

Semantic Jira - Semantic Expert Finder in the Bug Tracking Tool Jira

Velten Heyn and Adrian Paschke

Corporate Semantic Web, Institute of Computer Science,
Koenigin-Luise-Str. 24, 14195 Berlin, Germany
adrian.paschke@inf.fu.berlin.de
velten@zedat.fu-berlin.de

Technical Report TR-B-13-04 published in
http://edocs.fu-berlin.de/docs/receive/FUDOCs_series_00000000021

Abstract. The semantic expert recommender extension for the Jira bug tracking system semantically searches for similar tickets in Jira and recommends experts and links to existing organizational (Wiki) knowledge for each ticket. This helps to avoid redundant work and supports the search and collaboration with experts in the project management and maintenance phase based on semantically enriched tickets in Jira.

Keywords: Bug Tracking, Semantic Expert Finder, Semantic Web

1 Introduction

There is a huge economic potential in the use of Corporate Semantic Web tools in Software Engineering and project management. Bucktracking systems, such as Jira, can benefit from such semantic support. With the information about software bugs, issues, and project tasks cumulated by a bug tracking system and with integrated semantic techniques for transforming this information into meaningful knowledge, it becomes possible to automatically support employees in their daily tasks.

If they need help with a particular task they can browse through similar tickets in the bug tracking system and can reuse the documented solutions to solve their problem. However, with the growing information available in such an enterprise information system it becomes harder to find the relevant tickets.

In this paper we propose a semantic extension to Jira in order to overcome this problem. This Semantic Jira supports semantic search on the knowledge documented in Jira. Similar tickets are semantically inferred by the system and ranked by different metrics (tf-idf, freshness). Furthermore, the most active employees of the found similar tickets are extracted, ranked and proposed as knowledge experts. The underlying rationale for this approach is, that they are very likely experts in the field of knowledge, to which this ticket belongs to. The support of asking experts for help has the advantage, that the expert can directly communicate their knowledge and that they can abstract all unnecessary details if the help seeking colleague is not familiar with the domain.

Our expert recommender approach differs from the typical existing solutions which are expert finder search tools. The drawback of searching by users is, that the employees and project managers need to use the right search terms to describe the required skills of an expert for a given problem (bug, issue). This is a non trivial problem, in particular for non-IT persons. In our recommender approach an expert is recommended by a recommendation system, which uses the reported bug to infer the knowledge field and automatically find the experts in this field. These experts are recommended in the ticket by the Semantic Jira system together with information about similar tasks or tickets (documenting existing problem solutions) and matching wiki/wikipedia-articles which document existing enterprise knowledge.

The main research questions in the propose solution are:

1. How to find similar tickets in an automated way? What defines similarity?

2. How do we find possible experts and which information sources should we use for this?
3. Which possibilities do we have to request articles from enterprise Wiki systems which document relevant knowledge?

The benefits of the proposed solution, based on a semantic Jira approach, is that it increases the visibility of implicit and explicit knowledge in the company and that it links this knowledge to the ongoing activities of employees. This helps to avoid redundant work in a company and it helps to optimise the distribution of work on the general resources of employees.

The further paper is structured as follows: Section 2 describes the related work. Section 3 introduces the conceptual solution - a Semantic Jira back tracking system. Section 4 gives more details about the implementation of this system. The evaluation in section 5 is based on a user study performed in a company. Finally, section 6 concludes this work.

2 Related Work

“Who Knows about That Bug? Automatic Bug Report Assignment with a Vocabulary-Based Developer Expertise Model“ [1] from Dominique Urs Philipp Matter is the largest project we found in our state of art analysis. Their Expert Recommender uses source code analysis to find the most applicable expert for a given bug tracking item. The main drawback of this approach is, that it needs half a year active contribution to the source code from a developer to make proper recommendations. Their use case is to assign developers to new bug tracking items. The fact that they primarily consider source code for expert triage makes the system not usable for non-IT users. That’s why we decided to stick to the data available directly in the bug tracking system.

“Expertise Recommender: A Flexible Recommendation System and Architecture” [2] from David W. McDonald and Mark S. Ackerman is another paper which is making use of the change history of source code. A (proprietary) “Tech Support”-Database is used in their field study which was done in a company. With their approach it is only feasible to locate experts within the IT department. The paper further focuses on the architectural aspects of creating a reusable system which can take many different algorithms into account.

“Expert Recommender Systems in Practice: Evaluating Semi-automatic Profile Generation” [3] from Tim Reichelt and Volker Wulf is applying another source to tackle this problem. They are using a client program which examines the documents within a folder and subfolder which was selected by the user. This program sends these examined word statistics to the server and compares it with other statistics. It was also considered to use emails as data source but discarded because of privacy protection issues.

“Using Domain Ontologies for Finding Experts in Corporate Wikis” [4] from Ralph Schäfermeier and Adrian Paschke is taking the wiki entries of users as source. They are using the SEOntology to infer if the users’ Wiki contributions are using an experts’ knowledge language. They also include an author reputation metric to gain more precision.

All these previous works have the problem that they are not sufficiently integrated into the task workflow of an employee which is directly managed and supported by a bug tracking and project management system such as Jira.

3 Ticket Recommender and Expertfinder Model

Expertise represents the implicit knowledge and the competencies of an expert. The proposed conceptual solution allows inferring expertise from an expertise model which considers the knowledge fields of a ticket saved inside typical bug tracking systems such as Jira.

Typical roles involved in a ticket for such a model are:

- the ticket reporter / creator
- the ticket assignee / solver
- the ticket follower

Before the actual opening of a ticket and the specification of the involved roles, the ticket creator has to create the ticket in a relatively complex decision process as shown in figure 1.

Ticket Reporter The creator needs to have enough information and knowledge to distinguish if

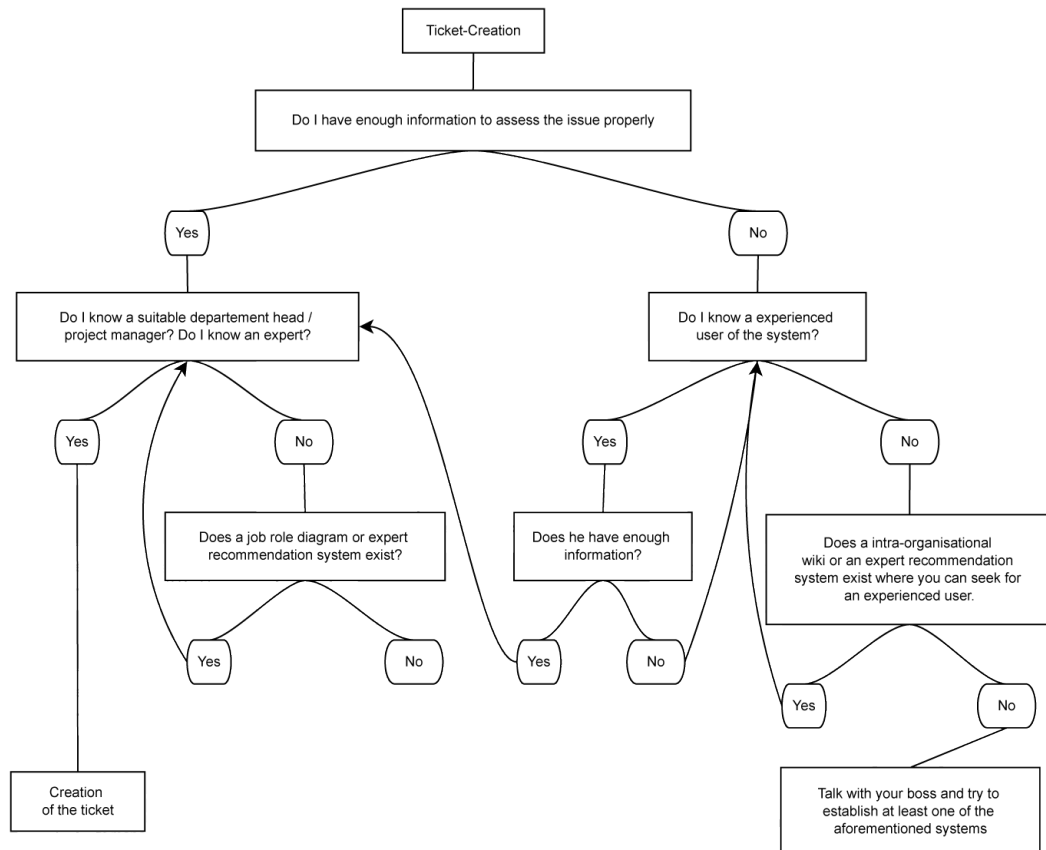


Fig. 1. Process of Ticket Creation

the occurred problem was a problem of misuse (could lead to a feature request) or a bug in the used system. If the creator can distinguish the problem, it either means that he can act as an expert for this reported part of the system or that he might know somebody who can act as contact person. That means, the more tickets a particular person has for the same area of work the more likely it is that he has knowledge either about the distribution of tasks and skills or about the integration within the system.

Ticket Assignee The ticket reporter and the assignee have different kinds of knowledge in a specific area. During the process of problem solving the assignee creates new artifacts (e.g. source code) and if it's not repetitive work new knowledge especially about specific details of the implementation. Thus the more tickets the person solves in a particularly area the more profound knowledge he should have.

It is likely that the words used within the description of a ticket capture the knowledge required to solve new tickets. That means, to find similar tickets and experts who can possibly help, it is required to classify the text in order to discover the knowledge field addressed by the ticket.

3.1 Retrieval and Recommendation Model

We apply a simple statistical tf-idf measure to get the relevancy of the words within the ticket description. After taking the most k relevant words of the ticket (where k is normally a value between 5 and 20) it is possible to search for other tickets which contain similar relevant words. The score for each ticket is then

$$similarity(s, d) = \frac{\sum_{w \in W} tw(w, s) \cdot tw(w, d)}{\sqrt{\sum_{w \in W} tw(w, s)^2 \cdot \sum_{w \in W} tw(w, d)^2}}$$

(where W is the set of relevant words, tw is the tf-idf weighting function, s is the source ticket and d the ticket of which you want to find out the similarity). The upper part of the division is the sum of the weights multiplied from both tickets. To normalise the results if they contain different amounts of words the upper part is divided by the second norm of each weight. If you look at W as a vector you can transform the upper formula to

$$similarity(s, d) = \frac{tW(W, s) \times tW(W, d)}{|tW(W, s)| \cdot |tW(W, d)|}$$

(where tW takes the vector of the words W and weights them with the tf-idf measure in regard to the ticket given in the second argument). We can either apply a linear scoring or an inverted exponential score for the retrieval and ranking of similar tickets to a given new ticket.

We further recommend similar tickets which are semantically similar according to a domain ontology representing the expert vocabulary used in the ticket domain. The underlying rationale for this approach is, that experts typically use a topic specific vocabulary which is modelled by the domain ontology. We apply a taxonomic ontology matcher which identifies the tickets' terms as resources from the expert taxonomy. The matcher computes the similarity between two concepts $c1$ and $c2$ based on the taxonomic distance $d(c1, c2)$ between them, which reflects their respective position in the concept hierarchy. The matcher is able to handle multiple inheritance of concepts at the leaf level of a taxonomy.

The concept similarity is defined as: $sim(c1, c2) = 1 - d_c(c1, c2)$. Every concept in a taxonomy is assigned a milestone value. Since the distance between two given concepts in a hierarchy represents the path over the *closest common parent* (ccp), the distance is calculated as:

$$\begin{aligned} d_c(c1, c2) &= d_c(c1, ccp) + d_c(c2, ccp) \\ d_c(c, ccp) &= milestone(ccp) - milestone(c) \end{aligned}$$

The milestone values of concepts in a taxonomy are calculated either

- with a *linear milestone decrease* $milestone(n) = 1 - [l(n)/l(N)]$, where $l(n)$ is the depth of the node n in the taxonomic hierarchy and $l(N)$ is the deepest hierarchy level, or
- with an *exponential milestone decrease* $milestone(n) = 0.5/k^{l(n)}$, where k is a factor greater than 1 indicating the rate at which the milestone values decrease along the hierarchy

After the retrieval and ranking of (statistically and semantically) similar tickets, in the next step the experts are identified. In general, the expertise score is a function $a \times t \mapsto e$ for a given author a and a topic t . The set of potential experts are the persons involved in these tickets either as creator, solver, or follower. For each person an experts score is calculated as the accumulated similarity measure (linear or inverted exponential). We distinguish two dimensions, the organisation score and the developer score and calculate the overall expert score $e = e_o + e_d$.

Additionally, we retrieve relevant Wikipages from an Enterprise Wiki documenting expert knowledge. We assume that an individual who contributes content relevant to a specific ticket topic to such a Wiki has expertise in this topic. We calculate a simple contribution based expertise score as follows:

$$expertise_{simple}(a, t) := \sum_{s \in S_{a,t}} weight_s(level(s)) + \sum_{w \in W_{a,t}} weight_w(w) \quad (1)$$

where by $S_{a,t}$ we denominate the set of sections that cover the topic t and under which author a has contributed content. t is again identified as a concept in the tickets' domain ontology. $weight_s$ is a weighting factor depending on the section level. We used a simple milestone metric in order to express the relevance of a section according to its level. The underlying assumption is that a section with a higher level is about a more general topic, and contributions to a highly specialized topic should be reflected with a higher weighting in the expertise score. Accordingly, by $W_{a,t}$ we denominate all occurrences of terms contributed by author a that can be mapped to ontology topic t , weighted by a relevance function $weight_w$.

We further semantically consolidate the expertise score by considering authors Wiki contributions about topics which are semantically similar to the topic of the addressed ticket topic. We utilize the class hierarchy established by the `owl:subClassOf` OWL property and other selectable subtypes of `owl:ObjectProperties` in order to capture concept relatedness.

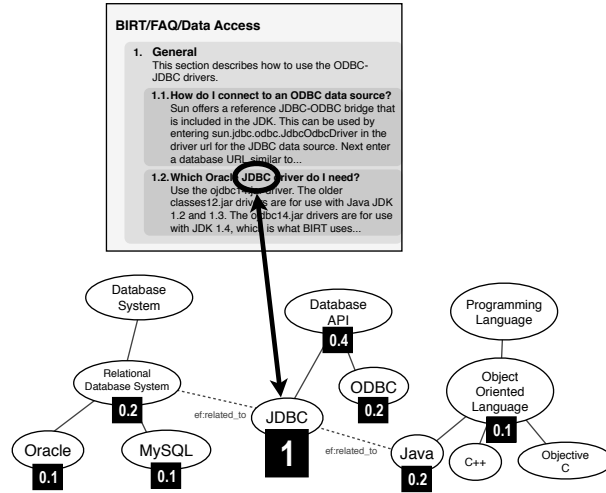


Fig. 2. Weighting of Detected and Related Features

Based on these relations, we calculate a relevance score using ontology based similarity measures until a defined threshold is reached (see figure 2). All concepts with a similarity value higher than this threshold are considered similar and added to the feature set, weighted by its similarity value. This yields a consolidated expertise score which is calculated as follows:

$$expertise(a, t) := \sum_{t_{sim} \in T, sim(t, t_{sim}) \geq sim_{min}} expertise_{simple}(a, t_{sim}) \cdot sim(t, t_{sim}). \quad (2)$$

Even if topic t is never referenced by author a , but neighbouring topics t_{sim} with a similarity to t greater than the threshold sim_{min} , then t benefits from a 's expertise in each topic t_{sim} . The more similar t and t_{sim} are, the higher the benefit for t . This score can be further adjusted by additionally

considering the authors reputation based on e.g. considering the revision history, stability, and life time of Wiki entries. [4]

4 Implementation

4.1 Basic features of the expert recommendation system

The developed plugin for Jira delivers the following basic features:

- Automatic indexing of all tickets with usage of
 - An adapter for the connection between the Index and the Jira ticket data structure
 - Automatic translation for ensuring a consistent language basis
- Recommendation of alternative tickets which are similar to the current ticket
 - Calculation of different rankings (administrational and development score)
 - Recommendation of experts
- Integration of matching Wiki articles and adapted expert score from the search engine and expert finder

4.2 System overview

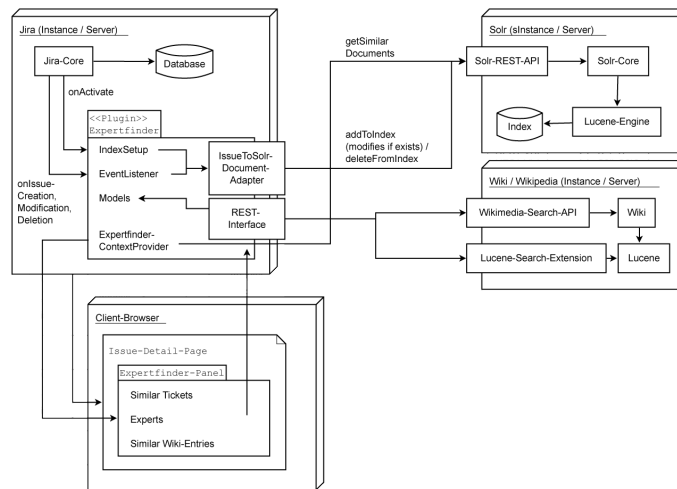


Fig. 3. Overview of the interaction between the different systems

As indexing and information retrieval platform Lucene with it's efficient indexing engine is used. On top a Solr server manages the Lucene index, supplies Lucene with data, and retrieves the data. Solr gives a REST connector, over which all established (CRUD-) operations are realized. For the semantic similarity computation we use the CSW Semantic Similarity Matchmaking Framework (SemF)¹. The framework allows taxonomic and non-taxonomic concept matching techniques to be applied to selected object properties.

¹ <http://www.corporate-semantic-web.de/technologies.html>

The Mediawiki software does have its own search engine to use for finding articles but their features are limited. It is not possible to search for more than one term at once otherwise all given words have to appear in the returned articles. With the extension LuceneSearch² it is possible to make more complex queries and give a ranking for the returned articles. To calculate the semantic expert score from the Wiki contributions we adapt and integrate the Wiki expert finder described in [4].

Figure 3 shows the interaction between the different systems and components. Jira is at the core of the system. From Jira the EventListener is activated to handle all operations on a ticket to send them to the Solr instance. The ExpertFinderContextProvider is triggered when a user requests a ticket details page and gathers all necessary information from Solr, processes it and fills the view.

5 Evaluation

For the evaluation employees from the IT department of a German mid-sized company with around 60 employees were asked to give estimates for best fitting experts for a self-chosen ticket they already worked on themselves. Afterwards all experts for the chosen tickets were collected from the system. Altogether 32 tickets have been evaluated, 98 experts have been given and 267 experts were proposed by the system. The results are shown in 5 and 4.

82.7%	recall
88.6%	of all experts are under the first ten results
74.3%	of all experts are under the first three results
45.7%	of all experts are at the first position of the results
10.03	possible experts are proposed by the recommender in average

Table 1. System evaluation

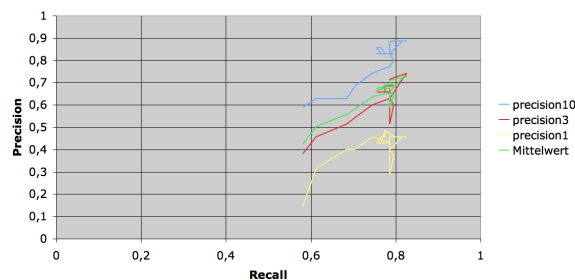


Fig. 4. Evaluation Precision-Recall diagram

As described in section 3.1 it is possible to configure how many words should be taken into account as relevant for the similarity measurement. This has a direct impact on the quality of the results as shown in figure 5.

² <https://www.mediawiki.org/wiki/Extension:Lucene-search>

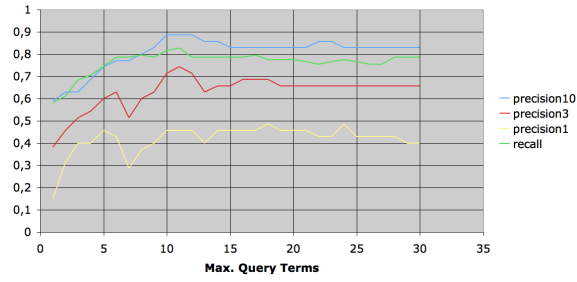


Fig. 5. Best parameter for the amount of relevant words / terms

The Wiki extension has been tested and evaluated using the project Wiki of the Eclipse Foundation³ and the Software Engineering Ontology (SEOntology) [5].

³ http://wiki.eclipse.org/Main_Page

6 Conclusion

The recommender already reaches relatively good results. Future work might additionally consider multi-lingual translations and wordnets together with larger background knowledge coming, e.g. from linked open data sources such as DBPedia. In future work a comparison on the basis of the the bug tracking data from the eclipse project which was used by some of the related works would be useful. A problem to solve is the import of this data into a Jira bug tracking instance. Furthermore, there are many other information stored or available in the bug tracking system which could be used to optimise the results, e.g. the history of tickets (the change of assignees, the change of the status or of the resolution).

7 Acknowledgements

This work has been partially supported by the InnoProfile project "Corporate Semantic Web" funded by the German Federal Ministry of Education and Research (BMBF).

References

1. D. Matter, A. Kuhn, and O. Nierstrasz. Assigning bug reports using a vocabulary-based expertise model of developers. In *Mining Software Repositories, 2009. MSR'09. 6th IEEE International Working Conference on. IEEE, 2009.*, pages 131–140, 2009.
2. David W. McDonald and Mark S. Ackerman. Expertise recommender: a flexible recommendation system and architecture. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work, CSCW '00*, pages 231–240, New York, NY, USA, 2000. ACM.
3. Tim Reichling and Volker Wulf. Expert recommender systems in practice: evaluating semi-automatic profile generation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09*, pages 59–68, New York, NY, USA, 2009. ACM.
4. Ralph Schäfermeier and Adrian Paschke. Using domain ontologies for finding experts in corporate wikis. In *Proceedings of the 7th International Conference on Semantic Systems, I-Semantics '11*, pages 63–70, New York, NY, USA, 2011. ACM.
5. P. Wongthongtham, E. Chang, and T. Dillon. Software Design Process Ontology Development. In *Proceedings of the 2nd IFIP WG 2.12 and WG 12.4 International Workshop on Web Semantics (SWWS) in conjunction with OTM 2006*, pages 1806–1813, 2006.