5 Conclusions

5.1 Aims of our project

5.1.1 General aims

To develop a system to give to the users the news that they want to read.
To make it in a user-friendly and attractive interface.
To use known technologies giving them a new use as a combo.
To make a project that looks professional and which could be used in reality.

5.1.2 Concrete aims

To develop a web that will offers a service of filtering the news and presenting to the users in a customizable formats and designs, that shout be attractive.

To make the system as much simple and user friendly possible but at the same time make it intelligent and capable of learn by him.

5.1.3 Others

To improve our English (reading, writing, speaking).
To improve our coordination and group working skills

5.2 Personal experience

At the beginning of this project we had different goals or objectives that we wanted to achieve. First of all we wanted to find an innovative project, not something very typical. Investigating on the net we discovered the concept “newsmastering”, based on the RSS feeders’ technology to create thematic information feeders. We found it interesting and analysing a little bit the web we can see that is not an exploited topic. After deciding the topic the project would be about we had to achieve another goal. The project should be technologically according to a master thesis project and we were ambitious about that so we started thinking about a way of applying different technologies in the same project.
After finishing News Hunter we think that the two main goals have been accomplished. The final result is a website programmed using PHP language and MySQL database system connected to an application generated with the .Net technologies and based on the RSS feeders’ technology that nowadays is starting to be used a lot in internet.

Moreover, we had the challenge to do this project as partners, so we wanted to see if it would be possible to work together knowing that there would be different points of view and that each one of us would like sometimes to take different decisions about how to solve a problem. We think that with this point the project is also successful because we’ve been able to deal and understand each other all the time. Especially, after the problems that obligate one of us to go back home in the middle of the project and that anyway hadn’t been problems enough to interfere in the achievement of the project.

That problems had added more difficulty to the project but at the same time allowed us to work on a project separated from thousands of kilometres, thing that gave us more experience working on that kind of projects.

Another handicap that we had at the beginning of the project was the problem of the language. Although the two of us can deal more or less in English language, at the beginning was a mystery if we would be able to finish a project of these characteristics in English. After finishing we are sure that our English could be better but we’re quite happy about our improving with it and we hope it will be enough to do a good presentation.
6 Bibliography

6.1 Used

- **What Is RSS** - by Mark Pilgrim December 18, 2002 (book)
- **Weaving the Web of News** (http://www.webreference.com/xml/column5/3.html)
  - The XML Xpert Michael Claben helps explain RSS and other similarly XML defined news format
- **RSS Tutorial** (http://www.mnot.net/rss/tutorial/)
  - This tutorial explains the features and benefits of a Web format called RSS, and gives a brief technical overview of it. It also includes information on a similar format called Atom. The reader is assumed to have some familiarity with XML and other Web technologies. It is not meant to be exhaustive.
- **XML.com's RSS feed** (http://www.xml.com/cs/xml/query/q/19)
- **Microsoft Developer Network** (http://msdn2.microsoft.com/en-us/default.aspx)
- **PHP.net** (http://php.net/)
- **MySQL 5.0 Reference Manual** (http://dev.mysql.com/doc/refman/5.0/es/index.html)
- **The Apache Software Foundation** (http://www.apache.org/)
- **WAMP Project** (http://www.wampserver.com/en/)
- **Universia gestión** (http://gestion.universia.es)
- **Web of Knowledge** (http://www.accesowok.fecyt.es)
- **Web of Science** (http://www.accesowok.fecyt.es)
- **Learn how to configure apache** (http://articles.techrepublic.com.com/5100-22-5076696.html)
- **Qué és un RSS feed** (http://www.microsiervos.com/archivo/internet/que-es-un-feed-rss.html)
- **Extensible Markup Language (XML)** (http://www.w3.org/XML/)
- **Aprendemas.com** (http://www.aprendemas.com)
- **Bibliotecana, la biblioteca digital de la UPC** (http://bibliotecana.upc.es)
- **Microsoft’s Most Valuable Professionals** (http://msmvp.com)
6.2 Essential Reading

- What Are Syndication Feeds - By Shelley Powers
- Usage Study, RSS—Crossing into the Mainstream (Ipsos Insight/Yahoo!) PDF (http://publisher.yahoo.com/rss/RSS_whitePaper1004.pdf)
- Los Motores De Newsmastering RSS Y Servidores De Sindicación Están A La Vuelta De Esquina: Expand Su Contenido RSS En Un Banco De Datos (http://www.masternewmedia.org/es/2005/06/15/los_motores_de_newsmastering_rss.htm)
- mySyndicaat (http://www.mysyndicaat.com/home)
- Newsmastering and advertising (http://libraryclips.blogsome.com/2006/03/20/newsmastering-and-advertising/)

6.3 Forum

- http://www.slideshare.net/kreaten/newsmastering (spanish).

6.4 Other pages

- Wikipedia (http://www.wikipedia.org/)
7 Appendices

7.1 The RSS NewsMaster

*The NewsMaster: A New Emergent Socio-Professional Role*

In the beginning was Yahoo and AltaVista. Then Google came... then the size and the needs grew to such a level of complexity and need for personalization that none of those tools was appropriate anymore.

We are soon to be under enough evolutionary pressure to improve our ability to access specialized news and information in ways and modalities that will make searching with Google or Yahoo appear like a primitive and much limited approach.

Indeed the key problem we are confronted with is this: how to navigate, filter and reduce to humanly manageable dimensions the gigantic flow of information coming at us in increasing amounts.

The entry of the RSS publishing and distribution format into the fray of media and technologies available for communication has only added to the tidal wave of news and information that has gradually been submerging us.

The proliferation of many alternative online independent news sources and the epidemic expansion of the bloggers community have created a sharp increase in information channels available to us outside the mainstream, established sources.

Further, the Web keeps growing at a steady pace, and the amount of content being made available on a daily basis is simply too much to be managed by our everyday Googles and RSS aggregators. (Nonetheless the interesting feedback and comments made by Jon Udell and the Scobleizer, my
argument does not really center about the stretching of our human abilities to scan thousands of titles while pretending that to be an effective way of searching and finding what is relevant. If they like to keep on doing that way, I am not against it.)

Jakob Nielsen wrote in June of 2003: "The easier it is to find places with good information, the less time users will spend visiting any individual website. This is one of many conclusions that follow from analyzing how people optimize their behavior in online information systems."

And you can bet that with the epidemic RSS is having, the exploding amount of content sources, the likely emergence of a multitude of capable newsmasters and innovative technologies supporting them, the number of information seekers, researchers and normal individuals that used to surf and access tens of sites per day, will, in relative terms, drastically decrease.

If information can come at you in such selected, delicious, ad-free fountains, why would you go surfing or searching the Web?

"Online publishers and e-commerce sites no longer need to see search engines simply as sources of traffic and advertising networks.
For the first time, publishers can use a search engine to filter and deliver relevant content to their readers within their own pages, including customized topic search and syndication"

We need something of an entirely new order of magnitude to manage all of this information.

Search engines, open directories, and millions of bloggers are not enough.

We need a multi-layered, self-organizing approach that allows the load to be highly distributed and the focus and depth to be guaranteed by the combined result of many highly focused individual efforts.

"The layered mechanism works because at no point is the entire weight of the filtering process concentrated in a single individual or a single resource.

It means that individual agents can work without the need for central control, with the only requirement for a functional system being an open set of connections between the agents.

What RSS does best is that it allows an individual to scan, filter, and pass forward.

That's all it ever has to do.
The network can and will do the rest."

7.1.1 The Issues Catalyzing The Change

- Too much information coming in.
- Too many sources too scan.
- General categorization of sources too limited. By author or publishing house is too limiting. I can't scan 100 authors a day to see if there is some interesting for me to read or write about.
- Too many new sources coming online daily. Which ones do I pick?
- Difficulty in separating relevant news from press releases, personal rants and job vacancy announcements
- Too much redundancy - feeds are replicated, news are duplicated several times, the same PR stories get republished over and over again.
- Dubious quality of many news sources.
- Selfish. Our search for relevant content and news is an individual activity, in general with no direct benefit for the community at large
- Non-reusable. Searching for information is mostly carried out as separated, non-contextual, just-in-time quick personal search.
- Demand for quality, filtered, spam and ad-free selected content on specialized topics is in very high demand it is at best pretty hard to find.
- RSS is the only place where you can start intelligently filtering the ocean of info, but the process and tools presently at hand are limited, and the time required to scan thousands of feeds and titles is non-trivial.
- Google, Yahoo, and all of the other major search engines are not sufficiently powerful to extract and provide quality content on specialized news topics. They are good at finding online resources and specific content but when it comes to producing ongoing specialized information channels these tools do not seem to stand the test of time.
- Bloggers are a valuable and interesting source of independent and alternative information but it is still very hard to extract valuable content from the thousands of personal rants, superficial notes and other trivial and research-lacking information items.

RSS newsreaders are still author-based aggregators of information requiring the news gatherer at bay with hundreds of interesting resources to scan. News seekers search information by focused topic and not by author. The approach is flawed at its foundations and needs to be rethought in ways that allow for the content that is relevant to me to surface more easily from the whole pool of sources I select.
7.1.2 The Network Starts To Organize Itself

As Stephen Downes mentions in his beautiful post about networks, the network of information needs to organize itself automatically, and we must help it in doing so.

Filtering of the online content happens at different levels and it needs to be a distributed task.

As Stephen points out no single resource is capable of handling the oceanic flow of information unaided. And if it does, like Google, it does so often ineffectively.

"Inherent in the structure of the Internet there are distinct layers of filtering mechanisms, each consisting of a "gather filter forward" mechanism.

At the most basic level the mechanism is fulfilled by a human agent, as in the case of blogs.

In others, it is fulfilled by automatic mechanisms, such in new search tools and online specialized aggregators...

The reason why this system works, while other approaches do not, is that there is no reasonable mechanism which can apply the vast requirements of filtering on a single resource.

If we use metadata, the indexing soon outweighs the content.

If we use search engines, each resource must be subject to extensive analysis to determine context.

What's important here is that each node of each layer need not worry about the rest, and need not be focused on the goal of the system.

The agent seeks what is available, the way a retinal cell gathers light, and passes on what is relevant, the way a neuron passes on a signal. The filtering occurs not in the individual node, but through the independent actions of the aggregation of nodes."

7.1.3 RSS Dolby

Bloggers play an important role in this autonomic filtering process, but they are also adding, in the process, lots of extra noise.

It may actually be, that to make the filtering process more accurate and precise, the level of noise needs to be increased rather than decreased.

I have been recently noticing how the universe of RSS-based content is also becoming extremely redundant and rich in duplicates and repetitions. I myself feel guilty, as an explorer and early tester of this polluted environment. But as I reason about it with detachment I realize that it is something indeed inevitable (many others will need to test) and probably beneficial.

The littering and polluting happens as you test how to create new RSS channels and feeds, and when this channels get picked up, exposed or shared inside a directory or RSS search engine, then it becomes very hard to prune and delete them from the RSS universe.
While apparently this would seem to baffle traditional logic, you have certainly heard about the Dolby sound system, which achieves though pre-emphasis a cleaner, hiss-free sound, by first adding more noise to the original source and then filtering out with greater ease what superfluous.

So it may indeed be that bloggers add an ergonomic hiss to the information flow which the upper level of news gatherers and aggregators will actually leverage to filter the actual "valuable content".

7.1.4 What So Great About This?

The discovery is that we can now start to really organize information in some meaningful, creative, very useful ways.

As smoothly as the words in the previous sentence traverse your mind, you must realize that this is quite an impressive turn we are about to make. Nor this is anything we have seen before happening during this civilization.

No Yahoo, Google, ODP, has or can do what we are about to see.

It is going to be a blossoming of specialized news channels like nothing we have seen before.

The discovery is in fact in the unlimited and yet untapped power we now have to search, filter, aggregate and create focused news/information channels with the only support of our know-how, culture, experience and a little unknown free technology called: RSS.

7.1.5 What's The Great Benefit?

The real value is in the fact that the effective widespread application of this filtering and redistribution process would be a highly valuable social endeavor.

As amazing as it may appear, the emergence of the newsmasters shows that individualistic and self-directed work may actually provide great social benefits.

The creation of dedicated information channels, originated by independent publishers and not by vested commercial interests or mainstream media conglomerates may create the opportunity for a true renaissance of culture, learning and to a multiplication of our abilities to manage large amounts of rapidly changing information.

Search specialist and librarians who will craft with time investments appropriate queries to get at the information they were looking for, will be generously rewarded with eternal fountains of relevant info for the time to come.

The act of search evolves from a mere set one-time shots at finding something just-in-time, to the start of a collective refined meta-filtering process that generates better information for everyone.
The more refined filtering process that the newmaster creates can benefit private enterprises but can also benefit, for extended time and with great value the research and learning needs of free individuals not connected to direct commercial interests.

Having a great flow of information that can be tapped, aggregated and filtered accurately is a benefit for all. Its uses and applications are limitless, and they can not only enrich our rapidly evolving global culture but can also give the opportunity for ethical sustainability to those who have the will and skill to leverage them.

Again:

"What's important here is that each node of each layer need not worry about the rest, and need not be focused on the goal of the system.

The agent seeks what is available, the way a retinal cell gathers light, and passes on what is relevant, the way a neuron passes on a signal.

The filtering occurs not in the individual node, but through the independent actions of the aggregation of nodes."

7.1.6 The Network Is Already Doing It!

At the most basic level, bloggers represent the first generation of self-organizers of the network. If we exclude Yahoo initial honorable effort an similar one (e.g.: ODP), there is nothing else, that comes close in scope, to what the universe of bloggers or noosphere, has been able to achieve.

Bloggers have been first to scout new independent sources, alternative voices, and who have pointed links to new ways of looking at any issue. They, more than any other group, have brought, among much useless noise, the true emergence of effective meta news sources that originate, filter and aggregate valuable content online.

"Decisions about selection and classification are made on a case by case basis using very coarse, and unregulated, mechanisms."

So, what it appears to the many superficial onlookers as a universe of mindless writings (blogs) is nothing less than the initial phase of a complex and orderly process whereby humanity at large takes control of filtering, gathering and re-organizing his own know-how an discoveries.

Newsmasters may indeed represent the second layer of filtering that we can now apply to this ocean of content and information. ...and Newsmasters, as it appears, have been indeed emerging and quietly working at this nouvelle craft, for quite some time now.

And with the above, the emergence of a new breed of tools, tuned to provide the specific facilities required by the fascinating job of the Newmaster, heralds the beginning of a popularization process which will allow, like in the past for DTP (desktop publishing), those who have the skill, to become the true competitors of any mainstream information source.
Now that we have stripped information from everything that was superfluous to it, the real value and skill is in the true ability to find, pick, select and bring together intelligently the best information bits available out there.

And due to the unique nature of RSS-based content and of its user-centric character, what Newsmasters create are much like geysers of information, open hoses that generate uninterrupted useful info on a specific topic. Could we say "fountains of knowledge"?

For what I can infer, it is also likely that in the near future, once Newsmasters are an established profession, end users themselves will start to apply on their own an extra level of more sophisticated aggregation and filtering to the highly specialized feeds they select.

7.1.7 So What Is A Newsmaster And What Does sHe Do?

NewsMaster definition:

The newsmaster is an individual capable of personally crafting RSS-based specialized information channels by utilizing technologies that allow him to select, aggregate, filter, exclude and identify quality news, information, content, tools and resources from the whole universe of content, news and information available on the Internet.

Newsmastering is the ability of a human being to concert, orchestrate, edit, and refine quality search formulas that tap into the whole Internet content universe and beyond, and that filter out relevant information through selected keywords, source selection, ranking, heuristics, and many other possible criteria.

The Newsmaster specializes in crafting uniquely powerful search/aggregation/filtering formulas generating continuous RSS feeds on narrowly selected topics by:

a) Resource identification. Selecting and aggregating valuable resources (like a normal aggregator does)

b) Query formulation at different levels and on different content bases.
   1) Creating advanced search queries in the blogosphere (a la Bloglines

2) Creating advanced search queries on the traditional Web content at large

3) Creating advanced search queries on the overall pot of content derived at points a), b) and c) and generating new highly filtered RSS feeds matching specific content and quality criteria.

4) Creating advanced search queries to the deep/invisible Web with dedicated tools and creating RSS content out of the results obtained.

f) Filtering. Applying exclusion filters based on keywords, keyphrases, date, language, and more.

g) Splicing (union): I want feed C to be the result of merging feeds A and B.
h) **Intersecting**: Given primary feeds A and B, I want feed C to consist of all items that appear in both primary feeds.

i) **Subtracting (difference)**: I want to remove from feed A all of the items that also appear in feed B. Put the result in feed C.

l) **Splitting (subset selection)**: I want to split feed D into feeds D1 and D2, according to some binary selection criterion on items.

g) **Heuristics**. utilization of heuristics allowing the system to learn from previous choices the patterns characterizing the most appropriate news items from the rest.

i) **Reputation**. possible utilization of real-time collaborative filtering mechanisms allowing the newsmaster to benefit also from the rating and reputation given to selected news/information items by other independent newsmasters.

f) **Personal selection**. The ability to integrate in the newsmastering workflow the ability to personally select information items or to continually improve and refine the governing search and aggregation formulas by way of personal evaluation.

### Some examples of newsmastering work available out there:

Here are some living examples of newsmastering work available out there:

1) Stephen Downes - Edu_RSS  
   [http://www.downes.ca/edu_rss.rss](http://www.downes.ca/edu_rss.rss)

2) Harry Potter channel  

3) MasterViews - a timid initial approach by the author to newsmastering with the automatic generation of a news page focusing on presentation, information design and PowerPoint resources and skills.  

4) [http://infoseedaily.net/](http://infoseedaily.net/)  
   Over 11 dozen news feeds and over 100 blogs aggregated to give a picture of the information security space. Nothing non-commercial met these needs, i don't think anything else does yet. (by Jose Nazario)

   A personal world news system, aggregating close to 100 news feeds. Through hueristics it identifies interesting news items and sorts the results in that order. Got sick of google news' fluffy feel? Here is the alternative. (by Jose Nazario)

6) Please suggest more! I know there must be thousands out there!
7.1.8 The NewsMaster Toolkit - An Opportunity To Support Robin Good

In an effort to offer you an elegant way to support my further research and reporting in this direction, I have made available a small report entitled the Newsmaster Toolkit.

It contains now over 40 different selected tools that can be used for experimenting and fully testing the thrill of newsmastering today. With your personal skills and these technologies you can start building dedicated news channels and focussed mini Web sites in little time.

Once again, don't get me wrong. The toolkit is an opportunity I am offering you to support my reporting in this area. If you want to thank, reward, encourage or simply acknowledge the value of this, please do get yourself a Toolkit today.

More information about the NewsMaster Toolkit here.

7.1.9 Personal Thoughts

Newsmasters and their role are central to my position as an evangelist of intelligent use of new media technologies. Here is a great opportunity to be taken: How to intelligently use the power of RSS-based content to facilitate more effective, learning and communication, by initiating a process of self-organization of all of the content available on the Internet.

I would suggest the ones of you that like what I am describing to start working and getting skilled at this as soon as possible.

Being there at beginning and understanding early in the game the principles of this fascinating mechanism can only provide you with greater ability to manage and craft more effective tools rather than having to follow what the market will make available for everyone.

At any rate, the desired asset will not be the one of only having a better or more evolved technological toolset, but rather at the combined personal ability and skill to filter, select, pick and edit/comment those news, ideas, information and resources that are most valuable to your specialized information channel(s).

While technology will certainly greatly contribute to our ability to rapidly select and manage while drawing from a large poll of sources, it will be our personal experience, know-how and intuition that will guide us in customizing the filtering mechanisms we will create.

So if you are into newsmastering you need both to learn and refine your research skills applied to this new vast set of information sources now available, as well as to facilitate the development and refinement of technologies that ease the tremendous scobleizing task at hand (picking through thousands of RSS-based content feeds and filtering out the pure junk from what may be just what you are looking for).
References:
How RSS Can Succeed by Stephen Downes
RSS Needs by Dina Metha
The algebra of feeds, or the amateurization of RSS bricolage by Sebastian Paquet
Research, inspiration and ideas-checking: Massimo Curatella

7.2 The Birth Of The NewsMaster: The Network Starts To Organize Itself

7.2.1 NewsMasters: RSS And The Opportunity For Sustainable Filtering And Aggregation Of Online Content Into Niche Websites And Dedicated Information Channels

An independent publisher opportunity and a socially valuable way of filtering and organizing news information and the Web at large.

Introduction

Yesterday, at the latest DEMO 2004 Conference in Scottsdale Arizona, Feedster "announced a new tool called FeedPaper, which combines syndicated Internet content (culled from the 500,000 newsfeeds continuously searched by Feedster), such as The New York Times, Washington Post, USA Today, Amazon.com, many IT publications, PRNewswire, and hundreds of thousands of blogs. A Feedpaper is a web-based and RSS-enabled micropublication on any topic of its creators choosing. Creating a Feedpaper enables its publisher to blend, track, and share information on any topic.

Online publishers and e-commerce sites no longer need to see search engines simply as sources of traffic and advertising networks.

For the first time, publishers can use a search engine to filter and deliver relevant content to their readers within their own pages, including customized topic search and syndication."

Scott Rafer, Feedster's CEO said.
(http://www.internetnews.com/dev-news/article.php/3313991)

Indeed the key problem we are confronted with is this: how to navigate, filter and reduce to humanly manageable dimensions the gigantic flow of information coming at us in increasing amounts.
The entry of the RSS publishing and distribution format into the fray of media and technologies available for communication has only added to the tidal wave of news and information that has gradually been submerging us.

The proliferation of many alternative online independent news sources and the epidemic expansion of the bloggers community have created a sharp increase in information channels available to us outside the mainstream, established sources.

Further, the Web keeps growing at a steady pace, and the amount of content being made available on a daily basis is simply too much to be managed by our everyday Googles and RSS aggregators.

We need something of an entirely new order to manage this.

An army of newsmasters.

7.2.2 What Is The Discovery?

The discovery is the unlimited and yet untapped power we now have to search, filter, aggregate and create focussed news/information channels with the only support of our know-how, culture, experience and a little unknown free technology called: RSS.

7.2.3 Where Is The Real Value?

The real value is in the fact that the effective widespread application of this filtering and redistribution process would be a highly valuable social endeavour and one that not only can enrich our rapidly evolving global culture but one which can also give sustainability to those who have the will and skill to ride it.

7.2.4 Summary: Critical Issues/ Problems At Hand

- Too much information coming in
- Too many sources too scan
- General categorization of sources too limited. By author or publishing house is too limiting. I can’t scan 100 authors a day to see if there is some interesting for me to read or write about
- Too many new sources coming online daily. Which ones do I pick?
- Difficulty in separating relevant news from press releases, personal rants and job vacancy announcements
- Dubious quality of many news sources
- Our search for relevant content and news is an individual activity, in general with no direct benefit for the community at large
- Searching for information is mostly carried out as separated, uncontextual, just-in-time quick personal search.

7.2.5 The Filtering Mechanisms Available To Us

As Stephen Downes mentions in his beautiful post about networks in his last issue of OL Daily, the network of information needs to organize itself automatically, and we must help it in doing so.
We must learn to perceive, to use, the internet in exactly the same way, as a self-organizing network.

Up until now we have organized and searched for information mostly by:

a) **Typical major search engines** (Google, Alltheweb, Teoma, Wisenut, etc.)

b) **Directories** (ODP, Yahoo, etc.)

c) **Meta search tools** (Copernic Meta, Vivissimo, Surf wax, Ixquick)

d) **Data Visualization tools** (Grokker, Kartoo, TouchGraph, Musicplasma)

We have also built internal directories inside our companies and organizations, databases and knowledgebases of all kinds. But of these last ones each, most of them remain like isolated islands of which few search tools can report the size, content and updated map.

More recently, the army of individual reporters, writers, journalists, academics and researchers who has adopted independent publishing systems like CMSs, weblogs, and other database-based content management solutions, has been proactively weaving, organizing and filtering, at an initial level of refinement the news flow and content coming at us daily.

**Bloggers foremost**, have taken a natural role of filtering and aggregating mechanisms of the newsflows running through the Internet.

**Bloggers have been first to scout** new independent sources, alternative voices, and who have pointed links to new ways of looking at any issue. They, more than any other group, have brought, among much useless noise, the true emergence of effective meta news sources that originate, filter and aggregate valuable content online.

**So, what it appears to the many superficial onlookers as a universe of mindless writings (blogs)** is nothing less than the initial phase of a complex and orderly process whereby humanity at large takes control of filtering, gathering and re-organizing his own know-how an discoveries.

The world becomes a global village and as such it starts taking care of itself in a global, dispersed, self-sufficient, synchronized fashion.

**A Google query is an individualistic just-in-time search with no benefit for society.**

A Google query is a one-time shot.

You type the query, you get the results and it is over.

You probably don't re-use that search anymore. You will retype it some day when you will need it again.

Nobody else outside of you benefits from that search.

Again, it is a one-time shot.
7.2.6 A Learning Voyage Into RSS

To better understand what this new RSS-based universe really entails and may likely involve, I have dived, soul and body, into it, to understand in first-person how much the new information publishing and distribution protocol could be effectively revolutionizing our immediate future.

The process has taken quite a few months, but the let me tell you that the view derived from this open training has been pretty amazing, to say the least.

In short order and for the sceptics out there, here is what I have done to "understand" RSS and to develop a good feel for where it is headed:

I have converted all of my online sites into content management systems that would be able to generate RSS feeds. I have organized content into categories and created individual RSS feeds for each one of them.

I have started using RSS newsreaders and aggregators and have reported extensively about my findings.

I have tested tools to convert RSS feeds into email, and to convert emails into RSS feeds.

I have started analysing sites and services syndicating other RSS feed content.

I have extensively delved into using all of the RSS monitoring tools like Technorati, DayPop, Blogdex, Popdex, Feedster, Blogstreet and many other ones.

Inspired by the work of Ari Paparo http://www.aripaparo.com/archive/000632.html and by the growing number of valuable directories and search engines aggregating and indexing valuable RSS-based content I went on to preparing the publicly accessible RSSTop55, a best and most recommended list of top RSS-based submission directories.

Early in this discovery voyage I got interested and fascinated by tools like MyRSS who would allow you to theoretically create an RSS feed out of anyone Web site page updating its content out there.

A love for "scraping" has followed right after. Scraping is the art of deciphering the underlying content structure of a Web page content, in order to be able to create a framework or container that can "extract" the content of that site at will in an organized fashion. Scraping a Web site allows anyone to create an RSS feed for any Web page. The feed is fuelled by the new content appearing on that page, and therefore not all pages lend themselves well to be scraped and to be sources for useful feeds.

I have then started syndicating RSS feeds generated by my different sites across the sites themselves and then later I have started creating valuable and unique news channels by aggregating RSS news feeds available out there.

Two good examples of this are: Kolabora, where you can see the home page populated by news coming from different RSS news feeds I have selected and MasterViews where once again the news page is completely generated by carefully filtering and aggregating selected RSS-based news feeds.
7.2.7 Possible Solutions: RSS Aggregators?

Many have touted the advent of RSS aggregators has one possible solution to the issue of information overflow, without yet realizing the intrinsic limitations of this approach when applied literally.

You can subscribe easily to hundreds of RSS feeds in no time at all. And there are a great number of people I like to read about.

And what I really want to read is what these people say about the things that interest me.

Jon Udell is a great guy. Of his posts though, only one in thirty is something that matches my level of interest and competence. The rest is too "geeky" for me. So, it is impossible for me to feel motivated to look at Jon's feed, though I know, that from time to time he will bring me some wonderful unique gems.

So, I give in, and I tend to scan only those sources that are more directly in tune with me. And by accepting this compromise I often loose all of the opportunity to discover Jon's magic realizations when they do really happen.

I have seen and heard of people subscribing to hundreds if not to thousands of feeds inside their RSS aggregators.

Is that manageable?
Do these people get better and more information than everyone else?

It is not.
They don't.

They don't because it is just not possible and not enjoyable at all to go through so many posts, in which, the greatest majority is not even of relevance to their interest.

It is just like doing a one keyword search on Google and then going through all of the results pages that come up. Yes, you can come up with some interesting discoveries but, would you really do that?

Like for Google, where the enormous amount of relevant results obscures the view of most of what could be of interest to me, we do need some new more effective way of filtering the river of news coming at us.

Visualization is one solution, and you will see many new fascinating applications of interface and information design to search mechanisms in the near future. Visualization is the quantum leap needed to perceive the forest from the trees before wanting to dive-in deeper into it.

7.2.8 How Do I Want To Find Information Relevant To Me

Overloaded and disappointed I have spent some thinking time testing many RSS tools and aggregators only to gradually increase my particular appreciation for Bloglines. Though Bloglines features are not completely unique and are available in parts across several other resources online, this
is the service that allowed me to come closest to the idea of news and content filtering that I was starting to build up in my head.

Bloglines in fact, not only allows you to add as many feeds as you like inside a neat set of organizable content folders, but it also allows you to subscribe to specific search queries you create. That is, Bloglines allows you to record any search you do across the RSS-based content it has indexed and to automatically rerun that search for you each time you check for new content.

As simple as it sounds this facility provides for a terrific resource of fresh news on any topic you may select!

The first time you experience the use of this you will feel some exhilarating vibe. The crafting of your own little search query becomes a key to generate valuable content on most any topic you can think of.

Unfortunately Bloglines indexes only less than one hundred thousands blogs (!), but think if you could do the same search query on Technorati where over one million blogs are tracked daily. Wouldn't that be interesting?

With the help of Bloglines' "subscribed searches" feature I am now a happy scout reading the "topics" I am interested in and free from NOT having to read as a robot the individual feeds of one hundred and more authors and sites I am interested in.

That brings by itself another important corollary. When creating a saved search, you are in fact creating a very custom and personal news feed, tailored to your own specific request.

If you could only save that Blogline feed as an actual RSS feed, you would start appreciating more directly the feeling of becoming an independent newsmaster yourself. (While Bloglines doesn't yet offer this option, other online services, Blogdigger, have started to do so.)

7.2.9 Emerging Technologies

The whole idea I am presenting in this essay is made possible by the rapid emergence of a new group of tools and technologies that allow scraping, search, filtering, aggregation and syndication of RSS-based content at costs and with ease of use that make them accessible to the layman.

My ideal future RSS aggregator gives me the possibility of not only subscribing to as many selected feeds as I like, but it also allows me to pick out my preferred content by filtering out through sophisticated queries and inclusion/exclusion features what is not relevant to my interests.

If you can select sources, formulate advance queries, create search feeds, and mix this combination in newly created information/news feeds you have just visualized the primitive profile of the professional role I have been visualizing.
7.2.10 A New Digital Professional Role: The NewsMaster

Figure now the ability to concert, orchestrate, edit, and refine quality search formulas that tap into the whole RSS universe and beyond, and that filter out relevant content based on selected keywords, sources, type of content, ranking and many other possible criteria.

Moreover, iSyndicate (today Yellowbrix) and others have done this before, and many sites still benefit from this simple automated mechanism.

But imagine now this mechanism being made available to each one who wants to use it.

Imagine having the power to select, aggregate, filter content while creating a new RSS feed that is as unique as the person who has built it.

And this set of "magic search formulas" that anyone can now build is not a one-time shot anymore.

This is a "hose" from which useful information around a tightly defined topic can be extracted every day or as soon as there is something of value out there.

It is a hose from which other people can quench their knowledge thirst.

My unique personal filtering effort becomes useful for everyone.

I have a new profession for information seekers, digital librarians, and knowledge management evangelists: the NewsMaster.

You may probably find a better term for this but the guy I am envisioning is a new type of webmaster who specializes in crafting uniquely powerful magic search formulas generating continuous RSS feed on narrowly selected topics by:

a) selecting and aggregating valuable resources (like a normal aggregator does)

b) creating advanced search queries in the blogosphere (a la Bloglines)

c) creating advanced search queries on the traditional Web content at large

d) creating advanced search queries on the overall pot of content derived at points a), b) and c) and generating new highly filtered RSS feeds matching specific content and quality criteria.

Searching and identifying key valuable sources and complex filtering formulas will be the outstanding job of the new information gatherers and publishers online: the NewsMasters.

Their ability and craft will be in identifying the query formulas that when cleverly combined and refined by trial and error will allow for powerful automatic news gathering formulas. These formulas need not only to identify the keywords/keyphrases that generate the best and most relevant set of results but need to grow through the intelligent use of multiple filtering criteria and precisely targeted inclusion/exclusion mechanisms.
The creation of dedicated information channels, originated by independent publishers and not by vested commercial interests or mainstream media conglomerates may create the opportunity for a true renaissance of culture, learning and to a multiplication of our abilities to manage large amounts of rapidly changing information.

**Search specialist and librarians who will craft** with time investments appropriate queries to get at the information they were looking for, will be generously rewarded with a eternal fountains of relevant info for the time to come.

The act of search evolves from a mere set one-time shots at finding something just-in-time, to the start of a collective refined meta-filtering process that generates better information for everyone.

### 7.2.11 Are Newsmasters Sustainable?

In the wake of all this I am sure some of you have seen also the opportunity to create niche Web site and information channels on very specific topics and with a possible significant rewarding economy. The opportunities are indeed in abundance and one need only to explore some of the first rough ideas I have listed here below, to get some directions on where to start.

**Here is an initial list** of possible revenue streams that could be ethically generated around the NewsMasters activity.

1. **Excerpt only RSS feeds** bring readers back online to check for interesting stories and allowing for online only contextual advertising (Google AdSense, Blogads)

2. **Low subscription rate to receive premium or "full content" feeds.**

3. **Editorial promotion inside RSS feeds.** Ethical review and inclusion of commercial reviews by independent publishers.

4. RSSAds and similar initiatives.

5. **Sponsorships!** – This is one of the best ways to make specialized RSS newsfeeds pay back. Have the feed description and reference image carry a valuable sponsorship.

6. **Donations**

7. **Micropayments**

8. **More traffic** to Web sites where other products/services are offered.

9. Custom newsfeed creation for third parties.

10. **Provision of high quality custom feeds** to major news syndication companies and organizations.

11. **Syndication of content** on commercial sites.
12. **Premium subscription** of feeds from a specific author or group of associates.

13. **More** (please suggest some in the bottom comments area)

### 7.2.12 Conclusions

**RSS puts a large quantity of content** de facto in the public domain and it offers an easy way to reuse, syndicate and aggregate such content in an unlimited number of ways.

If someone puts an RSS feed on the Internet it is most likely that she is doing this for others to read, and possibly further syndicate. So it is in the DNA of RSS to be wanting to be free, to be further reused, personalized and syndicated.

The above described ability in the above essay to have many individual users work as Newsmasters in organizing the very own content we are generating provides us all with a new horizon of access to news in ways and fashion that we had not even conceived of until yesterday.

**This is not Google News at the nth power. This is much, much more.** Specialized, focused information niches generated by the dedicated work of individual technology-enabled Newsmasters is something totally new for us to consider and reflect upon.

**The power of many people at work** seems to me irresistible and capable of creating profound changes in the way we operate and inform ourselves.

With this new search and filtering approach we can all start generating better and more refined information channels, as we become enabled to aggregate, select and filter the most relevant content according to extremely sophisticated criteria and long nurtured selected source lists.

The simplification of these new gathering and filtering tools, of which FeedPaper may be the first actual working example, is just around the corner. There is no way of avoiding this. It is just coming.

**Tools like FeedPaper** will allow capable thinkers to dedicate themselves to this new highly valuable socio-alchemical effort and to create fountains of information hopefully useful for much time to come.

With these technologies and with our applied intelligence many alternative new sources of info can be effectively created. This will naturally filter out the blog noise from the valuable blog/RSS info. True diversity will bloom as unlimited number of views and perspectives can be made available in such a possible information system.

**So, it does look to me as a renaissance of sorts is truly taking place.** And if can draw any conclusion from the first hand feelings I have had, I must say that this is going to be a fascinating, revolutionary turn.
Special Note from the author:

While we are only seeing the birth of the first primitive group of such tools, I expect many to become available soon and with an array of features and controls that will make the independent news reporters work a truly fascinating one.

Feedpaper from Feedster maybe the first one to become available.

For now, if you wish to pioneer this new exciting vastland, I have prepared a special toolkit just for this purpose.

It is called the The NewsMasters Toolkit and it contains all of the tools and resources you can start using now to build and syndicate your own selected news information channels.

7.2.12.1 Appendix:

The NewsMaster dream tool

(advanced aggregation/filtering tool – specification) Feb.17 2004

What I would want to have in the ideal tool I yet have to meet.

Characteristics:

- Search according to multiple criteria
- Ability to group multiple queries together
- Full Boolean expression control integration
- Ability to filter out selected keywords/ keyphrases
- Ability to pre-select predetermined sources
- Ability to ban selected news sources
- Ability to filter news sources based on several criteria:
  - Technorati Cosmos indicators
  - Google PageRank
  - Alexa absolute ranking
  - Google linkbacks (or other equivalent)
  - More to be defined

- Ability to filter news items based on: article length
- Ability to filter news items based on link exclusion and inclusion rules
- Control of frequency of updates
- Generation of RSS feed
- Generation Of OPML of news sources utilized
- Query formula export
- Ability include search results for specific queries on major search engines
- Export to multiple RSS formats