Semantic tools for improving software development in open source communities

Gregor Leban

Jožef Stefan Institute, Slovenia

ALERT

Institut "Jožef Stefan" Ljubljana, Slovenija

Introduction

ALERT is a system for helping people who participate in open source communities. It integrates communication channels used by a community and provides features such as semantic search, bug duplicate detection, custom notifications, etc.

Integration of communication channels

Open source communities use different communication channels, such as **mailing lists, forums, bug tracking systems and source code management systems**. These channels use different platforms. In ALERT we have sensors that collect all information posted in them and integrate them in a single knowledge base. All collected data is semantically annotated and interlinked.

Main features

- ✓ Integration of main communication channels
- ✓ Semantic annotation of text
- ✓ Source code info. extraction
- ✓ Semantic search
- ✓ Visualization of search results
- ✓ Developer expertise profiling
- ✓ Finding bug duplicates
- ✓ Suggesting bugs to fix for a developer
- ✓ Custom notifications

Demo accessible at:

http://aidemo.ijs.si/alert/
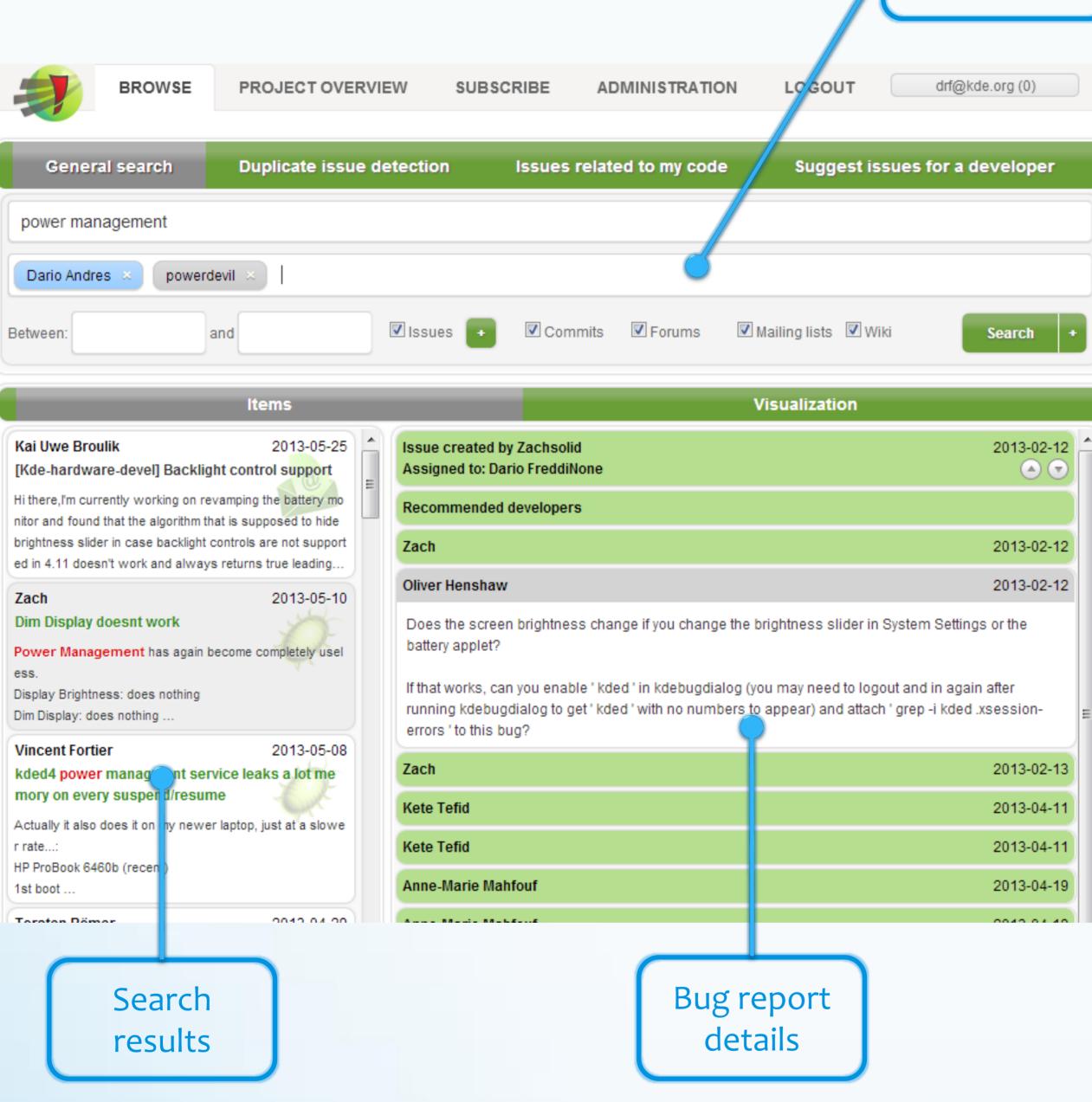
Visualization of search results

Along with listing the results, ALERT also provides a visual summary of them.

Social graph shows the people who are frequently participating in the results. It allows one to quickly identify main people for particular topics.

Timeline view shows distribution of search results over time. Useful for finding interesting patterns such as the time when a bug was likely introduced.

Topic summary displays main topics discussed in the results.

All visualizations are interactive and enable easy query updating.

search conditions

Search results
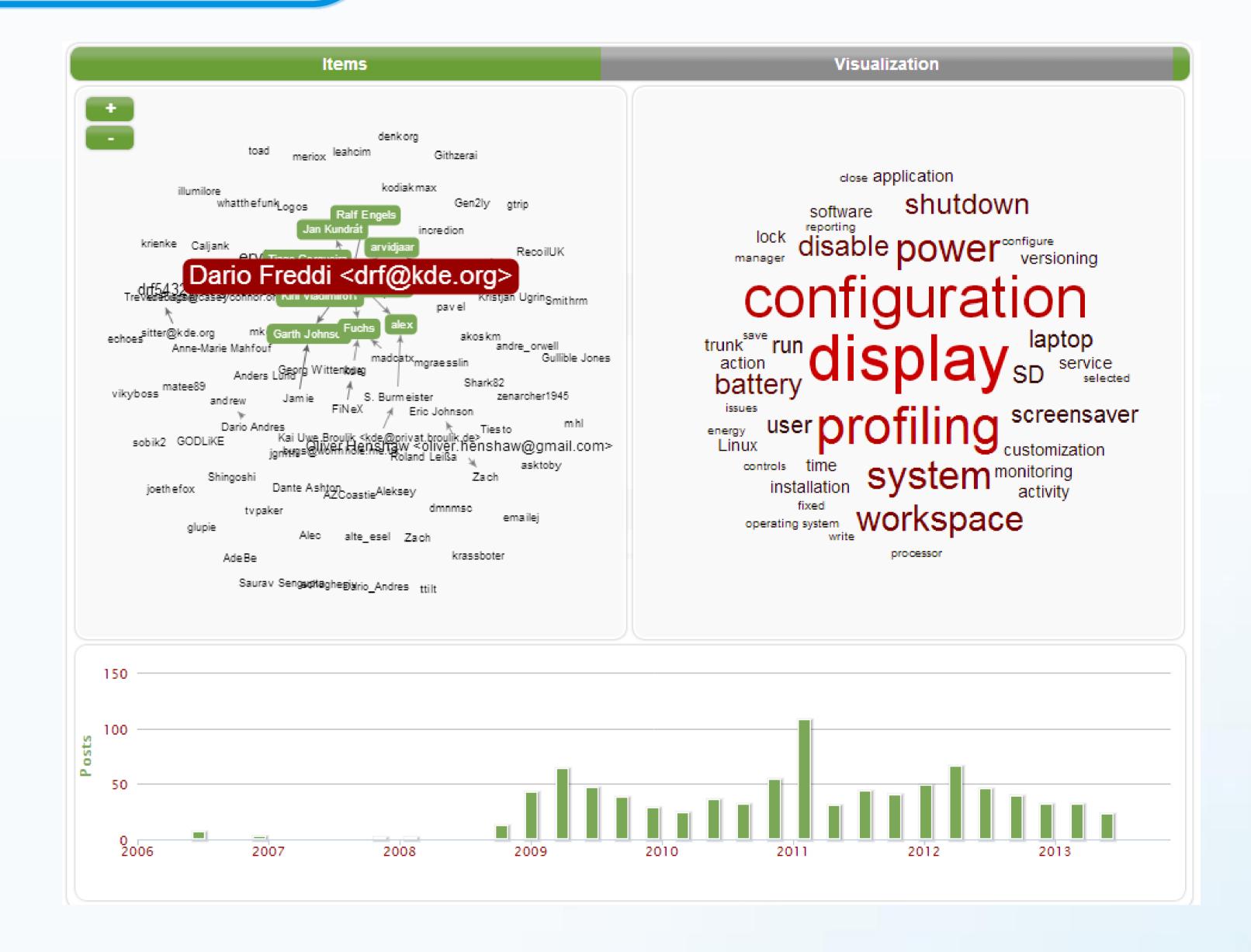
Bug report details

Semantic data annotation

Text is semantically annotated with a custom ontology related to computer science and software development. From **commited source** code we extract the files, classes and methods. References to methods and files are then annotated in bug report stack traces. Available meta-data (post date, author, product, component, …) is automatically added to the knowledge base.

Semantic search

Knowledge base allows us to search using various criteria: keywords, date, authors, products, methods , files, etc. Using the ontology we provide semantic search where the results would be the same whether searching for keywords „dialog", „window" or „form".

User expertise profiling

For each user we analyze the topics discussed by a user and use them to build his expertise profile. The profiles can be used to suggest developers who can fix an issue or to find a person who might be able to offer assistance on a particular topic.

Detecting bug duplicates

By analyzing the bug report similarities we can for each bug identify the most similar reports. Bug triager can use this to quickly determine if a bug is a duplicate of a previous bug.

Suggesting developers who can fix a bug

ALERT supports two ways of suggesting developers who can fix a bug described in a bug report. By analyzing the topics discussed in the report we can identify appropriate developers based on their expertise profile. Alternatively, if a bug report contains a stack trace we can suggest developers who have modified the methods mentioned in the stack trace.

Custom notifications

Developers usually have to repeatedly check the appropriate communication channels to see if new relevant information was posted. ALERT allows the users to specify patterns of interest and they can be automatically notified when information matching the pattern is created.

Using this the user can be notified when, for example, a new forum post mentioning „power management" is created, or when a bug that is more than 80% similar to an existing bug is created.

Acknowledgements