Objectives

• Create an app that allows for online collaboration and quick learning for hackerspaces, companies, governments, etc...

• Allow people from different backgrounds to more easily understand one another, and to relate their background to others
First off...
Have you seen one of these?
You can zoom out...
Or you can zoom in...
Or you can embed it in something else...

Adapted from Diaspora New Image.png sourced from joindiaspora.com
And perhaps put something slightly different....
But first, how do we do this???

And second, why is it useful???
To answer the first question, a technology called the semantic web could be used.
What is the Semantic Web?

As defined by the World Wide Web Consortium (W3C):

The Semantic Web is a web of data. There is lots of data we all use every day, and it is not part of the web. I can see my bank statements on the web, and my photographs, and I can see my appointments in a calendar. But can I see my photos in a calendar to see what I was doing when I took them? Can I see bank statement lines in a calendar?

Why not? Because we don't have a web of data. Because data is controlled by applications, and each application keeps it to itself.

The Semantic Web is about two things. It is about common formats for integration and combination of data drawn from diverse sources, where on the original Web mainly concentrated on the interchange of documents. It is also about language for recording how the data relates to real world objects. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing.
Okay, I am still not seeing it...
Think of the present state of the web as a web of documents.

I do a web search for say, John Doe and I get:

- A presentation that John gave
- A website that mentions some reward John received
- A website which represents the company John works for.
There may in fact be more than one John Doe. It is up to the user to decide which documents mention the John Doe that they are looking for.

In addition, I may find mention of John on Social Networking sites, or possibly my e-mail if I am able to search it.
The computer cannot tell me which John Doe is the correct one, but what I can do is set up a system that will link John Doe to all the documents that he has created. That way, the next time that I search for John Doe, I will see everything related to John Doe.
How do I create these links?

Use ontologies, in the form of RDF triples.

Example:

John likes Cake
A number of these triples together gives....

Produced with Gephi
Description

• Each dot in the graph is called a node

• Each line in the graph is called an edge
Which Identifier to Use?

• Identifiers
  - Digital Object Identifier (www.doi.org)
    In use at?
    - Universally Unique Identifier (http://en.wikipedia.org/wiki/Universally_Unique_Identifier)
      In use at: (http://www.semanticartifacts.com/)
    - and more for sure...
Now, how about another application...
Introducing H.E.L.P

H = Highly
E = Efficient
L = Learning
P = Platform
The project domain could also be replaced by a domain describing another person.

To accomplish this, we need to do something called data integration.


H.E.L.P (cont.)
Now read the rest of Liyang Yu’s book to figure out how to do this.
Consider this Diagram
What if I distributively fund?

I fund Project A, but I also fund projects B, J, H, E, and C
(cont.)

• If I evenly distribute part of the funding from project nodes, every linked node gets one-fifth of this part.

• Alternatively, if one node is used more than another I can disproportionately fund each of the five linked nodes.

• I will define this as 1st degree funding

• (Add some math....)
How about distributed funding to the 2nd degree?

I fund Project A, but I also fund projects B, J, H, E, C, F, D, B, L, K, and I.
I can follow the same rules as funding to the first degree, but for nodes E, C, J, K, and H I can use these nodes defined distribution of funding whether proportionate or disproportionate
How about distributed funding to the 3rd degree?

I fund Project A, but I also fund projects B, J, H, E, C, F, D, B, L, K, I, G, and M.
Integrity of Nodes / Malicious Nodes

• One might ask, what if someone creates a node and starts linking to other nodes pretending that it is useful?

• This is not a problem. Edges point from the subject to the object, not from the object to the subject (see Liyang Yu’s book). I will know whether the author(s) find a particular node useful by the way that the arrow points.
Integrity of Nodes (cont.)

- One might ask, how can I tell a priori if I want to fund the neighboring nodes?
- If I am funding a node, along with its surrounding nodes, the node of interest is only going to link to things that it likes or are useful to it. If I agree with what a particular node does, and what it stands for, I will likely fund other nodes. That is, I am trusting that the particular node checked out nodes it linked to. It seems reasonable that a node would not link to other nodes it did not wish to see funded.
Funding Another Node
This Funding Scheme is Like Flattr

• Well we may or may not use Flattr (possibly in conjunction), but it is a bit like the funding system amongst nodes. “….Everybody gets a slice.”

• You can check out their video at their learn more link at [http://flattr.com/](http://flattr.com/)
Why Diaspora?

- Like the semantic web, Diaspora has elements of being decentralized. In doing so, it implements Federated Social Web Technologies.
  
  (http://www.w3.org/2005/Incubator/federatedsocialweb/wiki/Main_Page)

- Each person owns their data, and has the option to have their own pod on their own server

- It allows users to cross-post to Facebook, Twitter, and Tumblr*. 

- It has a look that is similar to Facebook

*http://diasporaproject.org/
Pros from
http://diasporaproject.org/

• “Own your data”
  – “Connecting socially is human nature. You shouldn't have to trade away your personal information to participate.”

• “Creative Community”
  – “Meet people from all over the world who love the internet as much as you do.”

• “Best of the Web”
  – “Applications that enable you to discover & share cool stuff throughout the web. Cubbi.es is our first.”
Comparison to Diaspora

• The SIOC (Semantically-Interlinked Online Communities) project is like Diaspora in that data can be shared amongst communities

• It is included as a module for Drupal

You can find it at: http://sioc-project.org/
or on Wikipedia at: (http://en.wikipedia.org/wiki/SIOC)
or you can check out the Digital Enterprise Research Institute at the National University of Ireland for this and other exciting Semantic Web technologies

• There are also other distributed social networks like GNU Social, Appleseed, Friendica ....
  (see: http://en.wikipedia.org/wiki/Distributed_social_network)
How about Adding Bittorrent?

- Reduce load on servers (especially for downloading large files)
- Already used for Open Source and Free Software Projects
- Source: http://en.wikipedia.org/wiki/BitTorrent_(protocol)

- Is it compatible with other technologies such as the Semantic Web? Diaspora?
- See also:

  http://blog.martinemde.com/2011/05/bittorrent-distributed-social-network.html
How About Adding Grid Computing?

• Such a feature would help with computationally intensive projects.
• Like Diaspora, the network would be distributed
• BOINC seems to be a good choice

See:
http://en.wikipedia.org/wiki/Grid_computing
Why BOINC?
(Berkeley Open Infrastructure for Network Computing)

- Allows for distributed computing, so that a number of users on the network can form their own virtual supercomputer
- Has a LGPL license that may be compatible with Diaspora and Gephi. At least, getting them to work may not be that hard.

How About Programs to Run on the Grid?

• A lot of programs under the GPL are given at openinnovation.org.

• CAELinux also contains a lot of packages. In fact, its website advertises that it runs on the Amazon EC2 infrastructure. So why not a grid? At least, according to Wikipedia, stuff that could be broken up in parallel could work.
CAELinux (cont.)

• See: [http://www.caelinux.com/CMS/](http://www.caelinux.com/CMS/). Also take note of packages they don’t include such as LAMMPs, ...., and possibly SciGraphica

  See: “Getting Started with CAELinux 2011” by

• University sponsored BOINC project
  – http://volpex.cs.uh.edu/VCP/
  – SETI@home

• Hacker sponsored BOINC project
  http://aerospaceresearch.net/constellation/

• Even more BOINC projects:
  http://usa.lanex.com/
Why Gephi?

• Licensed under GPL
• Allows for “Network Visualization” such as the graphs produced above. [http://en.wikipedia.org/wiki/Network_visualization](http://en.wikipedia.org/wiki/Network_visualization)
• Loaded with features.
• Produces stunning graphs. “Like Photoshop but for data” (gephi.org)
• Uses an XML GEXF file format with extensive customization.
Gephi is a Desktop Application, How do I use it on the Web?

- This was proposed as a problem on the Gephi website: http://wiki.gephi.org/index.php/Google_Summer_Of_Code_2011

The problem was taken on by Urban Škudnik: https://gephi.org/index.php?s=GraphGL

He developed GraphGL: http://wiki.gephi.org/index.php/GraphGL

Compare this to OpenGL: www.khronos.org/webgl/wiki/Main_Page
Also see the “Red Book”, the authoritative source on OpenGL.
Gephi is a Desktop Application, How do I use it on the Web? (cont.)

- Gephi uses Java and OpenGL. HTML 5’s canvas element (if used) works well with JavaScript and WebGL.

- Java and JavaScript?

- Douglas Crockford gives a good explanation of the history of Java, JavaScript, JScript ….
  

- WebGL?
  - Based on OpenGL ES 2.0
    
    (http://www.khronos.org/webgl/wiki/WebGL_and_OpenGL)

  - A good place to learn it is here:
    
    http://learningwebgl.com/blog/
How about going further?

- The network diagram could link to documents and people, and license restrictions would be shown. Also symbols showing patents with royalties and without royalties could be shown.
- Nodes could be magnified when the mouse cursor moves over them
- Allow people to add nodes from the GUI (like Freeplane)

Tim Berners-Lee mentioned the idea of labeling somewhere.

*Note that this does not represent real data.*
CC Licenses

• The following licenses are at:
  
  http://creativecommons.org/licenses/

  which is published under the Creative Commons Attribution 3.0 License

![Attribution CC BY]

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.
CC Licenses (cont.)

Attribution-NoDerivs
CC BY-ND

This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

Attribution-NonCommercial-ShareAlike
CC BY-NC-SA

This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.
CC Licenses (cont.)

Attribution-ShareAlike
CC BY-SA

This license lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. This license is often compared to “copyleft” free and open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use. This is the license used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects.
CC Licenses (cont.)

Attribution-NonCommercial
CC BY-NC

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don’t have to license their derivative works on the same terms.

Attribution-NonCommercial-NoDerivs
CC BY-NC-ND

This license is the most restrictive of our six main licenses, only allowing others to download your works and share them with others as long as they credit you, but they can’t change them in any way or use them commercially.

• CC0 – Public Domain
Other Licenses

“The Free Art License (abbr.: FAL, French: Licence Art Libre) is a copyleft license that grants the right to freely copy, distribute, and transform creative works without needing the author's explicit permission.”

-Wikipedia (http://en.wikipedia.org/wiki/Free_Art_license)

- More Documentation Licenses can be found in this Wikipedia article: http://en.wikipedia.org/wiki/GFDL
Icons

• The Icons are from:
  http://openiconlibrary.sourceforge.net/

text-x-generic.png | Source: echo , License: CC-BY-SA-3.0

x-office-spreadsheet.png | Source: echo , License: CC-BY-SA-3.0

image-x-generic.png | Source: tango , License: PD
Comments

• Freeplane has labelling on mind maps, which are similar to network visualization diagrams
  http://freeplane.sourceforge.net/wiki/index.php/Main_Page

• Freeplane is offered under the GNU public license.

• Some of the proprietary softwares allow for collaborative editing (is there a shared patent/ license?).
  en.wikipedia.org/wiki/Mind_maps
Scalable Vector Graphics might be used to render images such as license labelling used in the graph

- http://www.w3.org/Graphics/SVG/

- Perhaps consider XHTML5 for WebGL like technologies (e.g. GraphGL). Maybe it would be better to keep everything in XML?
Why the Semantic Web?

• Another Quote:

"The goal of Semantic Web research is to transform the Web from a linked document repository into a distributed knowledge base and application platform, thus allowing the vast range of available information and services to be more effectively exploited."

W4A '07 Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A), ACM New York, NY, USA ©2007

We are building a “Giant Global Graph” – Tim B.L.
Semantic Web Browsing

• Conventional web browsers search for text strings within documents
• Semantic web browsers search for data.
• Allows for answers to questions like “Why did I blush when I was with a hot girl and it was raining”.

• Liyang Yu’s book has a chapter about this topic
How do I implement the Semantic Web?

Using these technologies:

To build what Tim Berner’s Lee calls the “Web of Data”

See also: http://en.wikipedia.org/wiki/Linked_Data
How about the Semantic Web with Diaspora?

- Diaspora uses MySQL, PostgreSQL, and SQLite3. Something like R2ML might be needed to communicate these to RDF.
- [http://www.w3.org/2001/sw/rdb2rdf/r2rml/](http://www.w3.org/2001/sw/rdb2rdf/r2rml/)
- See also: Milanovic, Milan V., Modelling Rules on the Semantic Web, Master Thesis, University of Belgrade: Faculty of Organizational Sciences, Belgrade, 2007
- See Drupal 7’s hack at the semantic web: [http://semantic-drupal.com/node/11](http://semantic-drupal.com/node/11)
- See the Presentation: “Some of the newest SW technologies at W3C: RDFa1.1 and R2ML, Benelux Semantic Web Meetup, Amsterdam 2010-11-24, Ivan Herman, W3C”
- And more??
What about Software Licenses?

- Diaspora – AGPL
- Gephi – GPL v3 (previously AGPL)
- Semantic Web – (by W3C – license?)
- BOINC - LGPL

See [http://www.gnu.org/licenses/gpl-faq.html](http://www.gnu.org/licenses/gpl-faq.html) and
Tools for Semantic Web Development

• There are more Java tools than Ruby tools for the semantic web. Jena (http://incubator.apache.org/jena/index.html) looks like an attractive option, except it is written for Java. Perhaps try JRuby?

• See more tools at: http://www.w3.org/2001/sw/wiki/Tools
• http://www.w3.org/2001/12/rubyrdf/intro.html
• http://blog.datagraph.org/2010/03/rdf-for-ruby
In Summary…

Build an Open Research / Innovation Platform

Something like Flattr
Web of trust for Chemicals

DIASPORA*

Product by 3d Printing from N Hackerspaces and Individuals

Bittorrent?

Packages from CAELinux

Do we need a Virtual Machine like Xen running distributed over a network to run CAELinux programs?
Additional Comments

• Sometimes we may have two identical things that need to be given the same URI, perhaps it is possible to clean up ‘messy data’ with Google Refine
Additional links of Interest

- For interesting databases consider:
  - http://www.data.gov/communities/semantic
  - http://www.freebase.com/
  - http://dbpedia.org/About

- Brad Fitzpatrick’s thoughts:
  - http://bradfitz.com/social-graph-problem/ (now at Google -- LJ)

- Companies:
  - http://www.slideshare.net/ABLVienna/poolparty-overview
  - http://www.revelytix.com/content/knoodlcom
Additional links of Interest (cont.)

• The Semantic Web and Drupal, part 1 of 2 : User - DERIGalway

  http://www.youtube.com/watch?v=dVADiTI5cHI&feature=related

• The Semantic Web and Drupal, part 2 of 2 : User - DERIGalway

  http://www.youtube.com/watch?v=fOBHRhkod90&feature=related

• The Semantic Web - An Overview : User - peterMankato

  http://www.youtube.com/watch?v=rhgUDGtT2EM&feature=related

• What is an Ontology? : SpryKnowledge

  http://www.youtube.com/watch?v=jfUPLuPL3Ho&feature=related
A very good presentation for distributed social networks with the semantic web:

http://www.slideshare.net/davidseth/semantic-web-for-distributed-social-networks-presentation

- This website seems to share a number of common goals:
  - http://www-sop.inria.fr/edelweiss/
Additional links of Interest (cont.)

Federated Social Web  (http://www.w3.org/2005/Incubator/federatedsocialweb/wiki/Main_Page)

• Ontological graphical mapping using W3C standards (similar to our graphics)
  http://snoggle.projects.semwebcentral.org/
  http://semweb.salzburgresearch.at/apps/rdf-gravity/index.html
  http://nodexl.codeplex.com/
Additional links of Interest (cont.)

• Semantic Web Search
  http://swoogle.umbc.edu/
  http://www.makeuseof.com/tag/top-7-semantic-search-engines-alternative-google-search/
  http://en.wikipedia.org/wiki/True_Knowledge
http://groups.csail.mit.edu/haystack/
http://www4.wiwiss.fu-berlin.de/bizer/ng4j/disco/
http://simile.mit.edu/wiki/Longwell
http://rdfgraphbrowser.sourceforge.net/
http://brownsauce.sourceforge.net/
http://tech.slashdot.org/story/04/11/10/0040223/Welkin-A-General-Purpose-RDF-Browser

• Presentation giving problems with RDF browsers (2008 or earlier ? – ask them if needed):
  http://simile.mit.edu/presentations/visualization.pdf
  The parent site is also very interesting. A lot of work here. http://simile.mit.edu/

• For comparison a Web Search Engine is:
  http://en.wikipedia.org/wiki/Web_search_engine
Additional links of Interest (cont.)

• A tool that allows the editing of search results:
  http://projects.csail.mit.edu/exhibit/Dido/

• Semantic web tool
  http://iglue.com/

• Graph Database tool (looks semantic), check out their blog and github
  http://dydra.com/

• Article mentioning patent in a similar area of the present idea:
  http://mashable.com/2012/01/03/google-ibm-patents-semantic-network/
  Except, it does not appear to do what we want to do. Perhaps there are other patents?
Additional links of Interest (cont.)

• The Dublin core project is a metadata project that has ties to the semantic web effort. http://dublincore.org/

• An interesting Semantic Web Blog:

  http://blog.semantic-web.at/tag/google/
Additional links of Interest (cont.)

• The present idea is very similar to:
  • [http://www.techquila.com/topicmapster_2.html](http://www.techquila.com/topicmapster_2.html) (see the TMShare Application)
    by Kalid Ahmed, kal@networkedplanet.com
  • [http://www.networkedplanet.com/](http://www.networkedplanet.com/) (which generates material for Microsoft SharePoint)
  
TMShare is mentioned in this application

Additional links of Interest (cont.)

- Ontopia is also similar to the present idea as well as Kalid Ahmed’s work.
  http://www.ontopia.net/index.jsp

- Ontopia’s site mentions a lot of interesting applications to topic maps such as
  Content Management Systems, including code released under the Apache 2.0 license (GPL compatible).

  Topic maps are essentially a superset of RDF maps since they can contain more information according to Wikipedia
  (http://en.wikipedia.org/wiki/Topic_maps)

  Ontopia mentions conversion of RDF to topic maps:
  http://www.ontopia.net/doc/current/doc/misc/rdf2tm.html

  For topic maps, something else to consider is:
  TMAPI - Common Topic Map Application Programming Interface

  "TMAPI is a programming interface for accessing and manipulating data held in a topic map."
  http://www.tmapi.org/
Additional links of Interest (cont.)

Also see:

- http://www.isotopicmaps.org/sam/sam-model/
Additional links of Interest (cont.)

• Visual Thesaurus is a lot like the present idea:
  http://www.visualthesaurus.com/
  from http://www.thinkmap.com/

• So is TouchGraph
  http://www.touchgraph.com/seo

• So is
  http://academic.research.microsoft.com

• So is Action Science Explorer (which is a lot like Gephi and NodeXL)

• So is Perl Trees (actually very much so)
  http://www.pearltrees.com/
• Perhaps the Semantic-MediaWiki project could be useful:

This following presentation mentions a bit about Semantic-MediaWiki (http://semantic mediawiki.org/).

Generating and Querying Semantic Metadata and Ontologies : GoogleTechTalks
http://www.youtube.com/watch?v=1YhE9WVZ55c&feature=related

• Perhaps their code is useful?
  They include semantic searching. It's written in PHP and licensed under GPL according to Wikipedia.
Additional Links (cont.)

• This idea uses collaborative innovation. There are multiple approaches. Here are two:

• Little or no restriction:
  http://en.wikipedia.org/wiki/Open_research

• Slightly more restrictive:
  http://en.wikipedia.org/wiki/Open_innovation
This book seems to be close to the project:

- T. Berners-Lee, W. Hall, J.A. Hendler, K. O’Hara,
  

Also see:

http://www.w3.org/2005/Incubator/socialweb/XGR-socialweb-20101206/
Additional Links (cont.)

• An interesting blogging site about the semantic web.

http://semanticweb.com/
Applications

• Partner with Maker Faire (http://makerfaire.com/)

• Partner with musicbrainz for semantic search (http://musicbrainz.org/). Inspired by a musician at OhmSpace.

• Use this sort of information with a Music Festival?

• Already in use (along with many other applications): http://rdfs.org/resume-rdf/ Express C.V. or Resume on the web (Source: http://en.wikipedia.org/wiki/Resource_Description_Framework)
End Here
First off, consider the semantic web to be one giant database (global graph....check the linux links)

Consider Liyang Yu’s definition of the semantic web:
What is the Semantic Web?

• Invented by Tim Berners-Lee, the inventor of the World-Wide Web (1)
• “The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries” (2)
• “The Semantic Web is a collection of technologies and standards that allow machines to understand the meaning (semantics) of information on the Web.” (3)
Try JSON

• http://flori.github.com/json/

• Http://en.wikipedia.org/wiki/Topic_Maps
Why CAELinux?

• Actually, the packages contained with the distribution are what we are interested in.

• CAE stands for Computer Aided Engineering

• A project with a lot of open source backing such as Polywell could benefit greatly from this. Hackerspaces could as well.

• See: http://www.caelinux.com/CMS/. Also take note of packages they don’t include such as LAMMPs, (find the others....)