



Use Case 4 in Pharmaceuticals and Health Managing biomedical research notes and raw data

KW Partner: Cognium systems

Business Case: Managing biomedical research notes and raw data

1. Overview

Challenge

Biomedical research notes and raw data are difficult to share because they are not understandable to others unless they are well structured.

Solution

A flexible, intuitive, and simple to use system for entering notes in a structured way (semantic markup) and then benefiting from the structure for sharing and managing the information.

Why a Semantic solution

Research information needs to be recorded in both structured and flexible manner, with complex interlinking. Semantic tagging paradigm corresponds best to all three of these requirements.

Key Business Benefits

Decreased costs, faster work cycle, as well as higher quality of collaboration, project management, and communication.

Business Partners

Institut Pasteur, NEPOMUK consortium (EC funded).

The product of biomedical research is new knowledge. Yet it is not enough to simply arrive at valuable new knowledge, it is equally as important to use it effectively. The shocking fact, however, is that the vast majority of information generated in the process of biomedical research is lost, irretrievably buried in the paper laboratory notebooks and private computer files where researchers record their experiments. Only a small percentage of this information ends up in shared repositories, e.g., scientific publications such as Nature or public databases like GenBank. Unfortunately, the lost information is important for the fight against such problems as disease, hunger, and environmental pollution. We have been working with researchers at Institut Pasteur to develop a Semantic Laboratory Notebook (iPad). iPad's goal is to prevent knowledge loss as well as address other problems of knowledge management in biomedical research. Our focus is on the "must-have" as opposed to "good-to-have" functionality, such as easy access to the important information (clarity, precise search, useful perspectives) and interconnectivity of information. iPad helps people to structure their information more thoroughly and to receive greater benefits from computer-aided information management. The existing systems force people to make a choice between flexible but unstructured data entry or rigid but useful organization. By contrast, data entry within iPad is as flexible and simple as using a regular document editor, while the result is a wellstructured information that is easy to manage and analyze. Consequently, iPad greatly improves and speeds up most aspects of the knowledge management process: organization, storage, retrieval, analysis, sharing, communication, publishing, etc. This, in turn, implies decreased costs, faster time to market, and higher quality of collaboration, project management, and communication.

Keys components

Existing Software

Domain ontologies

iPad: Semantic Laboratory Notebook

Research and development

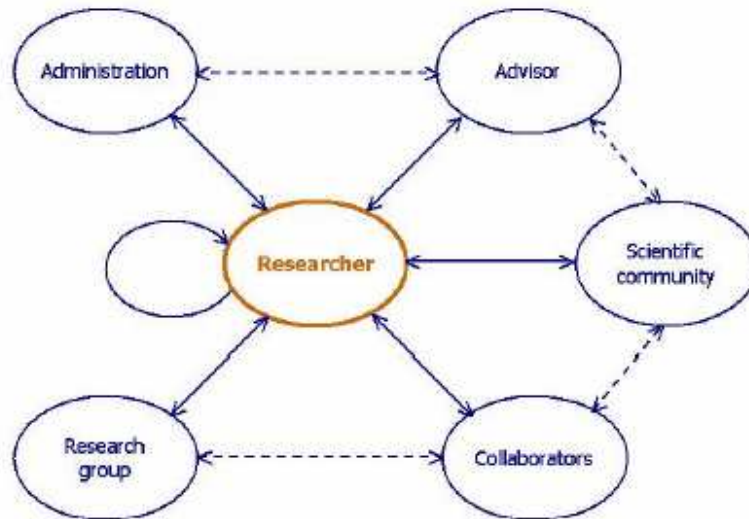
Semantic Wiki

Integrated Environment for structuring and working with structured information

Technology locks

2. Current Practices and Technologies

2.1. Information flow in a research organization



As illustrated in the above diagram, researchers have to communicate with several types of people that participate in the research process:

- Themselves, in the sense that they have retrieve information that they previously created
- Their collaborators, with whom they have to share methodologies and latest results
- Their research group, with whom they have to share resources in the lab (space, equipment, materials)
- Their advisor, to whom they have to report and from whom they need obtain advice
- The scientific community, by presenting at conferences and publishing articles, but also consuming the information produced by the community.

2.2. System requirements Analysis

- Central repository for information sharing
- Flexible system to enter the information. This is critical because research information is too complex to be fit into a form.
- Simple way to semantically tag information at the time of the data entry and afterwards. This has to be as simple as possible because researchers will not spend too much time on learning complex tools and tagging.
- Ability to interlink information using semantic links between tags
- Structure-based semantic search
- Access rights: on the tag level, strict security
- Stand-alone and Web-based interfaces
- Collaborative document editing
- Versioning
- Protection of intellectual property
- Report and perspective generation
- Ability to enter, structure, and work with any information type: simple tabular data, documents, network data, multimedia.
- Should use automatic or semi-automatic markup tools whenever possible
- Collaborative and dynamic ontology construction

- Integration with other software and equipment
- Mobile environment - being to able to use the software anywhere: in the office, at the experimental bench, at home, from the Web.

2.3. Review of the current systems

The existing software (e.g., Kalabie) for managing biomedical research notes is not widely used since it is either too restrictive (based on electronic forms) or it does not support a sufficient level of useful organization (based on unstructured documents). There exist a number of generic technologies based on approaches that are similar to ours, including software from Microsoft. However, they have not gained much popularity because their functionality is difficult to adapt to the needs of specific vertical markets.