



European Asylum Support Office

# Tools and Tips for Online COI Research

*EASO Practical Guides Series*

June 2014

SUPPORT IS OUR MISSION





European Asylum Support Office

# Tools and Tips for Online COI Research

June 2014

***Europe Direct is a service to help you find answers to your questions about the European Union.***

**Freephone number (\*):  
00 800 6 7 8 9 10 11**

(\* Certain mobile telephone operators do not allow access to 00800 numbers or these calls may be billed.

More information on the European Union is available on the Internet (<http://europa.eu>).

ISBN 978-92-9243-266-9

doi: 10.2847/31467

© European Asylum Support Office 2014

Neither EASO nor any person acting on its behalf may be held responsible for the use which may be made of the information contained herein.

## Acknowledgements

The present document is based mainly on information provided by the various speakers invited to the Online COI Research Conference organised by EASO's Centre for Information, Documentation and Analysis (CIDA) on 26 and 27 March 2014, in Malta. It is supplemented with other information where appropriate.

The information provided does not pretend to be exhaustive.

References to specific online technologies are provided as examples only; EASO does not recommend any specific tools.

If you wish to provide any feedback to this document, please contact EASO at: [EASO-CIDA@easo.europa.eu](mailto:EASO-CIDA@easo.europa.eu).

Information from the document should be quoted as:

EASO, *Tools and Tips for Online COI Research*, June 2014, available at: <http://easo.europa.eu/asylum-documentation/easo-publication-and-documentation>

## Contents

Acknowledgements.....	5
<b>1. Introduction .....</b>	<b>7</b>
<b>2. Searching / Information Retrieval.....</b>	<b>8</b>
2.1. Web Search Engines.....	8
2.2. COI Portals.....	9
2.2.1. Ecoinet.....	9
2.2.2. Refworld .....	9
2.2.3. The Common European COI Portal .....	10
2.3. Archived content.....	11
2.3.1. Wayback Machine .....	11
2.3.2. Screenshots.com.....	12
2.4. Searching social networks.....	12
2.4.1. Twitter .....	12
2.4.2. Facebook.....	12
2.4.3. Metasearch in social networks.....	13
2.4.4. Specific regional social media.....	13
<b>3. Organising information.....</b>	<b>14</b>
3.1. Automatisation and aggregation tools .....	14
3.1.1. Google and Yahoo Alerts .....	14
3.1.2. Monitoring changes on websites / blogs using RSS .....	14
3.1.3. Twitterfeed .....	15
3.1.4. Monitoring changes on websites without RSS .....	15
3.2. Social bookmarking tools .....	16
3.3. Reference / citation management tools .....	16
3.4. Collaborative online query system.....	17
<b>4. Privacy while searching the Web and data protection issues.....</b>	<b>17</b>
4.1. Basic security concepts .....	17
4.1.1. Browser settings .....	17
4.1.2. Security plugins.....	18
4.1.3. Using an anonymous search engine .....	18
4.2. Advanced concepts .....	18
4.2.1. VPN (Virtual Private Network).....	18
4.2.2. TOR (The Onion Router Network) .....	19
4.2.3. Securing Email by encryption .....	19
4.2.4. Further reading.....	19
Subject Index.....	20

## 1. Introduction

The emergence of new information technologies poses obvious challenges to traditional research methodologies. As new online information tools generate an increasing volume of information (some of which may be of dubious quality), COI researchers are increasingly faced with the difficult task of selecting specific information that is relevant and reliable.

Social media platforms, for example, allow a growing number of people to quickly and easily document events and to communicate this information around the world instantly. Given the unregulated nature of social media, the traditional COI quality criteria which guide the work of COI researchers (neutrality, objectivity, and protection of personal data) and their final products (relevance, reliability, balance, currency, accuracy, traceability and transparency) may not be easily applied in this new sphere. Indeed, the assessment of social media sources can be particularly complex and, at the same time, the wealth of information made available via various online tools may tend to create expectations among end-users as to the availability of information, even of a very local or specific nature, thereby posing an additional challenge to COI researchers.

However, as the tools presented in this document demonstrate, some of the new technologies available for collecting information also hold potential for COI researchers, such as increased research efficiency and privacy, providing better information management and monitoring, and greater knowledge-sharing. In this regard (while keeping in mind the specificities of country of origin information needs in the context of international protection) a lot can be learnt from the experiences of researchers in other fields such as journalism or intelligence work, who have been using such tools for some time now.

Thus, while new media present are a number of challenges, they also provide opportunities for the COI researcher, particularly in a context in which information on countries of origin is increasingly being shared, collected, filtered and presented using online tools and media.

For this reason, EASO organised a “Online COI Research Conference” on 26 and 27 March 2014, to which it invited a number of expert speakers and end-users. Their presentations and discussions form the basis of the present document, which aims to provide a high-level overview of some of the main tools and tips for online research that may be of use to COI researchers in this new and expanding sphere.

## 2. Searching / Information Retrieval

### 2.1. Web Search Engines

Web search engines are software systems made for searching the web. They do this via a [Web crawler](#)<sup>1</sup> or “spider” which runs through the internet at high speed and indexes web pages. The scope of a search engine is defined by this indexation, which can be different for the respective engines. Not all information on the web can be found via the web search engines. Information beyond the scope of the engines can be either in the deep/hidden web, or in social media.

In the hidden web, information is invisible to the search engines for various reasons, e.g.: the owner of a web site can block the access to web search engines; it can be an archive or in a database for which a subscription is required; or web crawlers might simply not have indexed the web page(s) (yet). Websites might also disappear, be removed and not available anymore on the current web (see chapter on [archived content](#)).

The market leader in web search engines is Google. It offers, like other engines do, advanced search possibilities and some applications which might help to dig deeper into the hidden web. **Some examples:**

- ✓ [Google Custom Search](#) is an example for such an application. It allows you to give a web link to Google thereby limiting the search to your selected web sources. By doing so, you create your personalised search engine.

A researcher on Afghanistan can create a customised search engine by selecting exclusively sources specialised on the country, for example, [Pajhwok Afghan News](#), a reliable Afghan newspaper; or the website of [UNAMA](#), the United Nations Mission in Afghanistan. By giving Google the command to only search a list of such web sites, the researcher limits the results to high-quality and specifically relevant sources, which makes the research more efficient and time-saving.

Figure 1 Example of a Google Custom search


The screenshot shows a Google Custom Search interface titled "Afghanistan Searcher". The search bar contains the word "Sikh" and the search button is labeled "Search". Below the search bar, it indicates "About 5,760,000 results (0.22 seconds)". The search results are listed as follows:

- Afghan Hindus and Sikhs | Culture, Religion, History and News**  
Feb 20, 2014 ... KABUL: The Wolesi Jirga on Saturday rejected a presidential decree, reserving a seat for the Hindu and Sikh minority in the lower house of ...  
[www.afghanhindu.info/](#)
- Marginalized Sikhs 'Don't Care' About Afghan Election**  
Apr 2, 2014 ... While many in Afghanistan are gripped by election fever, members of the country's tiny Sikh community say they couldn't care less.  
[www.rferl.org/content/sikhs-afghanistan-election.../25318893.html](#)
- Afghan Hindus, Sikhs Seek To Reclaim Their 'House Of Love'**  
Nov 11, 2011 ... If all goes well, a long-exiled community of Hindus and Sikhs will soon return to live in their village in Afghanistan known as a "House of Love."  
[www.rferl.org/content/hindus\\_sikhs\\_seek.../24388591.html](#)
- Hindus, Sikhs To Get Parliament Seat**  
Sep 4, 2013 ... Afghan President Hamid Karzai has issued a legislative order giving the country's Hindu and Sikhs a special seat in the lower house of the ...  
[www.rferl.org/content/afghan-hindu-sikh/2595408.html](#)
- Hindus, Sikhs of Afghanistan angered by Afghan parliament decision**  
Jul 31, 2013 ... Hindus and Sikhs of Afghanistan The Central Council of Hindus and Sikhs in Afghanistan along with members of civil society organization ...  
[www.ishana.com/hindus-sikhs-of-afghanistan-angered-by-afghan-parliament-decision-2253](#)
- Hindus, Sikhs warn of leaving Afghanistan | Pajhwok Afghan News**  
Jul 31, 2013 ... KABUL, (PAH) - The Council of Hindus and Sikhs in Afghanistan on Wednesday warned of leaving the country if they were not given proper ...  
[www.pajhwok.com/en/2013/.../hindus-sikhs-warn-leaving-afghanistan](#)
- Decree Reserves Seat for Hindus and Sikhs in Parliament**  
Sep 4, 2013 ... President Hamid Karzai issued a Legislative Decree that would jointly entitle the Hindu and Sikh populations of Afghanistan a single seat in the ...  
[www.tolnews.com/.../11907-presidential-decree-reserves-seats-for-hindu-and-sikh-in-parliament](#)
- Sikhs celebrating Vaisakhi in Nangarhar (Video) | Pajhwok Afghan News**  
Apr 12 2012 ... JALALABAD ( PAN ) - The long-oppressed Sikhs on Wednesday began celebrating Vaisakhi, the holiest day on their religious calendar, ...  
[www.pajhwok.com/en/.../sikhs-celebrating-vaisakhi-nangarhar-video](#)
- Sikhs throng temples to celebrate Vaisakhi | Pajhwok Afghan News**  
Apr 11, 2013 ... JALALABAD (PAN) - Sikh families on Thursday thronged holy temples in eastern Nangarhar province to celebrate Vaisakhi, the holiest day on ...  
[www.pajhwok.com/en/.../sikhs-throng-temples-celebrate-vaisakhi](#)

<sup>1</sup> A web crawler is a specialised software which systematically browses the internet with the aim of indexing the content



- ✓ [Google Maps](#) has a photo layer, which allows you to find pictures from an area and some information on who posted it.



If you have an interesting image about which you want to find further information, there is a camera button  in the Google image search which allows you to look for that image on the whole web (reversed image search).


Some search strategies also help you exploring the hidden web. For instance: don't try to find the specific answer directly on Google, but use Google to find a specialised database in which you can find the answer.

If you are looking for a company and you don't find it by searching the name on Google, try to think where the company could be registered which might be in a database that is available online, but possibly outside the scope of Google. For example, look up on Google for the website of a Chamber of Commerce and see whether it has a database with registered companies.

## 2.2. COI Portals

### 2.2.1. Ecoi.net

[Ecoi.net](#) is a publicly available portal gathering up-to-date COI relevant in procedures for international protection. It is maintained by the Austrian Red Cross (department ACCORD) and Informationsverbund Asyl & Migration (Germany). Ecoi.net contains more than 225.000 documents from more than 155 sources. Interestingly for researchers, ecoi.net provides a source description (clicking on the  icon) as well as a standardised citation format ( icon). It also offers English and German interfaces and a full-text (simple and advanced) search engine.

Amongst some of the features of ecoi.net are the **country pages** which display a selection of important documents, as well as country profiles, maps, national legislations and links. Personalised features, such as **My ecoi.net**, allow users to receive updates (new reports, other developments) on selected countries of origin after subscribing to a weekly email alert. Users can also assort relevant documents of a research in a basket in order to have them clearly arranged and to easily access them. Users have to click on the research basket icon  which appears on the top right of each entry. They can then copy and paste the list of documents into reports in form of a standardized reference list. In order to use these personalised features, users simply have to register by selecting the "registration" tab accessible from the Home Page.

Finally, the [ecoi.net blog](http://www.ecoi.net/blog) (<http://www.ecoi.net/blog>) informs users about ACCORD's work, but also important new publications and any relevant news in the field of COI. You can subscribe to the blog's RSS feed after registration.

### 2.2.2. Refworld

Similarly to ecoi.net, [Refworld](#) - UNHCR's public domain protection and information database - provides COI, but also legal and policy information. It contains more than 190,000 documents. In

2013, Refworld underwent an upgrade to improve its search capabilities and enhance the user experience. Amongst the changes are an improved keyword list and a Russian interface.

Like ecoinet, Refworld offers a personalised feature for subscription to weekly alerts. This feature allows receiving updates on selected countries of origin and selected publishers and document types. One may also save searches. To access this feature, users simply need to register by clicking on the **My Profile** tab located at the top right corner of the home page.

The Country News page can be particularly useful since it allows you to get a quick overview of what has been recently published on a particular country of origin.

You can find documents in Refworld by either searching or browsing. By default, Refworld uses assisted search to match spelling and other variations. Thus you do not need to know all possible spellings of a word in order to get maximum relevant results. So if you search for “bidoon citizenship”, for instance, the assisted search will look automatically for results with bidoon, bedoon, bidun, bedoun, etc. Assisted search also encompasses UK/US spelling variations (e.g. honour/honor), apostrophes (e.g. sharia/shar'ia) and many common hyphenations (e.g., subclan/sub-clan).

When you browse in Refworld, you click through:

- Categories
- Publishers
- Document types
- Topics
- Countries

If you retrieve a huge amount of documents, you can narrow the results using the Publisher, Country, or Topic tabs, or you can filter the results using the filter bar that appears above the document listing in the search.

### 2.2.3. The Common European COI Portal

The [Common European COI Portal](http://easo.europa.eu/) (<http://easo.europa.eu/>) is a platform that enables first instance asylum officials from EU+ Countries to access COI from various sources. Developed by the European Commission, with the support of EU+ countries, the Common European COI Portal is owned and maintained by EASO since the end of 2012.

Information available from the COI Portal originates from:

- Five **external COI databases** which are connected to the COI Portal (MILo/Germany; OFPRA/France; LANDINFO/Norway; LIFOS/Sweden; TELLUS/Finland)<sup>2</sup>. The connected COI systems deliver their own COI national products, as well as COI authored by third parties (NGOs, International organisations, media...), via the Portal.

---

<sup>2</sup> Additional COI databases are expected to connect in the future.

- The so-called “**Upload Area**” where EU+ Countries which do not have a COI database or which do not wish to connect can store their own national COI products (Country of Origin reports, FFM reports, responses to COI queries, etc.). The Upload Area is also used by EASO to store COI authored by EASO (COI reports, EASO Country Report methodology, EASO working documents, COI-related meeting reports). Additionally, the COI Specialist Networks<sup>3</sup> use the EASO Upload Area to share lists of sources, reports and workshops documentation.

Users of the Common COI Portal can search and retrieve information using a **simple or an advanced search engine**, regardless of where the information comes from (i.e., from the connected systems or the upload area). The “News” section allows users to be alerted about recent COI publications or events.

Within each EU+ country, the National Common Portal Administrator (NCPA) manages internal users access<sup>4</sup> and the national upload area (where relevant), with the cooperation of the EASO COI Portal Administrator.

## 2.3. Archived content

### 2.3.1. Wayback Machine

The US-based non-profit organisation Internet Archive provides a tool called [Wayback Machine](https://archive.org/web) (<https://archive.org/web>) preserving complete websites from 1996 until present. Once you enter a web site address (URL), Internet Archive displays an overview of previous versions of the particular website (focusing on textual content). You can then select a date and browse through the archived site (see Figure 2 for the example of the European Commission website).

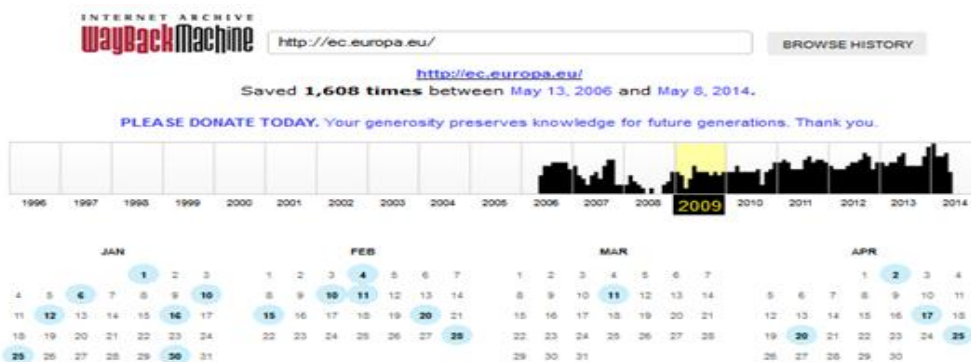


Figure 2 Print screen showing that the European Commission Website has been saved by the Wayback Machine 18 times between January and April 2009

<sup>3</sup> Currently, COI Specialist Networks have been set up for the following countries: Syria, Somalia, Pakistan, Iran, the Russian Federation, Iraq and Afghanistan. More COI Specialist Networks may be set up in the future. The COI Specialist networks consist of EU+ COI researchers who specialise or want to specialise on a specific country of origin. Members of the networks share information and sources, meet on a yearly basis to exchange expertise and generate knowledge. They may also contribute to the drafting of EASO COI publications.

<sup>4</sup> Only first instance asylum officials from EU+ Countries can access the COI Portal. To obtain more information about access to the Common European COI Portal, please contact: [COIPORTAL@easo.europa.eu](mailto:COIPORTAL@easo.europa.eu)

### 2.3.2. Screenshots.com

[Screenshots.com](http://www.screenshots.com) ([www.screenshots.com](http://www.screenshots.com)) has a different approach: it takes snapshot *images* of a website's front page. This can be useful to verify a website's existence or to compare changes in its layout and appearance over time.

## 2.4. Searching social networks

### 2.4.1. Twitter

[Twitter](https://twitter.com) can be compared to a public short text messages (SMS) service. "Following a twitterer" is similar to subscribing a specific news channel, or a very frequently updated blog (this is why Twitter and similar services are often referred to as *microblogging*). You can follow individuals like journalists or researchers. Also many established newspapers, broadcast stations (e.g., BBC), or news agencies (e.g., Reuters) offer additional services free of charge on this channel. All twitter services require signing up for a Twitter account.

#### *Topsy – Twitter Advanced Search and Archive*

The service company [Topsy](http://www.topsy.com) provides a real-time search interface for Twitter messages (= tweets). Apart from any type of keyword search, you can limit your query or results to a specific date range (from 2006 until today up to the minute) and language (currently English, Chinese, Japanese, Korean, Russian, German, Spanish, French, Portuguese, and Turkish).

A service called Topsy Alerts allows you to track specific ...

- **events and names**, e. g. **#EP2014** (the European Parliament Elections 2014)
- **topics**, e. g. **#Homs #Syria #Alawite** or **#erdoganblockedtwitter**
- **news media**, e. g., **@TimesOfMalta**

More information on microblogging is available at <http://en.wikipedia.org/wiki/Microblogging>

### 2.4.2. Facebook

[Facebook](https://www.facebook.com) is a powerful tool to find specific information about people, but also about places and institutions. The basic search in Facebook allows searching for and filtering based on names, e-mail address, employer or location. The quickest way to look up a person is to enter an email address in the search box, which uniquely identifies every user on Facebook.

More recently, Facebook introduced a semantic search engine called [Facebook Graph](https://www.facebook.com/graphsearch), which offers vastly extended search capabilities. This allows combining several building blocks such as personal details, places people visited, the things they "liked" or the photos they uploaded to make advanced queries. To write these queries, you need to use the specific terminology and syntax of the search engine – suggestions are provided as soon as you start typing. Some examples:

- People from **Nairobi, Kenya** who visited **Westgate Mall**
- People who are 30 years old and went to **University of Tehran** and like **Israel**
- Recent photos taken in **Kabul, Afghanistan** of people who live in **Kabul, Afghanistan**

Please note that Facebook Graph currently (Status 06-2014) is available in the US Web version only, which means that only users who choose “English / US” in their language settings get access to it. Facebook apps for mobile devices are also excluded so far.

More information is available on the official **Facebook page**:

[www.facebook.com/help/www/558823080813217](http://www.facebook.com/help/www/558823080813217)

### 2.4.3. Metasearch in social networks

Several tools exist to simplify searching across multiple social networks at once. These include for example:

- **Icerocket** (<http://www.icerocket.com/>) offers keyword search and some simple search operators across public content in Blogs, Facebook and Twitter. It also offers premium services such as advanced social media monitoring for a paid subscription.
- **Social Searcher** (<http://www.social-searcher.com/>), a similar tool to Icerocket, offers simultaneous searches in Facebook, Twitter, Google Plus, as well as email notifications for keywords.
- **Geofeedia** (<http://geofeedia.com/>) is an advanced location-based social media monitoring and analysis tool. It filters social media data by location and time. Based on a specific address or area on a map, the service shows the latest content of [Twitter](#), [Facebook](#), [YouTube](#), and others. Geofeedia is available on paid subscription basis only.

Generally, metasearch tools offer less specific search operators and filtering options than search tools for a specific network. This makes them more suitable for an initial overview, or as a starting point for research on a particular person or topic.

### 2.4.4. Specific regional social media

While Facebook, Twitter, YouTube and the like are dominating the social media scene globally, it may be worthwhile to include regional social media in the research, depending on the country of interest.

Some important examples include:

- Russian speaking world: **Vkontakte** ([www.vk.com](http://www.vk.com))
- China: **Renren** ([www.renren.com](http://www.renren.com)); **Qzone** ([www.qzone.qq.com](http://www.qzone.qq.com))
- India, Pakistan: **Orkut** ([www.orkut.com](http://www.orkut.com))

It should also be noted that activity in social networks can be very volatile and even networks with millions of users could almost disappear within months. An extensive overview on social networking websites is available at Wikipedia: [http://en.wikipedia.org/wiki/List\\_of\\_social\\_networking\\_websites](http://en.wikipedia.org/wiki/List_of_social_networking_websites)

## 3. Organising information

### 3.1. Automatisation and aggregation tools

#### 3.1.1. Google and Yahoo Alerts

An alert service can be used for monitoring a specific topic or country of interest. You can predefine queries with specific keywords, a region or country, and language. You will then receive regular updates (i.e., as soon as possible, daily or monthly) on pages recently added to the overall Google or Yahoo search engine index, optionally either delivered to your email address or via a so-called **RSS feed** (see below). It is also possible to limit your results to Yahoo's or Google's news channel, blogs, videos, discussions in chats, etc.

In Figure 3, you find an example of how to define a search query of specific interest, if you are a COI researcher focusing on Nigeria, and in particular on the question of gender.

#### Alerts

Search query:

Result type:

Language:

Region:

How often:

How many:

Your email:

Figure 3 Criteria for a Google Alert on the topic of gender in Nigeria

**Please note:** Selecting Nigeria as region here instead of “Any Region” would limit your search to the index of Google Nigeria (*google.ng*) only!

[Google Alerts](#) and [Yahoo Alerts](#) are available with a registered Google or Yahoo account only.

#### 3.1.2. Monitoring changes on websites / blogs using RSS

**RSS (Really Simple Syndication)** is a web format for publishing information. This allows users to subscribe to updates with so-called Feed Readers (RSS Readers), which automatically download updated content from websites. Feed readers free the user from checking manually the websites for updates. Many websites and blogs publish content in RSS format (usually called a *feed*). Feeds are usually identified by the universal RSS symbol (see Figure 4).

Feed readers also allow aggregating feeds from many different sources and organising them in different ways. For example, if you as a researcher need to closely follow up on the situation in Syria, you could be interested in news and publications by Human Rights Watch, UNHCR, and the International Crisis Group. You would then select the appropriate feed(s) from their respective websites, e.g.:

- <http://www.crisisgroup.org/en/Rss.aspx>
- <http://www.hrw.org/rss>
- <http://www.unhcr.org/pages/49c3646c23f.html>

and organise them in your feed reader under *Syria updates*. You would now automatically be notified on any updates without visiting the websites individually.

Feed readers, or feed-reader functionality, are available as desktop applications, as web-based services, or browser plugins<sup>5</sup>. The advantages of web-based services are ease of access regardless of the device and data synchronisation between devices, but you depend on the service provider and they require a user account.

Examples for feed readers in desktop applications are:

- Mozilla Thunderbird
- Microsoft Outlook



Figure 4 RSS icon

Examples for web-based RSS feed readers services are:

- [Feedly](http://www.feedly.com) ([www.feedly.com](http://www.feedly.com))
- [The old reader](http://www.theoldreader.com) ([www.theoldreader.com](http://www.theoldreader.com))
- [Feedspot](http://www.feedspot.com) ([www.feedspot.com](http://www.feedspot.com))

Browsers, such as Mozilla Firefox or Internet Explorer, have basic built-in functionalities for handling RSS feeds. Additional features to manage and display RSS feeds are often available via browser plugins (e.g., RSS Feed Reader for Google Chrome).

### 3.1.3. Twitterfeed

[Twitterfeed](http://twitterfeed.com) (<http://twitterfeed.com>) is a utility that allows you to automatically feed content to social media platforms like Twitter, Facebook, LinkedIn, etc. It takes your RSS feeds and sends them as updates (so-called postings) to the social platforms of your choice. You can also customise these postings.

It is a very useful tool to organise an automatised flow of information. It allows for aggregating and presenting news or new information from the sources you select on your social media platform. With this tool you tweet while you're sleeping.

### 3.1.4. Monitoring changes on websites without RSS

If a website does not offer RSS feeds, there are tools available which automatically monitor content-related changes on web pages. Monitoring changes on websites is possible using web-based services. They usually require creating an account, and automatically send out email messages specifying changed web pages. Examples of such services include:

- [Watch That Page](http://www.watchthatpage.com) ([www.watchthatpage.com](http://www.watchthatpage.com))
- [ChangeDetection.com](http://www.changedetection.com) ([www.changedetection.com](http://www.changedetection.com))

---

<sup>5</sup> A small application that extends a browser's functionality.

Some of these monitoring tools are available as browser plugins, they usually offer greater opportunities for customisation than their web-based counterparts (e.g., time interval for checking, monitoring only section of a page):

- Page Monitor (for Google Chrome)
- Update Scanner (for Mozilla Firefox)

### 3.2. Social bookmarking tools

So-called social bookmarking tools allow users to better organise bookmarks and share them with others. Compared with traditional bookmarks stored in a browser (e.g. Favourites in Microsoft Internet Explorer), these tools offer advantages such as:

- **Tagging** (a form of labelling) URLs with keywords offers a multi-dimensional way of organising and retrieving bookmarks in a more flexible way than a folder system.
- Bookmarks are **accessible from all devices** and are always synchronised.
- It is possible to **share** bookmarks, e.g., with a team of COI researchers or the public.
- Based on keywords (tags), it is possible to find public URLs saved by other users.

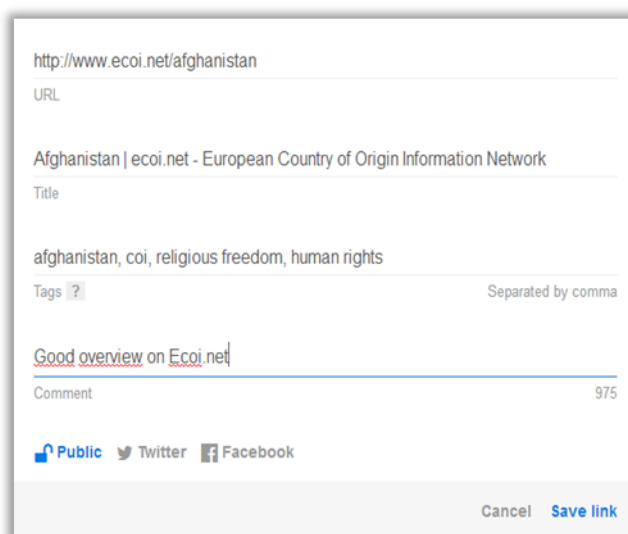


Figure 5 Example for tagging links with a social bookmarking tool (Delicious)

Popular examples of such tools (which often provide additional functionalities) include:

- [Delicious](http://www.delicious.com) ([www.delicious.com](http://www.delicious.com))
- [Diigo](http://www.diigo.com) ([www.diigo.com](http://www.diigo.com))
- [Evernote](http://www.evernote.com) ([www.evernote.com](http://www.evernote.com))
- [Bibsonomy](http://www.bibsonomy.org) ([www.bibsonomy.org](http://www.bibsonomy.org))

### 3.3. Reference / citation management tools

A **reference /citation management tool** can assist you in managing your personal electronic reference/bibliographic citation collection for systematic retrieval and reuse of citations for your publications. Most of these tools also allow for sharing your references with other individual users and/or within a larger community of interest.

You can select from a number of pre-set **citation styles** or **define your own institution's style**. Some tools are web-based and can be used free of charge within your browser (e. g., [Zotero](http://www.zotero.org) and [Mendeley](http://www.mendeley.com)) and also allow synchronising your references when working on different computers.



Others (as, e. g., [Endnote](#) or [citavi](#)) are commercial software products and require licensing and installation in your institution's ICT network and/or on your local PC.

An overview on reference/citation management tools is available at [en.wikipedia.org/wiki/Comparison\\_of\\_reference\\_management\\_software](http://en.wikipedia.org/wiki/Comparison_of_reference_management_software).

### 3.4. Collaborative online query system

Some members of the international volunteer network Asylos have created and are further developing a freely useable, community-based tool called **kollekt.io** (draft title). The tool is designed for collectively handling COI requests (posting queries and answers). After having posted a query, you send its web link via email to one or more persons whom you would expect to answer the query at least partly. For the time being, the tool does not provide for retrieval of former answers, but is suitable best for ad hoc queries to be handled by a group of users. The prototype requires signing up for a user account and is publicly available at <https://github.com/asylos/kollekt.io>.

## 4. Privacy while searching the Web and data protection issues

All online activities will usually leave traces. This is important to keep in mind when searching for sensitive information, e.g., about a particular subject or a person. It may happen that the researcher and/or the subject of research (e.g., an applicant) be identified by third parties. In certain cases, this might even compromise the researcher's, the source's, or the applicant's security.

Tools and strategies exist to mitigate risks related to data protection. It is, however, important to stress that tools as such are not sufficient to solve privacy issues in this complex field. It requires a holistic approach, including awareness and understanding of possible risks (e.g., viruses and malicious software, misuse of personal data, software security holes, hackers, intelligence agencies) and a good strategy in addressing these.

### 4.1. Basic security concepts

As basic prerequisites to staying anonymous, it is important to remember not to use any personal data or email addresses in web forms. Also, it is important to keep your browser up-to-date with the latest security updates.

#### 4.1.1. Browser settings

To avoid having sensitive information stored in your browser (and hence reduce the risk of potential security holes which could enable adversaries to read out this data), the web browser's history, cookies<sup>6</sup> and the cache<sup>7</sup> should be cleared after each session. Alternatively, you can also use the

---

<sup>6</sup> A small piece of data stored in the web browser, for identification and tracking purposes. More information at: [http://en.wikipedia.org/wiki/HTTP\\_cookie](http://en.wikipedia.org/wiki/HTTP_cookie).

<sup>7</sup> The browser cache is a temporary storage for data requested from a website (images, text, etc.), to speed up browsing when repeatedly visiting a website.

**Incognito or Invisible mode** offered by most browsers, which automatically erases history and cookies after closing the window.

#### 4.1.2. Security plugins

For major browsers such as Mozilla Firefox or Google Chrome, plugins exist which can improve some aspects of web browsing security. Some examples are:

- **Better Privacy:** Protection against long-term and special tracking cookies which cannot be easily removed
- **HTTPS everywhere:** Encrypts communication to major websites where https communication channels are offered
- **Ghostery :** Detects “invisible” trackers placed on web pages by advertisement networks and lets the user manage/block those
- **Noscript:** Removes all Javascript, Flash, and Java, which are frequently targeted for malicious attacks on user data

These little helpers can be found and installed via the browser's built-in extension/plugin system.

#### 4.1.3. Using an anonymous search engine

Today, Google is predominantly used for web searches. It is, however, tracking user behaviour by collecting data such as the IP address, platform and the search terms. Some alternative search engines promise not to store any information about their users, these include:

- **Startpage** (<https://startpage.com/>) actually delivers results from Google, but in an anonymized way, relaying the data entered by the user.
- **DuckDuckGo** (<https://duckduckgo.com>) is an independent search engine focusing on protecting the user's identity.

## 4.2. Advanced concepts

### 4.2.1. VPN (Virtual Private Network)

A Virtual Private Network (VPN) provides an encrypted<sup>8</sup> connection from the user's device to a remote proxy server<sup>9</sup> (which could be located in a different country) and from there to the internet. In this way a VPN can be used to protect personal identity and location (the user will appear to come from the location of the proxy). VPNs are especially useful to protect information in settings where internet access is only possible through an open, shared network (e.g. hotel lobbies, airports). Several commercial providers offer VPN services, usually for a monthly subscription fee.

---

<sup>8</sup> To protect information by encoding it so that only authorised people can access it.

<sup>9</sup> A computer system acting as an intermediary for requests from clients to the internet.

More information on VPNs can be found at <http://en.wikipedia.org/wiki/Vpn>

#### 4.2.2. TOR (The Onion Router Network)

**TOR** is a network of encrypted virtual tunnels, which routes all communication from the user to a website or service (e.g., chats, email) through a random path of nodes, which are part of a volunteer network. This system offers a very high level of anonymity in the web; it however comes at the cost of being quite slow for normal browsing.

For easy use, a pre-configured browser bundle exists, which can be installed and run from external drives. More information is available at <https://www.torproject.org>

#### 4.2.3. Securing Email by encryption

Email messages often travel through multiple secure and insecure networks, carrying the risk that the data is intercepted, copied or read out on the way. To secure sensitive email communication using encryption, different approaches exist. The most common one is a so-called “public-private key” approach where the sender uses the public key of the desired receiver to encrypt a message. The receiver then utilises his private key (which is always kept secret) to de-crypt the message (as shown in Figure 6).

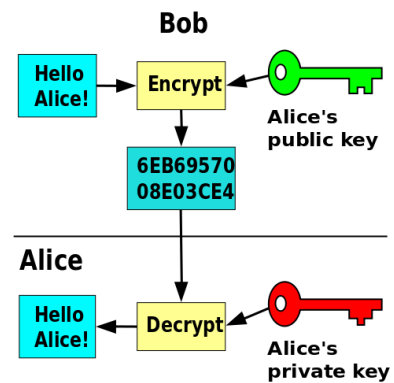


Figure 6 Public-private key approach (Source: [http://commons.wikimedia.org/wiki/File:Public\\_key\\_encryption.svg](http://commons.wikimedia.org/wiki/File:Public_key_encryption.svg))

In any case, encryption requires that the sender and receiver exchange information *before* the secure communication takes place, in order to allow the receiver to de-crypt the message.

Free tools exist which allow encryption independently from an email provider (more information can be found at [http://en.wikipedia.org/wiki/Pretty\\_Good\\_Privacy](http://en.wikipedia.org/wiki/Pretty_Good_Privacy) ). Several commercial email providers also offer automated encryption of messages.

#### 4.2.4. Further reading

For more information, the following websites may be consulted (*based on recommendations by F. Weinberg, HURIDOCs*):

- **Security-in-a-box**: Tools and tactics for your digital security <https://securityinabox.org>
- **Me and My Shadow**, <https://myshadow.org>
- **Digital Security and Privacy for Human Rights Defenders**, handbook <http://www.frontlinedefenders.org/eseaman>
- **EFF (Electronic Frontier Foundation) tutorials for risk analysis**: <https://ssd.eff.org/risk>

## Subject Index

<i>aggregation tools</i> .....	14	<i>Monitoring websites</i> .....	15
<i>anonymous search</i> .....	18	<i>Mozilla Firefox</i> .....	15
<i>automatisation tools</i> .....	14, 15	<i>Mozilla Thunderbird</i> .....	15
<i>Better Privacy</i> .....	18	<i>Noscript</i> .....	18
<i>Bibsonomy</i> .....	16	<i>Orkut</i> .....	13
<i>bookmarks</i> .....	16	<i>Page Monitor</i> .....	16
<i>Browser</i> .....	15	<i>Personalising</i> .....	9, 10
<i>cache</i> .....	17	<i>posting</i> .....	15
<i>ChangeDetection</i> .....	15	<i>Privacy</i> .....	17
<i>citation management</i> .....	16	<i>public-private key</i> .....	19
<i>citavi</i> .....	17	<i>Qzone</i> .....	13
<i>COI queries</i> .....	17	<i>Really Simple Syndication</i> .....	see RSS
<i>Common European COI Portal</i> .....	10	<i>reference management</i> .....	16
<i>cookies</i> .....	17	<i>Refworld</i> .....	9
<i>data protection</i> .....	17	<i>Renren</i> .....	13
<i>Delicious</i> .....	16	<i>RSS</i> .....	14
<i>Digital Security and Privacy</i> .....	19	<i>RSS feeds</i> .....	14, 15
<i>Diigo</i> .....	16	<i>RSS Readers</i> .....	14
<i>DuckDuckGo</i> .....	18	<i>Screenshots.com</i> .....	12
<i>ecoi.net</i> .....	9	<i>search engine</i> .....	8
<i>encryption</i> .....	19	<i>security</i> .....	17
<i>Endnote</i> .....	17	<i>Social bookmarking</i> .....	16
<i>Evernote</i> .....	16	<i>social media platform</i> .....	15
<i>Facebook</i> .....	12, 13, 15	<i>Social Searcher</i> .....	13
<i>Facebook Graph</i> .....	12	<i>Startpage</i> .....	18
<i>feed</i> .....	see RSS feeds	<i>tag</i> .....	16
<i>Feed Readers</i> .....	14	<i>tagging</i> .....	16
<i>Feedly</i> .....	15	<i>The old reader</i> .....	15
<i>Feedspot</i> .....	15	<i>The Onion Router Network</i> .....	See TOR
<i>Geofeedia</i> .....	13	<i>Topsy</i> .....	12
<i>Ghostery</i> .....	18	<i>TOR</i> .....	19
<i>Google</i> .....	8, 18	<i>Twitter</i> .....	12, 13, 15
<i>Google Alerts</i> .....	14	<i>Twitterfeed</i> .....	15
<i>Google Chrome</i> .....	15, 16	<i>Update Scanner</i> .....	16
<i>Google Custom Search</i> .....	8	<i>Virtual Private Network</i> .....	See VPN
<i>Google Maps</i> .....	9	<i>Vkontakte</i> .....	13
<i>HTTPS everywhere</i> .....	18	<i>VPN</i> .....	18
<i>Icerocket</i> .....	13	<i>Watch That Page</i> .....	15
<i>kollekt.io</i> .....	17	<i>Wayback Machine</i> .....	11
<i>LinkedIn</i> .....	15	<i>Web crawler</i> .....	8
<i>Mendeley</i> .....	16	<i>Web search engine</i> .....	see Search engine
<i>Metasearch</i> .....	13	<i>Yahoo Alerts</i> .....	14
<i>microblogging</i> .....	12	<i>YouTube</i> .....	13
<i>Microsoft Outlook</i> .....	15	<i>Zotero</i> .....	16

