Nudging to Expand User's Domain Knowledge while Exploring Linked Data

Marwan Al-Tawil, Dhavalkumar Thakker, Vania Dimitrova,
School of Computing, University of Leeds, United Kingdom.

Abstract. This paper investigates how a user could be aided to explore linked data in a way leading to expanding her domain knowledge. Earlier work has confirmed that users can gain knowledge while exploring information spaces generated from semantic databases. In such exploration, semantic links can play a key role. However, the learning effect of exploration through linked data has not been investigated and is usually unsupported. The prime goal of this paper is to design a way to nudge the user to paths which can have higher knowledge utility, and at the same time avoid known usability drawbacks (e.g. semantic links can provide an overwhelming amount of options leading to confusion and frustration). Three 'nudging' strategies have been proposed. A user study which examines how these strategies can affect the knowledge utility of an exploration path and suggests ways to combine them is presented. The work contributes to research in intelligent means to guide the user navigation through linked data to increase the effectiveness of exploration.

Keywords: Exploratory search, linked data, knowledge utility, nudging.

1 Introduction

There are growing arguments that Linked Data technologies can be utilised to enable user-oriented exploratory search systems for the future Internet [19]. In contrast to regular search, exploratory search is open-ended, multi-faceted, and iterative in nature, and is used in a broad range of applications [16] [20]. There are a wide range of tools available for offering exploratory search using semantic web technologies[1] (state-of-the-art in [21] and [22]). One class of such tools is semantic data browsers which operate on semantically tagged content and layout browsing trajectories using relationships in the underpinning ontologies. Earlier research has shown that semantic links can promote expansion of domain knowledge through serendipitous learning effect, which can enable adopting linked data exploration in learning applications [5, 10]. However, not all exploration paths are beneficial for knowledge expansion, and there are known usability drawbacks (e.g. while semantic links provides a structure for exploration, they can also give an overwhelming amount of options leading to confusion, frustration, and sense of being lost). Ways for influencing the user's navi-

[1] Workshops IESD'12@EKAW'12 and IESD'13@HT'13 (http://imash.leeds.ac.uk/event/2013/iesd.html)

gation behaviour (i.e. nudging) are required to aid the user's knowledge expansion. This calls for new intelligent support mechanisms which exploits the semantic graph.

Recent research is examining different ways to provide intelligent support in semantic data browsers. Personalised exploration based on user interests, where the exploration space is personalised by taking into account user interests, has been presented in [23]. Personalisation is offered based on closeness to the current entity, and thus favours browsing through familiar information spaces. Extracting semantic patterns from linked data sources to improve diversity in recommendation results to users has been proposed in [24]. Diversity is measured based on the semantic distance of topics and genres of the results. The concept of utility of statement has been presented in [25] to rank RDF statements with the expectation that some statements will be more valuable or interesting to users than other statements within some context. *Our work adds to this research stream by opening a new avenue which looks at the knowledge utility of the exploration path.* This can facilitate the adoption of linked data exploration in the learning domain, but can also be useful in other exploration applications where the user familiarity with the domain affects interaction.

Our ultimate goal is to design an intelligent mechanism to nudge the user in the information space to facilitate user's domain knowledge expansion. This will contribute to two dimensions that are underutilised by existing work. Firstly, each individual user can have different requirements in terms of knowledge expansion, and hence the nudging mechanism should be personalised to the individual user needs. Work carried out in [23] personalises exploration using user interests. One of the other key areas to personalise is to look into user's familiarity of the domain, where familiarity is related to understanding, and is often based on previous interactions, experiences, or learning. Secondly, an important dimension is to take into account richness of the graph as some nodes have more knowledge value than other. This criteria is similar to [25], however instead of considering inverse frequency of nodes, density can be utilised for ranking nodes. Here, density is associated with the level of knowledge details in the representation of a concept [17]. We hypothesise that both dimensions – user familiarity with the domain and semantic graph density – can underline nudging strategies.

This paper presents the first step in our research towards developing an intelligent mechanism for nudging the user through the information space to facilitate domain knowledge expansion. We propose here three strategies based on the semantic graph (referring the user to *dense* nodes) and user domain knowledge (referring the user to nodes that are either *familiar* or *unfamiliar* to her). A user study with a semantic data browser is conducted to investigate key benefits and limitations of the proposed strategies and to identify ways to combine and further improve them. Section 2 will outline the nudging strategies to promote knowledge expansion. A use case where these strategies are applied (semantic data browser MusicPinta) is presented in Section 3. A user study with MusicPinta is presented in Section 4, and the results are reported in Section 5. The paper concludes discussing the study findings, referring to future work.

2 Nudging Strategies for User Knowledge Expansion

The focus of this work is to utilise semantic data browsers for learning, i.e. expanding a user's knowledge via exploration. We refer to uni-focal semantic data browsers where exploration is often restricted to a single start point and uses 'a resource at a time' to navigate in a dataset [18]. At any time, the user focuses on one node (*focus entity*) from where she sees links to other directly connected nodes (*candidate entities*). At every juncture, when exploring a focus entity, the user has to make a decision about which candidate entities to select for further exploration. The aim of our research is to nudge the user to candidate entities which can lead to most valuable paths.

A nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options, and tries to influence choices in a way that will make choosers better off" [12]. In the context of semantic data browsers, *'nudging'* can be interpreted as a mechanism to support users during their exploration by suggesting candidate entities to move to, which ultimately reduce the user's exploration space. The expectation is that the reduced space will lead to exploration paths that are valuable to the user. In our research, value is measured by '*knowledge utility*', i.e. to what extend the user increases their domain knowledge while going through the path. We utilise two dimensions to identify candidate entities which can lead to paths with high knowledge utility: (i) *density* of nodes in the exploration space presented as a linked data graph, and (ii) user's *familiarity* with the domain. We utilise these dimensions and propose three nudging strategies.

Density exploration strategy (D-strategy) selects the candidate entity with the highest density. Density in the context of linked data graphs is associated with the level of knowledge details in the representation of a concept, and can measure the importance of a node [17]. We utilise Social Network Analysis (SNA) centrality metrics over linked data graph to measure the density of a node. SNA enables analysis based on node centrality which measures a node's importance in the network, i.e. a set of nodes that are linked with one another [9]. We use the most common centrality algorithm- 'Degree Centrality' [7] which considers importance based on connections.

Familiarity exploration strategy (F-strategy) selects a candidate entity which is familiar to the user. Familiarity is generally considered to be related to understanding, and it is often based on previous interactions, experiences, or learning [15]. This strategy will keep the user in exploration spaces with many familiar items.

Unfamiliarity exploration strategy (U-strategy) selects a candidate entity which is unfamiliar to the user. This is also based on the familiarity dimensions but assumes that going to unfamiliar nodes would have an impact on the knowledge utility of the resultant exploration path, as the user will be directed to explore new aspects.

To examine deploy these strategies and examine their usefulness for expanding the user's domain knowledge, we need to put them in the context of a semantic data browser. The selected use case, the implementation of the three strategies and an example of user interaction are presented in the next section.

3 Use Case – Semantic Data Browser MusicPinta

As a use case for examining the effect of nudging strategies on the user's exploration, we have selected the semantic data browser MusicPinta which was developed in our earlier research [10]. MusicPinta enables users to easily tap into facts and content in the music domain. The data sets used for MusicPinta comprise the following resources. *DBpedia*[2]**:** for musical instruments and artists. This dataset is extracted from dbpedia.org/sparql using CONSTRUCT queries. These queries along with a programming wrapper and additonal coding are made available as open source at the sourceforge[3]. *DBTune*[4]**:** for music-related structured data made available by the DBTune.org in linked data fashion. Among the datasets on DBTune.org we utilise: (i) *Jamendo* **-** a large repository of Creative Commons licensed music; (ii) *Megatune* **-** an independent music label; and (iii) *MusicBrainz* - a community-maintained open source encyclopedia of music information. All datasets, were available as RDF datasets and the *Music ontology*[5] was used as schema to interlink them.

The datasets provide an adequate setup (fairly large and diverse data set, yet of manageable size for experimentation) for examining strategies during exploration. It has 2.4M entities and 38M triple statements, taking 1.5GB physical space and includes 876 musical instruments ontology entities, 71k performances (albums, records, tracks) and 188k artists. The datasets coming from DBTune.org (such as MusicBrainz, Jamendo and Megatunes) already contain the "sameAs" links between them for linking same entities. We utilise the "sameAs" links provided by DBpedia to link MusicBrainz and DBpedia datasets. This way, the DBpedia is linked to the rest of the datasets from DBtune.org, enabling exploration via rich interconnected datasets.

Figures 1 and 2 show examples of the user interface in MusicPinta.

Fig 1. Description page of the focus entity 'Bouzouki'

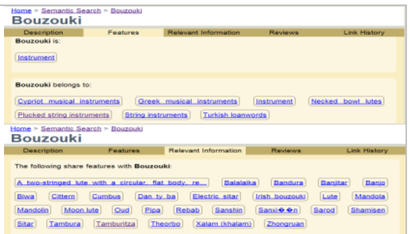

Fig 2. Semantic Links related to the focus entity 'Bouzouki' presented in Features and Relevant Information.

Implementation of the D-Strategy. To implement the D-strategy, the degree centrality measure is applied over the semantic datasets in MusicPinta. The implementation includes two steps. *The first step* extracts the sub-graph of all entities linked to a focus entity using Sesame[6]. Starting from a focus entity, we first extract the sub-graph including all entities that can be reached directly from a focus entity, and then re-

[2] http://dbpedia.org/About
[3] http://sourceforge.net/p/pinta/code/38/tree/
[4] http://dbtune.org/
[5] http://musicontology.com/
[6] http://www.openrdf.org/

peated this iteratively using all extracted entities as focus entities. In the current implementation, the iterative process is repeated five times starting from the focus entity, producing a sub-graph with radius of six nodes. This allows collecting mostly musical instruments, artists and reviews. The iterative repetition of five times is based on Miller's Law [13], which indicates the number of objects that an average human can hold in working memory is 7 ± 2. The output of this step provides two tables for the nodes and edges, respectively. *The second step* is uploading the sub-graph for the focus entity into Gephi[7] using the tables for nodes and edges. The statistical outputs for the degree centrality algorithm provided by Gephi are filtered to include the highest density *candidate entities* to be explored starting from the *focus entity*, ranked according to their degree centrality.

Implementation of the F-strategy and U-strategy. These strategies require identifying a user’s familiarity with the candidate entities which can be done implicitly (e.g. from the user’s interaction paths) or explicitly (by asking the user to specify their familiarity). At this stage of our research, we are examining whether familiarity could be useful and how to combine it with diversity. Hence, we have selected the easier option for familiarity–explicitly asking the user to select candidate entities that are familiar or unfamiliar, respectively.

User interaction with MusicPinta with nudging strategies. We adopt a ‘Wizard of Oz’ style of experimental design using MusicPinta, i.e. the strategies are not implemented directly in the system but are simulated with the help of a human. The user is ‘guided’ to select candidate nodes based on the calculated density (implemented outside MusicPinta) or stated familiarity (as declared by the user). Each strategy is followed independently, i.e. an exploration path follows either D-strategy, F-strategy, or U-strategy. This allows us to isolate the strategies in order to study their advantages and drawbacks. To ensure ‘equal’ start for each strategy, the user is directed to start from the most dense node. Then, the user follows one of the selected strategies. Figure 3 shows examples of the three strategies, as used in the study presented next.

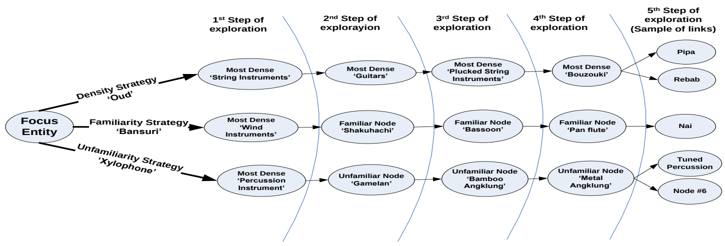

Fig. 3. Example of paths for the three exploration strategies.

4 User Study

The MusicPinta use case provide experimental set up for a user study to address the following **research questions**: *What are the likely benefits and drawbacks of each of the suggested nudging strategies, how they can be combined and improved?*

[7] https://gephi.org/

Measuring knowledge utility of an exploration path. To compare the *knowledge utility* of the exploration paths resulting from each strategy, we need a mechanism for measuring changes in the user's knowledge. For this, we follow the well-known classification for measuring knowledge by Bloom [2]. It suggests linking knowledge to six cognitive categories; the first two - remember and understand - are directly related to browsing and exploration activities[8]. Remember is about retrieving relevant knowledge from the long-term memory, including recognise (locating the knowledge) and recall (retrieving it). Understand is about constructing meaning, from which the most relevant to a semantic browser is categorise (determining that an entity belongs to a particular category) and compare (detecting similarities between entities). Based on this, *the knowledge utility of an exploration path is measured as the effect of exploration on the user's cognitive processes of remember, categorise, and compare.*

Participants. Twelve international postgraduates (age 18-50, mean=25) – non-native English speakers living in the UK - were recruited on a voluntary basis (a compensation of £10 Amazon voucher was offered). Users varied in Gender (7 males and 5 females), cultural background (1 Chinese, 1 Greek, 2 Jordanian, 2 Indian, 1 Iranian, 2 Malaysian, 1 Mexican, 1 Polish, and 1 Saudi Arabian).

Method. Each participant was given a study form and was provided with individual access session to MusicPinta via a URL[9]. Every session was *conducted separately* and observed by the first author. All participants were asked to provide feedback before, during, and after the interaction with MusicPinta, as follows:

Pre-study questionnaire [5 min] - collected information about participant's profiles, and their familiarity with the music domain, focusing on the music categories which would be explored – String instruments, Wind instruments, Percussion instruments. The participants' familiarity varied from none, low, medium, and high.

Introduction to MusicPinta [5 min] - the participants followed a script which introduced the main features of the system using 'Tabla' (an Arab percussion instrument).

Exploring three musical instruments [45min] - the users explored three musical instruments by following the three exploration strategies. Each instrument belongs to a particular instrument family and originates from a national culture; we used the GLOBE cultural clusters [6] for the national cultures, see summary in Table 1. The order of conditions was alternated to counter balance the impact on the results.

Table 1. Allocation of exploration strategies for the selected musical instruments.

Exploration Strategy	Instrument Name	Instrument Family	GLOBE Cultural Cluster
Density	Oud	String Instrument	Arab Cultures
Familiarity	Bansuri	Wind Instrument	Southern Asia
Unfamiliarity	Xylophone	Percussion Instruments	Eastern Europe

For each exploration strategy, we measured the degree of participants' cognitive processes of *remember*, *categorise*, and *compare* **before and after** the completion of each exploration, indicating the knowledge utility of the path (as discussed above). In addition, we considered the degree of recognition made by the participants for each

8 The remaining cognitive categories, which include apply, analyse, evaluate and create, require deeper learning activities which usually happen outside a browser.

9 http://imash.leeds.ac.uk/services/pinta/app/ *u/n:user18 pass:musicpinta18*

node during the exploration path to have an indication about participant's familiarity with the exploration domain. After each exploration, the participants were asked to fill a questionnaire about their exploration experience and the cognitive load (based on a modified version of the NASA-TLX questionnaire [14]). Participants were asked to think aloud; the experimenter kept notes of any interesting comments made.

Post-study questionnaire [5 min] - each participant was interviewed at the end of their session about their subjective feedback on the exploration strategies.

5 Results

To address the research question of the study, i.e. identifying benefits and drawbacks of each of the strategies and suggesting possible ways to combine/improve these strategies, we analysed the *user knowledge expansion* (which is the ultimate goal of the nudging mechanism we want to design) and *user exploration experience* (which is informed by usability aspects associated with exploratory search).

5.1. User Knowledge Expansion

For each exploration strategy, the user knowledge was measured before and after her exploration using three questions related to the focus entity X (Oud, Bansuri or Xylophone) and the selected cognitive processes related to knowledge utility (Section 2):

- **[Q1-remember]** *What comes in your mind when you hear the word X?*;
- **[Q2-categorise]** *What musical instrument categories does X belong to?*
- **[Q3-compare]** *What musical instruments are similar to X?*

The number of different entities mentioned in each user answer was counted. The difference between these numbers for each question before and after exploration is taken as an indication of the **effect of exploration** on the corresponding cognitive process. For example, if before the exploration a user could name 2 instruments similar to Oud (Q3) and after exploration the user named 6 instruments similar to Oud, the effect of the exploration on the cognitive process *compare* is indicated as 4 (i.e. as a result of the exploration the user learned 4 similar instruments to Oud). The effect of the three strategies is shown in Table 2.

Table 2. Effect of the three strategies on the user cognitive processes (median of the effect of exploration for all users).

Exploration Strategy	Effect of Exploration		
	Remember	Categorise	Compare
Density (D)	3	2.5	5
Familiar (F)	1	1	2
Unfamiliar (U)	1	1	1.5

Before exploration, the median values for the three questions were 0, as most users were not able to articulate many items linked to the three musical instruments.

The effect of the D-strategy on the three cognitive processes was higher than the effect of the F-strategy and the U-strategy; and this difference is significant (Table 3).

Table 3. Statistically significant differences of the values in Table 2 (Mann-Whitney, 1-tail, Na=Nb=12).

Difference in the Effect of Exploration (Table 2)	Cognitive Process	U	p
Density > Familiar	Remember	35.5	P<0.05
	Categorise	14	P<0.001
	Compare	18	P<0.001
Density > Unfamiliar	Remember	34	P<0.05
	Categorise	15	P<0.001
	Compare	27	P<0.001

Notably, for the cognitive processes *categorise* and *compare* the bigger effect on exploration of the D-strategy over the other strategies is highly significant ($p<0.001$).

To further analyse what caused that the D-strategy was better than the others, we looked at the data collected *during each user's exploration*. At every focus entity in an exploration path, the user was asked to click on both 'Features' and 'Relevant Information' (Fig 2) and name the entities she recognised from all entities MusicPinta associated with the focus entity. The recognised entities were recorded by the experimenter. The overall number of entities recognised along the user exploration paths is summarised in Table 4. The recognition along the whole path, which involves the initial search, the suggested first click (which was always the most dense entity to ensure the users started with the same conditions), and the following three clicks where the users explicitly followed the specified strategy (strategy-related part).

Table 4. Summary of the recognised entities along the users' exploration paths (median values).

Exploration Strategy	Recognised Entities	
	Whole Path	Strategy-related Part
Density (D)	40.5	21
Familiar (F)	20.5	7
Unfamiliar (U)	15	6
ALL	24	10

There is a statistically significant strong positive correlation between the number of recognised entities during the whole path and the effect of exploration on the *compare* cognitive process (Spearman;$R=0.67$;$p<0.0001$). There is a weak correlation between the recognition and the effect of the exploration on the *categorise* and *remember* cognitive processes (Spearman;$0<R<0.3$;$p>0.5$ in both cases). Hence, the *compare* cognitive process (i.e. indicating similar instruments) is influenced by the number of entities the users recognise during the exploration.

To further analyse why the F-strategy and U-strategy had smaller effect on the *compare* cognitive process than the D-strategy, we linked for each strategy the effect of exploration on the cognitive processes with the recognition only along the strategy-related path (given in the second column in Table 4). There is a statistically significant strong positive correlation between the recognition when the users followed the D-strategy and the effect on *remember* (Spearman;$R=0.82$;$p<0.001$) and *compare* (Spearman;$R=0.68$;$p<0.01$), while the observed moderate correlation for *categorise* was not statistically significant (Spearman;$R=0.43$; $p=0.08$). There was no correlation between the recognition based on the strategy and the effect on exploration (Spearman;$0<R<0.3$;$p>0.5$ in all). Hence, users recognised more during the D-strategy which led to a positive effect on the cognitive processes *remember* and *compare*.

Exploration cases with low knowledge utility. Further analysis of the individual cases when the effect of exploration was low identified several interesting situations:

- ***[FF]*** *Exploring familiar entities in a familiar domain*. The three users who were from the Southern Asia GLOBE cluster (i.e. familiar with banzuri) and followed the F-strategy (which was allocated to banzuri) did not improve their scores for *remember*, *categorise* and *compare*, despite the fact that many entities were recognized along the path. Hence, being familiar with the domain and sticking to familiar items had low knowledge utility, as the users did not notice new things.
- ***[FU]*** *Exploring familiar entities in an unfamiliar domain*. Two users who were not familiar with the Southern Asia GLOBE cluster and followed the F-strategy (for banzuri – Indian instrument) had poor scores for the three cognitive processes, as they stayed within the scope of what they knew and did not make any connection to any of the new things they were seeing on the exploration path. This indicates that even if the user is presented with something new, they may not be able to learn it as they may not be able to contextualise it.
- ***[UU]*** *Exploring unfamiliar entities in an unfamiliar domain*. Four of the users did not improve much their knowledge when following the U-strategy (which was allocated to xylophone). An analysis of the profiles of these users revealed that they had no knowledge of the instrument family (percussion instruments) and were not from the corresponding the Easter Europe GLOBE cluster (xylophone is Greek instrument). It was noted that the users recognized a fair bit of entities during the exploration path, yet they were not able to associate to xylophone.
- ***[DF]*** *Exploring dense entities in a familiar domain*. As a whole, the exploration paths which followed the D-strategy had the highest knowledge utility. However, there was one specific case when a user did not gain much about Oud (the entity for the D-strategy) from the exploration. A close examination of the profile of this user showed that she was both familiar with the instrument family (string instruments) and lived in the Arab Cultures GLOBE cluster where Oud is played. This is similar to the first case – although the user was recognising many things, they were not noticing new things and not expanding their knowledge.

One user gained most from all her exploration paths disregarding from the strategy she followed. She was looking for links between familiar and entities (e.g. starting from her national culture and picking instruments she did not know). This suggests that encouraging users to seek connections and form associations may increase the knowledge utility of their exploration paths.

5.2. User Exploration Experience

After each exploration path, the participants' feedback on the exploration experience during the path was collected including exploration complexity (adapted from NASA-TLX) and exploration usability (referring to aspects related to exploratory search over linked data, observed earlier [5]). Figures 4 and 5 give a summary of the feedback.

D-strategy. The exploration paths following density strategy were seen as *most interesting*; the difference is statistically significant (Mann Whitney; D-strategy > F-strategy, $U=42$, $p<0.05$; D-strategy > U-strategy, $U=36$, $p<0.05$). Most of the users noticed musical instruments from diverse cultures which enabled them to make connection and associations between musical instruments that were originating from their culture with musical instruments from other cultures. For example, one of the users stated that '*I saw new string instruments from China, India, Arabic world, Greek and Africa, and this cultural variation was very interesting for me*'. Also, users found D-strategy interesting since it led them to a mix of familiar and unfamiliar instruments, e.g. exploring string instruments users noticed entities with familiar instruments which led them to unfamiliar instruments. Overall, the users also found the D-strategy *least boring* and *most informative* (but these differences were not statistically significant). However, on a few occasions the users found the D-strategy confusing or frustrating, as it directed them to information spaces with many unfamiliar instruments.

F-strategy. The paths following familiarity strategy were also found *informative* since once directed Wind Instruments (start from a dense entity) the users were able to see many familiar instruments and make connections. However, two users found F-strategy *boring*, as they only explored familiar things and did not find new things.

U-strategy. More paths following unfamiliarity strategy were rated as frustrating comparing the other two strategies; the difference is significant (Mann Whitney; D-strategy < U-strategy, U=122.5, p<0.01; U-strategy < F-strategy, U=102.5, p<0.05). Furthermore, the U-strategy was rated as *least informative*; the difference is significant (Mann Whitney; D-strategy < U-strategy, U=36, p<0.05; U-strategy > F-strategy, U=42, p<0.05). In addition, half of the users indicated that they were confused since it was difficult for them to understand descriptions of unfamiliar instruments.

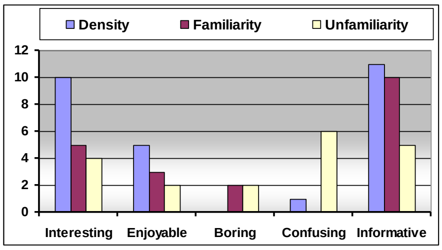

Fig 4. Users' exploration experience of the three exploration strategies (values show number of user paths rated with the corresponding characteristics**).**

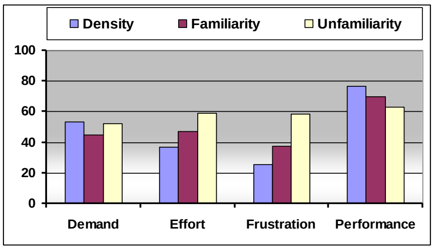

Fig 5. Users' subjective perception of the three exploration strategies, based on an adapted NASA-TLX questionnaire [14] (values are average rate in the range 1-100)

5.3. User Feedback

The individual interviews at the end of each user session provided additional feedback about possible ways to overcome observed problems and combine the three strategies. The users confirmed their preference for being directed to dense places, as they could see both familiar and unfamiliar things. The users elaborated several useful points:

- When the exploration goes through too familiar spaces (like cases FF and DF in Section 4.1), the user should be directed to something new.
- Newness is associated with unfamiliar entities, not seen during the exploration.
- Offering new things should be based on some aspects from the domain, e.g. in the case of MusicPinta the users suggested that new entities could be offered based on the cultural cluster or the instrument family. For example, a user stated

'*I would like to put Bansuri within Arabic Wind Instruments so it becomes easy to understand what bansuri is and to make useful associations.*'

- When the exploration goes through too unfamiliar spaces (like cases UU and FU in Section 5.1.) which can cause frustration and confusion, the user should be helped to make a connection between new things and what they are familiar with.

6 Discussion and Conclusion

In this work, we propose three nudging strategies to aid users when exploring linked data - based on the user's familiarity with the domain and the density of nodes within linked data graphs. The user study investigates benefits and limitations of each strategy based on the knowledge utility of the resultant exploration path, i.e. the degree to which the user's cognitive processes remember, categorise, and compare are invoked. Several observations about the strategies can be drawn from the study results.

D-Strategy used as the underpinning strategy. D-strategy has statistically significant higher cognitive effect on exploration compared to the other two strategies. In particular, participants were able to remember, recognize and compare more things when using D-strategy. Hence, Density strategy can play a key role as underpinning (default) strategy, which can be complimented/extended using the other strategies.

Recognition is a key enabler for user knowledge expansion. The study found that the more the participants recognised entities, the higher the effect on the cognitive processes was. Hence, nudges for triggering recognition (e.g. prompts asking to notice something familiar) should be provided after the user is directed to dense nodes.

Diversification to encourage connections. The cases when the F-Strategy and U-strategy performed poorly indicated interesting situations which could be detected. When the user stays mainly in familiar places, she may miss to notice new things. When such situations are detected, nudging should direct the user to something new. Similarly, when the user explores mainly unfamiliar spaces, a connection with something familiar can be pointed to increase the knowledge utility. Hence, diversification should be provided to help user discover connections and form associations by suggesting new things, and linking them with familiar or seen things. Diversification should be based on some domain aspects (e.g. instrument family or cultural category).

User profile to detect user's domain familiarity. To detect situations when prompts can be added, in addition to the exploration history, a mechanism for deriving a user profile is needed. Even a shallow profile (e.g. in here the profile was collected explicitly by asking the user about their cultural origin and familiarity with specific instrument families) would be helpful to detect that the user stays in too familiar or too unfamiliar information spaces, so an appropriate diversification prompt is given.

The paper presents the first step of work in progress towards deriving an intelligent mechanism for nudging users while exploring linked data graphs, with a focus on user knowledge expansion. Our immediate future work is to use the findings to design and implement a mechanism for nudging, which has D-strategy as a default and embeds prompts to encourage recognition and to diversify user's exploration space. In the long run, we intend to also extend the user profiling to include also implicit methods.

References

1. Alahmari, F., Thom, J., Magee, L. & Wong, W. Evaluating Semantic Browsers for Consuming Linked Data. In Proceedings of the 23rd of the ADC, 2012, Melbourne, Australia.
2. Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. Taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives (Complete edition), 2001. New York: Longma.
3. Brunetti, J., García R., Auer, S. From Overview to Facets and Pivoting for Interactive Exploration of Semantic Web Data. In the International Journal on Semantic Web and Information Systems, 2013.
4. Cheng, G., Tran, T. & Qu, Y. RELIN: Relatedness and Informativeness-Based Centrality for Entity Summarization. In ISWC 2011, pp. 114-129.
5. Dimitrova, V., Lau, L., Thakker, D., Yang-Turner, F. & Despotakis, D. (2013). Exploring Exploratory Search: A User Study with Linked Semantic Data. In IESD 2013 ACM 978-1-4503-2006-1/13/05.
6. Gupta, V., Hanges, P.J., Dorfman, P. (2002). Review Cultural clusters: methodology and findings. Journal of World Business, 37(2), 11-15.
7. Landherr, A., Friedl, B. & Heidemann, J. (2010). A Critical Review of Centrality Measures in Social Networks. In Business & Information Systems Engineering December 2010, Volume 2 (6), pp 371-385.
8. Nuzzolese, A., et al. Aemoo: exploring knowledge on the Web. In proceedings of the WebSci 2013.
9. Robert A. Hanneman & Mark Riddle (2005). Introduction to Social Network Methods. Book, 2005.
10. Thakker, D., Dimitrova, V., Lau, L., Yang-Turner, F. & Despotakis, D. Assisting User Browsing over Linked Data: Requirements Elicitation with a User Study. In proceedings of ICWE 2013, pp. 376-383.
11. White, R., Muresan & G., Marchionini, G. Evaluating Exploratory Search Systems. In SIGIR'06, USA.
12. Sunstein,C., Thaler,R: Nudge: Improving Decisions about Health, Wealth, & Happiness. Penguin (2009).
13. Miller, G. A. "The magical number seven, plus or minus two: Some limits on our capacity for processing information". Psychological Review 63 (2): 81–97. doi:10.1037/h0043158. PMID 13310704 (1956).
14. Hart, S.G. and Staveland, L.E. (1998). Development of the NASA-tlx (Task Load Index): Results of empirical and theoretical research. Human Mental Workload, 139-183.
15. Niklas Luhmann. Trust and power. Chichester, UK: Wiley. (1979).
16. Marchionini, G. Exploratory search: From Finding to understanding. In Communications of the ACM April 2006/Vol. 49, No. 4.
17. H. Alani and C. Brewster, "Ontology Ranking based on the Analysis of Concept Structures," Proceedings of the 3rd international conference on Knowledge capture KCAP 05, p. 51, 2005.
18. D. Schwabe. Explorator: a tool for exploring RDF data through direct manipulation. Framework, 2009.
19. Waitelonis, J., Knuth, M., Wolf, L., Hercher, J., and H. Sack. The Path is the Destination-Enabling a New Search Paradigm with Linked Data. In LD in the Future Internet @ Future Internet Assembly, 2010.
20. R. W. White, B. Kules, S. M. Drucker, and m.c. schraefel. Supporting Exploratory Search, Introduction, Special Issue, Communications of the ACM, vol.49, no.4, pp. 36–39. 2006.
21. A. Hermann, "Semantic Search : Reconciling Expressive Querying and Exploratory Search," in The Semantic Web – ISWC 2011, 2011, pp. 177–192.
22. I. Popov, M. Schraefel, W. Hall, and N. Shadbolt, "Connecting the Dots: A Multi-pivot Approach to Data Exploration," The Semantic Web ISWC 2011 10th International Semantic Web Conference Bonn Germany October 2327 2011, 2011.
23. Sah, M. & Wade, V. Personalized Concept-based Search and Exploration on the Web of Data using Results Categorization. In ESWC 2013.
24. Maccatrozzo , V., Aroyo , L., Robert , R. Crowdsourced Evaluation of Semantic Patterns for Recommen-dations. In UMAP 2013. LBR.
25. Dean, M., Basu, P., Carterette, B., Partridge, C. and Hendler, J. What to Send First? A Study of Utility in the Semantic Web. In Proceedings of the Joint Workshop on Large and Heterogeneous Data and Quantitative Formalization in the Semantic Web (in conjunction with ISWC 2012).