorsed:

- **'Field evaluation'**. TruQOs' performance should be evaluated in a Technology Enhanced Learning (TEL) platform with real users during a long period of time and compared to other recommendation mechanisms.

Model modifications:

As mentioned in the main contributions section, TruQOs is one of the multiple combinations of the different reputation's dimensions proposed in this thesis. These combinations should be selected to match the particular characteristics of the target domain.

⁵³ Dep. of Informatics home page: <http://www.kcl.ac.uk/nms/depts/informatics/index.aspx>

⁵⁴ Dr. Michael Luck home page: <http://www.dcs.kcl.ac.uk/staff/mml/>

Some examples of new combinations are:

- Using the “Quartiles Aggregation Method” to combine *Provider’s Evaluations Reputation* with *Provider’s Opinions Reputation*.
- Creating a flexible users’ ranking according to their *Provider’s Evaluations Reputation* and to be used in the calculation of *Structural Reputation*.
- Integrating *Resource Reputation* with *Social Reputation* by considering all the available opinions when assessing *Witnesses Reputation*.

VI.5.2. Concerning the simulation tool

Enhancing the existing tool/Solving issues:

The biggest issue discovered during the performed test is the exponential increase of time over the simulation cycles. As simulation progresses the Reputer needs more time to calculate and to offer a recommendation. The increasing number of resources and opinions in the system motivates this delay. Although users will create new resources and give opinions over the time for the whole lifecycle of the platform, the process of recommendation can be improved:

- **Optimization of the number of reputations to be calculated.** Within a single simulation cycle the same resource reputation is calculated for several users. Provider’s Evaluations and Resource dimensions are common for all the users. Thus, they should be computed only once and shared amongst users to speed up the process. For every user only Provider’s Opinions would be calculated and integrated with the common reputation values to assess the final TruQOs reputation.
- **Calculation of reputation in an iterative way.** Updating every simulation cycle the reputation with the new opinions and/or evaluations, and adding that value to the reputation values from previous simulation cycle, can boost the recommendation process.
- **Distribution of reputation tasks.** Creating several ReputationAgents to calculate the reputation of sets of users in parallel.

New functionalities:

The range of available functionalities of TST is conditioned by the time limitations of the PhD work and the purposes of the evaluation framework presented in this document. This range can be complimented with the addition of new capabilities:

- **Inclusion in the tool of other existing T&R and RS systems,** enabling the possibility of comparing the recommendation results between systems, including TruQOs. The recommendation engine should be adapted to the possibility of offering recommendation results from various algorithms.
- To improve the aforementioned comparison, a **new module for results visualization** needs to be implemented, allowing the representation in graphs and tables to facilitate the interpretation of the results.
- Creation of a **module for the design and integration of new recommendation engines** into the system.

Dissemination/Exploitation of results:

TST can serve as testbed to prove the adequacy and effectiveness of other Recommender Systems (RS) and/or Trust and Reputation (T&R) as a resources' and users' recommendation engine. In this sense a **freeware distribution of TST** will be created, protected with GNU GPL license, to be used by the community to integrate and to test their recommendation engines.

VI.5.3. Beyond TruQOs: portability and new research

There is no single trust and reputation model for all kinds of applications and environments. To build a suitable T&R system one shall first identify the application constraints, specifications and what it is desired to have in the system. Then, the appropriate features from the application domain need to be selected to fulfil pre-determined expectations (Noorian & Ulieru 2010). This section proposes some domains where TruQOs can be integrated without the need of changes, and some other where the model would need to be adapted to the specific characteristics of the target environment.

TEL platforms:

- **Platform for clinical knowledge exchange.** This platform is envisioned as a knowledge repository where clinical researchers can access to browse and discover relevant research papers. It is conceived as a Trusted ISI Web of Science. The impact factor of the journals could be used as objective evaluation value and researches would give opinions about the papers. Strong competitiveness amongst users is present in this platform and TruQOs would be used to combat possible misbehaviours and misleading opinions. Users' expertise, position within research institutions and/or values like the h-Index (Hirsch 2005) could serve as social trust to weight 3rd party opinions.
- **Laparoscopic surgical skills training platform.** The TELMA environment (Sánchez-González et al. 2011) is a TEL platform for minimally invasive surgery training. TELMA is based on a four-pillared architecture: (1) an authoring tool for the creation of didactic contents; (2) a learning content and knowledge management system that incorporates a modular and scalable system to capture, catalogue, search, and retrieve multimedia content; (3) an evaluation module that provides learning feedback to users; and (4) a professional network for collaborative learning between users.
The evaluation values provided by the module in (3) can be used as TruQOs' objective dimension, and combined with opinions expressed by the users after accessing a didactic content. The social trust could be assessed taking into account users' expertise as minimally invasive surgeons. All these values would then conform the information sources for the calculation of the reputation of the didactic contents. Moreover, a user recommendation could be also implemented using the same information.
- **Open eLearning environment for collaborative study.** Open eLearning environments can be defined as platforms where students can share their lecture notes or subjects' summaries, and independent teachers can create didactic resources or offer their mentorship. External professional could evaluate these summaries or didactic resources and users would give their personal opinions. In this case social trust could be estimated using the users' proactivity sharing resources and giving opinions.

Other medical application fields:

- **Tele-Consultation between caregivers.** Systems like eCONSULTA (Martínez-Sarriegui et al. 2011) designed to provide a platform for video-consultation between specialists and primary care can benefit from the integration of recommendation mechanisms like TruQOs. The recommendation of best physician to interact with provides security in the platform, enhancing interactions amongst users and improving the teleconsultation experience. Physicians can give opinions after those interactions and an objective evaluation can be generated using videoconferencing QoS (quality of service) parameters. Hierarchical and position information of the users within the institutions is available for social trust estimation. All these values can be used by TruQOs to calculate users' reputations and to provide the recommendation.
- **Wellbeing platforms.** These open platforms provide an online meeting point for users concerned about maintaining a healthy lifestyle, finding support and reinforcement in the social contact with other users with similar interests. The users share with other users their food diets and exercise plans. Furthermore, usually these platforms are also populated with professionals, nutritionists and personal trainers offering their services, as well as generating digital resources like the ones mentioned before. Providing users with trusted recommendations of which resources best suit their needs, or which professional they should contact to for personalised advice, can improve users' satisfaction with the system.

The characteristics of TruQOs make it a perfect candidate as recommendation engine. Diets can be evaluated according to the distribution of nutrients, and exercise plans according to the balance of groups of muscles involved in the plan. Another variable to include in the evaluation could be the adequacy of the diet or plan to their goal: losing weight, improving aerobic performance... And later users will give their personal opinion about the outcomes obtained by following the diet or the exercise plan. The level of adherence to the diets or plans by the users can be added as a measure of social trust, instead of structural information, in the process of weighting opinions from other users.

Other commercial application fields:

- **Films recommendation.** Traditionally movies recommendation is based only in personal or social opinions. This recommendation can benefit from the Provider's Reputation dimension in TruQOs, giving information of past movies by the same director, actors and/or producers. Objective information can be also added in the reputation assessment process, including awards of the specific films or from the filmography of the same director. Moreover, the reviews created by 3rd party trusted film critics could be considered as objective evaluations and thus included in the calculation of the final reputation with TruQOs. Social trust can be estimated by means of similarities in the type or genre of films between users.

VII. REFERENCES

- Abdul-Rahman, A. & Hailes, S., 2000. Supporting trust in virtual communities. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE Computer Society, p. 9.
- Abel, F. et al., 2008. A Rule-Based Recommender System for Online Discussion Forums W. Nejdl et al., eds. *Adaptive Hypermedia and Adaptive WebBased Systems*, 135(22), pp.12–21.
- Aberer, K. & Despotovic, Z., 2001. Managing Trust in a Peer-2-Peer Information System H. Paques, L. Liu, & D. Grossman, eds. *Proceedings of the tenth international conference on Information and knowledge management*, 3(2), pp.310–317.
- Adamopoulou, A.A. & Symeonidis, A.L., 2014. A simulation testbed for analyzing trust and reputation mechanisms in unreliable online markets. *Electronic Commerce Research and Applications*, 13(5), pp.368–386.
- Adomavicius, G. et al., 2005. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems*, 23(1), pp.103–145.
- Alahmadi, D.H. & Zeng, X.-J., 2015. ISTS: Implicit social trust and sentiment based approach to recommender systems. *Expert Systems with Applications*, 42(22), pp.8840–8849.
- Alhaqbani, B. & Fidge, C., 2009. A Time-Variant Medical Data Trustworthiness assessment model. In *2009 11th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE, pp. 130–137.
- Anand, S.S. & Mobasher, B., 2005. Intelligent techniques for Web Personalization B. Mobasher & S. S. Anand, eds. *Intelligent Techniques for Web Personalization*, 3169(LNAI 3169), pp.1–36.
- Anwar, M. & Greer, J., 2011. Facilitating Trust in Privacy-preserving E-learning Environments. *IEEE Transactions on Learning Technologies*, 5(1), pp.62–73.
- Arazy, O., Kumar, N. & Shapira, B., 2009. *Improving Social Recommender Systems*, IEEE Computer Society.
- Artz, D. & Gil, Y., 2007. A survey of trust in computer science and the Semantic Web. *Web Semantics Science Services and Agents on the World Wide Web*, 5(2), pp.58–71.
- Athanasiou, G. et al., 2015. A Trust Assessment mechanism for Ubiquitous Healthcare environment employing cloud theory. *Conference proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society.*, 2015, pp.1405–8.
- Au, R. & Croll, P., 2008. Consumer-Centric and Privacy-Preserving Identity Management for Distributed E-Health Systems. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE, pp. 234–234.
- Avancini, H. & Straccia, U., 2005. User recommendation for collaborative and personalised digital archives. *International Journal of Web Based Communities*, 1(2), p.163.
- Bahtiyar, Ş. & Çağlayan, M.U., 2014. Trust assessment of security for e-health systems. *Electronic Commerce Research and Applications*, 13(3), pp.164–177.

- Bailey, R.A., 2008. *Design of Comparative Experiments*, Cambridge University Press.
- Balabanović, M., 1997. An adaptive Web page recommendation service. In *Proceedings of the first international conference on Autonomous agents - AGENTS '97*. New York, New York, USA: ACM Press, pp. 378–385.
- Baldoni, M. et al., 2007. Reasoning-Based Curriculum Sequencing and Validation: Integration in a Service-Oriented Architecture E. Duval, R. Klamma, & M. Wolpers, eds. *Creating New Learning Experiences on a Global Scale*, 4753, pp.426–431.
- Basu, C., Hirsh, H. & Cohen, W., 1998. Recommendation as classification: Using social and content-based information in recommendation. *Communications*, pp, pp.714–720.
- Bellifemine, F. et al., 2002. *Jade programmer's guide*,
- Bellifemine, F., Caire, G. & Greenwood, D., 2007. *Developing Multi-Agent Systems with JADE*,
- Ben-Shimon, D. et al., 2007. Recommender system from personal social networks. In K. Wegrzyn-Wolska & P. Szczepaniak, eds. *Advances in Intelligent Web Mastering*. Advances in Soft Computing. Springer Berlin / Heidelberg, pp. 47–55.
- Berners-Lee, T., Hendler, J. & Lassila, O., 2001. The Semantic Web. *Scientific American*, pp.34–43.
- Berry, M.J. & Linoff, G.S., 2004. *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management* G. Linoff, ed., John Wiley & Sons.
- Biletskiy, Y. et al., 2009. An adjustable personalization of search and delivery of learning objects to learners. *Expert Systems with Applications*, 36(5), pp.9113–9120.
- Billsus, D. & Pazzani, M.J., 2000. User Modeling for Adaptive News Access. *User Modeling and UserAdapted Interaction*, 10(2), pp.147–180.
- Bittencourt, I.I. et al., 2009. A computational model for developing semantic web-based educational systems. *Knowledge-Based Systems*, 22(4), pp.302–315.
- Blanco-Fernandez, Y. et al., 2008. A flexible semantic inference methodology to reason about user preferences in knowledge-based recommender systems. *Knowledge-Based Systems*, 21(4), pp.305–320.
- Breslin, J.G., Passant, A. & Vrandec, D., 2011. *Handbook of Semantic Web Technologies* J. Domingue, D. Fensel, & J. A. Hendler, eds., Springer Berlin Heidelberg.
- Bridge, D. et al., 2005. Case-based recommender systems. *The Knowledge Engineering Review*, 20(03), p.315.
- Brockett, R.G. & Hiemstra, R., 1991. *Self-Direction in Adult Learning: Perspectives on Theory, Research, and Practice*. *Routledge Series on Theory and Practice of Adult Education in North America*., Routledge, Chapman and Hall, Inc.
- Buchegger, S. & Le Boudec, J.-Y., 2004. A Robust Reputation System for Peer-to-Peer and Mobile Ad-hoc Networks. In *Networks*.

- Burke, R., 2007. Hybrid Web Recommender Systems P. Brusilovsky, A. Kobsa, & W. Nejdl, eds. *The Adaptive Web*, 4321(THE ADAPTIVE WEB), pp.377 – 408.
- Burke, R., 2000. Knowledge-Based Recommender Systems. *Encyclopedia of Library and Information Systems*.
- Burke, R., 1999. The Wasabi Personal Shopper: A Case-Based Recommender System. In *Proceedings of the 11th National Conference on Innovative Applications of Artificial Intelligence*. pp. 844–849.
- Cai, D. et al., 2004. Block-level link analysis. *Proceedings of the 27th annual international conference on Research and development in information retrieval SIGIR 04*, p.440.
- Cantador, I., Bellogín, A. & Castells, P., 2008. A multilayer ontology-based hybrid recommendation model. *AI Communications*, 21(2-3), pp.203–210.
- Capuano, N. et al., 2014. Elicitation of latent learning needs through learning goals recommendation. *Computers in Human Behavior*, 30, pp.663–673.
- Carbo, J., Molina, J. & Davila, J., 2002. Comparing predictions of sporas vs. a fuzzy reputation agent system. In *Third International Conference on Fuzzy Sets and Fuzzy Systems*. Interlaken, pp. 147–153.
- Carbo, J., Molina, J. & Davila, J., 2003. Trust management through fuzzy reputation. *International Journal of Cooperative Information Systems*, 12(1), pp.135–155.
- Castelfranchi, C. & Falcone, R., 2001. Social Trust : A Cognitive Approach. In C. Castelfranchi & Y.-H. Tan, eds. *Trust and deception in virtual societies*. Kluwer Academic Publishers, pp. 55–90.
- Claypool, M. et al., 1999. Combining content-based and collaborative filters in an online newspaper. In *Proceedings of ACM SIGIR Workshop on Recommender Systems*. Citeseer, pp. 40–48.
- Cobos, C. et al., 2013. A hybrid system of pedagogical pattern recommendations based on singular value decomposition and variable data attributes. *Information Processing & Management*, 49(3), pp.607–625.
- Colley, H., Hodgkinson, P. & Malcolm, J., 2002. *Non-formal learning: mapping the conceptual terrain, a consultation report*, Leeds: University of Leeds, Lifelong Learning Institute.
- Collis, B. & Moonen, J., 2001. *Flexible learning in a digital world: Experiences and expectations*, London: Kogan Page Ltd.
- Cornelli, F. et al., 2002. Choosing reputable servants in a P2P network. In *Proceedings of the eleventh international conference on World Wide Web WWW 02*. WWW '02. ACM Press, pp. 376–386.
- Cristea, A.I., 2005. Authoring of Adaptive Hypermedia. *Educational Technology and Society*, 8(3), pp.6–8.
- Cukier, K. & Mayer-Schönberger, V., 2013. The Rise of Big Data: How It's Changing the Way We Think About the World. *Foreign Affairs*, 92(3), pp.1–7.

- Chae, Y., DiPippo, L.C. & Sun, Y.L., 2015. Trust Management for Defending On-Off Attacks. *IEEE Transactions on Parallel and Distributed Systems*, 26(4), pp.1178–1191.
- Chang, E., Dillon, T. & Hussain, F., 2006. *Trust and reputation for service-oriented environments: technologies for building business intelligence and consumer confidence*, Centre for Extended Enterprises and Business Intelligence.
- Chen, C.-M. & Duh, L.-J., 2008. Personalized web-based tutoring system based on fuzzy item response theory. *Expert Systems with Applications*, 34, pp.2298 – 2315.
- Chen, C.-M., Duh, L.-J. & Liu, C.-Y., 2004. A Personalized Courseware Recommendation System Based on Fuzzy Item Response Theory. In *IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE '04)*. IEEE Computer Society, pp. 305–308.
- Chen, G. & Chen, L., 2015. Augmenting service recommender systems by incorporating contextual opinions from user reviews. *User Modeling and User-Adapted Interaction*, 25(3), pp.295–329.
- Chen, L. & Sycara, K., 1998. WebMate : A Personal Agent for Browsing and Searching. In K. P. Sycara & M. Wooldridge, eds. *Knowledge Acquisition. AGENTS '98*. ACM Press, pp.132–139.
- Chen, M. & Singh, J.P., 2001. Computing and using reputations for internet ratings. *Proceedings of the 3rd ACM conference on Electronic Commerce EC 01*, 01, pp.154–162.
- Cho, Y.H., Kim, J.K. & Kim, S.H., 2002. A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications*, 23(3), pp.329–342.
- Choon-oh Lee et al., 2008. A framework for personalized Healthcare Service Recommendation. In *HealthCom 2008 - 10th International Conference on e-health Networking, Applications and Services*. IEEE, pp. 90–95.
- Christakou, C. & Stafylopatis, A., 2005. A hybrid movie recommender system based on neural networks. In *5th International Conference on Intelligent Systems Design and Applications (ISDA'05)*. IEEE, pp. 500–505.
- Damiani, E. et al., 2002. A reputation-based approach for choosing reliable resources in peer-to-peer networks A. V, ed. *Proceedings of the 9th ACM conference on Computer and communications security*, pages, pp.207–216.
- Dearden, A. et al., 2010. *Pervasive healthcare in lived experience: Thinking beyond the home*,
- Dellarocas, C., 2000. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. *Proceedings of the 2nd ACM conference on Electronic commerce EC 00*, 8(1999), pp.150–157.
- van Deursen, T., Koster, P. & Petković, M., 2008. Hedaquin: A Reputation-based Health Data Quality Indicator. *Electronic Notes in Theoretical Computer Science*, 197(2), pp.159–167.
- Devedzic, V., 2006. *Semantic Web and Education*, Springer-Verlag.
- Dingledine, R., Freedman, M. & Molnar, D., 2000. Accountability measures for peer-to-peer systems. In *Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. O'Reilly.

- Dolera Tormo, G. et al., 2013. Identity Management--In Privacy We Trust: Bridging the Trust Gap in eHealth Environments. *IEEE Security & Privacy*, 11(6), pp.34–41.
- Drachsler, H., Hummel, H., et al., 2009. Effects of the ISIS Recommender System for navigation support in self-organised Learning Networks. *Educational Technology & Society*, 12(3), pp.115–126.
- Drachsler, H. et al., 2010. ReMashed - An Usability Study of a Recommender System for Mash-Ups for Learning M. E. Auer, ed. *International Journal of Emerging Technologies in Learning iJET*, 5(S1), pp.7–11.
- Drachsler, H., Pecceu, D., et al., 2009. ReMashed – Recommendations for Mash-Up Personal Learning Environments U. Cress, V. Dimitrova, & M. Specht, eds. *Work*, 5794, pp.788–793.
- Drachsler, H., Hummel, H. & Koper, R., 2008. Using Simulations to Evaluate the Effects of Recommender Systems for Learners in Informal Learning Networks. *Learning*, 3(4), pp.404–423.
- Dron, J. et al., 2000. CoFIND: An experiment in n-dimensional collaborative filtering. *Journal of Network and Computer Applications*, 23(2), pp.131–142.
- Dwivedi, P. & Bharadwaj, K.K., 2013. Effective Trust-Aware E-learning Recommender System Based on Learning Styles and Knowledge Levels. *Educational Technology & Society*, 16(4), pp.201–216.
- Eirinaki, M., Louta, M.D. & Varlamis, I., 2014. A Trust-Aware System for Personalized User Recommendations in Social Networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(4), pp.409–421.
- Ekström M A, B.H.C., 2002. A rating system for AEC e-bidding that accounts for rater credibility. In *Proceedings of the CIB W65 Symposium*. pp. 736–766.
- Erdt, M., Fernandez, A. & Rensing, C., 2015. Evaluating Recommender Systems for Technology Enhanced Learning: A Quantitative Survey. *IEEE Transactions on Learning Technologies*, 8(4), pp.326–344.
- Eysenbach, G., 2008. Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness. *Journal of Medical Internet Research*, 10(3), p.e22.
- Fahrenholtz, D. & Lamersdorf, W., 2002. Transactional Security for a Distributed Reputation Management System. *University of Hamburg Computer Science Department*, 2455, pp.214–223.
- Farzan, R. & Brusilovsky, P., 2006. Social Navigation Support in a Course Recommendation System. In V. P. Wade, H. Ashman, & B. Smyth, eds. *Adaptive Hypermedia and Adaptive Web-Based Systems*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Fernandez-Luque, L., Karlsen, R. & Bonander, J., 2011. Review of extracting information from the Social Web for health personalization. *Journal of Medical Internet Research*, 13(1), p.e15.
- Frias-Martinez, E. et al., 2006. Automated user modeling for personalized digital libraries. *International Journal of Information Management*, 26(3), pp.234–248.

- Frias-Martinez, E., Chen, S.Y. & Liu, X., 2009. Evaluation of a personalized digital library based on cognitive styles: adaptivity vs. adaptability. *International Journal of Information Management*, 29(1), pp.48–56.
- Fullam, K.K., Klos, T.B., et al., 2005. A Specification of the Agent Reputation and Trust (ART) Testbed: Experimentation and Competition for Trust in Agent Societies. *Science And Technology*, pp.512–518.
- Fullam, K.K., Voss, M., et al., 2005. A specification of the Agent Reputation and Trust (ART) testbed. *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems AAMAS 05*, p.512.
- Fullam, K.; Barber, K.S., 2004. Evaluating approaches for trust and reputation research: exploring a competition testbed. In *Proceedings of The Workshop on Reputation in Agent Societies at Intelligent Agent Technology (IAT'2004)*. Beijing, pp. 20–23.
- Gama, J. et al., 2014. A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), pp.1–37.
- Gambetta, D., 2000. Can We Trust Trust? D. Gambetta, ed. *Trust Making and Breaking Cooperative Relations*, chapter13p, pp.213–237.
- Ghazanfar, M.A. & Prügel-Bennett, A., 2014. Leveraging clustering approaches to solve the gray-sheep users problem in recommender systems. *Expert Systems with Applications*, 41(7), pp.3261–3275.
- Golbeck, J., 2006. Generating predictive movie recommendations from trust in social networks K. Stølen et al., eds. *Trust Management*, 3986(x), pp.93–104.
- Golbeck, J., Parsia, B. & Hendler, J., 2003. Trust Networks on the Semantic Web M. Klusch et al., eds. *Cooperative Information Agents VII*, 1(1), pp.238–249.
- Goldberg, D. et al., 1992. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), pp.61–70.
- Gomez-Albarran, M. & Jimenez-Diaz, G., 2009. Recommendation and Students' Authoring in Repositories of Learning Objects: A Case-Based Reasoning Approach. *International Journal of Emerging Technologies in Learning*, 4, pp.227–232.
- Grandison, T. & Sloman, M., 2000. A survey of trust in internet applications. *IEEE Communications Surveys Tutorials*, 3(4), pp.2–16.
- Groh, G. & Ehlig, C., 2007. Recommendations in taste related domains: collaborative filtering vs. social filtering. In *Human Factors. GROUP '07*. ACM, pp. 127–136.
- Guha, R., 2003. Open Rating Systems. *1st Workshop on Friend of a Friend Social Networking*.
- Gupta, M., Judge, P. & Ammar, M., 2003. A reputation system for peer-to-peer networks. *Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video NOSSDAV 03*, pages, p.144.

- Gupta, R. & Singh, Y.N., 2015. Reputation Aggregation in Peer-to-Peer Network Using Differential Gossip Algorithm. *IEEE Transactions on Knowledge and Data Engineering*, 27(10), pp.2812–2823.
- Guy, I. et al., 2009. Personalized recommendation of social software items based on social relations. *Proceedings of the third ACM conference on Recommender systems RecSys 09*, New York,(7), p.53.
- Hage, P. & Harary, F., 1984. *Structural Models in Anthropology*, Cambridge University Press.
- Hamel, G., Doz, Y.L. & Prahalad, C.K., 1989. Collaborate with your competitors - and win. *Harvard Business Review*, 67(1), pp.133–139.
- Hang, C.-W., Kalia, A.K. & Singh, M.P., 2012. Behind the Curtain: Service Selection via Trust in Composite Services. In *2012 IEEE 19th International Conference on Web Services*. IEEE, pp.9–16.
- Hang, C.W., Wang, Y. & Singh, M.P., 2008. An adaptive probabilistic trust model and its evaluation. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems Volume 3*. International Foundation for Autonomous Agents and Multiagent Systems, pp. 1485–1488.
- Harbers, M. et al., 2007. The examination of an information-based approach to trust. In J. S. Sichman et al., eds. *Proc 5th European Symposium on Research in Computer Security*. Lecture Notes in Computer Science. Springer-Verlag, pp. 71–82.
- Herlocker, J.L. et al., 2004. Evaluating collaborative filtering recommender systems H. J. Scholl, ed. *ACM Transactions on Information Systems*, 22(1), pp.5–53.
- Hermoso, R., Billhardt, H. & Ossowski, S., 2006. Integrating trust in virtual organisations. In *Coordination, Organisation, Institution and Norms in Multiagent Systems (II)*. *Lectures Notes in Artificial Intelligence*. Springer Berlin Heidelberg, p. 13.
- Hermoso, R. & Centeno, R., 2007. Supporting Experimentation for Trust Models in Virtual Organizations: TOAST. In *Proceedings of 5th European Workshop on Multiagent Systems (EUMAS 2007)*. Hammamet, pp. 368–383.
- Hill, W. et al., 1995. Recommending and evaluating choices in a virtual community of use. In I. R. Katz et al., eds. *Proceedings of the ACM Conference on Human Factors in Computing Systems*. Papers: Using the Information of Others. ACM Press/Addison-Wesley Publishing Co., pp. 194–201.
- Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), pp.16569–72.
- Hoens, R., Blanton, M. & Chawla, N., 2010. Reliable medical recommendations systems with patient privacy. In *Proceedings of the International Health Informatics Symposium*.
- Hsu, S. et al., 2007. AIMED- A Personalized TV Recommendation System. In P. Cesar, K. Chorianopoulos, & J. F. Jensen, eds. *Interactive TV: a Shared Experience*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 166–174.

- Huang, Y. et al., 2009. A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web. *Learning*, 12(2), pp.144–162.
- Hummel, H.G.K. et al., 2007. Combining social-based and information-based approaches for personalised recommendation on sequencing learning activities. *International Journal of Learning Technology*, 3(2), p.152.
- Huynh, T.D., 2006. *Trust and Reputation in Open Multi-Agent Systems*. University of Southampton.
- Huynh, T.D., Jennings, N.R. & Shadbolt, N., 2004. Developing an integrated trust and reputation model for open multi-agent systems. *7th International Workshop on Trust in Agent Societies*, pp.65–74.
- Huynh, T.D., Jennings, N.R. & Shadbolt, N.R., 2006. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13(2), pp.119–154.
- Ibnkahla, M., 2000. Applications of neural networks to digital communications - a survey. *Signal Processing*, 80(7), pp.1185–1215.
- Jelenc, D. et al., 2013. Decision making matters: A better way to evaluate trust models. *Knowledge-Based Systems*, 52, pp.147–164.
- Jennings, A. & Higuchi, H., 1993. A User Model Neural Network for a Personal News Service. *User Modeling and UserAdapted Interaction*, 3(1), pp.1–25.
- Jøsang, A., 2001. A logic for uncertain probabilities. *International Journal of Uncertainty Fuzziness and KnowledgeBased Systems*, 9(3), pp.279–311.
- Jøsang, A. & Ismail, R., 2002. The Beta Reputation System. *Proceedings of the 15th Bled Electronic Commerce Conference*, 160(Citeseer), pp.324–337.
- Jøsang, A., Ismail, R. & Boyd, C., 2007. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2), pp.618–644.
- Jøsang, A. & Lo Presti, S., 2004. Analysing the Relationship Between Risk and Trust C. Jensen, S. Poslad, & T. Dimitrakos, eds. *Trust Management*, LNCS 2995, pp.135–145.
- Jurca, R. & Faltings, B., 2003. *An incentive compatible reputation mechanism* W. M. Y. M. Rosenschein J S Sandholm T, ed., IEEE Computer Society.
- Kai Yu, Volker Tresp, S.Y., 2004. A Nonparametric Hierarchical Bayesian Framework For Information Filtering. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. pp. 353–360.
- Kamvar, S.D., Schlosser, M.T. & Garcia-Molina, H., 2003. The Eigentrust algorithm for reputation management in P2P networks Y. F. R. Chen, L. Kovács, & S. Lawrence, eds. *Proceedings of the twelfth international conference on World Wide Web WWW 03*, 03, p.640.
- Karlins, M. & Abelson, H.I., 1970. *Persuasion: How opinions and attitudes are changed.*, Crosby Lockwood & Son.

- Katzenbeisser, S. & Petkovic, M., 2008. Privacy-Preserving Recommendation Systems for Consumer Healthcare Services. *2008 Third International Conference on Availability Reliability and Security*, pp.889–895.
- Kerr, R. & Cohen, R., 2009. Smart Cheaters Do Prosper : Defeating Trust and Reputation Systems. *Scenario*, pp.993–1000.
- Kerr, R. & Cohen, R., 2010. TREET: The Trust and Reputation Experimentation and Evaluation Testbed. *Electronic Commerce Research*, 10(3-4), pp.271–290.
- Kim, H.K.K.H.K., Kim, J.K.K.J.K. & Ryu, Y.U., 2009. *Personalized Recommendation over a Customer Network for Ubiquitous Shopping*, IEEE Computer Society.
- Kim, H.-N. et al., 2010. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. *Electronic Commerce Research and Applications*, 9(1), pp.73–83.
- Kim, J.-H. et al., 2009. Design of Diet Recommendation System for Healthcare Service Based on User Information. In *2009 Fourth International Conference on Computer Sciences and Convergence Information Technology*. IEEE, pp. 516–518.
- Kim, J.K. et al., 2002. A personalized recommendation procedure for Internet shopping support. *Electronic Commerce Research and Applications*, 1(3-4), pp.301–313.
- KIM, K. & AHN, H., 2008. A recommender system using GA K-means clustering in an online shopping market. *Expert Systems with Applications*, 34(2), pp.1200–1209.
- Kim, Y.A. & Phalak, R., 2012. A trust prediction framework in rating-based experience sharing social networks without a Web of Trust. *Information Sciences*, 191, pp.128–145.
- Koren, Y., 2009. Collaborative filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '09*. New York, New York, USA: ACM Press, p. 447.
- Koren, Y., Bell, R. & Volinsky, C., 2009. Matrix Factorization Techniques for Recommender Systems, *Computer*, 42(8), pp.30-37.
- Kravari, K. & Bassiliades, N., 2015. A Survey of Agent Platforms. *Journal of Artificial Societies and Social Simulation*, 18(1).
- Krishna, S. et al., 2014. A trust enhanced Recommender System for medicare applications. In *2014 Sixth International Conference on Advanced Computing (ICoAC)*. IEEE, pp. 324–328.
- Krulwich, B., 1997. Lifestyle Finder: Intelligent User Profiling Using Large-Scale Demographic Data. *Artificial Intelligence Magazine*, 18(2), pp.37–45.
- Lang, K., 1995. NewsWeeder: learning to filter netnews. In A. Prieditis & S. J. Russell, eds. *Proceedings of the 12th International Conference on Machine Learning*. Morgan Kaufmann, pp. 331–339.
- Larburu Rubio, N. et al., 2014. Quality of data computational models and telemedicine treatment effects. In *6th IEEE International Conference on e-Health Networking, Healthcom*.

- Lee, J.-S. & Zhu, D., 2012. Shilling Attack Detection—A New Approach for a Trustworthy Recommender System. *INFORMS Journal on Computing*, 24(1), pp.117–131.
- Lemire, D. et al., 2005. Collaborative filtering and inference rules for context-aware learning object recommendation. *Interactive Technology and Smart Education*, 2(3), pp.1–11.
- Lemire, D., 2005. Scale and Translation Invariant Collaborative Filtering Systems. *Information Retrieval*, 8(1), pp.129–150.
- Lemire, D. & Boley, H., 2003. RACOFI: A Rule-Aplying Collaborative Filtering System. *Database*, (October), pp.13–24.
- Levien, R., 2003. *Advogato Trust Metric*. UC Berkely.
- Li, J. & Zaman, N., 2014. Personalized Healthcare Recommender Based on Social Media. In *2014 IEEE 28th International Conference on Advanced Information Networking and Applications*. IEEE, pp. 993–1000.
- Li, X. & Ling, L., 2004. PeerTrust: Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities. *IEEE Transactions on Knowledge and Data Engineering*, 16(07), pp.843–857.
- Lihua, W. et al., 2005. Modeling user multiple interests by an improved GCS approach. *Expert Systems with Applications*, 29(4), pp.757–767.
- Lu, J., 2004. A personalized e-learning material recommender system. In *Proceedings of the 2nd International Conference on Information Technology and Applications*. Harbin, China.
- Lu, J. et al., 2015. Recommender system application developments: A survey. *Decision Support Systems*, 74, pp.12–32.
- Luke, S. et al., 2005. MASON: A Multiagent Simulation Environment. *SIMULATION*, 81(7), pp.517–527.
- Mahmood, T. & Ricci, F., 2009. Improving recommender systems with adaptive conversational strategies. *Proceedings of the 20th ACM conference on Hypertext and hypermedia HT 09*, p.73.
- Malhotra, N.K., 1999. *Marketing Research: An Applied Orientation* R. Grover & M. Vriens, eds., Prentice Hall.
- Malik, Z. & Bouguettaya, A., 2008. Rater Credibility Assessment in Web Services Interactions. *World Wide Web*, 12(1), pp.3–25.
- Malik, Z. & Bouguettaya, A., 2009. RATEWeb: Reputation Assessment for Trust Establishment among Web services. *V LDB Journal*, 18(4), pp.885–911.
- Manchala, D.W., 1998. Trust Metrics, Models and Protocols for Electronic Commerce Transactions. *Xerox Corporate Research and Technology*, pp.312–321.
- Manouselis, N. et al., 2011. Recommender Systems in Technology Enhanced Learning. In F. Ricci et al., eds. *Recommender Systems Handbook*. Boston, MA: Springer US, pp. 387–415.

- Manouselis, N. & Costopoulou, C., 2007. Experimental Analysis of Design Choices in Multiattribute Utility Collaborative Filtering. *International Journal of Pattern Recognition and Artificial Intelligence*, 21(2), p.311.
- Marsden, P. V & Lin, N., 1982. *Social structure and network analysis* P. V Marsden & N. Lin, eds., Sage Publications.
- Marsh, S.P., 1994. Formalising Trust as a Computational Concept. *Computing*, Doctoral Thesis, p.184.
- Martínez-Sarriegui, I. et al., 2011. eCONSULTA: Integración de un Sistema de Videoconsulta Web entre Asistencia Primaria y Atención Especializada. In *Libro de Actas del XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica, CASEIB 2001*. Cáceres.
- Martínez-Sarriegui, I. et al., 2012. TRHIOS : Trust and Reputation in Hierarchical and Quality-Oriented Societies. In *7th Iberian Conference on Information Systems and Technologies (CISTI'12)*. Madrid.
- Massa, P. & Avesani, P., 2004. Trust-aware collaborative filtering for recommender systems R. Meersman & Z. Tari, eds. *On the Move to Meaningful Internet Systems 2004 CoopIS DOA and ODBASE*, 3290(8), pp.492–508.
- Massa, P. & Avesani, P., 2007. Trust-aware recommender systems. In *Proceedings of the 2007 ACM conference on Recommender systems - RecSys '07*. New York, New York, USA: ACM Press, p. 17.
- Massa, P. & Bhattacharjee, B., 2004. Using trust in recommender systems: an experimental analysis C. D. Jensen, S. Poslad, & T. Dimitrakos, eds. *Trust Management*, 2995, pp.221–235.
- Maximilien, E.M. & Singh, M.P., 2005. Multiagent System for Dynamic Web Services Selection. In *Proceedings of 1st Workshop on Service-Oriented Computing and Agent-Based Engineering (SOCABE at AAMAS)*. pp. 25–29.
- Mcknight, D.H. & Chervany, N.L., 1996. The meanings of trust G. M. Williams & Jr., eds. *Measurement*, 55455(612), p.86.
- McNee, S.M., 2006. Meeting user information needs in recommender systems. *Library*.
- McSherry, D., 2002. Diversity-Conscious Retrieval S Craw & A Preece, eds. *Lecture Notes in Artificial Intelligence*, 2416, pp.219–233.
- McSherry, F. & Mironov, I., 2009. Differentially private recommender systems: building privacy into the net. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, pp. 627–636.
- Mehdi, M., Bouguila, N. & Bentahar, J., 2013. A QoS-Based Trust Approach for Service Selection and Composition via Bayesian Networks. In *2013 IEEE 20th International Conference on Web Services*. IEEE, pp. 211–218.
- Melville, P., Mooney, R.J. & Nagarajan, R., 2002. Content-boosted collaborative filtering for improved recommendations. *Machine Learning*, telligence(July), pp.187–192.

- Miller, N.H., Resnick, P. & Zeckhauser, R.J., 2002. Eliciting Honest Feedback in Electronic Markets. *SSRN Electronic Journal*.
- Mobasher, B., Jin, X. & Zhou, Y., 2004. Semantically Enhanced Collaborative Filtering on the Web B. Berendt et al., eds. *Information Systems Journal*, 3209(49), pp.57–76.
- Montaner, M., 2003. A Taxonomy of Recommender Agents on the Internet. *Artificial Intelligence Review*, 19(4), pp.285–330.
- Monteiro, E. et al., 2015. A recommender system for medical imaging diagnostic. *Studies in health technology and informatics*, 210, pp.461–3.
- Montuschi, P. et al., 2015. A Semantic Recommender System for Adaptive Learning. *IT Professional*, 17(5), pp.50–58.
- Mooney, R.J. & Roy, L., 1999. Content-Based Book Recommending Using Learning for Text Categorization N. Y. Acm Ny, ed. *Proceedings of the fifth ACM conference on Digital libraries DL 00*, p.8.
- Mui, L., Mohtashemi, M. & Ang, C., 2001. A Probabilistic Rating Framework for Pervasive Computing Environments. In *Proceedings of the MIT Student Oxygen Workshop (SOW'2001)*.
- Mui, L., Mohtashemi, M. & Halberstadt, A., 2002. *A computational model of trust and reputation*, IEEE Computer Society.
- Nakagawa, M. & Mobasher, B., 2003. A hybrid web personalization model based on site connectivity. *Proceedings of WebKDD*, pp.59–70.
- Nelson Minar, Roger Burkhart Y, C.L.Z., 1996. *The Swarm Simulation System: A Toolkit for Building Multi-Agent Simulations*, Santa Fe.
- Nguyen, H.T., Zhao, W. & Yang, J., 2010. A Trust and Reputation Model Based on Bayesian Network for Web Services. In *2010 IEEE International Conference on Web Services*. IEEE, pp. 251–258.
- Nilashi, M., Ibrahim, O. bin & Ithnin, N., 2014. Multi-criteria collaborative filtering with high accuracy using higher order singular value decomposition and Neuro-Fuzzy system. *Knowledge-Based Systems*, 60, pp.82–101.
- Noorian, Z. & Ulieru, M., 2010. The State of the Art in Trust and Reputation Systems: A Framework for Comparison. *Journal of theoretical and applied electronic commerce research*, 5(2).
- Oh, H.-K. et al., 2015. Can You Trust Online Ratings? A Mutual Reinforcement Model for Trustworthy Online Rating Systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(12), pp.1564–1576.
- Oropesa, I. et al., 2011. Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery. *Journal of Surgical Research*, 171(1), pp.e81–95.
- Page, F.S. et al., 2006. *Review of FIPA Specifications*,

- Page, L. et al., 1998. The PageRank Citation Ranking: Bringing Order to the Web. *World Wide Web Internet And Web Information Systems*, 54(2), pp.1–17.
- Pappas, N. & Popescu-Belis, A., 2015. Adaptive sentiment-aware one-class collaborative filtering. *Expert Systems with Applications*, 43, pp.23–41.
- Paradesi, S., Doshi, P. & Swaika, S., 2009. Integrating Behavioral Trust in Web Service Compositions. In *2009 IEEE International Conference on Web Services*. IEEE, pp. 453–460.
- Park, D.H. et al., 2012. A literature review and classification of recommender systems research. *Expert Systems with Applications*, 39(11), pp.10059–10072.
- Pazzani, M.J., 1999. A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, 13(5), pp.393–408.
- Pazzani, M.J., Muramatsu, J. & Billsus, D., 1996. Syskill & Webert: Identifying interesting web sites. In D. Weld & B. Clancey, eds. *AAAI/AAI Vol 1*. AAAI Press / MIT Press, pp. 54–61.
- Pearl, J., 1988. Probabilistic reasoning in intelligent systems: networks of plausible inference.
- Porcel, C., Moreno, J.M. & Herrera-Viedma, E., 2009. A multi-disciplinar recommender system to advice research resources in University Digital Libraries. *Expert Systems with Applications*, 36(10), pp.12520–12528.
- Rafaeli, S. et al., 2004. QSIA – a Web-based environment for learning , assessing and knowledge sharing in communities. *Computers & Education*, 43(3), pp.273–289.
- Rafaeli, S., Dan-Gur, Y. & Barak, M., 2005. Social Recommender Systems: Recommendations in Support of E-Learning. *International Journal of Distance Education Technologies*, 3(2), pp.30–47.
- Railsback, S.F., Lytinen, S.L. & Jackson, S.K., 2006. Agent-based Simulation Platforms: Review and Development Recommendations. *SIMULATION*, 82(9), pp.609–623.
- Ramchurn, S.D., Huynh, D. & Jennings, N.R., 2005. Trust in multi-agent systems. *The Knowledge Engineering Review*, 19(01), pp.1–25.
- Recker, M.M., Walker, A. & Lawless, K., 2003. What do you recommend? Implementation and analyses of collaborative information filtering of web resources for education. *Instructional Science*, 31(4), pp.299–316.
- Resnick, P. et al., 1994. GroupLens : An Open Architecture for Collaborative Filtering of Netnews. In R. Furuta & C. M. Neuwirth, eds. *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. CSCW '94. ACM, pp. 175–186.
- Resnick, P. et al., 2006. The value of reputation on eBay: A controlled experiment. *Experimental Economics*, 9(2), pp.79–101.
- Resnick, P. & Varian, H.R., 1997. Recommender systems. *Communications of the ACM*, 40(3), pp.56–58.
- Resnick, P. & Zeckhauser, R., 2002. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system M. R. Baye, ed. *Most*, 11(12), pp.127–157.

- Ricci, F., Rokach, L. & Shapira, B., 2011. Introduction to Recommender Systems Handbook. In F. Ricci et al., eds. *Recommender Systems Handbook*. Springer US, pp. 1–35.
- Romero, C. et al., 2009. Applying Web usage mining for personalizing hyperlinks in Web-based adaptive educational systems. *Computers & Education*, 53(3), pp.828–840.
- Romero Zaldivar, V.A. & Burgos, D., 2010. Meta-Mender: A meta-rule based recommendation system for educational applications N. Manouselis et al., eds. *Procedia Computer Science*, 1(2), pp.2877–2882.
- Rosa, R., Rodriguez, D. & Bressan, G., 2015. Music Recommendation System Based on User's Sentiments Extracted from Social Networks. *IEEE TRANSACTIONS ON CONSUMER ELECTRONICS*, 61(3), pp.359–367.
- Rowe, R. & Calnan, M., 2006. Trust relations in health care--the new agenda. *European journal of public health*, 16(1), pp.4–6.
- Ruohomaa, S., Kutvonen, L. & Koutrouli, E., 2007. Reputation Management Survey. In *The Second International Conference on Availability, Reliability and Security (ARES'07)*. IEEE, pp. 103–111.
- Sabater, J., 2004a. Evaluating the ReGreT system. *Applied Artificial Intelligence*, 18(9-10), pp.797–813.
- Sabater, J., 2004b. Toward a test-bed for trust and reputation models. In *Falcone et al.* pp. 101–105.
- Sabater, J. & Sierra, C., 2001a. REGRET: A reputation model for gregarious societies. In *4th Workshop on Deception, Fraud and Trust in Agent Societies*. ACM, pp. 61 – 69.
- Sabater, J. & Sierra, C., 2001b. REGRET: reputation in gregarious societies. In *Proceedings of the fifth international conference on Autonomous agents*. ACM, pp. 194–195.
- Sabater, J. & Sierra, C., 2002a. Reputation and Social Network Analysis In Multi-Agent Systems. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*. ACM, pp. 475–482.
- Sabater, J. & Sierra, C., 2002b. Social ReGreT, a reputation model based on social relations. *ACM SIGecom Exchanges - Chains of commitment*, 3(1), pp.44–56.
- Safran, D.G. et al., 1998. Linking primary care performance to outcomes of care. *The Journal of family practice*, 47(3), pp.213–220.
- Sami, A. et al., 2008. Design of physical activity recommendation system. IADIS European Conference on Data Mining.
- Sánchez-González, P. et al., 2011. TELMA: Technology enhanced learning environment for Minimally Invasive Surgery. *Procedia Computer Science*, 3(0), pp.1–9.
- Sánchez-Vera, M. del M. et al., 2012. Semantic Web technologies for generating feedback in online assessment environments. *Knowledge-Based Systems*, 33, pp.152–165.

- Santos, O.C., 2008. A recommender system to provide adaptive and inclusive standard-based support along the elearning life cycle. *Proceedings of the 2008 ACM conference on Recommender systems RecSys 08*, p.319.
- Santos, O.C., Boticario, J.G. & Pérez-Marín, D., 2014. Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle. *Science of Computer Programming*, 88, pp.92–109.
- Schafer, J. Ben, Konstan, J.A. & Riedl, J., 2001. E-Commerce Recommendation Applications. *Data Mining and Knowledge Discovery*, 5(1), pp.115–153.
- Schillo, M., Funk, P. & Rovatsos, M., 2000. Using trust for detecting deceitful agents in artificial societies. *Applied Artificial Intelligence*, 14(8), pp.825–848.
- Van Setten, M., 2005. *Supporting people in finding information: hybrid recommender systems and goal-based structuring*, Telematica Instituut.
- Sezgin, E. & Ozkan, S., 2013. A systematic literature review on Health Recommender Systems. In *2013 E-Health and Bioengineering Conference (EHB)*. IEEE, pp. 1–4.
- Shafer, G., 1976. Dempster-Shafer Theory S. C. Shapiro, ed. *International Journal of Approximate Reasoning*, 21(2), pp.1–2.
- Shambour, Q. & Lu, J., 2011. A hybrid trust-enhanced collaborative filtering recommendation approach for personalized government-to-business e-services. *International Journal of Intelligent Systems*, 26(9), pp.814–843.
- Shani, G. & Gunawardana, A., 2011. Evaluating Recommendation Systems F. Ricci et al., eds. *Recommender Systems Handbook*, 12(19), pp.1–41.
- Shardanand, U. & Maes, P., 1995. Social Information Filtering: Algorithms for Automating “Word of Mouth”. In *CHI 95 Proceedings of the SIGCHI conference on Human factors in computing systems*. CHI '95. ACM Press/Addison-Wesley Publishing Co., pp. 210–217.
- Shen, L. & Shen, R., 2004. Learning Content Recommendation Service Based-on Simple Sequencing Specification Liu Wy And Shi Yc And Li Q, ed. *Interchange*, 3143(60372078), pp.363–370.
- Sherchan, W., Loke, S.W. & Krishnaswamy, S., 2006. A fuzzy model for reasoning about reputation in web services. In *Proceedings of the 2006 ACM symposium on Applied computing - SAC '06*. New York, New York, USA: ACM Press, p. 1886.
- Shinde, S.K. & Kulkarni, U., 2012. Hybrid personalized recommender system using centering-bunching based clustering algorithm. *Expert Systems with Applications*, 39(1), pp.1381–1387.
- Shripad, K.V. & Vaidya, A.S., 2014. A Review on Trust-Aware and Privacy Preserving Profile Matching System for Personalized User Recommendations in Social networks. *International Journal of Computer Applications*, 104(12), pp.22–25.
- Singh, M.P. & Sycara, K., 2004. Developing trust in large-scale peer-to-peer systems. In *IEEE First Symposium on Multi-Agent Security and Survivability, 2004*. IEEE, pp. 1–10.

- Sinha, R. & Swearingen, K., 1999. Comparing Recommendations Made by Online Systems and Friends. *Interface*, 106.
- Smyth, B. & Cotter, P., 2000. A personalised TV listings service for the digital TV age. *Knowledge-Based Systems*, 13(2-3), pp.53–59.
- Stojanovic, L., Staab, S. & Studer, R., 2001. eLearning based on the Semantic Web W. A. L. Fowler & J. Hasebrook, eds. *WebNet2001World Conference on the WWW and Internet*, 4(Neidl), pp.23–27.
- Tadelis, S., 2003. Firm reputation with hidden information. *Economic Theory*, 21(2-3), pp.635–651.
- Tan, L., 2010. Future internet: The Internet of Things. *2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, pp.V5–376–V5–380.
- Tang, T. & McCalla, G., 2005. Smart Recommendation for an Evolving E-Learning System: Architecture and Experiment. *International Journal on E-Learning*, 4(1), pp.105–129.
- Teacy, W.T.L. et al., 2006. TRAVOS: Trust and reputation in the context of inaccurate information sources. *Journal of Autonomous Agents and Multi-Agent Systems*, 12(2), pp.183–198.
- Teece, D.J., 1992. Competition, cooperation, and innovation: Organizational arrangements for regimes of rapid technological process. *Journal of Economic Behaviour and Organization*, 18, pp.1–25.
- Thurrow, N.A. & Delano, J.D., 2010. Selection of Web Services Based on Opinion Mining of Free-Text User Reviews. In R. Sabherwal & M. Sumner, eds. *Proceedings of the International Conference on Information Systems*. Association for Information Systems, p. 42.
- Valarmathi, J. et al., 2011. A novel trust management scheme in pervasive healthcare. In *Recent Trends in Information Technology (ICRTIT)*, *International Conference on*. IEEE, pp. 503–508.
- Vavilis, S., Petković, M. & Zannone, N., 2014. A reference model for reputation systems. *Decision Support Systems*, 61, pp.147–154.
- Veer, E., 2011. Being Online: How the Internet is Changing Research for Consumers. *Journal of Research for Consumers*, (20), pp.1–7.
- Verbert, K. et al., 2011. Dataset-driven Research for Improving Recommender Systems for Learning. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pp.44–53.
- Vuorikari, R. & Berendt, B., 2009. Study on contexts in tracking usage and attention metadata in the field of multilingual Technology Enhanced Learning S. Fischer, E. Maehle, & R. Reischuk, eds. *Informatik 2009 im Focus das Leben*, In S. Fisc, pp.1654–1663.
- Wahab, O.A. et al., 2015. A survey on trust and reputation models for Web services: Single, composite, and communities. *Decision Support Systems*, 74, pp.121–134.
- Walker, A. et al., 2004. Collaborative Information Filtering: A Review and an Educational Application. *International Journal of Artificial Intelligence in Education*, 14(1), pp.3–28.

- Whitby, A., Jøsang, A. & Indulska, J., 2004. Filtering out unfair ratings in bayesian reputation systems. *Science*, 4(2)(2), pp.106–117.
- Wiesner, M. & Pfeifer, D., 2010. Adapting recommender systems to the requirements of personal health record systems. In *Proceedings of the ACM international conference on Health informatics - IHI '10*. New York, New York, USA: ACM Press, p. 410.
- Wilson, D.C., Smyth, B. & Sullivan, D.O., 2003. Sparsity Reduction in Collaborative Recommendation: A Case-Based Approach. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(05), pp.863–884.
- Wooldridge, M., 2002. *An Introduction to Multiagent Systems*, John Wiley & Sons.
- Wu, D., Lu, J. & Zhang, G., 2015. A Fuzzy Tree Matching-Based Personalized E-Learning Recommender System. *IEEE Transactions on Fuzzy Systems*, 23(6), pp.2412–2426.
- Xue, G.-R. et al., 2005. Scalable collaborative filtering using cluster-based smoothing. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '05*. New York, New York, USA: ACM Press, p. 114.
- Yahyaoui, H., 2012. A trust-based game theoretical model for Web services collaboration. *Knowledge-Based Systems*, 27, pp.162–169.
- Yahyaoui, H., 2010. Trust Assessment for Web Services Collaboration. In *2010 IEEE International Conference on Web Services*. IEEE, pp. 315–320.
- Yu, B. et al., 2002. An evidential model of distributed reputation management. *Proceedings of the first international joint conference on Autonomous agents and multiagent systems part 1 AAMAS 02*, 61(4), pp.294–301.
- Yu, B. & Singh, M.P., 2003. Detecting deception in reputation management. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems AAMAS 03*. ACM Press, p. 73.
- Yuan, W. et al., 2007. Using Reputation System in Ubiquitous Healthcare. In *9th International Conference on e-Health Networking, Application and Services*. IEEE, pp. 182–186.
- Yuan, W., Guan, D. & Lee, S., 2008. Trust management for ubiquitous healthcare. In *Parallel and Distributed Processing with Applications, 2008. ISPA'08. International Symposium on*. IEEE, pp. 63–70.
- Zacharia, G., Moukas, A. & Maes, P., 1999. Collaborative reputation mechanisms for electronic marketplaces. *Decision Support Systems*, 29(4), pp.371–388.
- Zaiane, O.R., 2002. Building a recommender agent for e-learning systems. In *International Conference on Computers in Education, 2002. Proceedings*. IEEE Comput. Soc, pp. 55–59.
- Zaman, N. & Li, J., 2014. Semantics-Enhanced Recommendation System for Social Healthcare. In *2014 IEEE 28th International Conference on Advanced Information Networking and Applications*. IEEE, pp. 765–770.

- Zhang, H.-R. & Min, F., 2016. Three-way recommender systems based on random forests. *Knowledge-Based Systems*, 91, pp.275–286.
- Ziegler, C.-N., 2004. Semantic Web Recommender Systems. In *Proceedings of the Joint ICDE/EDBT Ph.D. Workshop*. Heraklion.
- Ziegler, C.N. & Lausen, G., 2004. *Spreading activation models for trust propagation* S.-T. Yuan & J. Liu, eds., IEEE International Conference on eTechnology, eCommerce and eService 2004 EEE 04, pp.83-97.

VIII. ANNEXES

In this chapter the Annexes of the PhD Dissertation are presented, with additional or complementary information to the material presented in the body of the document.

VIII.1. Annex A: Dictionary of TruQOs and TST terms

Term	Definition
Agent	Each of the software entities of the multi-agent system conforming TruQOs Simulation Tool (TST). We distinguish three types: <ul style="list-style-type: none"> • jade-agents: Agents created automatically by JADE⁵⁵, the framework used to create the tool. • tst-agents: Internal agents of TST, created in this PhD thesis work to support the recommendation model and simulation engine. • user-agents: Agents created by TST as simulated users of the system in which the designer wants to test the performance of TruQOs.
Author (Auth)	Provider.
Consumer (C)	User requesting a recommendation. Demands and consumes resources available in the system of a desired Service, as well as generates direct inquiries to receive a Service from a Provider.
Designer	The final user of the simulation tool. The person who wants to evaluate and to analyse the performance of TruQOs as a recommendation engine with the ultimate intention of personalising the parameters of TruQOs to integrate it with the real open system he/she has modelled using TST for testing it.
Evaluation (ev)	Objective assessment of the quality of a resource generated by a provider; calculated automatically or by impartial experts.
Hierarchy	Users' classification or organization in levels inside an Institution, based on their authority, experience, professional standing or ability; and the dependency relation among them. In TruQOs it has a tree shape: each node, except the higher, can have one or more child in lower levels.
Institution (Inst)	Group of Users sharing common abilities, knowledge, services or just legal entity. Affiliation of every User: in TruQOs all the Users belong to, and only to, one Institution.
Opinion (op)	Consumers' personal (subjective) perception of the quality of a resource or a Service provided by a provider.
Provider (P)	User whose reputation is being assessed to provide recommendation to a consumer. Creates resources of a specific Service. Additionally answers to direct request from consumers.

⁵⁵ <http://jade.cselt.it/>

Term	Definition
Ranking	Institutions' Ranking: classification based on the quality of the results offered by each institution's providers for a specific Service. The position in the ranking varies over time according to the variations in the Reputation generating resources of the providers of each Institution. Unlike Hierarchy, each node can only have one child, thus conforming a linear tree. For every Service a different tree is calculated.
Rating	Numeric value assigned to an opinion.
Reliability (RL)	Indicates the level of trustworthiness of the calculated Reputation. It reflects the ability of a provider to generate a resource or offer a Service under stated conditions for a specified period of time.
Reputation (R)	The estimation in which a provider or a resource is held by the consumers of the system. It is associated to the quality of the resource or the Service provided.
Reputation' (R')	Expresses the final estimation of trustworthiness of a resource or a provider, calculated aggregating the Reputation and its Reliability
Service	In TruQOs, each of the different type of resources generated by the providers; it can be considered as the final node of a taxonomy of resources.
Trust	Confidence in the certainty of the quality of future resources or Services offered by providers based on the quality of past resources or past experiences of consumers. Assessed through Reputation's calculation.
Witness (Wit, w)	Consumer that has previously consumed a resource generated by a provider, or asked to the provider for a Service, and whose opinion about that interaction is requested in order to build up the Reputation of the provider; used to provide a recommendation to another consumer.

VIII.2. Annex B: Summary of equations

VIII.2.1. TruQOs

VIII.2.1.1. Resource Reputation

$$R_{Resource} = \sum_N \rho(t, t_j) \cdot op_j^{ResX} \in [0, 1]$$

$$RL_{Resource} = (1 - 2 \cdot \sigma)^q \in [0, 1]$$

$$\rho(t, t_j) = \frac{t_j}{t}$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (op_i^{ResX} - \overline{op^{ResX}})^2}$$

VIII.2.1.2. Provider Reputations

VIII.2.1.2.1. Reputation based on provider's past evaluations

$$R_{ProviderEv} = \sum_N \rho(t, t_j) \cdot ev_j \in [0, 1]$$

$$RL_{ProviderEv} = (1 - 2 \cdot \sigma)^q \in [0, 1]$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (ev_i - \overline{ev})^2}$$

VIII.2.1.2.2. Reputations based on past opinions about the provider

VIII.2.1.2.2.1. Reputation base on consumer's personal experience

$$R_{Personal} = R_{personal}(C \rightarrow P) = \sum_N \rho(t, t_j) \cdot op_j \in [0, 1]$$

$$RL_{Personal} = RL_{personal}(C \rightarrow P) = (1 - \mu) \cdot N_i(IDB) + \mu \cdot D_t(IDB) \in [0, 1]$$

$$N_i(IDB) = \begin{cases} \sin\left(\frac{\pi}{2 \cdot itm} |IDB|\right) & |IDB| \in [0, itm] \\ 1 & otherwise \end{cases}$$

$$D_t(IDB) = 1 - \sum_{impressions} \rho(t, t_j) \cdot |op_i - R_{Personal}| \quad)$$

VIII.2.1.2.2. Reputation based on social information

$$R_W = \frac{1}{N_W} \cdot \left\{ \omega \cdot \left[\sum_{j \in |W|} G(C \rightarrow w_j) \right] + (1 - \omega) \cdot \left[\sum_{j \in |W|} G(w_j \rightarrow P) \right] \right\} \in [0,1]$$

$$G(a \rightarrow b) \equiv f(\text{socialTrust}(a \rightarrow b), R_{\text{Personal}}(a \rightarrow b))$$

$$f(x, y) = 1 - e^{-c(x+y)}$$

$$\text{socialTrust}(a \rightarrow b) = \frac{R_{\text{Structural}}(a \rightarrow b)}{\varphi}$$

$$R_{\text{Social}} = f(\text{socialTrust}(C \rightarrow P), R_W) \in [0, 1]$$

$$RL_{\text{Social}} = \frac{1}{N_W} \cdot \left\{ \omega \cdot \left[\sum_{j \in |W|} RL_{\text{Personal}}(C \rightarrow w_j) \right] + (1 - \omega) \cdot \left[\sum_{j \in |W|} RL_{\text{Personal}}(w_j \rightarrow P) \right] \right\} \in [0, 1]$$

VIII.2.1.3. Reputation based on system's structure

VIII.2.1.3.1. Reputation based on Institutions' providers' evaluations

$$R_{\text{Institution}} = \frac{1}{N} \sum_N R_{\text{ProviderEv}}(j) \in [0, 1]$$

$$RL_{\text{Institution}} = \frac{1}{N} \sum_N RL_{\text{ProviderEv}}(j) \in [0, 1]$$

VIII.2.1.3.2. Structural Reputation

$$R_{\text{Structural}} = R_{\text{Structural}}(a \rightarrow b) = \frac{1}{1 + e^{-(HDS+\Delta)}} \in [0, 1]$$

$$RL_{\text{Structural}} = RL_{\text{Structural}}(a \rightarrow b) = 1$$

$$\Delta = \begin{cases} \frac{l_B - l_A}{\max(l) - 1} & I_A = I_B \\ \frac{L(I_B) - L(I_A)}{\max(L) - 1} & I_A \neq I_B \end{cases}$$

VIII.2.1.4. Final Reputations

VIII.2.1.4.1. Opinions-based provider Reputation

$$R_{ProviderOp} = k_{Personal} \cdot R_{Personal} + k_{Social} \cdot R_{Social} \in [0,1]$$

$$RL_{ProviderOp} = k_{Personal} \cdot RL_{Personal} + k_{Social} \cdot RL_{Social} \in [0,1]$$

$$k_{Personal} + k_{Social} = 1$$

VIII.2.1.4.2. Provider Reputation

$$R_{Provider} = k_{ProviderEv} \cdot R_{ProviderEv} + k_{ProviderOp} \cdot R_{ProviderOp} \in [0,1]$$

$$RL_{Provider} = k_{ProviderEv} \cdot RL_{ProviderEv} + k_{ProviderOp} \cdot RL_{ProviderOp} \in [0,1]$$

$$k_{ProviderEv} + k_{ProviderOp} = 1$$

VIII.2.1.4.3. TruQOs Reputation

VIII.2.1.4.3.1. Resource Recommendation

$$R_{TruQOs} = k_{Resource} \cdot R_{Resource} + k_{Provider} \cdot R_{Provider} \in [0,1]$$

$$RL_{TruQOs} = k_{Resource} \cdot RL_{Resource} + k_{Provider} \cdot RL_{Provider} \in [0,1]$$

$$k_{Resource} + k_{Provider} = 1$$

VIII.2.1.4.3.2. User Recommendation

$$R_{TruQOs} = R_{Provider} \in [0,1]$$

$$RL_{TruQOs} = RL_{Provider} \in [0,1]$$

VIII.2.1.5. Newcomers

VIII.2.1.5.1. Resource Recommendation

$$R_{TruQOs}(ResRec) = K \cdot R_{ProviderEv} \in [0,1]$$

$$RL_{TruQOs}(ResRec) = RL_{ProviderEv}$$

VIII.2.1.5.2. User Recommendation

$$R_{TruQOs}(UserRec) = K \in [0,1]$$

$$RL_{TruQOs}(UserRec) = 1$$

VIII.2.1.5.3. Weighting parameter K

$$K = \begin{cases} 1 & 0 \leq \Delta_t \leq \left\lceil \frac{prom}{2} \right\rceil \\ 1 - \frac{1}{prom - \left\lceil \frac{prom}{2} \right\rceil + 1} \left(\Delta_t - \left\lceil \frac{prom}{2} \right\rceil \right) & \left\lceil \frac{prom}{2} \right\rceil < \Delta_t \leq prom \end{cases}$$

$$\Delta_t = t - t_{Debut}$$

VIII.2.2. TST

VIII.2.2.1. Probability distributions of Resource/Service quality

$$\text{Uniform}(a, b) \equiv F(x; a, b) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1} \quad a, b \in (0, 1)$$

$$\text{Normal}(\mu, \sigma) \equiv F(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \mu, \sigma \in (0, 1)$$

$$\text{Beta}(\alpha, \beta) \equiv F(x; \alpha, \beta) = I_x(\alpha, \beta) = \frac{B(x; \alpha, \beta)}{B(\alpha, \beta)} \quad \alpha, \beta > 0$$

$$B(x; \alpha, \beta) = \int_0^x u^{\alpha-1} \cdot (1-u)^{\beta-1} du$$

$$B(\alpha, \beta) = \frac{(\alpha-1)! \cdot (\beta-1)!}{(\alpha+\beta-1)!}$$

VIII.2.2.2. OnlyOpinions Reputation

$$R_{\text{OnlyOpinions}}(\text{ResRec}) = \frac{1}{N} \sum_N op_j^{\text{ResX}} \in [0, 1]$$

$$R_{\text{OnlyOpinions}}(\text{UserRec}) = \frac{1}{N} \sum_N op_j \in [0, 1]$$

VIII.2.2.3. UserAgent behaviours

VIII.2.2.3.1. Consumer

$$nSteps_{\text{RequestRec}} = \text{round}\left(\frac{1 - \text{talkativeness}}{\text{talkativeness}}\right)$$

$$f(\text{opining}; \text{laziness}) = \begin{cases} 1 - \text{laziness} & \text{giving opinion} \\ \text{laziness} & \text{not opining} \end{cases}$$

VIII.2.2.3.2. Provider

$$nSteps_{CreateRes} = \text{round}\left(\frac{1 - \text{productivity}}{\text{productivity}}\right)$$

$$f(\text{answering}; \text{availability}) = \begin{cases} \text{availability} & \text{answering} \\ 1 - \text{availability} & \text{not answering} \end{cases}$$

VIII.2.3. Evaluation

VIII.2.3.1. Ratios of recommendation

$$Ratio_{ResRec} = \frac{Resources_{Accessed}}{Resources_{Generated}}$$

$$Ratio_{UserRec} = \frac{Interactions_{Answered}}{nSteps}$$

VIII.2.3.2. Ratio of benefits

$$Ratio_{Benefits} = \frac{Resources_{Accessed}}{Interactions_{Answered}}$$

VIII.3. Annex C: Summary of user roles and parameters

Consumer	
	Consumers request recommendation of resources and/or users', consume resources and interact with providers to obtain results of a desired Service.	
	Parameters:	
	Talkativeness	Indicates how often a consumer requests a recommendation of a certain type.
	Resources:Users	Number of resource recommendations before requesting a user recommendation.
	Opinions Behaviour	The behaviour of the consumer when giving opinions about a resource or a provider.
	Target	The provider to whom the opinions' behaviour is applied.
	Laziness	Reflects the probability that a consumer gives an opinion about a resource or a provider.
	Service	The type of resources or direct interactions' results the consumer is searching for.
Provider	
	Providers generate resources into the system and satisfy consumers' requests for the provision of results of a desired Service.	
	Parameters:	
	Service	The type of generated resources or direct interactions' results the provider offers.
	Productivity	Indicates how often a provider generates a resource.
	Availability	Reflects the probability that a provider answers a direct request from a consumer.
Polyvalent	
	These users act as providers and consumers, not at the same time.	
	Parameters:	
	They share the parameters of the role they perform for each simulation round. They use the same Service as provider and consumer for consuming or generating resources.	

VIII.4. Annex D: Scenarios of evaluation

TmDQs FUNCTIONALITY TESTS									
Scenario	Name	Description	Recommendation type	Reputation Mode	Cycles	N of Tests			
						Total	New	Repeated	
S0	S0	Initial state	Only Resources	No reputation	20	1	1	0	
	S1	Evaluations' effect "Ur" decreases temporarily its Q. Analysis of evaluation and resource reputations impact.	Only Resources	No Rep, Only ops, TmDQs	30	6	4	2	
	S1a	No Ev Temporal weight	Only Resources	Param. Variation	5:5:10:5:5	3	3	0	
S0	S1b	Ev Temporal weight	Only Resources	Param. Variation	5:5:10:5:5	3	1	2	
	S2	Options' effect "Ur" increases Q temporally and Users of "Ix, Iy" attack "Ur". Analysis of opinions and resource reputation impact.	Only Resources	No Rep, Only ops, TmDQs	30	18	16	2	
	S2a	No Op Temporal weight	Only Resources	Param. Variation	5:5:10:5:5	9	9	0	
S0	S2b	Op Temporal weight	Only Resources	Param. Variation	5:5:10:5:5	9	7	2	
	S3	Combined effect Repetition of P1 and P2 with all reputations impact + combined P1 and P2 effect.	Only Resources	No Rep, Only ops, TmDQs	30	78	68	10	
S0	S3a	Quality decrease	Only Resources	Param. Variation	5:5:10:5:5	23	20	3	
	S3b	Opinion attack	Only Resources	Param. Variation	5:5:10:5:5	23	16	7	
	S3c	Q-Idp variation	Only Resources	Param. Variation	5:5:10:5:5	32	32	0	
S0	S4	Attackers size Effect Different number of attackers start criticizing "Ur". Attack maintained in time	Only Resources	Only ops, TmDQs	10:20	12	12	0	
	S5	Ranking's effect "Ur" changes its Q. Users of "Ix, Iy" start lying. Ranking renovation and changes in HDs. None.	Only Resources	No Rep, Only ops, TmDQs	30	51	51	0	
S0	S5a	Competition	Only Resources	Param. Variation	10:20	17	17	0	
	S5b	Discredit	Only Resources	Param. Variation	5:5:20	17	17	0	
	S5c	Promotion	Only Resources	Param. Variation	5:5:10:10	17	17	0	
S6	S6	Hierarchy's effect Simpler hierarchy. All Users Q=0.7. At some point "Ug" changes its Q and Users start lying. Changes in HDs. H/M/L.	Only Resources	No Rep, Only ops, TmDQs	40	33	33	0	
	S6a	Rivalry	Only Resources	Param. Variation	20:20	11	11	0	
	S6b	Competitiveness	Only Resources	Param. Variation	20:20	11	11	0	
S6c	S6c	Nepotism	Only Resources	Param. Variation	20:20	11	11	0	
S7	S7	Reliability's effect Simpler hierarchy. All Users Q=0.7. "Ug" changes its Reliability (Q=0.69, Q=0.71, Q=0.69 ...) and then back to 0.7.	Only Resources	No Rep, Only ops, TmDQs	30	25	20	5	
	S7a	q-Factor effect	Only Resources	Param. Variation	10:10:10	7	7	0	
	S7b	Rumor's effect	Only Resources	Param. Variation	10:10:10	11	9	2	
S7c	S7c	Rep-Hel Effect	Only Resources	Param. Variation	10:10:10	7	4	3	
S8	S8	Newcomer's effect "U3T" enters with higher quality (Q=0.9) in an existing institution or in a new one. Click "U3T" and "Ur" belonging to "I4T".	Resources and Users	No Rep, Only ops, TmDQs	30	24	20	4	
	S8a	ReRecommendaion	Only Resources	Param. Variation	10:10	12	10	2	
	S8a1	Existing Institution	Only Resources	Param. Variation	10:10	6	6	0	
S0	S8a2	New Institution	Only Resources	Param. Variation	10:10	6	4	2	
	S8b	UserRecommendation	Only Users	Param. Variation	10:10	12	10	2	
	S8b1	Existing Institution	Only Users	Param. Variation	10:10	6	6	0	
S8b2	S8b2	New Institution	Only Users	Param. Variation	10:10	6	4	2	
S9	S9	Direct requests' effect Consumers Results -1. Changes in resources' and direct interactions' Q. Click (red/crossed/directRequests)	Resources and Users	No Rep, Only ops, TmDQs	32/48	24	24	0	
	S9a	Producer service	Resources and Users	Param. Variation	16:16	12	12	0	
	S9a1	Personal service	Resources and Users	Param. Variation	32:16	12	12	0	
S10	S10	Initial dispersion	Only Resources	Only ops, TmDQs	20	3	3	0	
						275	252	23	