

OSBR.CA

The Open Source Business Resource

Editorial

Dru Lavigne

Data Access in Canada:

CivicAccess.ca

Tracey P. Lauriault & Hugh McGuire

How is Copyright Relevant to Source Data and Source Code?

Joseph Potvin

Implementing Open Data: The Open Data Commons Project

Jordan Hatcher

The Personal Research Portal

Ismael Peña-López

Call for Proposals

Recent Reports

Upcoming Events

Contribute

FEBRUARY
2008



FEBRUARY 2008

PUBLISHER:

The Open Source Business Resource is a monthly publication of the Talent First Network. Archives are available at the website: <http://www.osbr.ca>

EDITOR:

Dru Lavigne
dru@osbr.ca

ISSN:

1913-6102

ADVERTISING:

Rowland Few
rowland@osbr.ca

GRAPHICS:

Ryan May

ADVISORY BOARD:

Tony Bailetti
John Callahan
Kevin Goheen
Peter Hoddinott
Thomas Kunz
Steven Muegge
Trevor Pearce
Stoyan Tanev
Michael Weiss

© 2008 Talent First Network

Editorial

Dru Lavigne on open data.

3

Data Access in Canada: CivicAccess.ca

Tracey P. Lauriault and Hugh McGuire from Citizens for Open Access to Civic information and Data describe the factors that limit access to public data.

4

How is Copyright Relevant to Source Data and Source Code?

Joseph Potvin discusses aspects of copyright law which impact software and database professionals.

11

Implementing Open Data: The Open Data Commons Project

Jordan Hatcher introduces legal tools which can be applied to the open content movement.

17

The Personal Research Portal

Ismael Peña-López from the Open University of Catalonia proposes the Personal Research Portal as a means to create a digital identity for knowledge workers.

23

Call for Proposals

28

Recent Reports

30

Upcoming Events

31

Contribute

32



*"So imagine a mirror
Bigger than the room it was placed in
Imagine my wish for a future
that cannot hold my wish
Imagine the want to hold a rod
that cannot hold the fish"
Imagine a rod that cannot hold the fish"*
Paul Heaton, The Beautiful South

Digital technologies have forever changed the way that knowledge is disseminated and accessed. Yet, the main problem knowledge workers face is invisibility: if people don't know that you know, and people are not aware of what you know, you do not exist.

Governments and institutions are being pushed to foster Open Access (OA) literature as a way to achieve universal reach of research diffusion at inexpensive and immediate levels. Most efforts have been made at the institutional level, dedicating little energy to what the individual can do to contribute. The philosophy and tools around web 2.0 bring clear opportunities for individuals to contribute and to build a broader personal presence on the Internet and a better diffusion for their work, interests or publications.

We propose the concept of the personal research portal (PRP) as a means to create a digital identity for knowledge workers – tied to one's digital public notebook and personal repository – and a virtual network of colleagues working in the same field. Complementary to formal publishing or taking part in offline meetings, the PRP would be a knowledge management system that would enhance reading, storing, and creation at both the private and public levels, and contribute to create an online identity that, in turn, will help to create a network whose currency is knowledge.

Beyond the e-Portfolio

The approach we present is closely related to the concept of e-portfolio (<http://en.wikipedia.org/wiki/E-portfolio>), but from the researcher's or practitioner's point of view rather than the learner's. As e-portfolios are usually associated with students and teaching, we here propose the term personal research portal (PRP) to avoid confusion.

The goals of the PRP should be: i) to gain more access to international knowledge output from other knowledge workers; ii) to give more international access to research generated by a knowledge worker; iii) higher promotion of institutional knowledge output; iv) improved citation and research impact; v) improved access to subsidiary data; and vi) a strongly facilitated peer review.

To do so, the PRP should be a low cost, highly flexible virtual space in order to:

- host a public repository for interlinked personal production which includes past and present (work in progress) information and documentation
- gather digital resources, news, general information and materials which are accessible from each and every computer
- self-archive and self-publish ongoing research while avoiding waits and delays
- increase one's visibility while enabling networking and knowledge sharing

All in all, the PRP should track the read-think-write routine performed by practitioners and scientists involved in research.

The big difference from publishing is that the PRP should not only keep record of stock knowledge, formal knowledge that lasts or should last, but also flow knowledge, non-structured knowledge that is not intended to remain permanent because it is devoted to foster exchange.

While many PRPs could potentially be built from a mesh of different applications, we here propose a prototype built from these components:

- a static web site with personal and professional information drawing the researcher profile
- a blog for recording news, reflections, and flow knowledge arising from readings, research results and hypotheses
- a blogroll to provide both a live reader for the researcher and a live bibliography of bookmarks for the community
- a wiki to store stock knowledge which evolves over time with the collaboration of third parties
- a bibliographic manager with online access to all or most records
- a personal repository to self-archive published papers as well as self-publish preprints, working papers, presentations, and syllabuses
- social bookmarking tools and file stores for image, sound, and video
- RSS feeds for every dynamic page

Social Software

In the last few years, new and easy to use web tools became available which provide interconnectivity allowing for communication and collaboration, with the only pre-requisite being a personal computer connected to a network.

These technologies represent inexpensive, highly accessible means allowing anyone to share their knowledge with their peers. At least three immediate reflections arise:

- they provide a way for people to easily share, publicize, and diffuse their findings as well as who is behind them
- information published in this manner is easily available from anywhere
- the more everyone engages in a community, the richer it becomes

A possible barrier for using such applications is capacity building in users. Web 2.0 applications and social software are designed for non-technical users, thus only a low level of digital literacy is required. From our point of view, the major requirements to enter the conversation are some degree of e-awareness, a minimum of digital culture, and the ability to change.

To overcome this last barrier, we believe that the network of peers itself, boosted by social software, can help stewarding technology. Wikis, forums, blogs, and other tools provide perfect companions to take the newcomer by the hand in his way into web 2.0. Of course, the conversation is also affected by different cultural backgrounds and different mother tongues, but this is not exclusive from online interaction and, moreover, local communities can form without the mediation of formal literature.

A PRP Prototype

Here we suggest a prototype of a PRP in order to provide a background image of reference. Our philosophy in building is not coding from scratch, but combining existing tools.

We also think that mastering some of these tools will soon become basic literacy skills, much as typewriting, writing an essay, or imparting a live presentation.

Our first consideration would be obtaining one's own domain and hosting. The former, because a domain name is automatically associated with a specific content and its managers. The latter, to retain autonomy of the services, shape, and content on the site.

Static pages and most of the dynamic ones can be built using an open source content management system (CMS) such as Drupal (<http://drupal.org/>) or Joomla (<http://www.joomla.org/>). WordPress (<http://wordpress.org/>) is a blog engine that can also be used as a CMS. Examples of e-portfolio applications include Elgg (<http://elgg.org/>) and OSPI (<http://www.theospi.org/>).

With regards to collaboration: if the expected output is content, a wiki is the best option. If the goal is the process and the debate itself, then forums are required. Some of the preceding applications include wikis and message boards. Mediawiki (<http://mediawiki.org/>) for the wiki, and phpBB (<http://phpbb.com>) for the message board, are also good choices.

Concerning bibliographic managers, Refbase (<http://refbase.sourceforge.net/>) and BibCiter (<http://bibciter.sourceforge.net/>) are web based and provide RSS output. EPrints (<http://www.eprints.org/>) and Open Journal Systems (<http://pkp.sfu.ca/ojs/>) provide for self-archiving and self-publishing, respectively.

There are many other applications to share bookmarks, photos and slideshows, to publish podcasts or vodcasts.

However, most are online services which are provided and hosted by third parties. Their use should be based on the availability to import and export one's data and should be properly linked on the PRP. RSS output, the glue, is a must.

When connectivity is not available and working locally should be made possible, XAMPP (<http://sourceforge.net/projects/xampp/>) provides the ability to (re)install all the social software applications to the local hard drive or a USB pen drive. Indeed, it can work as a backup for the PRP as well as make it portable across different operating systems.

Digital Identity

In an age of information overload, one of the main problems that knowledge workers face is invisibility. This invisibility causes, at least, two major consequences:

- minimum awareness and recognition of one's findings, fields of work, interests and even existence
- difficult access to mainstream circuits in one's field

It is thus important that researchers gain visibility so that they and their work become known in academic and practitioner circles at the international level.

Setting up a PRP should hence be understood, at a primary level, as the creation of a personal home page. Notwithstanding, this digital identity, or the researcher's presence on the Internet, is juxtaposed to the identity shown by authorship in paper journals and conference speeches, with each identity complementing the other. While the latter identity is strongly tied to a handful of concepts exposed in a determinate paper, the digital identity should give further information on the following aspects:

- who and where am I?
- what do I do?
- what interests me?
- what have I done?

If mainstream systems such as congresses, journals, and seminars act as diffusion hubs for offline identities, search engines, portals, blogs, institutional pages, and signature files in e-mails act as diffusion hubs for online identities.

Nevertheless, there are, in our opinion, two main differences among both channels: the higher potential reach of online media and the always updated information provided by PRPs, especially those provided with an RSS feed and correctly meta-tagged data.

Summing up, the main component of a PRP should be evolving, up-to-date information of one's work. Search engines are web 2.0-friendly and highly score live pages with rich and focused content. Descriptions about one's research and interests, side-by-side to documents and other materials and links to and from other people with similar interests enhances the possibility of being found under determinate keywords. This information should be created through static pages by means of simple HTML documents or, better, using a CMS or CMS-like features from other applications such as blogs. The blogroll should play, among others, a great part in the linking role.

Reinforcing Digital Identity

In the process of gathering information to increase one's knowledge and prepare research, it is usual to take notes, highlights of what has been read, reflections that arise after the reading, or just a notation of the fact that something has been read.

Social software empowers researchers as their notes can be published automatically and, even more, link the people and documents that generated these notes. Thus, the digital notebook makes the read, write, analyze, reflect, and learn process fully public. Another immediate consequence is that a live digital store is created daily, a store that is categorized, searchable and fully accessible, with absolute immediacy and no filtering other than your own criterion.

Joining the blog as a collector of flow knowledge, a wiki allows all sorts of content interlinking, tagging and categorizing with the aim of increasing the information available as a whole and enabling collaboration.

A last tool worth looking at is a bibliographic manager. The evident use of a bibliographic manager is keeping all one's references properly sorted and, in some cases, providing tools to ease the task of citing while writing. Some are web applications installed on a web server which allows not only managing but publishing one's references and bibliographies. This feature contributes to both building one's digital identity, by wrapping all the PRP with names and references belonging to the same area of knowledge that reinforce the identity's framework, as well as shared content to the PRP, which was one of the main goals of the PRP. It is also more attractive to Internet search engines, again reinforcing the achievement of the visibility goal.

Network Building

After identity, meeting other colleagues, exchanging impressions, and working together is what social software is all about.

We want to stress the point that, more than search engines, RSS feeds enable knowledge sharing in real time.

RSS feeds allow subscription, that can be selective through tags, syndication and aggregation to new knowledge created around the world.

Fostering community building will be enhanced by citations and their corresponding links, pingbacks and trackbacks. This interaction can be reinforced by comments on others' PRPs or the creation of friend of a friend (FOAF, <http://www.foaf-project.org/>) files and blogrolls. These last two shape a virtual research network around the PRP and its creator. The extension of this behavior among other researchers helps invisible knowledge workers become present in the relevant, virtual forums.

The PRP reduces contact time as one is findable, and can enrich this contact time because all the information is already there for anyone to read, thus enabling peer review. A higher exposure allows for more highly informed dialogues to take place, paves the path to future collaborations, and shifts a cultural change towards openness.

Of course, no conversation takes place by only speaking, so a feedreader will also become a perfect companion to one's blog.

Self-archiving, Self-publishing

We have talked so far about virtual identity, the digital notebook, the collection of content and explicit knowledge, and the creation of networks. We should not forget that sometimes we create knowledge that should be published, not as notes, but as a finished work.

The PRP provides for self-archiving one's preprints and published works in a personal repository. This does not solve the problem of access to journal publishing itself, but it does solve access to published works.

The author can self-publish and obtain an ISSN or ISBN for each published work within the PRP. Such publications contribute to increase the visibility of the author, shape his digital identity, enrich the content of the site, and make it more appealing to users and search engines.

Conclusions

While increasing, there is still underuse of wikis, social bookmarking, social networking, file sharing, RSS feeds, discussion forums and blogs within academic circles. Researchers and practitioners, faculty and non-scholars, experts and learners, managers and engineers have the ability to provide plenty of knowledge in their lives and works. If shared, this knowledge will be a part of a network worth keeping. If not, these circles will be disconnected and starve. The PRP could help these knowledge workers both as a personal knowledge manager and as a rich knowledge network weaver. Costs are few and benefits are many.

This article is based on the author's paper: The Personal Research Portal: Web 2.0 Driven Individual Commitment with Open Access for Development (<http://www.km4dev.org/journal/index.php/km4dj/article/view/92>).

Ismael Peña-López is Public Policies for Development and ICT4D lecturer at the Open University of Catalonia, Spain. His main research interests are the digital divide, e-readiness, ICT4D and digital inclusion. As a teacher in virtual learning environments, he is also interested in Open Access and its relationship with e-learning and development. He is the lead (and only) developer of Bibciter, a GPL Bibliography Manager. His Personal Research Portal can be accessed at ICTlogy (<http://ictlogy.net/>).