The Social Semantic Desktop

NEPOMUK

Mozilla Community Involvement

Deliverable D7.4

Version 1.1

18.01.2008

Dissemination level: PU

Authors

Arnaud Laprévote, EDGE-IT S.A.R.L
Stéphane Laurière, XPertNet
Marcelo Leitner, EDGE-IT S.A.R.L
Anne Nicolas, EDGE-IT S.A.R.L

Mentors

Ansgar Bernardi, DFKI
Gunnar Aastrand Grimnes, DFKI
Markus Junker, Irion Management Consulting GmbH
Brendan Murray, IBM
Alexander Troussov, IBM

Project Co-ordinator

Dr. Ansgar Bernardi
German Research Center for Artificial Intelligence (DFKI) GmbH
Trippstadter Str. 122
67663 Kaiserslautern
Germany
E-Mail: bernardi@dfki.uni-kl.de, phone: +49 631 205 75 105

Partners

Versions

Version	Date	Reason
0.1	28.12.2007	First draft by Stéphane Laurière
0.2	08.01.2008	Integration of input by Edge-IT
0.3	12.01.2008	Enhancement of the state of the art section
0.4	14.01.2008	Input from the mentors taken into account
1.0	15.01.2008	Final version; minor streamlining by IMC
1.1	18.01.2008	Integration of input from A.Troussov. Fixed an issue with the Mozilla Foundation logo.

Explanations of abbreviations on front page

Nature
R: Report
P: Prototype
R/P: Report and Prototype
O: Other

Dissemination level
PU: Public
PP: Restricted to other FP6 participants
RE: Restricted to specified group
CO: Confidential, only for Nepomuk partners

Executive summary

To be really successful, the Nepomuk project needs not only to realize and deploy an innovative approach for collaborative knowledge work, but also to integrate the Nepomuk framework into mainstream desktop environments and development platforms. The Nepomuk community initiative focuses on involving at least the KDE, Eclipse and Mozilla communities of users and developers into the Nepomuk project.

The main goal of the work described in this report is the involvement of the Mozilla community into the Nepomuk project. To this end, a strategy similar to the one adopted for the KDE and Eclipse communities has been followed. An analysis of the existing community and of the relevant projects already ongoing has been carried out, a Nepomuk-Mozilla community platform has been started, and the specification of a reference implementation of Nepomuk core components based on Mozilla XPCOM architecture has been laid down. These activities all together comprise the Nepomuk-Mozilla sub-project.

Given the importance of the Mozilla platform eco-system and of its growing importance on the desktop, it appeared crucial from the outset to the Nepomuk Consortium to provide a demonstrator of the Nepomuk capabilities on top of the Mozilla platform.

This report follows the reports "D7.2 - KDE community involvement" and "D7.3 - Eclipse community involvement", of which it reuses the global structure: the work carried out and the strategy for involving external contributors are similar to the Nepomuk-KDE and Nepomuk-Eclipse initiatives indeed, even though the target communities are different in nature.

In order to present the project and to raise interest across the community, a dedicated Web site and a public mailing-list have been set up at the following URL: http://nepomuk-mozilla.semanticdesktop.org. The Nepomuk-Mozilla project will be submitted to the FOSDEM 2008 Mozilla developer room for widening the number of partakers in the project and for increasing synergies with other projects that share similar concerns.

The specification and use cases of the target Nepomuk-Mozilla prototype have been defined and are presented in this report. The objective of the prototype is to implement the Nepomuk social semantic desktop vision on top of XPCOM and XUL components. This prototype will make it possible to annotate any and all Web data (including parts of Web pages) or local files from Firefox, emails from Thunderbird, to store the annotations in the Nepomuk metadata store and to share them using the Nepomuk P2P social layer. The advantage of using a common metadata store on the desktop lies in sharing annotations and ontologies across the applications instead of having them burried into each application information system. This paves the way for a Mozilla instrumented "Personal Semantic Web".

In parallel to the development of the Nepomuk-Mozilla prototype, the focus will be brought in year 2008 on the involvement of more Mozilla developers into the Nepomuk-Mozilla project, on presenting the project to a wider audience on the occasion of Mozilla related events, and on increasing the presence of the project on the Web.

The availability of the three implementations of the Nepomuk specification on top of the three frameworks KDE4, Eclipse RCP and Mozilla paves the way for fruitful comparisons and hopefully for cross-fertilization between the three communities, hence bolstering the adoption and collaborative enhancements of the Nepomuk standards.

Table of contents

1 Introduction

Nepomuk intends to realize and deploy a comprehensive solution – methods, data structures, and a set of tools – for extending the personal computer into a collaborative environment, which improves the state of art in online collaboration and personal data management and augments the intellect of people by providing and organizing information created by single or group efforts.[1]

Community involvement

To be really successful, the Nepomuk project needs not only to realize and deploy an innovative approach for collaborative knowledge work, but also to integrate the Nepomuk framework into mainstream desktop environments and development platforms. The Nepomuk community initiative focuses on involving at least the KDE, Eclipse and Mozilla communities of users and developers into the Nepomuk project.

Mozilla community

The main goal of the work described in this report is the involvement of the Mozilla community of users and developers into the Nepomuk project. To this end, a strategy similar to the one adopted for the KDE and Eclipse communities has been followed. An analysis of the existing community and of the relevant projects already ongoing has been carried out, a Nepomuk-Mozilla community platform has been started, and a reference implementation of Nepomuk core components based on Mozilla XPCOM architecture [5] has been developed. These activities all together comprise the Nepomuk-Mozilla sub-project.

Involving several key desktop communities represented in figure 1 into the initiative is considered by the Nepomuk Consortium as a mean for guaranteeing that the Nepomuk standards, ontologies and APIs will be sustained beyond the end of the project itself. It is expected that the implementations of the standards on top of different frameworks will encourage a competitive spirit between the targeted platforms, thus bolstering cross-fertilization of ideas between the communities, to the benefit of the future versions of the Nepomuk specification.

Figure 1: Nepomuk targeted communities

Given the importance of the Mozilla platform eco-system and of its growing importance on the desktop, it appeared crucial from the outset to the Nepomuk Consortium to provide a demonstrator of the Nepomuk capabilities on top of the Mozilla platform.

Overview

This document presents what has been accomplished in the Nepomuk-Mozilla project in 2006 and 2007. It consists of the following sections:

1. Section 2 gives an introduction to Mozilla in general and presents specific Mozilla projects that are of interest to the Nepomuk-Mozilla initiative;

---

[1]Source: Nepomuk description of work, Annex I

2. In section 3, the steps that have been undertaken to raise awareness around the project are presented;

3. Section 4 introduces the Nepomuk-Mozilla prototype specification, its objectives and its architecture;

4. Section 5 makes concluding remarks on the progress in the first phase of the project.

This report follows the reports "D7.2 - KDE community involvement" [2] and "D7.3 - Eclipse community involvement" [3], of which it reuses the global structure: the work carried out and the strategy for involving external contributors are similar to the Nepomuk-KDE and Nepomuk-Eclipse initiatives indeed, even though the target communities are different in nature.

---

[2]`http://nepomuk.semanticdesktop.org/xwiki/bin/Main1/D7-2`
[3]`http://nepomuk.semanticdesktop.org/xwiki/bin/Main1/D7-2`

2 Mozilla community overview and relevant projects

This section reminds of the main facts about the Mozilla eco-system, it underlines the importance of the Mozilla community for the Nepomuk project and it presents an overview of the main Mozilla projects that provide semantic or social features being of interest to the Nepomuk initiative.

2.1 The Mozilla project

As stated on the Mozilla.org Web site[4], the name "Mozilla" is used as the principal trademark representing the Mozilla Foundation and the official releases of internet client software developed through [the Mozilla] open source project"[5]. The Mozilla project was born in 1998, following two important announcements by Netscape Communications:

> First, that the Netscape Communicator product would be available free of charge;
> Second, that the source code for Communicator would also be free.

The first developer release of the source code to Communicator was made available in 1998. The group mozilla.org has been established by Netscape in 1998 "to act as the virtual meeting place for the Mozilla code. That group is overseen by the Mozilla Foundation". The Mozilla Foundation, whose logo is represented on figure 2, "was established in July 2003 as a California not-for-profit corporation dedicated to the public benefit. The Mozilla Corporation was subsequently established in August 2005 as a wholly-owned subsidiary of the Foundation to coordinate the development and marketing of Mozilla technologies and products."[6]

The Mozilla Foundation is a California non-profit corporation exempt from Federal income taxation. It is governed by its Board of Directors. In 2004, the association Mozilla Europe was born in France; its mission is "to support the Mozilla Project by developing, deploying and promoting the free Mozilla software, its by-products and related project within Europe."[7]

The Mozilla Foundation defines itself as follows:

Mozilla is

- The producer and provider of the award-winning Firefox web browser and Thunderbird e-mail software.
- A software development tools provider.
- An open source community of developers and testers.
- A partner for the technology industry.
- An advocate for standards on the Net who provides tools for developing standard web content.
- Educators, providing educational materials about Internet technologies and software development.

As of 2008, the Mozilla foundation maintains the following software among others[8]:

---

[4] http://www.mozilla.org
[5] http://www.mozilla.org/mission.html
[6] http://www.mozilla.org/foundation/about.html
[7] http://www.mozilla-europe.org/en/about/
[8] http://www.mozilla.org/projects/

Figure 2: The Mozilla Foundation logo

- Firefox: the well known Internet browser,

- Thunderbird: a standalone e-mail and newsgroup client can be used as a companion to Mozilla Firefox or by itself,

- Bugzilla: a bug tracking Web tool,

- Bonsai: a Web interface to the checkin log that you can query for check-ins between certain dates, of certain files, or by certain developers,

- Tinderbox: a 24/7 build and test Web tool.

Core Mozilla project source code is licensed under a disjunctive tri-license giving the choice of one of the three following sets of open-source licensing terms: Mozilla Public License version 1.1 or later, GNU General Public License version 2.0 or later, GNU Lesser General Public License version 2.1 or later.

Beside providing end user open-source products, Mozilla authored two development frameworks for building cross operating systems components and user interfaces: `XPCOM` and `XUL`, which are presented in the following sections. Mozilla has established a vibrant network of partners in the industry for sustaining the long term development and use of these key technologies[9].

2.2 Mozilla `XPCOM`

As stated on the Mozilla Developer Center pages,

> `XPCOM` is a cross platform component object model, similar to Microsoft `COM`. It has multiple language bindings, letting the `XPCOM` components be used and implemented in `JavaScript`, `Java`, and `Python` in addition to `C++`. Interfaces in `XPCOM` are defined in a dialect of `IDL` called `XPIDL`.

> `XPCOM` itself provides a set of core components and classes, e.g. file and memory management, threads, basic data structures (strings, arrays, variants), etc. The majority of `XPCOM` components is not

---

[9]`http://www.mozilla.org/about/`

part of this core set and is provided by other parts of the platform (e.g. Gecko or Necko) or by an application or even by an extension.[10]

XPCOM supports any platform that hosts a good C++ compiler, including: Microsoft Windows, Linux, HP-UX, AIX, Solaris, OpenVMS, MacOS, BSD. This agility makes it a platform of choice for implementing and disseminating the Nepomuk vision.

2.3 Mozilla XUL

The Mozilla XUL framework is described as follows on the Mozilla development pages:

> XUL (XML User Interface Language) is Mozilla's XML-based language that lets you build feature-rich cross platform applications that can run connected or disconnected from the Internet. These applications are easily customized with alternative text, graphics and layout so they can be readily branded or localized for various markets. Web developers already familiar with Dynamic HTML (DHTML) will learn XUL quickly and can start building applications right away.[11]

The main benefit of XUL is that it provides a simple and portable definition of common widgets. This reduces the software development effort in a way analogous to the savings offered by 4th generation programming languages tools. The article "The Joy of XUL" [7] on developer.mozilla.org helps to better understand XUL and why the Mozilla platform is built with it.

In order to foster the use and development of the XUL technologies, Mozilla released in 2006 the tool XULRunner: XULRunner is a runtime environment developed by the Mozilla Foundation for providing a common back-end for XUL applications as a shared run-time environment. As stated on the Mozilla developer pages,

> XULRunner can be used to bootstrap XUL and XPCOM applications that are as rich as Firefox and Thunderbird. It provides mechanisms for installing, upgrading, and uninstalling these applications. XULRunner is to provide libxul, a solution which allows the embedding of Mozilla technologies in other projects and products.[12]

In that sense, XULRunner shares common concerns with the Eclipse Rich Client Platform [1].

2.4 Mozilla projects related to social or semantic technologies

Several existing projects building upon the Mozilla platform are addressing the need for an enhanced semantic and social workbench.

2.4.1 Operator

Operator is a Microformat [4] detection extension for Firefox 2 and later. It was developed by Michael Kaply at IBM and is officialy endorsed by the Mozilla

---

[10]http://developer.mozilla.org/en/docs/XPCOM
[11]http://developer.mozilla.org/en/docs/XUL
[12]http://developer.mozilla.org/en/docs/XULRunner

Labs as stated on a Mozilla Labs page introducing the tool (December 2006)[13]. Taking advantage of Microformats, Operator typically makes it possible to add an event information from Upcoming.org to a Calendar or to map the location of the event without typing any data. As stated on the Mozilla Labs pages, "Operator requires information on the Web to be encoded using microformats, and since this method for semantically encoding information is relatively new, not all sites are using microformats yet. However, Operator works great with any blog that uses rel-tag, and the sites Yahoo! Local, Flickr, and Upcoming.org, all of which contain millions of pieces of information expressed using microformats".

The Operator addon can be downloaded from the addons.mozilla.org Web site[14].

The Microformats libraries developed for Operator and the tool itself will be considered for being integrated into the Mozilla based semantic desktop that is being developed by the Nepomuk team in 2008, with the objective to harness as much as possible the Microformats in the Nepomuk personal semantic Web environment. Detailed use cases on how Microformats will be put into practice in the Nepomuk platform will be produced in the course of the Nepomuk-Mozilla development, in 2008.

2.4.2   Firefox 3 tagging features

Firefox 3[15] is the next major release of the Firefox browser. The official release is expected to be out in 2008 (beta versions are already available in January 2008). Among other innovative features, Firefox 3 brings the ability to annotate a URI with tags, similarlingly to what's possible on social bookmarking sites such as Del.icio.us[16] or Bibsonomy[17].

More information on these upcoming features is available from the Firefox Bugzilla Web site. The capabilities planned comprise the following:

- Provide some user interface that allows users to annotate a URI with tags[18],

- Add full text index to bookmark and/or history search[19],

- Allows users to search on all available URI metadata,[20]

- Provide some user interface that allows users to annotate a URI with free-form notes[21].

These capabilities share many concern with the ones meant for the Nepomuk platform. For this reason, part of the work envisaged in the Nepomuk-Mozilla prototype will consist of providing the features above directly from the Nepomuk infrastructure. The way the integration between the Firefox XPCOM components and the Nepomuk hub is meant to be achieved is described in section 4. The objective is to make the native Nepomuk metadata store interoperable with the Firefox store, so that the user can run queries against metadata managed either by Firefox or by other applications.

---

[13]http://labs.mozilla.com/2006/12/introducing-operator/
[14]https://addons.mozilla.org/fr/firefox/addon/4106
[15]http://wiki.mozilla.org/Firefox3
[16]http://del.icio.us
[17]http://www.bibsonomy.org/
[18]https://bugzilla.mozilla.org/show_bug.cgi?id=374524
[19]https://bugzilla.mozilla.org/show_bug.cgi?id=342913
[20]https://bugzilla.mozilla.org/show_bug.cgi?id=374530
[21]https://bugzilla.mozilla.org/show_bug.cgi?id=374525

2.4.3   Firefox 3 desktop integration

Firefox 3 introduces deeper desktop integration through the availability of Web-based protocol handlers, described as follows by the Mozilla Foundation:

> Web-based protocol handlers: Web applications, such as your favorite webmail provider, can now be used instead of desktop applications for handling mailto: links from other sites. Similar support is available for other protocols (Web applications will have to first enable this by registering as handlers with Firefox).[22]

Those Web-based protocol handlers will contribute to blur the frontier of the users' personal semantic Web across her desktop and her Web applications. In the frame of Nepomuk, this Firefox integration in the desktop will help creating a personal semantic Web encompassing metadata that is managed both by local or remote applications.

2.4.4   Prism

Prism[23] is another technology currently worked on by the Mozilla Labs team. It is described by its authors as follows:

> Personal computing is currently in a state of transition. While traditionally users have interacted mostly with desktop applications, more and more of them are using Web applications. But the latter often fit awkwardly into the document-centric interface of web browsers. And they are surrounded with controls-like back and forward buttons and a location bar-that have nothing to do with interacting with the application itself.
>
> Mozilla Labs is launching a series of experiments to bridge the divide in the user experience between web applications and desktop apps and to explore new usability models as the line between traditional desktop and new Web applications continues to blur.
>
> Unlike Adobe AIR and Microsoft Silverlight, we're not building a proprietary platform to replace the web. We think the web is a powerful and open platform for this sort of innovation, so our goal is to identify and facilitate the development of enhancements that bring the advantages of desktop apps to the Web platform.
>
> The first of these experiments is based on Webrunner, which we've moved into the Mozilla Labs code repository and renamed to Prism.

As illustrated by figure 3, Prism bridges the user's desktop and the Web. Since the Nepomuk social semantic desktop vision encompasses data that is stored both on the local computer and on the Web, the Prism framework will be considered in 2008 for integrating further the Mozilla version of a Nepomuk integrated personal semantic Web with the Web.

The Prism vision shares some aspects with the PyroDesktop[24], "a new kind of desktop environment for Linux built on Mozilla Firefox [aiming at] true integration between the Web and modern desktop computing".

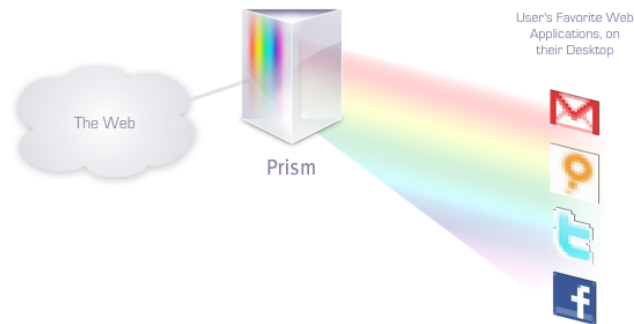[22]http://www.mozilla.com/en-US/firefox/3.0b2/releasenotes/
[23]http://labs.mozilla.com/2007/10/prism/
[24]http://pyrodesktop.org/

Figure 3: Mozilla Prism

2.4.5   Zotero

Zotero[25] is a Firefox extension for collecting and managing research sources. It is said by Zotero team to include "the best parts of older reference manager software (like EndNote) – the ability to store author, title, and publication fields and to export that information as formatted references – and the best parts of modern software and web applications (like iTunes and del.icio.us), such as the ability to interact, tag, and search in advanced ways. [...] Since it lives in the Web browser, it can effortlessly transmit information to, and receive information from, other web services and applications; since it runs on one's personal computer, it can also communicate with software running there (such as Microsoft Word). And it can be used offline as well."

Zotero is a production of the Center for History and New Media at George Mason University. It is funded by the United States Institute of Museum and Library Services, the Andrew W. Mellon Foundation, and the Alfred P. Sloan Foundation.

Zotero features are described as follows on the project's Web page:

- Automatic capture of citation information from Web pages,
- Storage of PDFs, files, images, links, and whole Web pages,
- Flexible notetaking with autosave,
- Fast, as-you-type search through your materials,
- Playlist-like library organization, including saved searches (smart collections) and tags,
- Platform for new forms of digital research that can be extended with other Web tools and services,
- Formatted citation export (style list to grow rapidly),
- Integration with Microsoft Word and OpenOffice.org,
- Offline work capabilities,
- Annotation features (sticky notes), with the possibility to generate reports in various formats from the entered annotations.

Since 2007, Zotero also consists of a server that makes it possible to share collections with other users, including annotations such as recommendations. The server comes with a set of APIs for other services to interact with Zotero.

By providing both semantic and social capabilities from the Firefox browser, Zotero is one of the most advanced tool to date to our knowledge in the

---
[25]http://www.zotero.org

Mozilla realm that realizes partly the Nepomuk social semantic Web vision. Collaboration with the Zotero development team will be considered in 2008 for sharing ideas, APIs and possibly code. Zotero has reached a very impressive level of maturity and a large number of astonishing features.

Figure 4 illustrates how the user can annotate a remote Web page using the Zotero user interface. The tabs titled "Tags" and "Related" let the user draw relations between the current resource and others managed by the Zotero database.
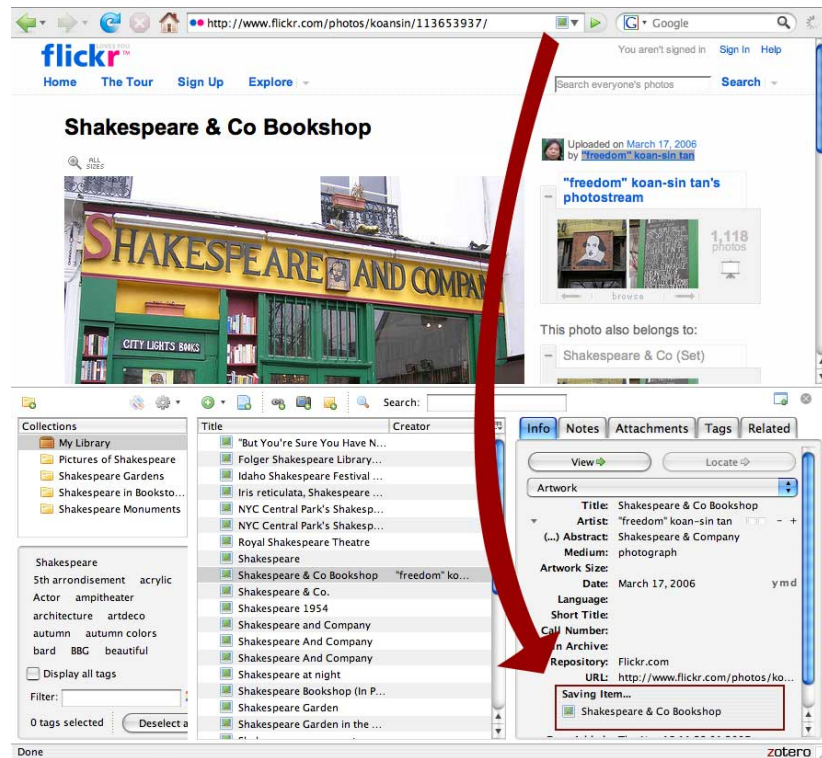

Figure 4: Zotero annotation UI

The Nepomuk-Mozilla prototype is expected to differ from the Zotero 2007 platform in various respects, but it seems clear that the two projects could cross-fertilize fruitfuly and share several services or APIs. The semi-automatic extraction of metadata from documents in Zotero could for instance rely on the Nepomuk data extractors services, building upon either Strigi[26] extractors (`KDE4`) or on the Aperture framework[27] (Java enabled platforms). Nepomuk-Mozilla will come with built-in P2P capabilities which will make it possible for users to share part of their semantic database with others in a distributed manner; the annotation ontology in Nepomuk-Mozilla will rely on the Nepomuk standards, already adopted by two other implementations, i.e. Nepomuk `KDE4` and Nepomuk Eclipse, and which shares many similarities with the Xesam[28] ontology that is being instrumented as well.

2.4.6   ClearForest Gnosis

ClearForest Gnosis[29] is another Firefox extension, described as "the cutting edge of real time semantic processing for the Web. By evaluating the pages

---

[26] http://en.wikipedia.org/wiki/Strigi_(KDE)
[27] http://aperture.sourceforge.net
[28] http://xesam.org/
[29] http://sws.clearforest.com/Blog/?page_id=32/

you read – as you read them – Gnosis immediately locates key information such as people, organizations, companies, products and geographies hidden within the text."

Gnosis automatically processes pages from various data sources such as CNN, Google Finance, Wikipedia, the New York Times, and others. Figure 5 illustrates the automatic classification of terms in a Web page using Gnosis.
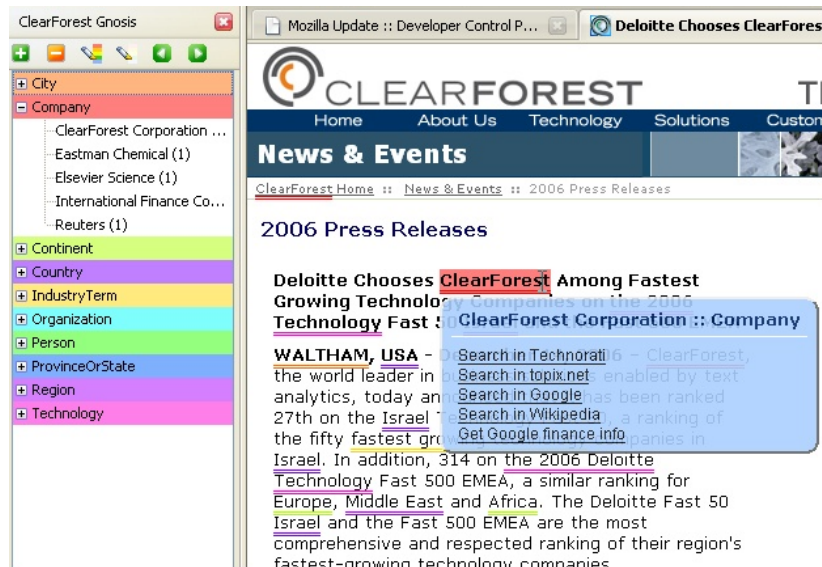

Figure 5: ClearForest Gnosis

Technically, the addon sends the text browsed to a Web service managed by ClearForest, whose `WSDL` is available publicly[30]. The Web service returns a tagged `XML` file containing semantically enhanced information, as described on the technical description of the service[31]. This Web service is part of the "ClearForest Semantic Web Services (SWS)" which is "built on top of a subset of ClearForest's natural language processing toolkit"[32].

The use case put forth by the ClearForest Gnosis plugin is typically a use case envisaged by the Nepomuk platform. Text analysis capabilities including intelligent, personalized, semantic text annotation and disambiguation are being developed in Nepomuk within WP1 and WP2 [8] [10] [3].

2.4.7   TiddlySnip

TiddlySnip is a Firefox extension for the popular TiddlyWiki[33] wiki system.

TiddlyWiki is described as A reusable non-linear personal Web notebook. More concretely, it is described as follows on the project's Web site:

> TiddlyWiki is a free wiki engine created by Jeremy Ruston and a community of independent developers. It's written in `HTML`, `CSS` and `JavaScript` to run on any modern browser without needing any server side logic. It allows anyone to create personal self contained hypertext documents that can be posted to a Web server, sent by email or kept on a USB thumb drive to make a wiki on a stick.

---

[30]`http://sws.clearforest.com/ws/sws.asmx?wsdl`
[31]`http://sws.clearforest.com/Blog/wp-content/CLFDocs/ClearForest%20SWS1%20API%20V01.rtf`
[32]`http://sws.clearforest.com/Blog/?p=13`
[33]`http://www.tiddlywiki.com/`

TiddlySnip integrates Firefox with TiddlyWiki by letting users simply select text, right click and choose TiddlySnip selection for adding tagged snippets to the wiki engine in a few clicks.

The Nepomuk-Mozilla prototype will draw inspiration from the TiddlySnip plugin for bringing the capability to enhance semantically Web pages so easily to the integrated wiki-like Nepomuk note taking system.

2.4.8   AllPeers

AllPeers[34] is a Mozilla application for sharing large files from a browser using P2P technology. It is an open-source application, tri-licensed under the Mozilla Public License, the GNU General Public License, and the GNU Lesser General Public License.

As stated in the AllPeers hacker's getting started guide[35],

> AllPeers is divided into a number of high-level modules [. . .] making heavy use of `XPCOM` to encapsulate functionality into components and use these components across modules. Most of the code is written in `C++`, with some portions (especially those relating to the user interface) in `JavaScript`. The major high-level modules are: core functionality, resource framework, low-level networking, resource-level networking, application-level networking, data transfer, user interface.

Since the AllPeers platform is available under an open-source license, and comes with an advanced framework for dealing with files on a P2P network, an integration of the Nepomuk social layer with AllPeers will be considered in the course of 2008.

2.4.9   Other tools of interest

This section presents briefly a set of tools that may be of interest while designing and developing further the Nepomuk-Mozilla prototype. The list below is incomplete, and it is likely that several other tools will be considered in 2008.

Scrapbook
Scrapbook[36] is another note taking extension for Firefox, which is much less advanced than Zotero to date though.

Annozilla
The Annozilla project[37] is "designed to view and create annotations associated with a Web page, as defined by the W3C Annotea project[38]. The idea is to store annotations as `RDF` on a server, using XPointer (or at least XPointer-like constructs) to identify the region of the document being annotated.". As of January 2008, Matthew Wilson, head developer of the project, is porting the tool to the Firefox 3 platform, as witnessed by one of his postings to the Annonzilla mailing-list[39].

The Coop
The Coop[40] is a Firefox addon in development at the Mozilla Labs, that will "let users keep track of what their friends are doing online, and share new and interesting content with one or more of those friends. It will integrate with popular web services, using their existing data feeds as a transport mechanism. [. . .] Perhaps the most common social interaction on the Web today is

---

[34]`http://www.allpeers.com`
[35]`http://developer.allpeers.com/src:hack:starting`
[36]`http://amb.vis.ne.jp/mozilla/scrapbook/`
[37]`http://annozilla.mozdev.org/`
[38]`http://www.w3.org/2001/Annotea/`
[39]`http://mozdev.org/pipermail/annozilla/2008-January/000170.html`
[40]`http://wiki.mozilla.org/Labs/The_Coop`

sending someone a link. It's done over IM, email, weblogs, RSS feeds from ag-
gregator sites, bookmark sharing sites like del.icio.us, social networking sites
like Facebook or MySpace, and even over the phone. The desire is the same:
"hey, friend, go check out this neat thing and then let's talk about it!"". The
goal of The Coop is to ease this interaction, and merge it with similar tools
provided by a large number of popular Web services.

Outfoxed
Outfoxed is the implementation side of the master's thesis by Stan James at
the University of Osnabrück, Germany. The thesis title is on Trusted Metadata
Distribution Using Social Networks.

> In a nutshell, [S.James] is exploring ways for you to use your net-
> work of trusted friends to determine what's good, bad, and danger-
> ous on the internet. Outfoxed does this by adding functionality to
> the Firefox web browser. Coding began on Dec 27th, 2004. Many
> things found on the internet are low-quality, false, or dangerous.
> Web surfers are often asked to make decisions of trust without
> any background information. Meta-information about what is good
> or trustworthy is traditionally spread by word of mouth or proac-
> tive inquiry, but the vast quantity of resources encountered online
> demands a faster means of distribution. This paper describes a
> method for using a social network of defined trust relations for in-
> dividualized collection of metadata. Additionally, the applications
> of this method are demonstrated in a software prototype which
> reorders search results based on metadata, and presents trusted
> evaluations of web pages, downloaded files, and running processes
> as they are encountered. Preliminary results from test users is pre-
> sented. Finally, it is argued that the chains of trust inherent in
> social networks should be applied at lower levels of computation,
> even at the level of the operating system.

As of January 2008, Outfoxed is part of the Lijit company[41].

Piggy Bank
Piggy Bank[42] is a Firefox extension that turns the browser into a mashup plat-
form, by allowing you to extract data from different web sites and mix them
together. Piggy Bank also allows you to store this extracted information locally
for you to search later and to exchange at need the collected information with
others.

---

[41] http://www.lijit.com
[42] http://simile.mit.edu/wiki/Piggy_Bank

3   Mozilla community involvement

Beside defining a roadmap for implementing a Mozilla demonstrator show-casing an implementation of the Nepomuk services, one of the main tasks of the Nepomuk-Mozilla project is to involve Mozilla developers into the Nepomuk project. This section gives an overview of the Mozilla community activities in the Nepomuk-Mozilla project.

The strategy has been similar to the one followed in the two other main communities targeted by the Nepomuk project, i.e. KDE and Eclipse: a communication platform has been set up, face-to-face meetings have been organized with key members of the Mozilla community, and steps toward fruitful collaboration with other activities have been undertaken, as described in the following paragraphs.

3.1   The Nepomuk-Mozilla Web portal

A Web site has been created at the following location, with the goal to present the project's objectives and to keep the community abreast of the latest developments: `http://nepomuk-mozilla.semanticdesktop.org`.

Figure 6 shows the home page of the Nepomuk-Mozilla Web site.

Figure 6: Nepomuk-Mozilla Web site

The Web site comprises a wiki where registered members can contribute ideas and documents. The wiki is open to the public, inviting the community to take part in the initiative.

In addition to the Nepomuk-Mozilla Web site, a public mailing-list has been set up. The mailing-list is located at the following URL: `http://lists.semanticdesktop.org/mailman/listinfo/nepomuk-mozilla`. Anyone can subscribe to the list either using the Web interface of Mailman. A few persons beside the Nepomuk internal participants have subscribed to the list. It is expected that the list audience will increase significantly in 2008 through a wider promotion of the project.

3.2   Mozilla-Europe head interview

From the organizational point of view, Mozilla consists of a corporation – Mozilla Corporation – and of two main associations – the Mozilla foundation and the Mozilla-Europe association. Mozilla-Europe offices are located in the same building as Edge-IT's, and the two companies have good relations. This proximity made it easy to present the Nepomuk-Mozilla project to the head of the association, Tristan Nitot, on the occasion of a discussion that took place in December 2007, and to discuss with him about potential synergies between Mozilla and the Nepomuk-Mozilla initiative.

Tristan Nitot pointed out existing projects within the Mozilla ecosystem that would be of interest to Nepomuk-Mozilla in terms of interaction with a Mozilla based semantic desktop: Operator, Prism and Firefox 3 Web-based protocol handlers, which are described in the state of the art section 2.

Following that discussion, the Nepomuk-Mozilla team will get in touch with the Mozilla Labs people to investigate possible co-developments around Operator in particular.

3.3   FOSDEM 2008 – Mozilla room

The Free and Open-source Software Developers' European Meeting (FOS-DEM)[43] is taking place annually in Brussels, Belgium.

The FOSDEM was born in 2000 as Raphael Bauduin, a fan of the Linux movement in Belgium, decided to organize a small meeting for developers of Open Source software, which he called "Open Source Developers' European Meeting" (OSDEM). Many years later, OSDEM has grown to the event today knows as FOSDEM. The meeting gathers around 4000 visitors at the University Libre of Brussels' campus every year during the last week-end of February.

As several other large scale open-source projects, Mozilla has a dedicated developer room[44] at the FOSDEM 2008, like the KDE, GNOME, X.org, OpenOffice.org projects and others[45].

The Nepomuk-Mozilla team will take part in the Mozilla developer discussions on that occasion, and will submit a talk proposal to the Mozilla FOSDEM 2008 organizing committee about the Nepomuk-Mozilla plans.

---

[43] http://www.fosdem.org
[44] http://www.fosdem.org/2008/schedule/devroom/mozilla
[45] http://www.fosdem.org/2008/schedule/devrooms

4   Nepomuk-Mozilla prototype: a roadmap

4.1   Introduction

To make the Nepomuk ideas graspable in the context of the Mozilla applications, a prototype is needed. This section presents a first set of specifications of the target prototype and a preliminary architecture. Based on this prototype, the related communities will be able to see how to embed and access Nepomuk in their environment.

4.2   Objectives

The objective of the prototype is to create a universal metadata management workbench with social capabilities on top of the Mozilla `XUL XPCOM` technologies. This workbench will allow to annotate any and all Web data (including parts of Web pages) or local files from Firefox, emails from Thunderbird and to store the annotations in the Nepomuk metadata store. The advantage of using a common metadata store is to share annotations and ontologies across the applications instead of having them buried into each application information system. This paves the way for a "Personal Semantic Web" as targeted by Nepomuk. Once this metadata mangement layer will be available from Firefox and Thunderbird, a social layer will be added for sharing the stored metadata with colleagues and acquaintances in P2P.

The targeted workbench will rely on the Nepomuk core services. Nepomuk will profit from building such a workbench on top of Mozilla technology by disseminating the Nepomuk APIs, ontologies and core services to the wide community of Mozilla developers and users. Conversely, it is expected that the Mozilla community will benefit from the Nepomuk-Mozilla initiative because the Nepomuk approach shows an elegant way to integrate deeper in the desktop infrastructure the burgeoning semantic features that are showing up in several Mozilla tools.

The integration of the Nepomuk framework with the Mozilla technologies will increase the conceptual and operational efficiency of the semantic desktop vision. The Nepomuk Information Element ontology `NIE` [6] for instance already foresees the linking of arbitrary Web pages taken from any browser; a deep instrumentation of `NIE` within Firefox is likely to result in an easy-to-use tool that will help spreading the Nepomuk vision.

4.3   Work plan

The prototype development workplan will consist of two main steps:

- A first step will aim at integrating Firefox and Thunderbird with the Nepomuk communication service. The path toward the implementation of this step is described in the section "`XPCOM` integration with the Nepomuk communication hub" 4.3.1 below.

- A second step, described in the section "Towards a Mozilla social semantic desktop" 4.3.2, consists of harnessing this integration by developping addons for dealing with metadata, and to involve the community of Mozilla developers in the use of this shared desktop metadata store.

Once the metadata backend and interfaces will be available indeed, the sharing layer will be added, possibly in cooperation with existing projects such as

AllPeers[46].

4.3.1 `XPCOM` integration with the Nepomuk communication hub

The first step is to make it possible to use the Nepomuk metadata store (the Nepomuk `RDF` repository component[47]) from the Mozilla products such as Firefox and Thunderbird. This metadata store is available as a native service in the case of Linux operating systems running `KDE4`, and can also be accessed through `HTTP REST` or `HTTP SOAP` calls on operating systems such as Microsoft Windows, Apple, Solaris and Linux. The bridge between the Nepomuk services and `XPCOM` components can hence be realized at least in two ways: either through native dbus calls, or over an `HTTP` bus.

On Linux operating systems running `KDE4`, the Nepomuk `RDF` repository service is directly provided by `KDE4`. It gives access to `libnepomuk`, a library for managing metadata from various applications. The metadata is then stored in Soprano [13], which is described as follows on the Soprano Web site[48]:

> Soprano is an open and pluggable `RDF` resource framework which is build on top of `QT4`. It provides `RDF` storage, `RDF` parsing, serialization, inference, and full text indexing in a `C++` API.

A documentation of the Nepomuk metadata is available from the `KDE4` documentation Web site[49].

The integration of Firefox and Thunderbird with the Nepomuk services consists in issuing calls to the Nepomuk services from Firefox/Thunderbird either over dbus or over `HTTP`. A first step has consisted in gaining insight on whether a dbus/`XPCOM` bridge was already available. From the research carried out by January 2008, it seems there isn't such a bridge yet. The detailed specifications of the bridge are being worked on. In parallel, the idea of a `REST` bus system is being specified and investigated, since it brings the advantage of simplicity and operating systems interoperability. In a 2005 blog post that has been much referenced on the Web since then, Miguel de Icaza, creator of the GNOME[50] and Mono[51] projects, points out that such a bus system would bring serveral advantages indeed [9]:

- `HTTP` is a well known protocol.
- There are plenty of `HTTP` client and server implementations that can be employed.
- The protocol can be as simple or as complex as needed by the applications. From passing all the arguments on a `REST` header as arguments to the tool-aware `SOAP`.
- `HTTP` works well with large volumes of data and large numbers of clients.
- Scaling `HTTP` is a well understood problem.
- Users can start with a `REST` API, but can easily move to SOAP if they need to.
- `HTTP` includes format negotiations: clients can request a number of different formats and the server can send the best possible match, it is already part of the protocol. This colors the request with a second dimension if they choose to.

---

[46]http://www.allpeers.com
[47]http://dev.nepomuk.semanticdesktop.org/wiki/RdfRepository
[48]http//soprano.sourceforge.net
[49]http://api.kde.org/4.0-api/kdelibs-apidocs/nepomuk/html/index.html
[50]http://www.gnome.org
[51]http://www.mono-project.com/

- Servers can take advantage of `HTTP` frameworks to implement their functionality on the desktop.

- It is not another dependency on an obscure Linux library.

- The possibility of easily capturing, proxying and caching any requests.

- Redirects could be used to redirect request to remote hosts transparently and securily.

- And finally, makes the desktop "Web 2.0 ready" :-) [52]

Within the EPOS project [2], the EPOS DFKI team already prototyped an `HTTP` bridge between `XUL` components and an `RDF` repository, using the `XML-RPC` protocol: this is the Dragontalk prototype [12], which provides a set of extensions for the Mozilla products Thunderbird and Firefox that capture the user's activity[53]. The Nepomuk-Mozilla team will gain insight into this existing prototype for building the `REST` bridge between Mozilla products and Nepomuk services.

On the side of an integration of dbus with `XPCOM` components, the two following links indicate that such an integration is currently a matter of interest to several communities:

- Message discussing an `XPCOM` and dbus integration on the freedesktop-d-bus list[54],

- Wiki page describing paths towards a dbus and Mozilla integration[55].

4.3.2   Towards a Mozilla social semantic desktop

Once the integration of Firefox and Thunderbird with dbus will be achieved, a set of addons for Firefox and Thunderbird will be developed: they will harness the Nepomuk `RDF` repository service, then the Nepomuk social capabilities. Below is a proposal of a set of addons building on this integration. All the proposals below are meant to be developed in close collaboration with Mozilla team whenever possible.

- Develop a `XUL` user interface for associating metadata to Web pages or selections of pages from Firefox (the existing user interfaces of Firefox 3 for associating metadata with `URI`s will be considered, and reused when possible). The metadata should then be stored in the `KDE4` metadata store. A typical use case for such an addon can be summarized as follows:

  ○ user right clicks on a Web page, this fires a contextual menu pop-up,

  ○ among other actions, the menu contains the following one: "Tag this page",

  ○ if the user clicks on "Tag this page", a wizard pops up, displaying existing tags for that page if there are some already linked with that page, and letting the user check other tags, which are retrieved from the metadata database as items having the type "Tag" in the Nepomuk Annotation Ontology, or associate a type with the page – available types are the types having the type `rdf:type` in the

---

[52]http://tirania.org/blog/archive/2005/Nov-26-2.html
[53]http://dragontalk.opendfki.de/
[54]http://marc.info/?l=freedesktop-dbus&m=117323275132196&w=2
[55]http://zenit.senecac.on.ca/wiki/index.php/D-Bus_and_other_Linux_desktop_integration_improvements

metadata database. Such types can be for instance Event, City, etc. − or associate properties with the page, such as the ones from the `NAO`[56] and `PIMO` ontologies.

- Set up the same type of user interface for Thunderbird messages. Metadata has to be updated when an email is moved from one folder to another, or deleted.

- Design and implement a Thunderbird extension for browsing emails by types or properties.

- Implement a Firefox extension for browsing Web pages by types or properties.

- Implement an integrated `XUL` user interface for browsing Web pages, emails and local files by types or properties, and for annotating them.

- Develop export/import and synchronization scripts for several Firefox applications dealing with metadata such as the ones mentionned in the state of the art section of this report 2 (Zotero, Scrapbook, ClearForest Gnosis, Firefox 3 tag manager, TiddlySnip, Annozilla and others) so that the user can import existing metadata into the common `KDE4` metadata store. While developing those synchronization mecanisms, the authors of the addons will be contacted for discussing a possible native storage of the addon metadata in the native operating system metadata store (`KDE4` Soprano in the case of `KDE4`, Java implementation of the Nepomuk `RDF` repository in the case of other operating systems).

- Add a social layer for sharing metadata and files with friends and colleagues. A study on how this can be achieved in details will be conducted. This involves gaining insight in the AllPeers project and other similar data sharing projects.

- Make it possible to store a full text index of all browsed pages contents in the Nepomuk metadata store.

- Develop a Firefox addon for extracting structured information from Web pages containing information elements expressed using the Microformats [4] standard. This addon may build upon existing tools such as Operator [11] mentionned in section 2.4.1.

- Implement a service for letting the Nepomuk metadata store access the Firefox 3 store and conversely: the Firefox 3 store is used in particular for managing the users' annotations, as described in section 2.4.2.

Once all these steps will be realized, it is considered that a first version of the Nepomuk-Mozilla social semantic desktop will be born.

4.4   Time schedule

The table 1 below proposes a roadmap for the releases of the various key components toward an implementation of a Nepomuk-Mozilla social semantic desktop.

---

[56]`http://www.semanticdesktop.org/ontologies/nao/`

Planned date	Task
31/1/2008	Detailed technical description on how Firefox and Thunderbird will be integrated with the Nepomuk dbus server on the one hand, and with the Nepomuk `HTTP` services on the other hand.
31/3/2008	First prototype of a Firefox and Thunderbird integration with the Nepomuk dbus server.
31/3/2008	First `XUL` user interfaces for metadata additions from Firefox and from Thunderbird (enhancement of what's possible in terms of tagging in Firefox 3).
15/4/2008	First version of an integrated `XUL` user interface for dealing with metadata on the desktop across files, emails, Web pages, selection of Web pages.
31/5/2008	Have several external application developers use the `XPCOM` metadata dbus service integration.
30/6/2008	- Second version of the `XUL` user interface for metadata, - Specification of the social layer features and of the social layer architecture.
30/9/2008	- Social layer implemented: share metadata and associated files between desktops, - Third version of the communication bridge and of the user interface, - Third version of the `XUL` user interface for metadata, - Have many other external addons use the metadata service.

Table 1: Roadmap toward a Mozilla social semantic desktop

5   Conclusion

During the first two years of the project, the involvement of the Mozilla community into the Nepomuk project has followed an approach similar to the Nepomuk-KDE and Nepomuk-Eclipse initiatives. The analysis of the existing community has led to a better understanding of the ongoing projects, the Nepomuk-Mozilla community platform has started raising awareness on the topic, and a roadmap has been defined for releasing a set of prototypes paving the way for a Mozilla implementation of the Nepomuk social semantic desktop.

In parallel to the development of the Nepomuk-Mozilla libraries and user interfaces, the focus will be brought in 2008 on the involvement of Mozilla developers into the project, by presenting the project to a wider audience on the occasion of Mozilla related events, and by increasing the presence of the project on the Web.

References

[1] Eclipse Rich Client Platform. `http://www.eclipse.org/rcp/`.

[2] EPOS - Evolving Personal to Organizational Knowledge Spaces. `http://www3.dfki.uni-kl.de/epos/`.

[3] IBM Languageware Miner for Multidimensional Socio-Semantic Networks. `http://www.alphaworks.ibm.com/tech/galaxy`.

[4] Microformats. `http://www.microformats.org`.

[5] Mozilla `XPCOM`. `http://www.mozilla.org/projects/xpcom/`.

[6] Antoni Mylka, Leo Sauermann, Michael Sintek, and Ludger van Elst (DFKI). Nepomuk Information Element Ontology. `http://www.semanticdesktop.org/ontologies/2007/01/19/nie/`.

[7] Peter Bojanic. The Joy of `XUL`. `http://developer.mozilla.org/en/docs/The_Joy_of_XUL`.

[8] Mikhail Kotelnikov, Alexander Polonsky, Malte Kiesel, Max Volkel, Heiko Haller, Mikhail Sogrin, Par Lannero, and Brian Davis. Nepomuk deliverable d1.1: Interactive semantic wikis. Technical report, Cognium Systems, DFKI, FZI, IBM, KTH, NUIG, 2006.

[9] Miguel de Icaza. Replacing `D-BUS` and `CORBA` with a simple `HTTP` framework. `http://tirania.org/blog/archive/2005/Nov-26-2.html`.

[10] Enrico Minack and Leo Sauermann. Nepomuk deliverable d2.1 - adapters, extractors, and knowledge structure services. Technical report, L3S, DFKI, 2006.

[11] Mozilla Labs. Operator. `http://labs.mozilla.com/2006/12/introducing-operator/`.

[12] Sven Schwarz, Moritz Ploessl, and Emmanuel Gasne (DFKI). Dragontalk. `http://dragontalk.opendfki.de/`.

[13] Sebastian Trüg. Soprano. `http://soprano.sourceforge.net`.