

# MAPPING CONVERSATIONS ON THE WEB



using semantic web technology

## **ERKLÄRUNG**

---

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig erstellt und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Soweit ich auf fremde Materialien, Texte oder Gedankengänge zurückgegriffen habe, enthalten meine Ausführungen vollständige und eindeutige Verweise auf die Urheber und Quellen.

Alle weiteren Inhalte der vorgelegten Arbeit stammen von mir im urheberrechtlichen Sinn, soweit keine Verweise und Zitate erfolgen.

Mir ist bekannt, dass ein Täuschungsversuch vorliegt, wenn die vorstehende Erklärung sich als unrichtig erweist.

Sebastian Gerhard, 713 036

<b>INTRODUCTION</b>	<b>2</b>
<b>The Talmud, Connected Conversations</b>	<b>3</b>
<b>Letter Exchanges, Remote Conversations</b>	<b>5</b>
<b>The MEMEX, Associative Thoughts</b>	<b>7</b>
<b>Same Conversations, Evolving Media</b>	<b>11</b>
<b>Perceiving Documented Conversations</b>	<b>12</b>
Testing Reading Behaviour	17
<b>Conversations on Today's Web</b>	<b>19</b>
<b>Comments, Low Level Conversations</b>	<b>22</b>
Comment Services and Solutions	24
<b>Handling &amp; Organizing Conversational Data</b>	<b>30</b>
<b>Linked Data And The Semantic Web</b>	<b>32</b>
Defining A Social Graph: FOAF & XFN	37
The Facebook Way: Open Graph	39
<b>My Approach To Reinvigorate Conversations On The Web</b>	<b>42</b>
Identifying the element and defining a vocabulary	43
<b>Crafting An Interface For Data</b>	<b>47</b>
<b>The Prototype</b>	<b>50</b>
<b>Appendix</b>	<b>54</b>

## Introduction

In my thesis on “mapping conversations on the Web” I will underpin the core concept that my studies on “digital media” follow. Being trained in a variety of disciplines, I will elaborate the technological, social and visual context in which this thesis circulates.

[In the first part](#), we will explore the role of conversations as the essence of any kind of human interaction. We will briefly dive into history and have a look at examples that show us how fruitful cultivated conversations are and how they led to cultural progress.

We will see that conversations play an important role in our modern society and how the conversations alongside the ways we developed to document them, shaped the way we process information.

We will talk about the idea of a public memory storage, a universal library and ways to put huge piles of data in order.

[In the second part](#), we will learn that the conversational style of the Web has come a long way. How it has developed and how it could look like in a near future. How the Web and media in general have shaped our behavior and that “new” concepts and ideas on the Web are older than what we know as the world wide Web. The conversations we will focus on, are not the comments on Facebook status updates or any other small social gestures. In this first step it is about dialogues between people who put thought in their pieces, who run some platform to publish their content.

[In the third part](#), we will learn that Semantic Web technology does not have to be rocket science and why the social space could be the breeding ground for a new kind of Web, a net of linked and related data.

A prototype will prove the concept and show how a down-to-earth approach could look like that takes the existing infrastructures into account.

## The Talmud, Documented Conversations

The concept of connected and documented conversations found its first prominent use in the Talmud, the central piece of text of the Judaism. It dates back around the year 200 CE. After transferring knowledge and scholarship orally for hundreds of years, the situation changed after the romans tore down the city of Jerusalem and defeated the Jewish commonwealth, leading to an ongoing displacement of the Jewish people.

Formerly, the Jewish scholars, mostly rabbis, were required to face a new situation as Judea had not only lost its autonomy, but also the temples, which served as a central communication hub for teaching and study. Their established system refused to work in this new scenario and people began to write down the eclectic discourses on legal and spiritual matters.

The oral law was authored and maintained by many rabbis at many different places—probably at the same time. So when the first laws of the Talmud were written down, the authors developed a concept to structure and organize the content. This formal framework was able to handle the input of multiple authors, subsequent changes and newly gathered insights and discussions that needed to be included in the Talmud over nearly a thousand years.

[http://www.come-and-hear.com/  
structure.html](http://www.come-and-hear.com/structure.html)

In order to keep the work structured, the rabbis developed a hierarchy based on subject matter and not by biblical verse, a novelty at this time. It kept the various thoughts, interpretations and sometimes dialectic passages separated, which furthered understanding and analysis. The structure allows to read “between the lines”, to follow the evolution of laws and scientific findings and how they made their way into society. It helps to understand it, giving you the hows and whys.

The Talmud’s two main books, the “Mishna” and the “Gemara” are not split in two parts but may occur multiple times interchangeably within a chapter.

This inner structure of the Talmud is continued in the physical structure of the series of books. According to a cer-

tain printed edition (the Vilna printing, around 1800), the chapters are divided in folios, semantically divided spans of pages, which are numbered and thereby easy to reference.



A talmud page, beginning of the berachoth chapter. [Mishna](#) and [Gemara](#) in the center of the type area. The right margin holds the comments of the last important editor (Raschi) and the left margin younger comments.

The Talmud in its role of a collaborative manifest of cultural heritage, makes the jewish people “People of the Book”, Author Jonathan Rosen states in his book “The Talmud & the Internet”. He continues by comparing reading the Talmud to browsing the Web:

*"The Talmud and  
the Internet—  
A Journey between  
Worlds"—  
Jonathan Rosen,  
Farrar, Straus &  
Giroux, 2000*

“

*“the Internet is similar in that respect, as one  
can wander in and out of the internet through  
research and searching, and feel the displace-  
ment that the Jews felt during their journeys  
throughout the centuries.”*

*—Jonathan Rosen*

He links that to the iterative writing process and the genesis of the book that is comparable to the ongoing redevelopment of the Web.

## Letter Exchanges: Remote Conversations

Another interesting type of documented conversations are letter exchanges among people. A letter exchange is a series of letters that two individuals send each other over a period of time. Besides personal friendships, letter exchanges played an important role in science. Especially in the 18<sup>th</sup> and 19<sup>th</sup> century, letter exchanges provided a flexible platform for ongoing discussions and to elaborate on certain topics.

*Goethe an Winckelmann.  
In: J.W.v. Goethe. Werke.  
Weimarer Ausgabe, 1. Abt.,  
46. Band, 1891, S.11 f.*

“

*...denn oft wird ein Freund, an den man  
schreibt, mehr der Anlaß als der Gegenstand  
des Briefes...*

*—J.W.v. Goethe*

<http://www.briefwechsel-schiller-goethe.de/> brings the Goethe & Schiller letter exchange to the Web

This quote reveals the motivation of many authors: Having an obvious audience, being forced to name and quote sources and to respond in a timely manner, was a driving motivation for many writers, scientists and philosophers.

*ARD Abenteuer Reisen  
<http://bit.ly/dwIOJQ>*

The scientific remains of the german polymath Alexander von Humboldt for example, are not less than *13 000 letters* he exchanged with about *2 700 people*. Categorizing the letters has turned into task a team of historians has been trying to accomplish for more than 30 years. They recognized that a chronological system won't work.

*"Once the consciousness was hooked  
on busyness and external stimuli,  
Thoreau saw, it was hard to break  
the habit. Never mind the telegraph,  
even the post office could become an  
addiction..."*

*Hamlet's BlackBerry,  
William Powers, Harper 2010*

*Der Brief und das Blatt.  
Die Entstehung wissenschaftlicher  
Zeitschriften aus der  
Gelehrtenkorrespondenz",  
Wissenschaftsforschung 2009*

So they focussed on grouping the letters that form individual conversations, dealing each with a specific topic. Using this concept, they realized that they create interesting documentations of the authors thought processes that revealed a lot about the motivations behind their work. A letter exchange furthers associative thinking. Both partners, during the exchange, falsify and validate hypotheses of each other, doing a collaborative research and delivering its essence in this specific media.

The fact that people exchanging letters are commonly not sitting in the same spot, lets them think and write freely, without interruptions and interjections.

It is interesting that people had to deal with interruptions even in a time long before telephones and the internet.

“

***"Letters are "linguistically fixated interactions, iterative correlations with cognitive and emotional functions."***

***—Regine Zott***

These series of letters formed their own kind of literature, diaries and memoirs and compiled letters, written with a certain audience in mind.

Many letter exchanges between famous people were printed and sold as books.

Later on, scientific journals emerged from letter exchanges, seeking for a broader audience. This meant also a shift from *informal* (letter) to *formal* communication (journal).

German astronomer Franz Xaver von Zach founded the first german journal. His brisk correspondence with Carl Friedrich Gauß (mathematician & physicists), Friedrich Besse (astronomer & cartographer), Johann Hieronmus Schröter (jurist and astronomer) and various people in public service, is a rich source for historians that helps to understand correlations between society, politics and science in the 19<sup>th</sup> century.



## The MEMEX, Associative Thoughts

The replicable thought processes and conversations found use in the MEMEX concept as well. Vannevar Bush, an influential american engineer, developed one of the first analog computers, called the "Differential Analyzer" in the late 1920ies.

Besides being a professor for mathematics and electrical engineering, he became an advisor for the U.S. government during the presidencies of Roosevelt and Truman. He served as the "Director of the Office of Scientific Research and Development" during the second world war, he oversaw and directed the work of more than six thousand leading american scientists. As an influential strategist, he managed to align military strategies with scientific interests.

[http://www.theatlantic.com/  
magazine/archive/1969/12/  
as-we-may-think/3881/](http://www.theatlantic.com/magazine/archive/1969/12/as-we-may-think/3881/)

The second world war was, in his opinion, an inspiration and the driving force for all disciplines of scientific work. In his famous essay "As We May Think", published in the "Atlantic Monthly" magazine in july 1945, he encourages scientists to continue their fruitful work and to share the huge amount of insights and knowledge they obtained, working closely together, during the war period. He coined the phrase "Science, the endless frontier".

His essay is not concerned with the individual knowledge of individual scientists, or how they discover it, it is concerned with how we save it, how we transmit it, and how we get access to it in a broader, societal sense, thinking beyond books and traditional libraries. He sketches an idea of another universal library.

“

*“There is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered by the findings and conclusions of thousands of other workers -- conclusions which he cannot find time to grasp, much less to remember, as they appear. Yet specialization becomes increasingly necessary for progress, and the effort to bridge between disciplines is correspondingly superficial.”*

— *Vannevar Bush*

He points out that the all methods to record and archive data fail, when it comes to selecting and finding the relevant pieces of data that could be relevant for our own work. That is mostly the case because all classification and filing systems are artificial (using letters or numbers, allowing items to be stores in only one place at a time),

“

*“the extraction of parallel material from the existing record, and the final insertion of new material into the general body of the common record....”*

— *Vannevar Bush*

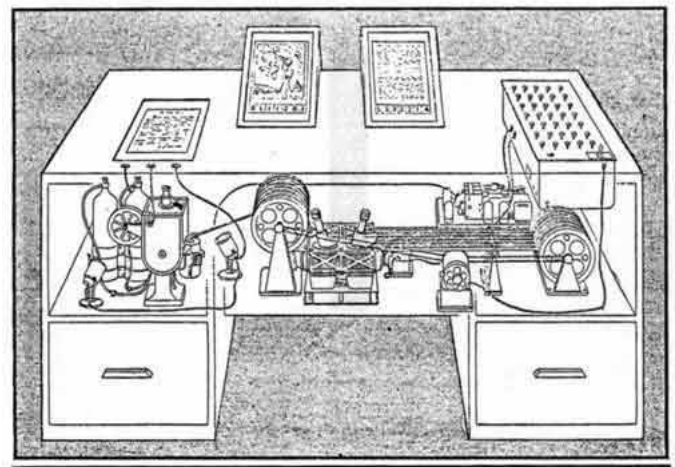
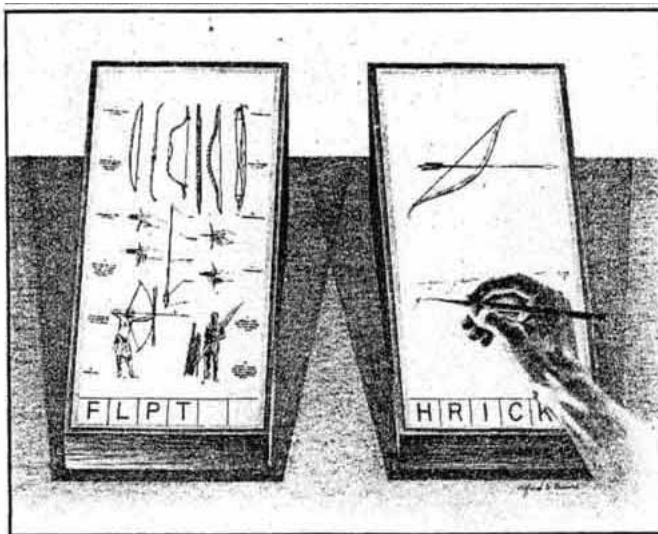
The extraction of parallel material is the challenge that needed to be solved. His proposal is a system he called MEMEX. It is probably derived from “memory extension”, which boils the system down to the core concept.

Bush emphasizes the ability to create and record “trails” through the mass of knowledge. He describes these trails as “...any item can be joined into numerous trails.” His vision includes new encyclopedias, “ready made with a mesh of associative trails...”. He yet mentions a variety of applications, i.e. medical and chemical databases.

The “trails” he describes as connections between data, are mapped and saved thought processes of users, made available for users to follow and iterate. Loading

up the created set of trails would give the user not solely all the important pieces of data, but also their inherent connections, which imply how and why these trails were established.

Exporting research processes on microfilm for other user's systems, was a feature he proposed.



MEMEX Slates &  
complete setup

It was meant to document conversations during which a series of search queries was conducted, when scientists or MEMEX users met in real life. These exported microfilms allowed the users to follow up the discussion and to understand the chain of reasoning that led to an agreement or disagreement.

Due to his past in analog computing, his vision of the MEMEX system is an electromechanical data storage system, using micro photography. Much of what he describes had yet to be invented, but to some extent, he created the concepts for hypertext, personal computers, networks, wikis and handwriting- and speech recognition that took more than 30 years to blaze the trail to the market

Bush's thesis runs that society had a problem with sharing knowledge, using the existing sources and integrating these results in their own work.

But and that's ironic, his proposed solution, the MEMEX, is a single user device that augments an individuals memory and furthers research processes. He refuses to answer the question of how we can use and cultivate a collective memory or storage of data.

That is another visionary aspect of his essay: The development of the MEMEX blueprints the way of computer history, taking nearly 50 years to move from mainframes and terminals to personal computers and the World Wide Web. All kind of data structure is struggling with the enormous growth of data.

<http://sloan.stanford.edu/MouseSite/1968Demo.html>

Bush's essay inspired a lot of engineers and conceptual thinkers. Douglas Engelbart, in his so called "mother of all demos" created the first prototype running on a computer, demoing what became the mouse, word processors, copy&paste, networked collaboration and many other interaction paradigms that shaped the way we interacted with computers for more than three decades.

<http://en.wikipedia.org/wiki/Wiki>

The term hypertext, describing the concept of mesh-like connected pieces of information using nodes and links, was coined by Ted Nelson in 1965 and is the technological concept behind the largest woven mesh, Wikipedia, created by Ward Cunningham in 1995. He built a tool he called "wikiwikiweb" what was actually the first Wiki software around. A Wiki gives users a set of tool to easily edit pages within a single Wiki, offering back-link functionality and even the possibility to create links to pages that do not yet exist. The latter is an essential feature needed to weave the mesh of data in a MEMEX manner.

## Same Conversations—Evolving Media

*The phaidros, written by Plato,  
is a dialogue between Socrates  
and Phaedrus, an interlocutor in  
several dialogues. The Phaidros  
was presumably composed around  
370 BC.*

*"During [the twentieth] century  
we have for the first time been  
dominated by non-interactive forms  
of entertainment: cinema, radio,  
recorded music and television.  
Before they came along all enter-  
tainment was interactive: theatre,  
music, sport—the performers  
and audience were there together,  
and even a respectfully silent  
audience exerted a powerful shap-  
ing presence on the unfolding of  
whatever drama they were there  
for. We didn't need a special word  
for interactivity in the same way  
that we don't (yet) need a special  
word for people with only one  
head"*

*Douglas Adams,  
How to Stop Worrying and  
Learn to Love the Internet.*

As with the advent of any new media, there have always been theorists around, bemoaning the downfall of mankind. As *Socrates*, in *Plato's* Phaidros feared that by relying on the written word as a substitute for knowledge formerly carried inside their heads, people would lose their ability to reflect on knowledge and become forgetful. And yes, in some perspective he was right, in the same way as the italian scholar *Hieronimo Squarciafico* most prominently worried that the easy availability of printed books would lead to mental laziness.

And yes, it seems both were right. We again experienced a change with any new type of media like radio and television. Especially TV and radio led our society in an era of one-dimensional communication and its the Web that is beginning to change that again.

Although we still await neurological prove that using media rewires our synapses and re-shapes or mind but its obvious that at least in some way, our cognition has adapted to the way we are consuming information on any new medium.

That links to what *Marshal McLuhan* pointed out in the 1960s, when he said that media are not just passive sources of information. They supply the stuff of thought but also shape the process of thought. The german philosopher *Friedrich Nietzsche*, whose style of writing changed completely when he switched from ink and paper to a typewriter, proves *McLuhan's* thesis already in the 19<sup>th</sup> century. *Nietzsche's* style of writing turned "from arguments to aphorism, from thoughts to puns and from rhetoric to telegram style."

The semantic structure of hypertext is said to reflect the idea of associative thinking. And yes, it they does. But they hold a new challenge for us: The Web, easily available with just a mouse click or finger tip, introduces us to the anamnesis of a phenomena that is called "Wikipedia syndrome".

It describes:

[http://www.urbandictionary.com/  
define.php?term=Wikipedia%20  
Syndrome](http://www.urbandictionary.com/define.php?term=Wikipedia%20Syndrome)

“The act of going to a site with the intention of looking up one piece of information, and instead finding yourself on one or more articles that have nothing to do with what you had originally come for. Stems from Wikipedia linking to articles within articles.”

In a short user testing ran alongside this thesis, the test subjects were asked to browse through their feedreader, their Twitter- and Facebookstream and it became evident that keeping in line with the initial activity was a hard task. All of them, being asked to focus on the primary activity got lost in the linked sources.

(see videos on the accompanying CD)

## Perceiving Documented Conversations

Another aspect that plays a substantial role when we consume media is the visual perception. Due to the physiology of the human eye, with the retina and its central spot, the *fovea centralis*, we are able to get a clear sight on an area of about 1,5 cm. When we read, our eyeballs need to move quickly to adjust the *fovea centralis* to the spot in focus. These rapid and directed movements are called *saccades*. They are *ballistic*, meaning that they, once started, are irreversible. During these movements, we are frankly blind and our attention is paused.

fixations (cyan) &  
saccades (grey)

The condition of American journalism in the first decade of the twenty- first century can be expressed in a single unhappy word: crisis. Whether it's a plagiarism scandal at a leading newspaper,

The pauses between the *saccades* are called *fixations*. During the fixations we “pick up” the information. Besides that, we seem to have a peripheral awareness, also catching the areas around the fixation points. This

peripheral awareness seems to be responsible for targeting the fixations. It helps us to preprocess the words ahead of the fixations, allowing us to notice information that is not in focus. This natural behavior, implied by the physiology of our eyes is coherent to the way we process the collected information.

*The more complex a subject is,  
the more saccades and fixations  
we need. When reading text,  
one saccade embraces approx.  
3–4 words*

When reading, we do not read letters or words but rather identify patterns and structures, consisting of multiple words. When we read a hypertext document with a handful of links inserted in each paragraph, the amount of information within reach increases enormously. The links, spanning around those patterns, break the flow of perception.

Hyperlinks, embedded in the content, don't just point towards related resources, they propel us toward them. Our peripheral awareness is triggered when a hyperlink is identified.

*In this example hyper-  
links are highlighted in  
orange and underlined.  
Selecting colors and 'text  
decoration' for hyperlinks  
is an important choice  
for the overall design of  
web content.*

example from [http://  
en.wikipedia.org/wiki/  
Jef\\_Raskin](http://en.wikipedia.org/wiki/Jef_Raskin)

**Raskin was born in New York City. He received degrees in mathematics (B.S. 1964) and philosophy (B.A. 1965) at the State University of New York at Stony Brook. In 1967 he earned a master's degree in computer science at Pennsylvania State University. His first computer program, a music program, was part of his master's thesis.**

**Raskin later enrolled in a graduate music program at the University of California, San Diego (UCSD), but stopped to teach art, photography and computer science there, working as an assistant professor from 1970 until 1974.**

In combination with curiosity, an emotion related to natural inquisitive behavior of the human species is responsible for the way we interact with hyperlinks. Hyperlinks in the source text don't merely point toward their included reference, no—they propel us toward them. Making us think that we need to know the referenced source to fully understand the original text.



Context in visual perception, describes two closely connected things. On the one hand it describes the prior knowledge we have of a certain matter and our expectation towards the content, on the other hand, it is used to describe the surrounding of the content. Context allows us to get sentences right that use words carrying multiple meanings like “bank” or “band” and to understand the social background. Even when we look at images, the context plays an important role in the process of understanding. As many psychological studies proved, we extract information embedded in context faster and are able to memorize much better.

When we align content, carving out the visual context, we visualize hierarchic and semantic coherences.

We develop adaptive filters that help us to keep focus. The attenuation model” of perception that the american psychologists Treisman and Riley described in 1964 was related to acoustic perception but can be transferred to the Web with all the different streams and layers of information many websites or other types of media they offer. In their paper “Is selective attention selective perception or selective response? A further test, 1969” Treisman and Riley proved that we are able to process information that is not in the center of our attention. The selective response, our conscious reaction on any signal, triggers the cognitive process in our mind.

Current visual representations of conversational texts or text in general, do not support our physiologic and cognitive processes. Advertising formats do not align with these patterns and live alongside and sometimes above the original copy of the text. The loss of print advertising revenue in magazines has not yet been caught up by online ads and brings up always new banner formats that break the user experience.

[www.stateofthemedias.org/  
2010/index.php](http://www.stateofthemedias.org/2010/index.php)



*The whole page is cluttered and disrupted with unrelated content*



a popular sharing widget, as seen on [www.abduzeedo.com](http://www.abduzeedo.com)

an article on the website of the german newspaper "FAZ"

<http://bit.ly/9QWWmk>



- Advertising
- Navigation
- Promotion for other sections
- sharing tools

Copytext, read: content and advertisings are no longer distinguishable, the borders are blurring.

*A nice approach of the NYT on an human interface for a news website:*

<http://www.nytimes.com/timesskimmer/>

<http://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/6868/>

Webpublishers offer us "tools" to share the URL and the site's content, decorate text with so called "related articles" and add preview widgets to the embedded links, creating another layer of reading obstacles.

Apple's webkit based browser Safari 5 comes with feature called "Reader" that strips out all elements that do not belong to the articles body text. Thereby it gives us an environment that allows us to focus on reading, consuming the article. It gives us a chance to get us in, what media critic Nicholas Carr calls, "deep reading", being alone with a text freeing our minds to follow thoughts and ideas. Filtering out all the decoration, Safari Reader shows us how little content is there is on actually stuffed, long websites.

The Web does not have any interface guidelines like Apple introduced for their mobile devices, or rules for setting type and designing information on the Web.

[www.instapaper.com](http://www.instapaper.com)

Safari Reader and apps like Marco Arment's Instapaper, create a common reading interface. Not only because they just strip out ads and distracting site elements, they succeed in successful content delivery because they format it always the same way.

We do not have to get accustomed to typographic parameters like line height and line length or font styles that are used to highlight important passages.



the same article, on the left in its original layout on the website and on the right in the silenced reading mode in Safari Reader

*A nice approach to a realigned guardian on the Web:*

<http://www.gyford.com/phil/writing/2010/06/09/todays-guardian.php>

## Testing Reading Behavior

In order to prove my assumptions on reading on the Web, I created a test, comparing reading behavior and cognitive post processing.

Ten selected test subjects were asked to read a text in its original context on a newspaper website and in a second round, another text, comparable in complexity and length, isolated using the Safari Reader. Both texts, similar regarding content, had a length of 1 200 words. Having read respective texts, they were asked two comparable questions about the content and to summarize the text in two sentences.

*see the following table for detailed results. The dots indicate correct/wrong answers to the comprehension questions*

In the first round, reading the text on the website, it took the subjects an average time of 5:44. 6 out of 10 were able to answer the questions and to summarize the the core statement of the text.

In the second round, reading the text in Safari Reader, it took them about 6:22. 9 out of 10 subjects were able to answer all the questions and name the core statement.

	Colonia Dignidad Safari Reader	The Cheatles webpage layout
person 01	3:45* ● ●	3:29 ● ○
person 02	7:32 ● ●	6:18 ● ○
person 03	7:40 ● ●	7:20 ○ ●
person 04	6:48 ○ ●	7:03 ○ ○
person 05	6:57 ● ●	6:39 ● ○
person 06	4:02 ● ●	3:38 ○ ●
person 07	4:24 ● ●	5:10 ● ○
person 08	5:37 ● ○	4:42 ● ●
person 09	6:40 ● ●	5:24 ○ ○
person 10	7:03 ● ○	6:15 ● ○

\* all values are noted in minutes

● correct answer

○ wrong answer

*"On the average Web page, users have time to read at most 28% of the words during an average visit; 20% is more likely."*

*The NNG observes reading behavior on the internet for about 15 years. Their initial assumptions regarding reading behavior remain true. But design concepts alongside with new advertising formats change peoples individual patterns.*

*People do not read very much, scanning is still the extremely common behavior.*

<http://www.useit.com/alertbox/percent-text-read.html>

*Understanding text and interpreting reading speed is a complex process. A long-term study by Harald Weinreich, conducted at the University of Hamburg, elaborates on that topic and goes into detail on navigation between hypertext documents.*

<http://portal.acm.org/citation.cfm?id=1326566>

*"A study of people reading long-form text on tablets finds higher reading speeds than in the past, but they're still slower than reading print."*

<http://www.useit.com/alertbox/ipad-kindle-reading.html>

Its obvious that reading in the isolated Reader took an average of 10 % longer than reading in the original newspaper layout. But accompanied with the results of the short questions it seems to be evident that they read the text much more carefully. All of them appreciated the interface without distractions and the linear layout.

The test shows a relationship between well structured and unobstructed text and how our mind is able to process the content.

A recently published study by the Nielsen Norman Group has compared "reading interfaces" on Amazons Kindle, the iPad and traditional books revealed that readers reading a paper book read faster than those test subjects who used an electronic device. That is mostly because of the well known interface of a book. Readers can almost feel how many pages are left in the book or in a chapter and know how ink on paper looks like. Books provide a consisting experience throughout the book—the Web does not.

Publishing has written of the Web. Magazine- and book-sellers have moved toward selling encapsulated content for devices like the iPad, iPhone or the Kindle. And not because the Web as a platform fails at delivering content, but because the search for business models matching their classic approach known from print media failed. Text on the Web is open to reuse and recycle it. Its is searchable and findable on the net, furthering the ideas of linked conversations and a universal library. It will still be accessible shareable and the true sense of the word whereas the encapsulated content on proprietary platforms will not.

## Conversations On Today's Web

*Interestingly, the HTTP protocol holds a reserved error status code "402 Payment Required" that originated 15 years ago.*

*Innovation—  
The Missing Dimension  
Richard K. Lester, &  
Michael J. Piore*

*Harvard University Press*

**Gordon Pask's Conversation Theory** originated from a cybernetics framework and attempts to explain learning in both living organisms and machines. The fundamental idea of the theory was that learning occurs through conversations about a subject matter which serve to make knowledge explicit. Conversations can be conducted at a number of different levels: natural language (general discussion), object languages (for discussing the subject matter), and metalanguages (for talking about learning/ language).

As we learned, conversations have been and are till today, the backbone of social interactions. People built places to cultivate them, like the roman and greek did more than two thousand years ago. Early books like the Talmud used the dialectic style to document and further cultural development. The history of conversations and their role in our modern world is the subject of the book „Innovation—the missing dimension“ written by R.K.Lester and M.Piore, two MIT professors.

They conclude that conversations play an important role in the process of innovation. Using case studies, they prove that conversations are essential basis and tool for innovative thinking.

As in any medium, conversation describe “recursive interactions” between people. A conversation establishes connections between individuals with different views on a certain shared topic to learn from others and to gain insights. Leading a conversation is a social skill, differing among the worlds cultural circles.

Conversations take place when people gather, as letter exchanges and nowadays, mostly on the Web. In the early days of “networked communication”, on dialup bulletin boards systems, the usenet’s newsgroups and later on IRC servers, communication was organized in channels or rooms, moderated and curated by an operator. On the Web, it is easy to find people sharing a common interest. Internet forums originated in the mid-nineties and even today, they have a right to exist and serve as a platform for communication. Probably closer to the idea of the conversation theory, which describes a framework that leads to the “construction of knowledge”, as proposed by Gordon Pask in the 1970ies.

Those forums that have a focus on a specific area of interest, are a source for tips and tricks. They are a growing resource of answered questions, build for problem-solving. Through the linear arrangement of posts in a thread, users are able to identify the steps and important pieces that led

to a solution, letting them adapt the provided solution to their own individual problem. Its obvious that the content generated by a smart community is more valuable than any redactional content, or at least, due to the social proof, feels more authentic and valuable.

Forum communities grow around the discussions on shared interests or needs and mostly not on the basis of existing friendships and connections. Forums play an important role in product support. Apple for instance, known as a brand that keeps new products a secret and a CEO who sometimes sends strange mails by night, runs a successful support forum for their users. And they do not do it out of generosity, no—their forums are an important tool for customer research, giving them insights on user behavior and on the satisfaction of users.

[http://www.eltern.de/  
community](http://www.eltern.de/community)

<http://www.motor-talk.de>

Another well working example and proof for the value of clustered conversations as a resource are forums like the forums accompanying the “Eltern” magazine or the “motor-talk.de” forums, serving as a platform for automobile enthusiasts. The latter examples are valuable to advertisers as well by congregating a target group par excellence.

In the business perspective, Charlene Li in her book “Groundswell” states:

[http://www.forrester.com/  
Groundswell](http://www.forrester.com/Groundswell)



***“listening to customers is the best research method a company can do to gain customer or user insights. No focus group or survey lets you collect thoughts and opinions from people when they are in their “natural habitats”.***

***—Charlene Li***

Those discussions and opinions exist within a certain context which provides the researcher with additional information meaning that furthers understanding of individual motivation. Its not design research and the principles of user centered design that help us to identify the problems and glitches users have with products and



*People involved in a conversation  
can be classified as:*

*those that want to add something  
to the conversation (**constructors**),  
those that have a strong contrary  
opinion (**objectors**) and those that  
just want attention (**detractors**)*

*Friendfeed has been acquired  
by Facebook in 2009*

*"The Open Graph protocol enables  
any web page to become a rich  
object in a social graph. For  
instance, this is used on Face-  
book to enable any web page to  
have the same functionality as a  
Facebook Page".*

*<http://opengraphprotocol.org/>*

services but their conversations that started outside of the incubator we create when we start a focus group or a survey. Brands start to recover the value of conversations with users and customers that scared advertisers and marketers to death.

Volatile conversations on today's Web are split up over a dozen platforms and technologies. Even when we use twitter to engage with people in the first step, we sooner or later have to move on to another platform that holds more space for our conversation, creating a gap that is not yet filled.

We have an ever-growing amount of applications, social networks and websites where we can upload items we would like to share. And its not only about lengthy articles, conversations may consist of different formats like videos, photos, a list of links, a series of twitter messages and other "social moments" that we collect using a web application.

Due to the sharing, one of the main concepts that drive Social Web applications, the degree of clutter and granulation takes on a new dimension. We get and create thousands of elements and thereby a bunch of unattached "social graphs"—mapped relationships between us, our friends and all the elements we share across the Web.

Lifestream applications like Friendfeed or lifestream.tv are focussed on collecting and sharing content created elsewhere. The collected data stream can be syndicated itself and read in a newsreader application or can be imported again into another applications.

Facebook's API set "Open Graph" lets, besides a lot of other features, a Facebook user sign in on various platforms using their Facebook identity. Among other things, it allows users to add or invite their existing Facebook friends to the new service or to publish new material on their Facebook page. But the content itself remains in the database of the respective service where it has been originally published. An interesting development concerning the visual representation of these data streams can be observed on the personalized

<http://digg.com/>

start pages of social Web applications. Facebook, Twitter, XING and the brand new redesign of the crowd sourced newsportal digg.com, are following this linear approach, using a recurring visual pattern.

As most people have diverging “networks” on the various platforms, we may, in order to promote the URL of our blog post to the relevant audience, send a short excerpt via twitter and our side-blog on tumblr to our followers and subscribers. We may publish the post in full text as an update on our Facebook wall and last but not least, the post may be syndicated in an RSS or ATOM feed.

Having one piece of information, like our blog post that has been distributed over our networks, means that the conversation it may trigger will be spread the same way.

We get a bunch of independent discussion trees, arguing and discussing the original articles content, but not directly connected to the source or the sources mentioned and elaborated in the article. We create another mass of loose snippets of information.

*“Topsy is a new kind of search engine, with a new way of looking at the Internet. Topsy doesn’t think the Internet is a collection of documents. Or even a web of documents. Topsy sees the Internet as a stream of conversations...”*

<http://labs.topsy.com/about/>

A reader trying to reproduce our reasoning throughout the conversation, is forced to collect the pieces on his own, or due to privacy and our closed networks, even will not be able to access them.

Search engine topsy.com states in their manifesto that they do not see the Web as a pile of data, but as a stream of conversations they are trying to map. The search results topsy delivers are the links that people included in their tweets.

## Comments, Low Level Conversations

Comments were the initial approach to turn linear websites into conversations. A blog post, essay or article on a website shows a recurring pattern. They mostly consist of a headline, a short introduction, the body text which is structured using paragraphs and sub headlines and embedded images. Below the text and ,tools’ to share



and redistribute the content, we find the comments section. Comments are simple linear threads. They do not have a contextual structure. They are nothing more than an unordered list of blocks of text, featuring the commenting text itself, the comment-authors avatar, and the date it was posted.

[http://en.wikipedia.org/wiki/  
Eternal\\_September](http://en.wikipedia.org/wiki/Eternal_September)

*"The web is turning writing into  
a conversation. Twenty years ago,  
writers wrote and readers read.  
The web lets readers respond, and  
increasingly they do—in comment  
threads, on forums, and in their  
own blog posts."*

[http://www.paulgraham.com/  
disagree.html](http://www.paulgraham.com/disagree.html)

In the early days of the Social Web, comments worked quite well. The amount of comments even on populated sites was manageable. A phenomena coined in september 1993 called "eternal september" describes the phenomena that every year in september, when a new class of freshmen arrived at the colleges, the protocols and the netiquette users had to obey on the usenet, had to be taught again. Once a year, the audience reinvigorated. But having this influx just once a year, the situation calmed down after a couple of weeks. But starting in 1993, AOL offered usenet access to all its customers, flooding the usenets chatrooms with tens of thousands of 'freshmen'—not only in september but every single day. This phenomena has not come to an end on the Web and its platforms.

Also commenting is a skill that needs to be trained and needs the commenters to obey conventions, valid in all kind of conversations. Tech blogger and entrepreneur Paul Graham nails it down in an article named, "how to disagree":

It turns out that comments do not scale. In the endless list of comments, we do not find the interesting pieces or the important parts of the conversation. Just think about an update to the article following an elaborated comment or an event that took place since the article was originally posted. Or insights that appear when some bright people pick up the idea explained in the post. There is no hierarchy that clarifies the status of single threads.

An interesting study concerning the value of comments was conducted by Lukas Mathis, a swiss computer engineer. He set up a testing environment to prove that restricting the size of comments will not lead to a higher quality of the comments. His testers rated about 9 000 comments and it turned out that indeed, there is a tendency showing that the longer the comments were, the higher was their relevance and the more likely they were interesting. Where we

[http://ignorethecode.net/  
blog/2009/09/29/  
comments\\_size\\_does\\_matter/](http://ignorethecode.net/blog/2009/09/29/comments_size_does_matter/)

need conversation, externalized thoughts shouts or small gestures are not able to carry meaning.

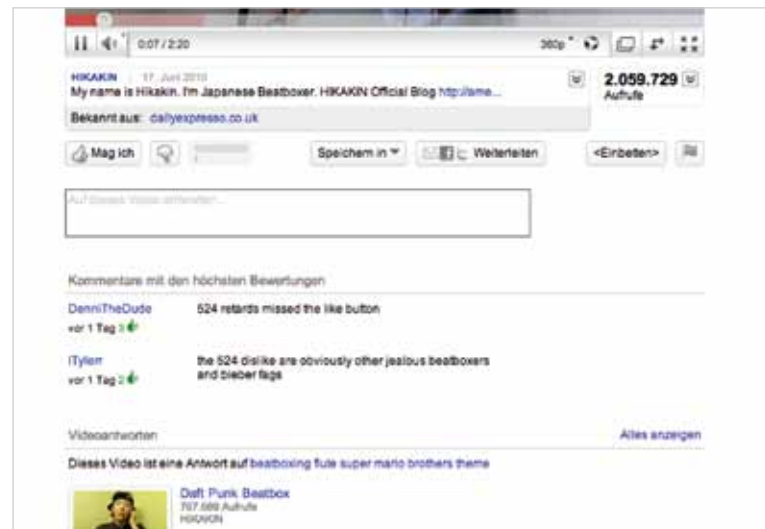
Also huge websites like *Youtube* are working on the comment issue. Their latest redesign hides or moves the comments down the page, featuring only the two best rated comments by default. And that is a major glitch:

*a brilliant rant on  
Youtube comments:*

<http://xkcd.com/202/>

Voting & Ranking form elements  
and the two promoted comments

[http://www.youtube.com/  
user/HIKAKIN](http://www.youtube.com/user/HIKAKIN)



Always promoting the most rated comments will make them an even more likely choice and keep them in a position where they get most attention.

## Comment Services and Solutions

Another aspect of comments is that they are deployed to serve as a system to show bloggers that their writing is actually read and that the content they write and produce, matters (at least in some way). Regarding the theory Paul Graham develops, it does not quite make sense. As soon as the first blogging software solutions were introduced, another approach on reader feedback occurred.

The so called "linkback", "trackback" or "pingback" systems. Integrated as a plugin for blogging software, they "ping back", when other websites or software linked back to the

original article, giving a short notification (requiring both communicating sites to have the plugin installed and to communicate with a server) when an article is mentioned elsewhere.

*"a guide to constructive  
commenting"*

<http://www.geeknewscentral.com/2008/03/31/a-guide-to-constructive-commenting/>

What started as a good idea to enable the communication among websites became another spam sling, requiring "anti-trackback-spam" filters. Many blogs and blogging software developers have stopped using them because it became too much of a burden.

Just ask someone who maintains a couple of websites based on wordpress if he would rather switch to a service that takes care about all the spam/security matters or spend hours and days hacking. Most system would include a list of trackbacks, similar to comments, below the article or would even combine them with comments.

Techcrunch calls  
trackbacks "Responses"

#### Responses

- [Steve Jobs Loves the Tablet. Duh.](#)  
January 24th, 2010 at 4:43 pm
- [Jobs: Tablet "Most Important Thing I've Ever Done" | Tasty Slate](#)  
January 24th, 2010 at 5:00 pm
- [Steve Jobs considera a Tablet a coisa mais importante que já fez. « Apple Addicted](#)  
January 24th, 2010 at 5:20 pm
- [iDrive Online Backup » Steve Job's best creation - Coffee Table Gadget?](#)  
January 24th, 2010 at 5:20 pm
- [Steve Jobs over de Tablet "This will be the most important thing I've ever done"](#)  
January 24th, 2010 at 5:40 pm
- [Steve Jobs on Apple Tablet: "Most Important Thing I've Ever Done"](#)  
January 24th, 2010 at 6:01 pm
- [スティーブ・ジョブズがタブレット機について語る「生涯で一番大事な仕事」 | ギーク流英語活用術](#)  
January 24th, 2010 at 6:03 pm
- [Overheard: Steve Jobs Says Apple Tablet « My Rubicon](#)  
January 24th, 2010 at 6:26 pm
- [Steve Jobs afirma que la tablet será lo más importante que ha hecho en la vida | Appleismo](#)  
January 24th, 2010 at 6:52 pm
- [Jobs diz sobre o tablet: "Esta será a coisa mais importante que já fiz" « Blog MacManiacos](#)  
January 24th, 2010 at 7:16 pm

[View All Responses](#)

Handling comments does not only mean sorting out spam but listening to your audience and engaging with them takes its time. Jon Hicks, blogger at [hicksdesign.co.uk/](http://hicksdesign.co.uk/) says: "As a discussion system blog comments feel broken" when he explains why he has disabled comments on his blog.

Besides the technical and spam related problems, many bloggers feel the need to be able to have a spot on the Web that they own. Where they do not have to justify



} article height  
600 px

} comments  
height  
22 980 px

screenshot height  
24 920 px

22 980 px means  
scrolling down  
30 screens on a  
common notenbook  
display

example taken from  
<http://techcrunch.com>

themselves for their writing or the way they interpret a topic. The Web as a place that offers space for anyone to unfold personality and ideas, sometimes is not willing to accept the freedom of individuals.

“

*“I turned off comments in the last redesign of powazek.com because I needed a place online that was just for me. With comments on, when I sat down to write, I’d preemptively hear the comments I’d inevitably get. It made writing a chore, and eventually I stopped writing altogether. Turning comments off was like taking a weight off my shoulders. It freed me to write again.”*

—Derek Powazek

Innermost ownership of data and ideas plays an important role in social software development. We design apps to help people feel ownership over the content they create and share. That is why social Web applications like Myspace, Flickr, Facebook and others continually use both “me/my” or “you” to attribute to the content or sections of the page. The „Endowment Effect“ describes this phenomena that people tend to value things more once a sense of ownership has been established.

Some companies jumped on the comments bandwagon and developed something they consider a solution by porting the comments into a closed system. We will have a brief look at them on the next pages:

### The Disqus Comment Platform

The developers of “DISQUS” state “We’re big believers in the conversations and communities that form on blogs and other sites.” They offer external hosting of comments for websites. They enable you to link the conversation to search- and tracking engines like Backtype and connect it with the search results on Facebook, Twitter, Friendfeed, Digg and Youtube. “Bring lively, real-time chatter onto your page with real-time posting and updating. With rich interac-

[www.disqus.com](http://www.disqus.com)

the DISQUS dashboard showing latest comments, replies to previously posted comments and comment subscriptions

---



tive features such as threaded replies and comment replies, your discussions never feel dull..” And again, its just about speed, chatter and entertainment, not about substance and content. Even though their interface comes with a “see in context” button, they do not provide the original context of a comment and all parts of the conversation.

## Echo

<http://aboutecho.com/>

Echo is another tool to manage comments on website. They serve huge business websites like [www.time.com](http://www.time.com), [forbes.com](http://forbes.com), the blogsearch tool [technorati](http://technorati.com) and [Forbes.com](http://Forbes.com), to name a few. But although they pretend to care about data portability, its all about feeding the same old business model “Increased Engagement” in their corporate language means adding videos and photos to your comments. “Increase Time Spent” means increasing the time users spend on the site to perceive advertising.

## Tumblr Sharing Platform

<http://www.tumblr.org>

Tumblr is a posh blogging platform, able to handle any kind of media. A tumblr blog can be fed using a phone call, using a text message, email, bookmarklet or using the web interface. tumblr never supported comments. Since their launch nearly 3 years ago, the tumblr engineers are constantly trying to improve the conversation experience among blogs running on the tumblr platform. They offer a set of tools to engage communication among authors and readers. I will briefly summarize some of the ideas:

**reblogging** Create a quick copy of a post of a foreign weblog and paste it in your own. That addresses a common practice to cite a paragraph from an article published on a foreign website. Thereby bloggers create duplicates of paragraphs, creating a hardcopy of a single paragraph that might be taken completely out of context.

**likes** Express your liking, generates a short “username likes this” below the post.

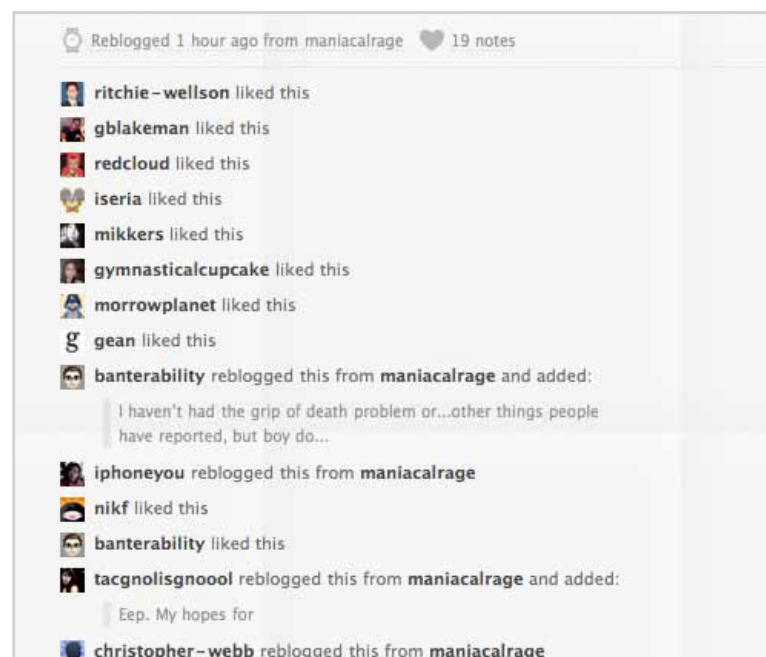
**submit** Lets registered users submit content to a blog.

**replies** Nice approach, connects two posts on different weblogs with each other.

**questions** Lets users ask questions, similar to what the “formspring.me” web service offers. But like all other services, tumblr keeps the data in their closed system. If you do not run a tumblr blog, you will not enter any conversation.

This is the appendix that tumblr attaches to an article that has been liked, reblogged...

---



Some bloggers are currently testing twitter as a comment platform but again, its not a solution but another layer of information added, attempting to fix the mesh.

Another interesting attempt that substantiates the need for elaborated conversations and discussions among website is squarespace, a hosted full service platform for publishing websites.





The Squarespace article post interface provides a set of options to include references and related content. It accepts URLs of related posts or websites, a source article and short excerpt. The comment tool allows the commenter either to post a “classic” comment above the content he wants to comment or to use their own trackback system to post an answer on their own platform.


an article on Squarespace with  
the differentiated references and  
comments


## the reference tool

WEDNESDAY, JULY 7, 2010 AT 05:49AM

You can drag and drop anything on your site. Want to move a page link? Enter structure mode and drag the page link to a new location. You can even drag links between columns or from a column on to your top navigation bar. Need to edit a journal entry? Enter content mode (the big 'A') and click 'modify' on the entry you'd like to edit.


 [Your Name Here] |  Post a Comment |  3 References |  Share Article


 View Printer Friendly Version


 Email Article to Friend

### References (3)

References allow you to track sources for this article, as well as articles that were written in response to this article.

 **Response: Press the Magic Button**  
by Derek Powazek on June 28, 2010  
My "one strike" rule for Twitter/Flickr and why you shouldn't be offended when someone breaks it on you.

 **Related: Songfever**

 **Related: der Spiegel**

### Reader Comments

There are no comments for this journal entry. To create a new comment, use the form below.



*An open source approach to  
collaboratively annotate, edit and  
“shift” the web is shiftspace:*

<http://www.shiftspace.org/>

## Googles “Sidewiki”

Google’s Sidewiki is another approach to enable conversations and discussions on websites. Google promotes sidewiki asking the potential users „What if you could learn from others who have visited a page before you?” What sounds like an universal attempt to fix usability or user experience on just any website, is just an annotation and discussion tool that offers its services as well on those websites that do not support comments. All sidewiki entries, provided in a browser sidebar, are collected on the users Google profile page and can be forwarded and published on Twitter, Facebook and blogs hosted at Google’s own Blogger service.

## Handling & Organizing Conversational Data

The web is growing at an enormous speed. Eric Schmidt revealed at activate2010:

“

*„From dawn of our civilization to 2003,  
5 exabytes of data were created. The same  
amount was created in the last 2 days.“*

*—Eric Schmidt*

[http://twitter.com/iRowan/  
status/17507179613](http://twitter.com/iRowan/status/17507179613)

Today’s web and the search algorithms will scale to that amount of data only to some extent. We generate far more information than we are able to handle. Due to the structures mentioned before, we get an exploding number of access points to our data. Features to aggregate our content and distribute it to other platforms as well as importing data from external databases is an ever growing multiplier.

Due to the loose structures and the massive amount of data, users have started to use systems to take control over data. They use tags to create a taxonomy to keep data under control and keep track of the increasing number of feed flow rates. But with another system, as error prone as the data structure itself, we will not solve the problem of complexity and access. The talmud and the MEMEX, as any



collection of data, struggle with categorization of the single elements. Think about huge libraries like the Library of Congress in Washington or the Russian State Library in Moscow, both storing about 42 million pieces. In order to create hierarchies to keep these works accessible and categorized, librarians used the metaphor of the shelves where the books were sitting in. In the same way as a shelf is limited in its capability to store books, the shelf-like hierarchy is.

A nice example for such a system on the Web is Yahoo!'s former catalog-based search engine:

*List of items on the shelves were  
(external) metadata, existing  
meta = alongside the books.*

"A little over fifteen years ago, a couple of guys out of Stanford launched a service called Yahoo that offered a list of things available on the Web. It was the first really significant attempt to bring order to the Web. As the Web expanded, the Yahoo list grew into a hierarchy with categories. As the Web expanded more they realized that, to maintain the value in the directory, they were going to have to systematize, so they hired a professional ontologist, and they developed their now-familiar top-level categories, which go to subcategories, each subcategory contains links to still other subcategories, and so on. Now we have this ontologically managed list of what's out there."

The guys at Yahoo! were overwhelmed by the growing amount of sites that had to fit in the respective categories. They hired an ontologist to create the top level categories. During the refinement of the categories, they came up with an "@" as a suffix to categories which indicated that the prefixed subcategory is not really part of this category and therefore just a link to the "right" place. But because the "information architects" at Yahoo! knew due to user testing and their own expectations that people expect it in the respective category.

(As an example: Although you would expect to find "Books" in the "Entertainment" category, it's located in "Arts/Humanities" and therefore, in the "Entertainment"-category, you find it suffixed with an "@" as "Books@".

[http://www.nielsen-online.com/  
pr/pr\\_060330.pdf](http://www.nielsen-online.com/pr/pr_060330.pdf)

That is one reason why Google's search engine was so quickly adopted in the 2000's by a majority of users. Google did and does not use a metaphor for the

representation of search results or for predicting what we are looking for when we run a query on the website. They did not have to predict the future of what future objects will have to fit into any categories. In retrospect it's kind of interesting that it was *Yahoo!* who bought the social bookmarking service *delicious.com* in 2005. Delicious was the service that promoted a non-hierarchical classification system using tags using freely chosen index terms.

*For the content users create in the Social Web, hierarchies work only to some extent.*

With the concepts forming the Semantic Web we will see previously undisclosed solutions tackling this issue mankind is dealing with for ages.

## Linked Data and the Semantic Web

*Resource Description Framework  
in attributes*

Both, the Web of hypertext and the Web of data are built of documents. The Semantic Web consists of documents as well. But its idea goes beyond just connecting documents using hyperlinks. It is not about connecting whole documents using hyperlinks but about links between data elements, described using RDFa or other any other metadata description framework. Linked data leads us to related data and thereby weaves a net that is both understandable by humans and machines. Instead of document anchors, RDF is used to describe the relationship between things, identified using URIs.

“URI” stands for “Uniform Resource Identifier” and describes a string of characters that identify a resource over a network according to the related protocol. A nice example of mapping an URI to an object is the way the FOAF protocol gives URIs to persons, using the “person” class.

Identifying data with URIs and serving information against a URI alongside with the RDF triplets, are the major component of the Semantic Web.

Making connections between data is as essential for weaving the Web of data as it is on the Web of documents. Value of information is very much a function of what it links to, what points toward it and its included information.

Browsing linked data using for example a query against a node in the RDF graph, the server would return the outgoing and oncoming arcs of the respective node, frankly all the RDF statements in which the queried statement appears as either subject or object.

In order to find the correlation between two pieces of linked data that is spread over documents, the respective statement must be noted in both documents.

TBLs vision of a Web of linked data is as old as the first incarnation of the www. In his book "Weaving the Web", published in 1999 he declared:

Tim Berners-Lee

<http://www.w3.org/DesignIssues/LinkedData.html>



***„The first step is putting data on the Web in a form that machines can naturally understand, or converting it to that form. This creates what I call a Semantic Web – a web of data that can be processed directly or indirectly by machines.“***

**– Tim Berners-Lee**

In order to identify the properties of our elements and to follow links, we need a common vocabulary. Different approaches and opinions toward that emerged:

*for a short best practice guide,  
see "how to put your data online"  
in the appendix*

The Semantic Web had a huge problem catch on and to find supporters who implement and push the technology over the last couple of years. The "top-down" approach was and is mostly driven by the academic world, focussing too much on the "semantic" part and not enough on the 'Web'. The "Linked Data" faction is more down to earth, more pragmatic. But the developed technologies like RDF and the "Web Ontology Language" OWL are valuable and a vital part of the concept.

Computer scientists who have dedicated their work to the Semantic Web, are leading a curiously long discussion about ontologies. Ontologies are well thought “taxonomies” that describes the relationships between thing more detailed.

In this ongoing discussion its all about structures. It seems that there are two types of ontologies around. Whereas one faction’s “ontology” means a “explicitly recorded, machine readable description of the ontology of data”, the other faction’s interpretation is more lively, referring to an „implicit“ structure all data comes with. (Like a spreadsheet which is defined by rows, columns cells and their obvious relationship between them.)

<http://groups.csail.mit.edu/haystack/blog/2009/11/03/does-the-semantic-web-need-ontologies/>

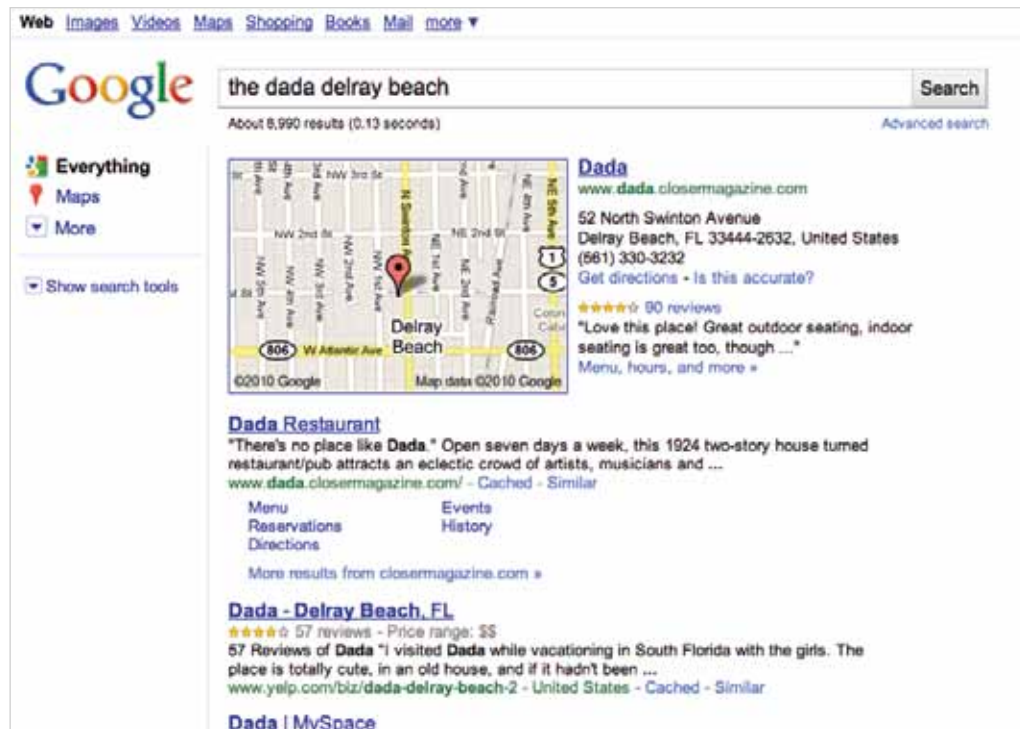
Folks like David Karger, a professor for computer science and AI at the MIT in Boston, omit that explicit ontologies are more focussed on human beings and the fact that humans eventually consume and process the data.

*“An ontology is a formal specification of a shared conceptualization”*

The Semantic Web in the same way as the World Wide Web, do both share the same background in science and engineering. It’s obvious that it may take its time and an down to earth approach to succeed in a Web thats populated many users who do not feel a need to store scientific data.

TBL calls the success of the “Linked Data” movement a “grassroots success”. The DBpedia project, driven by the Freie Universität Berlin and the Universität Leipzig is one of those “grassroot” elements. The DBpedia knowledge base accesses the wikipedia knowledge base as linked data on the Web and serves itself as a central data hub on the emerging web of data.

Although the Semantic Web is great in theory, we are still not seeing it in large use. But in the last twelve months, some of the big players on the Web begin to push some implementations forward. Search engines Google and Yahoo! display properly marked up metadata using RDFa



search results for  
“the dada delray beach” enriched  
with RDFa elements

in the search results, featuring a set of almost “tangible”  
pieces of data. Interestingly both search engines focussed  
on information that’s user generated.

Among the first implemented data types that Google  
includes in their search results are:

**personal and business data** like friend, contact and  
acquaintance properties to identify social relationships.  
(Using XFN, including name, nickname, photo, title, role,  
affiliation or FOAF and DOAP, to name a few.)

**Ratings** for restaurants, movies, videos, products etc..  
including authors name, item, rating, date and a summary.

**Events** including start and end-date, location, photos and  
geodata.

<http://googlewebmastercentral.blogspot.com/2009/05/introducing-rich-snippets.html>

Google calls their approach “rich snippets”. Besides RDFa  
they are aware of data marked up with Microformats and  
the html5 sibling microdata. Microformats are even a step  
closer to xhtml, using existing tags and css selectors to  
mark up data. They support a set of formats similar to RDFa

formats and are continuously extended.

Since the launch of the rich snippets in June 2009, the amount of rich snippets impressions grew four-fold, as Google's Pravir Gupta stated at the Semantic Technology Conference in San Francisco in June 2010.

*An example using the "hCard"  
Element created to mark up  
address data on a website.*

```
<div id="hcard-Sebastian-Gerhard" class="vcard">  
  <span class="fn">Sebastian Gerhard</span>  
  <div class="adr">  
    <div class="street-address">Lichtenbergweg 3</div>  
    <span class="locality">Griesheim</span>  
  </div>  
</div>
```

Besides Google and Facebook following their own business objectives, an interesting ontology for e-commerce websites emerges. The "GoodRelations" ontology describes itself as follows:

“

*“GoodRelations is a standardized vocabulary for product, price, and company data that can be embedded into existing static and dynamic Web pages and that can be processed by other computers. This increases the visibility of your products and services in the latest generation of search engines, recommender systems, and other novel applications.”*

<http://www.heppnetz.de/projects/goodrelations/>

A very small branch of the idea of the Semantic Web, again with a very down to earth approach. Delivering value to both search engines and customers to exchange data across platforms and services, making database-based services for comparing prices obsolete. Search engines get their right to exist their business model from the "chaos" that we generated on the web and are therefore not keen to reinvent themselves when we move to a structured Web.

## Defining a Social Graph: FOAF & XFN

“

*“The Web is more a social creation than a technical one. I designed it for a social effect—to help people work together—and not as a technical toy. The ultimate goal of the Web is to support and improve our weblike existence in the world. We clump into families, associations, and companies. We develop trust across the miles and distrust around the corner.”*

*—Tim Berners-Lee, Weaving The Web*

As mentioned earlier the, Semantic Web and the ideas and expectations that fuel it, suddenly feel familiar. People start to recognize that real life relationships between people and objects and tasks that connect them, making the Semantic Web the reason of what insights the web 2.0 brought us.

Social relations and interactions can be displayed and described in the same way any other type of non-personal data can. Interactions with social objects, the elements social web applications are taking care of, are described using objects and actions—the same grammar we use in RDF triplets.

The *FOAF protocol* (friend of a friend) started in mid-2000, can be considered as the first social Semantic Web app. It is a machine-readable ontology, describing persons and their social network in a semantic way, without any centralised database. The current iteration of FOAF Again, FOAF focusses on people, since people are the things that link together most of the other things that link together most of the other things we create and describe on the web. FOAF is based on RDF, allowing the vocabulary to be used in combination with any other RDF vocabulary.

See an example of FOAF data below that describes me:

*an example foaf.rdf file that should  
be stored in your websites root  
folder*

```
<rdf:RDF xmlns:foaf="http://xmlns.com/foaf/0.1/">

<foaf:PersonalProfileDocument rdf:about="">
  <foaf:maker rdf:resource="#me"/>
  <foaf:primaryTopic rdf:resource="#me"/>
</foaf:PersonalProfileDocument>

<foaf:Person rdf:ID="me">
<foaf:name>Sebastian Gerhard</foaf:name>
<foaf:title>Herr</foaf:title>
<foaf:givenname>Sebastian</foaf:givenname>
<foaf:family_name>Gerhard</foaf:family_name>
<foaf:nick>Woy</foaf:nick>
</rdf:RDF>
```

This snippet should be included in a "foaf.rdf" file on the respective web server where the Google bot can find it. But retrieving and finding other users' foaf.rdf files is still in discussion.

One suggested solution is to provide a link pointing towards it:

```
<link rel="meta" type="application/rdf+xml"
title="FOAF" href="foaf.rdf" />
```

Weaving the social network therefore means doing it by hand, adding descriptive information about relationships to the links.

The XFN standard „puts a human face on linking“. The profile includes attributes that describe the authors personal relationship to the individuals behind the links. It uses the "rel" attribute to add details about the relationship between the two persons interlinking their websites.



After referencing the profile in the html documents <head> section using

```
<head profile="http://gmpg.org/xfn/11">
```

one can use the referenced „rel=“ elements:

```
<a href="http://phoboslab.org/" rel="met  
friend">Dominic</a>
```

```
<a href="http://henrifrederik.de/"  
rel="child">Henri</a>
```

```
<a href="http://stevevoss.de" rel="met  
colleague">Steve</a>
```

The creators decided to use existing markup instead of an new xml based syntax to keep the solution simple and close to the infrastructure of the web. They kept it easy to hand-author and both machine- and human-readable.

[http://code.google.com/intl/  
de-DE/apis/socialgraph/](http://code.google.com/intl/de-DE/apis/socialgraph/)

The “Social Graph API” designed by the founder of the blogging/social networking platform “Liverjournal” Brad Fitzpatrick, combines FOAF and XFN into a powerful API that serves as an information “broker” between social network sites. The core is an open source database that accepts informations from multiple social networks and then provides it so end-users via the API or an dedicated interface.

The database does also grant access to profiles on the social network sites. The problem is that Facebook for instance, would never grant full access to their database. The author’s suggestion is that today’s leading social applications should offer an export feature that lets people take their data with them, when they move along to another service.

### **The Facebook Way: Open Graph**

Let us now have a look at Facebook and their Open Graph API. Facebook made a giant leap ahead regarding their use of metadata. Their Open Graph API is all about publishing and sharing snippets of data, forming the most impressive marketing tool ever.

<http://www.leveragingideas.com/2010/05/04/the-rise-of-the-social-gesture/>

In order to make every webpage, also those that not part of Facebook “likable” and shareable, the Open Graph API offers publishers tools to mark objects of likes and interests, movies, bands, venues, books, etc. up as objects, telling Facebook what is on the single pages. By collecting this information, and that is what they do, they create a huge library of peoples tastes and likes. This could become a major threat on Amazon’s dominance over the recommendation/review space by turning the whole web into such a system by using social gestures as “likes” on the web.

They start to build bridges between services that are at their core still data silos, but when a band, let us say “*Broken Bells*”, is marked up properly on websites like last.fm and spotify, will be identified both as the same “like” and pulls it out of the silo structure.

We do not only tell Facebook a lot about us, but we hand them over a large stack of data that build a very detailed profile of our interests. Whereas the technological solution is kind of elegant, one can argue against Facebook’s master plan, namely striving to become a huge semantic search engine that is primarily furthered not by Facebook itself but the publishers using Facebook’s social widgets and their proprietary “like” gesture.

We will see in how far publishers all over the web will be willing to give up ownership of data and how all those platforms and social networks that are focussed on building communities around the „likable“ items like music, photos and bookmarks are willing to cooperate (what means, in plain english, to get out of business), when Facebook unifies all platforms, being the only repository of the data.

*opengraph notation*

```
<html xmlns:og="http://opengraphprotocol.org/
  schema/">

<head>
  <title>Contact Me</title>
  <meta property="og:type" content="person" />
  <meta property="og:street-address"
    content="Lichtenbergweg 3" />
  <meta property="og:locality"
    content="Griesheim"
  <meta property="og:image"
    content="www.example.org/sgerhard.jpg" />
  ...
</head>
...
</html>
```

Linked data confronts us with a paradigm shift all over again. We need to adopt a new vision of open data that go beyond the superficial debate on privacy and copyright. The linked data design will someday seriously affect today's common business models.

Conversations on the web consist of the correlations between a written piece and its context, framed by some well thought comments and the referenced resources. Thereby we establish a complex system of trails and nodes as Vannevar Bush dreamt about. Let us have a look at an elegant solution that can be implemented today, using a set of standards that lets it interoperate with common platforms.

## **My approach to reinvigorate conversations on the web**

In order to re-enable fruitful conversations on the social web, we have to cope with a lot of issues and concerns.

### **Let us summarize what research and user testing brought up**

Conversations are driving innovation and are a valuable resource to get insights in processes and overall development. The fact that the two or more conversationalists engage in the conversation and going into it, gives them a lively character.

Conversations that help to clarify or elaborate an issue mostly take place in forums where they are threaded and easy to find. The thread lets users follow up a conversation that took place a while ago and shows the approaches to a problem. Especially in the realm of customer relationship management, connected conversations are very valuable.

*CRM is all about conversations*

Articles on websites, pieces where the authors took time to write down ideas, doing research beforehand, are nowadays only connected using hyperlinks that do not clarify why and how two loosely connected articles belong together. They are not threaded and force the user/reader to follow and rebuild the string of the conversations. Furthermore, the ever emerging new platforms split off our graph and thereby our conversations. Our data is encapsulated and our identities can not be reused on the majority of platforms.

Common business models on the web influence the design of modern websites and do not further the user experience. Publishers create obstacles that users/readers have to cope with when they want to consume their content. Human curiosity and the physiology of the eye collide with the way we consume hypertext documents and the included elements.

This thesis proposes a framework that lets us connect valuable conversations on the web, alongside with interesting resources and—most important—the people who are involved in the conversation.

The technologies that let us build a net of data are available, but not widely adopted. “Hybrid iterations” like RDFa and microformats that intertwine concepts like RDF and xhtml, act as a door opener and make people aware of the benefits of linked data. That happens by using common infrastructures and interaction patterns. The surface still feels like what we are used to, but the technology under the hood is new.

Most people, even though some of those who dedicate their work to the Semantic Web/Linked Data, do not have an idea how we will interact with data.

So this approach is focussing on text, externalized thoughts that are actually related to each other, like many other real world objects that give the semantic its tangible character.

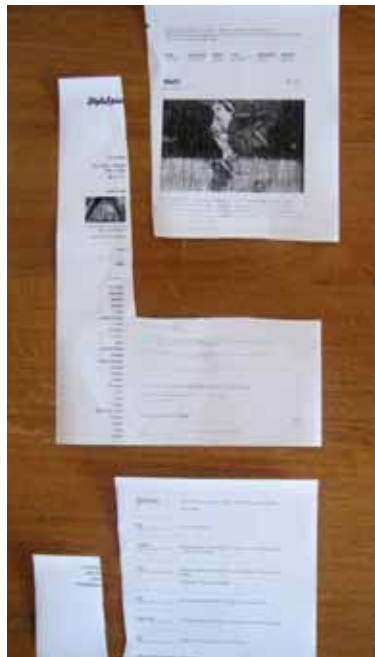
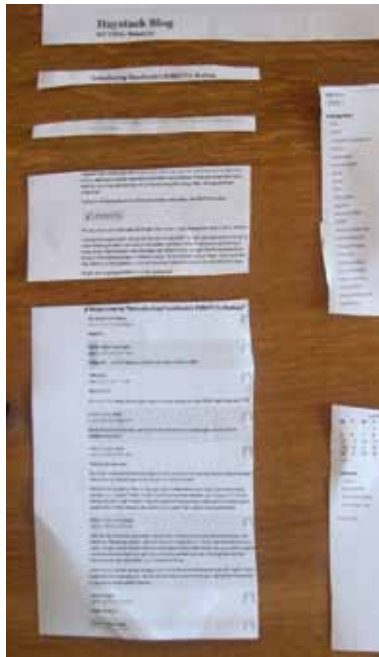
Identifying the patterns and embedding in an interface that is crafted entirely to adapt to today’s practices on the web is what shaped the application.

### **Identifying the element and defining a vocabulary**

After the user testing in the research phase, I used a paper prototype to let test subjects identify the elements of conversation on websites. One initial idea was to break the conversations down into the single paragraphs. Paragraphs are the elements most writers go back to and refer to when they reply to a post or when they use a citation to underpin an argument.

*see test results on the following pag*

*Interestingly, test subjects who are  
writing texts on a regular basis,  
granulated the printed articles more  
fine.*



cuttin test results. Websites used  
from left to right:

---

[http://groups.csail.mit.edu/  
haystack/blog/](http://groups.csail.mit.edu/haystack/blog/)

---

<http://www.stylespion.de>

---

<http://www.zeldman.com>

---

What the videos and the paper prototyping showed, is that conversations on the web are already very granulated and fragmented enough. Although citing or copying a paragraph into one's own post may already skew context and meaning. I is not beneficial to understanding and internalization of the content.

*more on the test's results on page 50*

**The important elements needed to be mapped onto a vocabulary are:**

**Autor, Website, Article, Date, Topic** and **References**, the latter bidirectional, both referenced and „referenced by“.

Initially I thought about developing a simple but strong vocabulary for conversations. But as many attempts on the web that aim to unify processes like sign-ins or whole identities like oAuth and OpenID show, it is more desirable to use existing tools and structures. Every movement needs a sufficient momentum known as the critical mass to succeed and spreading peoples attention over a dozen of good but related things will not be successful at all.

Keeping in mind that many conversations happen on Facebook, I investigated the possibilities of the opengraph vocabulary. It seemed quite obvious to chose their vocabulary to keep the doors open. Being able to join a conversation that started on Facebook and is continued on any other website would be exactly what the concept intended.

But the standard opengraph protocol's vocabulary is limited when it comes to conversations between people. Most of the elements are made to describe products to "feed the like machine".

The aforementioned XFN protocol and the FOAF vocabulary seemed appropriate to create an URI for the author and to describe links that we call "blogroll", a list websites that an author recommends.

[http://dublincore.org/  
documents/2000/07/11/  
dcmes-qualifiers/](http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/)

When it comes to describing the article, I opted for the *Dublin Core* standard. Dublin Core, named after Dublin/Ohio in the midwestern USA, where the first working group met in 1995, was designed as set of metadata for libraries. Dublin Core has a basic set of 15 elements and extending qualifiers that are continuously maintained and adapted. The Dublin Core vocabulary can describe any kind of media like video, image, sound, text or web pages. For setting the relations between two articles the dc:relation element was chosen. The qualifiers "*isReferencedBy*" and "*references*" are added to the core element, creating the following syntax: *dc:relation.isReferencedBy*.

Dublin Core combined with RDFa gives us a very powerful tool to mark up our content with metadata. RDFa is due to the fact that it is embedded in html very comprehensible. Metadata can be added in line with the content—making it easy to understand relations.



After referencing the respective prefix mapping in the html header or parent element

```
xmlns:dc="http://purl.org/dc/elements/1.1/"
```

we can start to include the metadata:

```
<h3><span property="dc:date">2010-09-15  
  </span> &mdash;<span property="dc:creator"  
  content="Sebastian Gerhard">posted by  
  sgerhard</span></h3>
```

```
<h2><span property="foaf:primaryTopic"  
  content="conversations on the web">My take on  
  conversations on the web</span></h2>
```

```

```

```
<p>Handling comments does not only mean to sort out  
  spam but listening to your audience and engaging  
  with them takes its time. <a href="http://www.  
  hicksdesign.co.uk/journal" property="dc:relation.  
  references" content="http://www.hicksde-  
  sign.co.uk/journal">Jon Hicks</a> says: "As a  
  discussion system blog comments feel broken" when  
  he explains why he has disabled comments on his  
  blog.  
</p>
```

All attributes added using RDFa relate to the element they are attached to, always in the context of the parent. If we have a website that features multiple articles or authors, we can use the *"@about"* to set the context to specific elements.

```
<div about="/blog/articles/first_article">  
  <h2 property="dc:title">I started a blog</h2>  
  ...  
</div>  
<div about="/blog/articles/second_article">  
  <h2 property="dc:title">first week recap</h2>  
  ...  
</div>  
  
</html>
```

We can use the RDFa's @typeof attribute to declare new data items like persons listed in our blogroll:

```
<div xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <ul>
    <li typeof="foaf:Person">
      <a property="foaf:name" rel="foaf:homepage"
        href="http://example.com/bob/">Bob</a>
    </li>

    <li typeof="foaf:Person">
      <a property="foaf:name" rel="foaf:homepage"
        href="http://example.com/eve/">Eve</a>
    </li>

    <li typeof="foaf:Person">
      <a property="foaf:name" rel="foaf:homepage"
        href="http://example.com/manu/">Manu</a>
    </li>
  </ul>
</div>
```

We have created 3 data object of type "person" and have added properties and values for name and the links to individual websites.

This data structure is simple but powerful. Due to the nature of RDFa, its both machine- and human readable and reflects the superficial data structure of online communication we are accustomed to.

## Crafting an interface for data

Given this data structure, lets have a look at he interface for our data. How can we combine the character of linked data, the consumption behaviour and the human factor in a system that addresses the issues we elaborated earlier?

The visual structure and all major design decision are drawn from the results of the research process. Each element's purpose is transporting the data structure throughout the layers, furthering the understanding of the content and its structures. Think about the concept of RDFa—combining the initial information with metadata. In the same way the

metadata adds meaning to the content, the visual design can clarify semantic coherences and semantic markups to the information. In this thesis' context, typography serves as a suitable tool to do that.

Reading and understanding text is something literate people do on a daily basis. The mental models on how to process and manage it, differ slightly, but the cognitive processes are the same. One idea that came up during research and investigation of documented conversations was to enhance interaction with them.

But the user testings and video sessions left no doubt that linear conversations should stay linear. Any kind of unexpected behavior distracts users and hinders them to get in touch with the material. The recursive character of conversations needs to be reflected and usability terms like "compliance with expectations" and "learning curve" are especially important, because following conversations is something people are used to, in face-to-face conversations, on the phone and in many ways on the internet.

[http://www.useit.com/  
alertbox/9710a.html](http://www.useit.com/alertbox/9710a.html)

In order to verify the assumptions on structures and elements of the content we are dealing with in this thesis, five testers were each handed scissors and three printed articles found on the web.

*see images on page 46*

They were asked to dissect the printouts and to label them according to their expectations. Only three of five testers said they read blogs, the other two did not exactly know what "blog" means, but all five dissected them in the same way. The two testers who had never encountered a blog before said: "Well, it's just like an article in a magazine, each with a list of letters to the editors attached."

Furthering and using strong metaphors and mental models is a powerful method to reduce barriers and to help people learn how to use a system. People do have "internal understanding" how written text "works" and therefore breaking with this mental model is not recommended.

[http://developer.yahoo.com/  
ypatterns/](http://developer.yahoo.com/ypatterns/)

[http://www.flickr.com/photos/  
factoryjoe/  
sets/72157604068027525/](http://www.flickr.com/photos/<br/>factoryjoe/<br/>sets/72157604068027525/)

Interaction design patterns, emerged from recurring problems and solutions in a comparable context, transfer these discrete mental models to the web. Designers have libraries of established interface patterns at hand that are a valuable resource when the *nouns/objects and verbs/actions* that shape an application need to be translated into an user interface. These libraries help to establish de-facto standards for web interfaces.

A design practice that has risen to popularity are widgets. Widgets are small chunks of code that can easily be integrated in any html document. The technology is based on html/javascript, or sometimes flash. Among the most popular widgets are those that display content from social networks like Facebook, Flickr, Dopplr and others. They embed your photos, friends or planned trips in your personal website.

But: The idea behind these content centered widgets “works” hand in glove with the way the publishers store and handle data: encapsulated and closed in their data silos. So these widgets are just a interim or transitional solution, bridging the gap between closed data and the ease of embedding content in any website, also for those who do not know a single line of html or javascript.

*some popular widgets*



trivadvisor widget on Facebook

Facebook widget displaying fans of a certain group on a non-Facebook page



Digg widget displaying content related to mashable on any other site




Another kind of widgets are the aforementioned widgets or interfaces that encourage users to share content.

And that's how I propose to implement the "linked conversations" plugin on websites that are about to join an open conversation. Sharing means connecting to people and spreading the word about something. And that's what we are about to do.

---

the textfaces widget

This post joins a 

## The prototype

For demonstration purposes, I mocked up a small weblog. It features the standard pattern, as elaborated before: a blog title, the articles title, the date the article was posted, the author's name, a sample image and some paragraphs of text, including a quotation from a referenced site. All elements are marked up using dublicore & friend-of-a-friend vocabulary.

I propose to embed a piece of javascript in the page that shows the "conversation" bubble as soon as the script detects RDFa data in the page. Clicking the bubble brings up a hud on top of the website. Hud means „head-up-display“, presenting information „in place“, not requiring the user to look away from his or her usual viewpoint. The hud display goes along with the design concept mentioned earlier:

The surface is all well known technology and infrastructure, plain html pages like any weblog, known for more than a decade. The embedded metadata lives on another layer, invisible to users. When the hud is brought up, we make these extra pieces of information visible and comprehensible. And not turning by them into graphical objects that create mere infoporn but by creating obeying the rule of the information system we are dealing with.

### infoporn

*Sexy and visually appealing visualizations of data, often using "Web 2.0" interactivity to suck the viewer in."*

<http://www.urbandictionary.com/define.php?term=infoporn>

Instead of mocking up a vision of user agents that aggregate information according to a profile, we just put together those pieces of data that belong together but had no relation due to the way we connect information across documents on the web.

The hud, alongside with the metadata, features three buttons. Besides "close" to dismiss the hud, we can "load the conversation" or "send conversation to instapaper."

the hud, crafted using html5 and css3 elements. Only the icons are pixel data.

blog comments feel broken" when he explains why he has disabled comments on his blog.

**Follow the conversation**

Besides the technical and spam related problems, many bloggers like Sebastian Gerhard be able to have a spot on the web that in "My take on conversations on the web"ify themselves for the 2010-09-15 way they interpret a topic. The web as a place that if anyone is willing to unfold personality and ideas, sometimes is not willing to accept the freedom of individuals.

Today, many sites are using Infinite Scroll...

**Load Conversation** **Send to Instapaper™** **Close**

In Powazek's opinion, it might get better. "I may enable comments again someday. But what I really want to do is

"Load the conversation" does the magic, loading up the related data:

Using a javascript library like rdfquery, we are able to process the embedded metadata directly in the browser. We can query all referenced sources (in both ways) and the topic of the conversation. The matches could be stored in a client side database, if needed.

The “conversation view” loads up the gathered data. It creates an environment that lets the user focus on the “bead chain” that the conversation establishes.

The concept re-introduces a mental model to connected and documented thoughts on the web. In a similar way as journals did with letter exchanges, it furthers the understanding and allows reader or “thought followers to keep up with conversations.

On the other hand, it creates an encyclopedia of thoughts that lives alongside with Wikipedia or other hypertext systems used to collect knowledge.

The visual interface is designed for readability and attention, creating space to concentrate. Links that push readers outwards, are moved to a footnote element. Those elements that are marked up using RDFa and represent the core of the information, get promoted and the used quotation is highlighted.

The interface lets users follow the relations that the triplets describe. It lets them follow the FOAF-element that defines a person and browse along the associated data.

Following these trials shows the user both the relations between people and their publications. One can follow up other conversations that deal with other topics, discover associated sources and, that is the most interesting part, discover logical relations that semantic reasoner software build by drawing logical consequences and conclusions from the base of existing relationships.



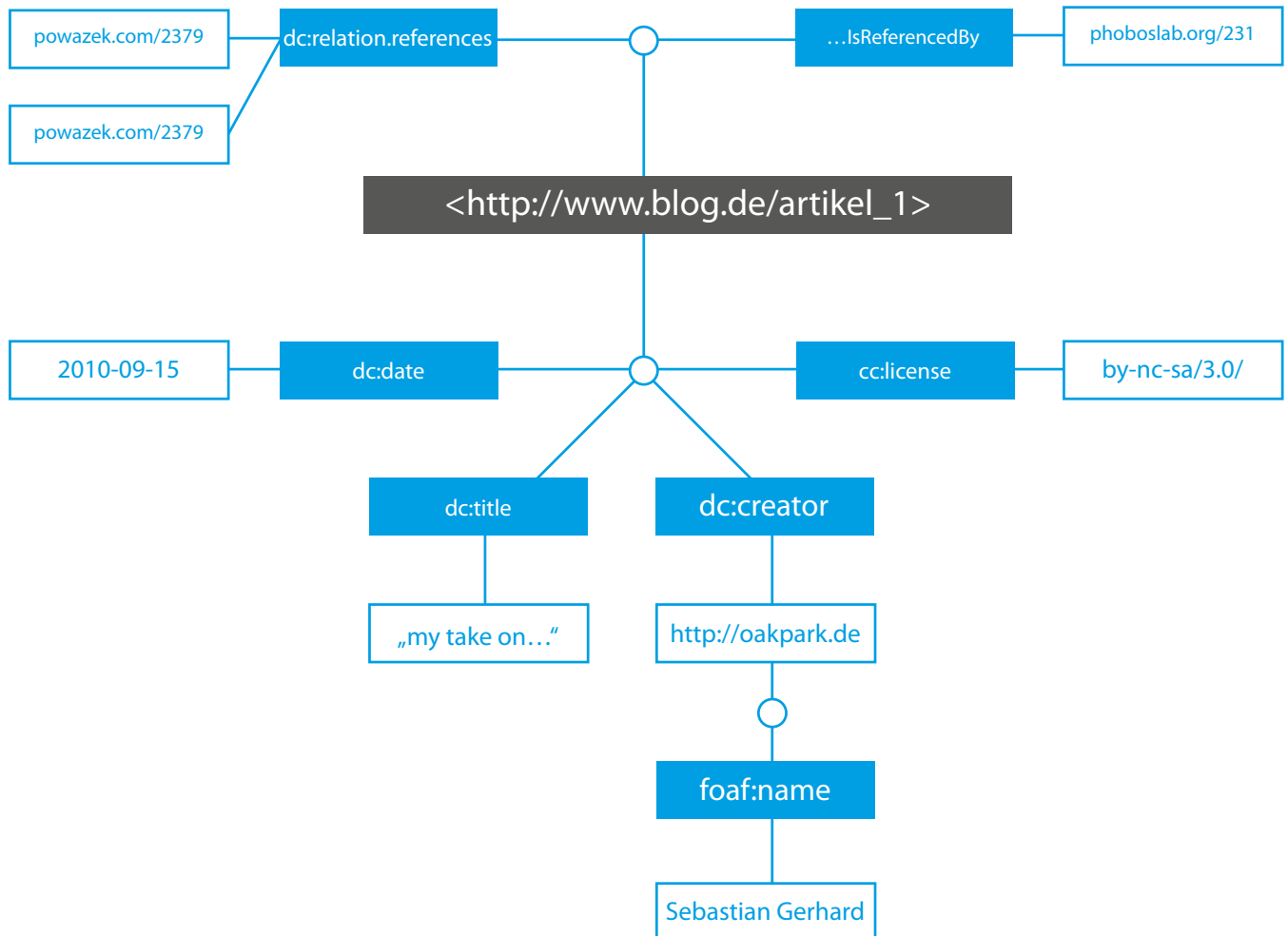
A user could for example discover conversational behavior of an author, how he enters a conversation and how the conversation continues from this point on. It could serve as a measure for genuine relevance that is not fixed to the humming noise of masses that use social gestures to “promote” content. Instead, it creates the relevance by promoting the authors expertise and his associative talent and unrolling the context.

The conversational graph that develops over time fits to Vannevar Bush’s concept behind the MEMEX of creating open trails through knowledge and serving them as open data to anyone who has access to the system.

This sensible overall approach that takes the whole image into account uses established standards and creates an open system that can easily be joined to any other social graph we created on the net.

## **TBL's Guide to well designed data**

- ★ make your stuff available on the web (whatever format)
- ★ ★ make it available as structured data  
(e.g. excel instead of image scan of a table)
- ★ ★ ★ non-proprietary format (e.g. csv instead of excel)
- ★ ★ ★ ★ use URLs to identify things, so that people can point  
at your stuff
- ★ ★ ★ ★ ★ link your data to other people's data to provide context



graphical representation of  
the RDFa data

**Find the transcripts of the video taped  
reading sessions on the following pages:**

**Video 1: [abduzeedo.com/design community](http://abduzeedo.com/design-community)**

Browsing the inspiration community

00:40 curated vs. crowdsourced content

01:30 collecting interesting sites in browser tabs

01:50 pulling up the sharing tools by accident

02:10 comments are ignored

03:25 the “blink” moment, deciding if a link is  
worth a visit in a second

04:20 draw on unlimited resources behind  
each browser tab

04:50 Categorizing a digital scrapbook

**Video 2: Hackernews**

browsing threaded news and discussions on Hackernews

00:45 poor readability and finishability  
due to improper layout

01:08 valuable comments, sometimes more  
important than the initial article

02:30 Links pointing to Wikipedia force reader to make the  
way back to the initial source

02:56 Skimming over a text without going into detail,  
judging content in a „blink“.

- 03:00 Evaluating content with a glimpse  
Identifying editorial patterns
- 03:35 Blog comments are useless aggregated and wisely  
opinionated comments are more interesting
- 04:58 credible vs. critical mass  
voting mechanisms work best among caring and  
interested users
- 05:15 Visual hierarches further understanding  
or at least, die trying

### **Video 3: Google Reader, the link katapult**

- 00:15 “Berry picking”, sorting out interesting content
- 01:10 Categorizing and filtering information
- 01:50 Harvesting the open browser tabs, quitting Google  
Reader
- 02:05 Dismissing an article, no semantic structure,  
“nicht genug Bilder”
- 02:25 Lost between tabs, “how did I get here?”
- 04:00 Comments are not relevant on certain websites

## Video 4: Fever Feedreader

- 00:15 Lost on a landing page, [www.bounceapp.com](http://www.bounceapp.com)
- 00:50 Back in Fever's "Hot View"
- 03:10 Switching to Safari, harvesting the tabs opened beforehand
- 03:31 Sending a large article to Instapaper to read it later
- 03:55 Catching up on a topic opened before reading further material
- 04:20 Lost in a video, facial expression aligns with video content
- 06:06 Liking an article on tumblr using the UI elements in the top right corner
- 06:16 Done with the session, watched one video, marked two articles to read later, articles read: zero.

## Video 5: Reddit

comments, long form and recursive discussions. Those comments who have a reply or a at least a couple of lines long, attract the reader

- 01:10 going out to Wikipedia