Idea Relationship Analysis in Open Innovation Crowdsourcing Systems

Adam Westerski
Universidad Politécnica de Madrid
Avenida Complutense 30
28040 Madrid, Spain
westerski@dit.upm.es

Carlos A. Iglesias
Universidad Politécnica de Madrid
Avenida Complutense 30
28040 Madrid, Spain
cif@gsi.dit.upm.es

Javier Espinosa Garcia
Universidad Politécnica de Madrid
Avenida Complutense 30
28040 Madrid, Spain
javier.espinosag@alumnos.upm.es

***Abstract*—Idea Management Systems are an implementation of open innovation notion in the Web environment with the use of crowdsourcing techniques. In this area, one of the popular methods for coping with large amounts of data is duplicate detection. With our research, we answer a question if there is room to introduce more relationship types and in what degree would this change affect the amount of idea metadata and its diversity. Furthermore, based on hierarchical dependencies between idea relationships and relationship transitivity we propose a number of methods for dataset summarization. To evaluate our hypotheses we annotate idea datasets with new relationships using the contemporary methods of Idea Management Systems to detect idea similarity. Having datasets with relationship annotations at our disposal, we determine if idea features not related to idea topic (e.g. innovation size) have any relation to how annotators perceive types of idea similarity or dissimilarity.**

***Index Terms*—idea management, relationship, clustering, crowdsourcing, open innovation, community**

I. Introduction

The burst of popularity of Web 2.0 technologies and social networking has led to their use in many domains of business for customer communication as well as for knowledge management inside large enterprises [20]. One of such practices is the implementation of open innovation paradigm [10] with the use of social collaborative Web platforms. Traditionally, open innovation concept involved asking parties not directly involved in product development for ideas in a suggestion box-like fashion (e.g. as practised by Toyota for over 50 years, much before open innovation term or Internet were born [32]). Nowadays, the huge popularity of social networking platforms and increasing literacy of consumers with Web tools allows to extend those practices with crowdsourcing activities [16] that not only gather ideas from consumers on mass scale but also to make them aware of each others innovations and involve them in collaboration on idea improvement and idea evaluation (e.g. via voting and commenting on ideas).

In relation to this new open innovation methodology, we set our research in the area of Idea Management Systems (IMS). Those web applications aggregate user feedback submitted via popular social media portals like Twitter or Facebook (e.g. IdeaScale [4]) as well as allow users to directly post ideas and collaborate through a specialised front-end. Aside of gathering ideas and engaging the community in a collaborative process, the IMSes focus on knowledge management in the back-end to support decision makers in idea assessment and selection of best ideas. However, in the contemporary systems, this promise of smart idea screening has encountered problems related to processing huge data volumes and the characteristics of collected data. According to case studies of various vendors [8], [18] those problems are: redundancy of ideas, triviality of ideas, sudden peaks of submitted ideas related to certain events (e.g. new product release).

In the research presented in the following paper, we refer to the large data volume problem by discussing the concept of idea relationships and summarization of idea datasets based on types of relationships that connect ideas. In the contemporary systems this problem is typically handled by duplicate detection in conjunction with crowdsourcing methodologies that employ users in submitting duplicate reports rather than utilizing fully automatic approaches. With reference to this state, we investigate if there is room to introduce new types of relationships between ideas and if this change would allow to make a meaningful impact on downsizing the idea dataset in instances of Idea Management Systems that contain tens of thousands of ideas.

To achieve our goal, firstly we investigated the research already done on describing knowledge relationships and discussed how it relates to our specific domain of open innovation (see Sec. II). Next, we formulated a number of hypotheses that focused our research on solving very specific problems of introducing new idea relationships in open innovation (see Sec. III). Afterwards, we proposed an extension of idea relationships (see Sec. IV) and finally made experiments that aimed to evaluate if the previously stated hypotheses could be met or not (see Sec. V).

II. Related Work

Semantics of relationships is a highly investigated topic in the area of linguistics and psychology. It's application in the computer science has been reviewed in a number of works for domains such as information retrieval or information extraction [19], [15], [24]. Myaeng et al. [24] reviewed the created classifications of relationships in those areas and split them into pragmatic and linguistic. During our research on idea relationships we used those relationship hierarchies as a

978-1-936968-36-7 © 2012 ICST
DOI 10.4108/icst.collaboratecom.2012.250412

reference. In particular, we analysed a taxonomy of relationships proposed by Bejar et al. [7] and attempted to transform the language relationships into idea relationships. In many cases our conclusion was that relationships applicable for language constructs either did not make sense when applied for innovation or intersected with each other making classification of ideas a difficult task.

Such a debate about relevance of linguistic relationships in other areas has been the topic of interest of knowledge management research [6], [28]. The concept of relationships has been investigated and modelled on the level of entire knowledge objects rather than just language constructs. A number of works in the ontology research (e.g. Cyc [21]) and Semantic Web in particular (e.g. OWL [26]) attempt to define such knowledge relationships on a generic level. Additionally, in many cases researchers have analysed semantics of relationships for specific narrow domains. For instance, the Learning Object Model (LOM) specification [1] defines a simplified model that has been argued and extended in a number of works [27], [13]. We did not encounter similar studies of relationships for Idea Management Systems in particular, however we used the achievements from other domains such as aforementioned e-learning or multimedia [22] to recognize how information objects can be linked for delivering a more complete overview of the entire knowledge repository. Additionally, in comparison to related work on relationship hierarchies in both knowledge management and earlier described linguistics, our research does not attempt to find a most complete or suitable relationship taxonomy for Idea Management but to determine if there is any point of introducing such.

With reference to the domain of Idea Management in general, there has been a number of attempts to cope with the problem of information summerization other than with the use of idea relationships. In contemporary industrial solutions the idea assessment is aided most often by automatically generated community metrics (comment count, view count, vote count etc.) or by manual expert reviews. However, Harstinski et al. [17] as well as Gangi et al. [14] report of poor performance of those methods in terms of relevancy to amount of implemented ideas and impact on idea selections. Such state has triggered researchers to investigate other possibilities for idea assessment such as: prediction markets [9], problem solving approaches [5], calculating new types of metrics [12] or using data from other enterprise systems to automatically assess ideas [25], [29]. The research presented in this paper, similarity as prediction markets attempts to employ crowds to aid idea assessment but also determines relevancy of some of previously developed metrics to aid selection of similar ideas.

III. Hypothesis

In the previous section we have shown that relationships between knowledge have been investigated in many different domains and with different results. In the following paper, we focus only on certain aspects of introducing a relationship hierarchy into Idea Management Systems and therefore to make our goals more clear we have defined the following hypotheses:

H1. The semantics of idea relationships are more complex than duplicate relationship.

H2. Wider range of relationships can be used to summarize datasets better than with just duplicate relationship.

H3. Apart of idea topic there are idea characteristics (e.g. innovation type) and idea text features that have an impact on how idea annotators perceive the type of relationship between ideas.

With **H1** we put forward a hypothesis that duplicate relationship is insufficient to describe all relationships between ideas stored in the Idea Management Systems. To verify this hypothesis we propose to annotate a subset of ideas using a broad set of relationships identified during our research and compare the results to annotations done only when a duplicate relationship was available.

With **H2** we suggest that the newly proposed relationships, once applied as annotations to ideas, can help in information summerization and would allow to aggregate more ideas than it was possible before when just using the duplicate relationship. To evaluate that hypothesis we propose to reuse the annotations from the previous experiment and aggregate ideas based on their similarities, inheritance of relationship types and transitivity of relationships.

Finally, with **H3** we suggest that annotators pick relationships for ideas not only as a function of similarity on the topic level (e.g. discussing the same product) but also based on the scale of innovation that an idea proposes, how detailed the description is etc. To verity this hypothesis, we refer to previous research on innovation taxonomies and compare the idea relationship annotations from previous hypotheses experiments with idea similarity expressed with metrics derived from annotations with taxonomy terms for describing idea characteristics.

IV. Idea Relationship Hierarchy

To facilitate the aforementioned experiments with the stated hypotheses we have created a hierarchy of relationships between ideas in an Idea Management System. The preliminary version of the hierarchy has been created based on our past experiences in the Gi2MO project with various idea datasets [31], [30]. Later, we refined the hierarchy by running a number of test annotation experiments with various datasets and by referring to the earlier described related work. The final version of hierarchy proposal used during our experiments is presented in Table I.

The relationships that can be established between ideas have been separated into two categories: those that can be identified by analysis of the text of two ideas (A - Based on knowledge) and relationships that are created based on user interactions with the system (B - Based on Action). This state is a result of experiments with applying the presented

Table I
PROPOSED IDEA RELATIONSHIP HIERARCHY FOR IDEA MANAGEMENT SYSTEMS.

A Based on knowledge	Relationship existing between knowledge content of ideas created independently
A1 Similar	Ideas similar to each other
A1.1 Describes Same Object	Ideas that propose a similar innovation for the same item
A1.1.1 Extends	One idea extends other
A1.1.1.1.1 Complementary	Ideas that can work together
A1.1.1.1.1.1 Details	One idea focuses on part that other neglects
A1.1.1.1.1.2 Generalizes	One idea describes a more broad vision of other
A1.1.1.1.2 Excluding	Implementations of ideas exclude each other
A1.1.1.1.2.1 Alternative Solution	Ideas refer to the same object and problem but solved in different ways
A1.1.1.1.2.1 Alternative Idea	Two completely distinct ideas that in effect are impossible to implement together
A1.1.2 Duplicates	Ideas describe exactly the same innovation
A1.2 Describes Related Object	Ideas that propose innovation for different objects that are somehow related to each other
A2 Disjoint	Ideas not having any meaningful similarities
B Based on Action	The relationship is created by an action operating on both ideas by a user of the system
B1 Based on Moderator Action	Action taken by moderator of the system in reaction to submitted ideas and relationship annotations
B1.1 Follows	Implementation of an idea should follow some other idea
B1.2 Proceeds	Implementation of an idea should proceed some other idea
B1.3 Merged	Two ideas merged into a single one
B2 Based on Innovator Action	Relationships created based on user interaction with ideas
B2.1 Originates	Ideas created by extending some other idea
B2.2 Is version	Created by updating an idea (e.g. in reaction to community feedback)
B2.3 References	One idea referencing other idea (or resource from outside the system)

relationships to various idea datasets. Although relationship models referenced before (e.g. LOM [1]) do not apply such distinction, we identified that annotators were unable to put any of the relationships from group (B) just based on the idea text and without the contextual knowledge of the entire idea repository, including history of the examined ideas.

During our experiments for evaluation of the hypotheses we focused on annotation of already created ideas collected from existing online public systems. The person providing annotations did not have an option to create new ideas, therefore we utilized only the relationships from branch (A). Furthermore, during the experiments we interpreted the presented relationship hierarchy with the following rules:

- *similar*, *disjoint*, *describesRelatedObject*, and all *excluding* relationships are symmetric
- *details* relationship is inverse of *generealizes* relationship
- *extends* and *duplicate* relationships are not symmetric and during annotation we provided additional *is extended* and *is duplicated* relationships being inverse to the aforementioned

Table II
UBUNTU BRAINSTORM DATASET STATISTICS

Metric	Metric Value
Idea number	21690
Comments number	133090
Users number	10062
Implemented Ideas number	576
Amount of Votes cast	2608917

V. EVALUATION

During the evaluation stage we conducted three experiments, one per each hypothesis. The content used for all experiments was taken from Ubuntu Brainstorm Idea Management System (see Table II). The distinctive characteristics of this system are:

- the topic of all ideas is improvement or introduction of innovations into an open-source linux operating system distribution, related products and services.
- users submit not only ideas but also solutions. The original creator of an idea posts the first solution and afterwards any member of the community is allowed to add his own solution for implementing the same innovation.

During the experiment we collected all data of the Ubuntu BrainStrom instance and imported into our own system [2]. Next, a single annotator was asked to provide relationships for 200 ideas that included: 120 random selected ideas, 40 ideas that have been implemented, 10 top rated ideas, 10 lowest rated ideas, 10 top commented ideas, 10 least commented ideas. The annotator was only presented the idea text (without the complementary solutions). Per each idea the annotator was presented with 5 similar ideas for which he had to specify the relationships (see Fig. 1). The similar ideas were selected by the system based on Lucene keyword similarity algorithm [23] run over the index of all 21690 Ubuntu ideas. As a result, we obtained annotations for 1000 idea relationships. This data was used in each of the following hypotheses evaluations.

A. Relationship amount comparison

In order to determine if any other relationship apart of duplicate is valid we checked if among the obtained relationships were any other than duplicates as well as compared the amount of particular relationships. Apart of data obtained during our own annotation experiment, we also compared our results with the duplicate annotations already present in the online version of Ubuntu Brainstorm (limiting the maximal amount of duplicates to 5 per idea just like we did in our own experiment). As

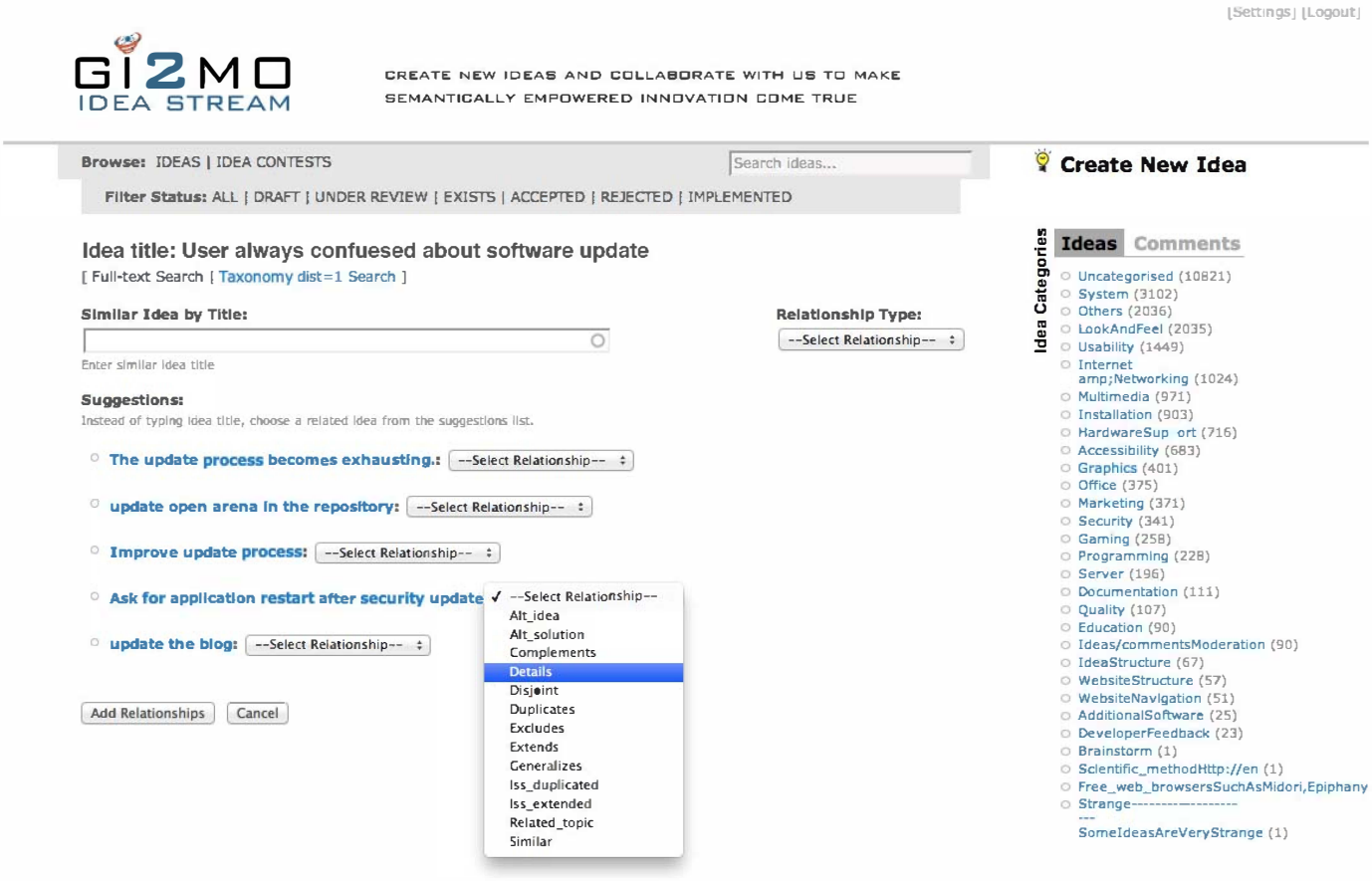

Figure 1. Annotator using a tool for similar idea detection.

a result, within our own annotations the duplicate relationship accounted for only 25% of all annotations. In comparison to the Ubuntu community annotations we got an increase of 76.7% in relationship count in favour of our solution with regard to what was available before. The detailed results can be observed in Table III.

Table III
COMPARISON OF RELATIONSHIP COUNT ACROSS DIFFERENT EXPERIMENTS (200 IDEAS ANNOTATED, 5 RELATIONSHIPS MAX. PER EACH, NO INHERITANCE OR TRANSITIVE RELATIONSHIPS REASONING)

Ubuntu Brainstorm			
Duplicate	249		
Gi2MO Relationships			
total 440 (328 without duplicates)			
similar	4	disjoint	558
related object	111	extends	2
is extended	3	complements	0
details	136	generalizes	27
excludes	0	alternative solution	19
alternative idea	26	duplicates	112
is duplicated	0		

As an outcome of this experiment we can confirm that introducing new relationships resulted in more metadata and annotators taking advantage of the new power they were given. Therefore, based on the presented results, hypothesis **H1** is supported.

B. Idea aggregation for information summarization

In the previous experiment we have shown that by introducing a more broad set of relationships we were able to obtain a much bigger amount of annotations. Nevertheless, this does not imply that the amount of unique similar ideas would rise in the same degree (different relationships can point to the same ideas).

Therefore, to answer a question if annotations made with the new set of relationships would allow to summarize the data more than just the previously present duplicate relationship, we processed the annotated Ubuntu dataset by aggregating all similar ideas into a single one (just as it is done in the contemporary systems when duplicate ideas are detected). In contrast to the previous experiment the main difference is that we count the amount of unique ideas that can be aggregated rather than total number of relationships obtained. In particular, we analysed the amount of unique ideas that could be aggregated in relation to entire dataset size. The results were: 1.13 % of dataset were duplicates that could be aggregated based on Ubuntu community annotations, 0.5 % of dataset aggregated based on duplicate relationships from our experiment and finally 1.95 % of dataset aggregated while

Table IV
IDEA AGGREGATION RATIO IN DIFFERENT INFERENCING SCENARIOS

Relationship	No Inheritance		Inheritance	
	No Transitivity	Transitivity	No Transitivity	Transitivity
Similar	0.02	0.02	2.85	**3.37**
Related Object	0.72	0.75	0.72	**0.75**
Extends	0.01	0.01	1.39	**1.52**
Complements	0	0	1.06	**1.18**
Details	0.87	0.90	0.87	**0.90**
Generalizes	0.18	0.21	0.20	**0.21**
Excludes	0	0	0.29	**0.30**
Alternative Solution	0.12	0.12	0.12	**0.12**
Alternative Idea	0.17	0.18	0.17	**0.18**
Duplicates	0.71	0.71	0.71	**0.73**

using the full relationship hierarchy and aggregating all similar ideas. This indicates that the summerization of our solution with respect to Ubuntu community annotations gave a 95% increase.

Additionally, to see if the summierization ratio of our solution could be futher improved, we analysed two ways of extending the knowledge base using inference of:

- inheritances between relationship types, e.g. when aggregating *complementary* ideas also *details* and *generalizes* annotations are taken into account
- transitive relationships, e.g. if A *extends* B and C *extends* B, than both B and C are aggregated into A

To compare all three options (user made annotations, inherited relationships, transitivity of relationships), we defined **idea aggregation rating** metric that states how many unique ideas have been aggregated per a single idea in the system (see Table IV). Observing the results, it can be seen that using transitivity gives a significant increase of ideas aggregated which can be particularly seen for top relationships in the hierarchy when relationship inheritance is applied (e.g. amount for similar ideas aggregated change from 0.02 per idea to 3.37 after applying inheritance and inferring related ideas via transitive relationships).

Taking into account quite a considerable difference in dataset summerization that is mainly the result introducing new relationships but also the application of logic operators for relationships, we can state hypothesis **H2** as supported in the conditions of our experiment.

C. Idea metric similarity and idea relationships

Looking at the results of previous experiments, we can see that the amount of relationships has risen in comparison to legacy solutions but it can be also observed that over half of ideas initially identified as discussing a similar topic (558 out of 1000) were not related by any relationship. In our last experiment we investigated if this state could be attributed to ideas being different on a level of characteristics other than topic and not related to domain information (e.g. if ideas proposing innovations for exactly the same product but with different originality are less likely to be similar, or if ideas of that discuss the same innovation but with different amount of details are also perceived as different by annotators). We attempted to identify those non-domain characteristics and see if any of them would help to determine similarity or at least particular similarity type for ideas that were already annotated as similar.

To define idea characteristics we referred to our previous research on idea classification and reused the Gi2MO Types taxonomy [3] that advocates the use of 4 main idea characteristic areas:

- trigger type (type of event or action that caused the creation of idea, e.g. faulty experience or lack of feature)
- innovation type (how much innovative is the idea, e.g. radical innovation vs. incremental innovation)
- proposal type (how broad is the description, e.g. full solution vs. bug report)
- object type (is innovation proposed for object or service, is it a complete change or element change etc.)

All together Gi2MO Types delivers 74 terms grouped into the above 4 categories. Based on presence or absence of particular terms in idea annotations Gi2MO Types defines 14 metrics that characterise an idea. Those metrics are an interpretation of particular taxonomy branches and aim to summarize the annotations made with taxonomy and facilitate idea comparison for further analysis:

- Idea Completeness - how many characteristics could be defined for an idea
- Trigger Experience Completeness - how complete was the experience with an object that triggered the idea (e.g. faulty experience or just a lack of feature)
- Trigger Situational Dependence - how much was the trigger dependent on occurrence of some particular event
- Trigger Relatedness - how closely related are the object of innovation and the object that triggered the idea
- Idea Adaptiveness - is the idea meant for new markets or for existing ones
- Idea Originality - how original is the idea (i.e. new, incremental, no innovation at all etc.)
- Idea Originality Scope - how far does the originality of the idea reach. Is it only the particular element of the organisation, entire organisation or the entire market ?
- Community Cooperativeness - how well did user formulate and communicate his idea and in what detail did he describe the implementation of the idea
- Implementation Freshness - how much the implementation of the idea is related to what has been already created

Table V
BIVARIATE CORRELATIONS BETWEEN THE GI2MO TYPES METRIC VALUE SIMILARITIES AND TOP LEVEL RELATIONSHIPS FROM THE PROPOSED HIERARCHY

Metric/Relationship	Similar	Disjoint	Duplicate	Related Object	Extends	Complements	Excludes
Completeness	**0.18**	**-0.19**	0.08	0.03	**0.11**	0.08	0.07
Experience Completeness	-0.04	0.06	0.07	0.01	-0.09	-0.07	-0.04
Situational Dependence	0.01	-0.03	0.02	0.05	-0.04	-0.05	0.01
Relatedness	0.01	-0.02	0.11	-0.04	-0.07	-0.02	-0.08
Adaptiveness	**0.18**	**-0.20**	0.01	0.01	**0.18**	**0.12**	**0.11**
Originality	**0.15**	**-0.17**	0.06	0.05	0.09	0.05	0.07
Originality Scope	0.06	-0.01	0.08	**-0.14**	**0.13**	0.06	-0.05
Cooperativeness	**0.14**	**-0.14**	**0.12**	0.03	0.04	0.06	-0.02
Freshness	0.09	-0.09	0.01	**-0.11**	**0.15**	**0.10**	0.09
Integrability	**0.24**	**-0.25**	0.08	0.02	**0.19**	**0.16**	0.07
Applicability Scope	**0.14**	**-0.14**	-0.01	0.01	**0.10**	-0.06	0.08
Constructiveness	-0.02	-0.01	0.01	-0.05	0.00	**0.14**	0.08
Scope	**0.11**	**0.13**	-0.01	0.09	0.07	0.09	-0.02
Object Dependability	**0.24**	**-0.21**	-0.01	0.01	**0.25**	**0.21**	**0.11**

in a former version of the same product
- Implementation Integrity - how tangible is the object of innovation (a product, service or a process change)
- Implementation Applicability Scope - how broad is the application of the idea (e.g. is it a type of products or just a specific product)
- Implementation Constructiveness - to what degree an idea is creating a new product or just improving an old one
- Implementation Scope - the scale of modification that idea proposes (e.g. complete change, element change or characteristic change)
- Implementation Dependability - how much does the introduction of innovation impact other products

For our evaluation, we annotated ideas with Gi2MO Types taxonomy terms and calculated the aforementioned metrics. The metric similarity S between two related ideas i_A and i_B was calculated individually per each metric M_x as an absolute difference of a given metric value for two related ideas:

$$S(i_A, i_B, M_x) = 1 - |M_x(i_A) - M_x(i_B)|$$

The calculations, as described above, were made for a subset of 50 ideas taken from the previous evaluations, in particular: 10 implemented, 10 top rated, 10 most down-ranked, 10 top commented, 10 least commented (but having at least 1 comment). Using the taxonomy we annotated those 50 ideas as well as all their related ideas with Gi2MO Types terms that identified their characteristics. As a result we got 250 annotated ideas (each of the 50 ideas had 5 related ideas).

Having such dataset we analysed the correlations between presence or absence of particular idea relationships and the idea characteristic similarity expressed with Gi2MO Types metrics. The results are presented in Table V.

To analyse the results we used Cohen correlation scale for social sciences [11]. According to that scale, 48 correlations out of 180 can be described as small (between 0.1 and 0.3), while the rest as insignificant (smaller than 0.1). Most of those correlations that can be labelled as small for a single relationship have been observed in case of disjoint and similar relationships (8 metrics out of 14 possible).

Nevertheless, taking into account the presented results the final hypothesis **H3** can not be proven as supported in the conditions of the conducted experiment.

Since our hypothesis assumed a contrary situation, we investigated further the characteristics of ideas in order to seek a probable cause of correlation values such as discovered in the experiment (e.g. why originality scale did not play any role in choosing idea similarity). In particular, we focused on comparing the characteristics of similar and disjoint idea subsets that had most correlations with the analysed metrics: we examined the average values of metrics and the deviation from this value in the respectable datasets (see Fig. 2). Interestingly, as a result, we noticed that diversity of data from the point of view of used metrics was very small overall. Both the similar and disjoint ideas had very big metric similarity and in many cases those values where very close to each other for similar and dissimilar datasets.

VI. CONCLUSIONS

The research presented in the following paper aimed to verify the usefulness of applying a relationship hierarchy for open innovation data stored in Idea Management Systems. The presented experiments have proven that there is indeed room to introduce a more sophisticated set of relationships in comparison to what exists in the contemporary state of the art. We have proposed a new hierarchy of relationships and have shown that using it can more significantly increase the amount of relationships obtained when putting in the same annotation effort. Furthermore, we observed that the most frequent relationship that even exceeds the currently used *duplicate* relationship is an *extension* of idea that *details* the proposed earlier innovation.

Additionally, based on the proposed relationship hierarchy, we have presented a number of methods for dataset summerization and shown that introduction of new idea relationships as well as tools such as relationship inheritance and relationship transitivity can lead to aggregating twice as much similar ideas in comparison to the contemporary duplicate detection solution.

Finally, in search for methodologies that could aid detection of similar ideas, we compared the idea relationship types

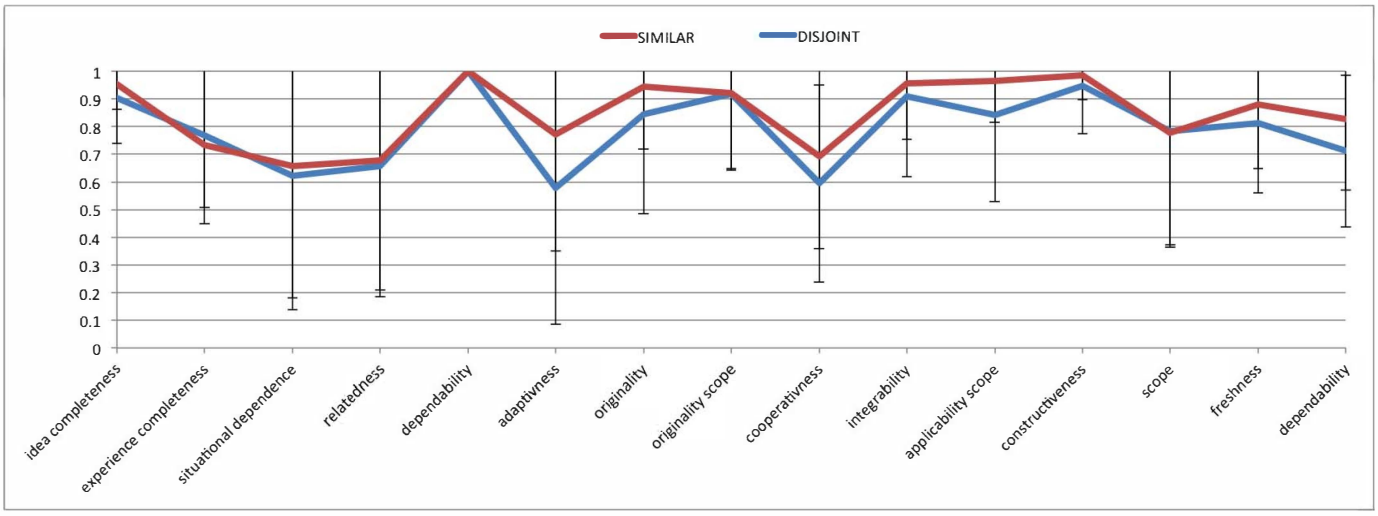

Figure 2. Average and standard deviation of idea characteristic metrics similarity.

with characteristics of related ideas such as: innovation type, innovation trigger, object type, and proposal type. As an outcome, we did not find any significant correlations that would indicate that annotators use those features of ideas for recognizing relationship type. This can lead to a conclusion that while introducing new idea relationships is useful for Idea Management Systems, the type of those relationships in case of similar ideas cannot be determined by detection of features other than idea topic.

In terms of future work we would envision confirming the obtained results by applying the same experiments but in an environment of different Idea Management Systems (e.g. Dell IdeaStorm or myStarBucks Ideas). All experiments run for the needs of research presented in this paper have been done using Ubuntu Brainstorm instance which fits the requirements in terms of data volume and problems discussed in the introduction, however it is also very specific dataset due to the open-source community characteristics and very narrow scope of products that the collected ideas involve. Additionally, it would be desirable to evaluate the proposed relationship hierarchy with a bigger amount of annotators to verity if they reach an agreement in terms of their annotations.

ACKNOWLEDGMENTS

This research has been partly funded by the Spanish CENIT project THOFU. We express our gratitude to Atos Origin R&D for their support and assistance as well as providing us access to their Idea Management Platform.

REFERENCES

[1] Draft standard for learning object metadata. technical report 1484.12.1, ieee inc. Technical report, LTCS WG12: IEEE Learning Technology Standards Committee., 2002.

[2] Gi2mo ideastream idea management system project page. http://www.gi2mo.org/apps/ideastream, 2012.

[3] Gi2mo types specification. http://www.gi2mo.org/taxonomy/, 2012.

[4] Ideascale idea management system homepage. http://www.ideascale.com, 2012.

[5] E. D. Adamides and N. Karacapilidis. Information technology support for the knowledge and social processes of innovation management. *Technovation*, 26(1):50–59, 2006.

[6] C. A. Bean and R. Green. *Relationships in the Organization of Knowledge. Information Science and Knowledge Management*. Kluwer Academic Publishers, 2001.

[7] I. I. Bejar, R. Chaffin, and S. Embretson. *Cognitive and Psychometric Analysis of Analogical Problem Solving*. Springer, 1990.

[8] R. Belecheanu. D6.3.1 sap living lab environment. Technical report, Laboranova Collaboration Environment for Strategic Innovation, 2009.

[9] E. Bothos, D. Apostolou, and G. Mentzas. Idem: A prediction market for idea management. designing e-business systems. markets, services, and networks. In *7th Workshop on E-Business, WeB 2008*, Paris, France, 2008.

[10] H. W. Chesbrough. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business Review Press, 2003.

[11] J. Cohen. *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum, 1988.

[12] S. J. Conn, M. T. Torkkeli, and I. Bitran. Assessing the management of innovation with software tools: an application of innovationenterprizer. *International Journal of Technology Management*, 45(3/4):323–336, 2009.

[13] S. Fischer. Course and exercise sequencing using metadata in adaptive hypermedia learning systems. *ACM Journal of Educational Resources in Computing*, 1(1), 2001.

[14] P. M. D. Gangi and M. Wasko. Steal my idea! organizational adoption of user innovations from a user innovation community: A case study of dell ideastorm. *Decision Support Systems*, 48:303– 312, 2009.

[15] R. Green, C. A. Bean, and S. H. Myaeng. *The Semantics of Relationships: An Interdisciplinary Perspective*. Kluwer Academic Publishers, 2002.

[16] J. Howe. The rise of crowdsourcing. *Wired*, 14(6), 2004.

[17] S. Hrastinski, N. Z. Kviselius, and M. Edenius. A review of technologies for open innovaton: Characteristics and future trends. In *Proceedings of the 43rd Hawaii International Conference on System Sceiences*, 2010.

[18] G. Jouret. Inside ciscos search for the next big idea. *Harvard Business Review*, September 2009.

[19] C. S. G. Khoo and J.-C. Na. Semantic relations in information science. *Annual Review of Information Science and Technology*, 40(1), 2006.

[20] R. Koplowitz. Enterprise social networking 2010market overview. Technical report, Forrester, 2010.

[21] D. Lenat. Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38:33–38, 1995.

[22] E. E. Marsh and M. D. White. A taxonomy of relationships between images and text. *Journal of Documentation*, 56(6), 2003.

[23] M. McCandless, E. Hatcher, and O. Gospodnetic. *Lucene in Action,*. Manning Publications, 2010.

[24] S. H. Myaeng and M. L. McHal. Toward a relation hierarchy for information retrieval. In *2nd ASIS SIG/CR Classification Research Workshop*, 1991.

[25] K. Ning, D. O'Sullivan, Q. Zhu, and S. Decker. Semantic innovation management across the extended enterprise. *International Journal on Industrial and Systems Engineering*, 1, 2006.

[26] P. F. Patel-Schneider, P. Hayes, and I. Horrocks. Ontology language semantics and abstract syntax, w3c recommendation. Technical report, OWL Web, 2004.

[27] M. E. Rodriguez, J. Conesa, and M. A. Sicilia. Clarifying the semantics of relationships between learning objects. In *Third International Conference on Metadata and Semantics Research (MTSR'09)*, 2009.

[28] L. M. Stephens, L. M. Stephens, Y. F. Chen, and Y. F. Chen. Principles for organizing semantic relations in large knowledge bases. *IEEE Transactions on Knowledge and Data Engineering*, 8:492–496, 1995.

[29] A. Westerski and C. A. Iglesias. Exploiting structured linked data in enterprise knowledge management systems: An idea management case study. In *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2011 15th IEEE International*, 2011.

[30] A. Westerski and C. A. Iglesias. Mining sentiments in idea management systems as a tool for rating ideas. In *Large-Scale Idea Management and Deliberation workshop. 10th International Conference on the Design of Cooperative Systems (COOP2012)*, Marseille, France, 2012.

[31] A. Westerski, C. A. Iglesias, and F. T. Rico. A model for integration and interlinking of idea management. In *Metadata and Semantic Research: 4th International Conference, MTSR 2010*, Alcalá de Henares, Spain, 2010.

[32] Y. Yasuda. *40 Years, 20 Million Ideas: The Toyota Suggestion System.* Productivity Pr, 1990.