eLucidate

Vol. 12 Issue 2, Summer 2015

ISSN: 1742-5921

This issue of eLucidate has been edited by Gary Horrocks email: gary.horrocks@gmail.com

eLucidate is published by UKeiG, the UK eInformation Group. Membership and newsletter subscriptions are handled by the UKeiG Administrator, Val Skelton, 4 Friendly Street Mews, London, SE8 4DU, tel: 020 8691 0246, e-mail: val.skelton@ukeig.org.uk

UKeiG is a Special Interest Group of CILIP: the Chartered Institute of Library and Information Professionals. CILIP Registered Office: 7 Ridgmount Street, London, WC1E 7AE, Registered Charity No. 313014

Contents

New Google, New Challenges

Karen Blakeman, RBA Information Services

karen.blakeman@rba.co.uk

Just when you think you've worked out how Google works it all changes, not just the commands but the whole way Google searches and displays results. Semantic search, artificial intelligence, extreme personalisation and quick answers that save you the bother of having to read any further are just the beginning. Add a dash of EU legislation and things can start to go very wrong.

It all started to go awry with the closure of Google labs and many innovative search experiments. Then slowly but surely services and search options were dropped from Google itself: starred results, page previews, time-lines, synonym search, the plus sign, translated foreign pages search and Google Reader to name but a few. The most recent to get the chop was Reading Level. Although I did not use it that often it was a nice little tool that enabled you, when necessary, to change the emphasis of results from basic, introductory information to more research biased documents. All done, as Google often says, to provide us with "a simpler, more intuitive Google experience". (Toward a simpler, more beautiful Google.) In reality, if a function is rarely used, doesn't bring in much revenue and requires significant human time and effort to maintain then it goes.

Remember that Google is a stock exchange listed company and its main aim is to make money for its shareholders. Much of that revenue comes from advertising and, yes, plenty of people do click on those ads. Google is not interested in those of us who use their services to conduct in-depth research; we rarely click on the ads, many of us have ad blockers and anyway some of our searches are so esoteric that no ads match what we are looking for anyway! Google is interested in the person looking for cheap flights to Prague; the person looking for cheap hotels in Manchester; the group of people who are attending a conference and are searching on their mobile devices for a local pub to meet up in. These are the people who will click on an ad or sponsored link occupying most of the screen that is "good enough" for their needs. It is these types of searches that make up the bulk of Google's enquiries and create revenue for Google.

What serious researchers are left with is a reduced core of commands for refining searches. Getting the most out of this limited toolkit often requires lateral thinking but it is well worth the effort. Judging from the Top Tips suggested by participants of the UKeiG Google workshops, 'site:' and 'filetype:' commands are top of the list enabling us to filter results by type of organisation (academic, NHS, government) and focusing on document type (PDFs for research papers and official documents, spreadsheets for statistics and raw data). Getting to know how to use the Google commands and search filters is the key to better, more reliable results.

For Google, mobile is where it is all happening and that is where it is concentrating most of its efforts. (It's Official: Google Says More Searches Now On Mobile Than On Desktop.) Obviously the layout of the information is different depending on whether you are using a desktop computer or a smartphone, but so is the actual content. This can present problems when trying to talk someone through a Google search because as well as past searches and browsing behaviour the results are personalised according to the type of device and operating system being used. Now Cards also seem to be a major part of Google's future mobile strategy. There are already over one hundred Cards for android users with more in the pipeline (Google Adds 70 More Now Cards.) Now Cards remind users of local traffic, the weather, special offers from local businesses, public transport departures, delays and arrival times and seemingly just about everything else without being prompted.

Some mobile search features eventually migrate to the desktop. "Facts" and information that immediately answer your question (Quick Answers) now regularly appear at the top and to the side of results. The topics range from local football results to symptoms of medical conditions. It is not clear how Google chooses the sources for these and they can change from day to day. The quality of some of them is dreadful and sometimes dangerous. I once came across a quick answer from an appalling site on food allergies. The person responsible for the site openly stated that they were not medically qualified and proceeded to give "advice" that was misleading and could have caused a life threatening reaction in an allergic individual. Looking at the site in more detail it became clear that its main purpose was to sell food diaries, books, and allergy testing kits. At least with that Quick Answer it was possible to click through to the original website.

For many answers no source is given and when tackled about this Google said that it doesn't provide a source link when the information is basic factual data and can be found in multiple places on the Internet. (When Google Shows A Source Or Credit For Quick Answers & Knowledge Graph.) Unfortunately, repeating something many times does not necessarily make it true and there are many examples of Google getting it wrong. (The quality of Google's results is very strained.) Never,

never trust Google's Quick Answers or so-called facts. Always check them with another source.

We will probably be seeing more "facts" in the future and it has been suggested Google could start to use them as an additional ranking factor. If a page contains data and information that is repeated across the web then it will be deemed to be authoritative. If it has conflicting statements it could be regarded as untrustworthy. So where does that leave a comprehensive scientific review article that tries to give a balanced overview of a topic and cites contradictory evidence? (Google wants to rank websites based on facts not links - 28 February 2015 - New Scientist and Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources.)

Then there is Google's habit of dropping terms from our searches without asking and, on occasion, completely rewriting your search. I wrote a blog posting about the problem that was picked up by Google's Dan Russell (Dear Google, stop messing with my search.) He commented:

"When you do a multi-term query on Google (even with quoted terms), the algorithm sometimes backs-off from hard ANDing all of the terms together. It's a kind of "soft" backoff. Why? Because it's clear that people will often write long queries (with anywhere from 5 to 10 terms) for which there are no results. Google will then selectively remove the terms that are the lowest frequency to give you some results (rather than none). Bear in mind that 99% of searchers have no idea why they'd want to hard AND, and just get frustrated when they get no results... But I see what you mean about wanting to know if there are NO hits to a given query. I'll pass this information along to the Google design team and see if we can't do something with this."

Google's reaction was to introduce Verbatim, which runs your search with no omissions and no variations on your terms. It is one of the most powerful tools for regaining control of the search process and essential as Google's search rewrites become increasingly bizarre.

There are times when no matter how many advanced commands you use you could still be missing important information as court cases and legislation begin to dictate what can and cannot be displayed. The EU ruling on the so-called right to be forgotten gives individuals the right to ask for links to information about themselves that is inaccurate, excessive or out of date to be removed from search results. This affects searches that include a person's name and all search engines that have offices or operate in EU countries plus Norway, Switzerland, Iceland and Lichtenstein. An individual has to send in a request to the search engine stating which links they want removed and why. It is then up to the search engine to decide whether or not the request is valid. If the search engine refuses the request the individual is entitled to take the matter further by applying to the relevant Data Protection agency.

According to its own transparency report Google has so far received 265,571 requests to remove about 964,533 links, of which 41.3% have been removed from its search results. Remember, it is just the links to the information that are removed; the information itself still remains somewhere on the Internet. You just have to know how and where to look for it. The easiest way to circumvent the ruling is to use a non-European version of Google such as Google.com or Google.ca that currently give you access to the full list of results. Do not assume that this will always be an option as there are some who would like the delisting to be worldwide.

Should you be using Google or general search tools in the first place? For example if you are interested in directorships held by a person, then the official company registers such as Companies House should be top of your list and won't be affected by the right to be forgotten. And, as was said earlier, never trust Google's facts and Quick Answers. Find out about the specialist sites that cover your subject area and type of information (news, statistics, company financials, government data, research papers etc.).

So what of the future? Google's Eric Schmidt recently said that the Internet will disappear by which he meant that it will become such an integral part of every device and appliance, and everything we do that we won't consider it as being something separate. Cars, washing machines, refrigerators that warn you when something starts to go off, and home environmental controls are just a few examples of this Internet of Things. Google has invested heavily in advanced technologies such as artificial intelligence and it is already being used in some areas of search, as part of robotics projects, driverless cars and even in the home. Nest, for example, is not just a home thermostat controller. It learns and adapts to your routine just as Google search personalises results by looking at your browsing patterns. If you are worried now about how much information the search giant has on you, imagine how much more it could collect through the Internet of Things.

Disconnecting and limiting the amount of data being gathered will no doubt be as difficult as it is now. "Are you REALLY sure you no longer wish to enjoy an enhanced, refrigerator experience by sharing the state of its contents with your friends?" Google will ask as I delete the fridge's profile from my Google dashboard. I await the day when, as I go past Tesco on my way home, Google flashes an alert on my Android phone warning me that the cheese in my fridge has gone off accompanied by ads for special offers on Gorgonzola. Perhaps you'd like a cheeky little merlot to go with that (£3 off)? Then Google Maps tells me the roads near my house are gridlocked, Google News informs me that several fire engines are in attendance outside my house, and my Google room thermostats report unusually high temperatures. Cue adverts for air conditioning units, ice cream machines and fire extinguishers all connected to the Internet and, of course, powered by Google.

The Internet of Things (IoT): Creating Really Big Data

Martin De Saulles, Principal Lecturer in Digital Marketing, University of Brighton

Information Matters

The term, Internet of Things (IoT), has been around since 1999 when brand manager Kevin Ashton at Proctor and Gamble applied RFID technology to streamline his company's massive supply chain. Since then it has grown to encompass the deployment of Internet-connected sensors and trackers across a range of industries and business processes. Essentially, it is an evolution of closed, proprietary telemetry systems that go back to the 1970s. On an industrial scale it includes connected sensors to track the wear and tear of a jet engine and water quality sensors to measure the safety of drinking water. At the city level it includes initiatives such as the AirSensa project which is installing 10,000 air quality sensors around London, each of them providing accurate, street-level data in real time for third parties to analyse. Individually many of us are already contributing data to the IoT via smartphone apps or devices such as Fitbit that track our movements, exercise regimes and health status. Google's Nest thermostat is a good example of how everyday household devices are being connected to allow the remote monitoring and control of domestic lighting and temperatures. Google's announcement in late May of its Brillo operating system for smart devices reveals its plans to extend data gathering beyond search and Android smartphones.

As these systems are rolled out and integrated into our daily lives the biggest challenge is going to be making sense of the data thrown off the myriad of devices in our pockets, houses, workplaces and cities. Cisco estimate that the number of Internet-connected devices overtook the number of people on the planet in 2008 and that by 2020 there will be 50 billion 'things' transmitting information. By 2018, they claim, the data created by IoT devices will be 277 times higher than the amount of data generated by smartphones and PCs. With a Boeing 787 generating 40 terabytes (TB) of data per hour of flight and the mining operations of a company like Rio Tinto generating up to 2.4 TB per minute it is easy to see how the IoT presents huge opportunities for companies able to create value from it all.

As with any new technology there is the danger of getting carried away with its potential for radical change. Last August, Gartner put the IoT at the top of the Peak of Inflated Expectations on its Hype Cycle of emerging technologies with Big Data heading into the Trough of Disillusionment (if you're not familiar with Gartner's Hype Cycle, this Guardian article will be useful.) However, this is not a good reason to write the IoT off as a fad that will soon pass. IoT systems are being developed and deployed and, despite some inevitable hiccups on the way, in a decade I believe the embedded connectivity of many everyday items will be the reality. The ability for cost savings at the household and industrial levels are too compelling for it to go away.

However, there are a number of issues to be resolved before that happens. At a personal level, privacy is a major concern. Much of the personal data being uploaded into the "cloud" from IoT devices is of a sensitive nature. For example:

- Health status data (smartphone fitness trackers, smart watch health monitors)
- Eating habits data (smart fridges - Yes, the mythical smart fridge does exist)
- Quality of our driving data (smart boxes connected to cars are becoming common to reduce insurance premiums, particularly for younger drivers)
- Household status data (Google's Nest thermostat is able to tell whether anyone is in the house by tracking movement)

It is not difficult to see how this data, particularly when combined with other information about our lifestyles, could be used in ways not necessarily in our best interests. Therefore, the security of how data flows from IoT devices to third parties and how secure they keep the data once it arrives is also crucial. This is of particular concern with the firmware powering many devices that, unless closely monitored, is prone to third party attacks if not updated regularly. Managing updates on a PC is probably difficult enough for most non-technical consumers so extending this across to smart fridges, door locks and thermostats, for example, is another matter altogether.

Another major challenge for IoT developers is how to make money from their initiatives. At the bottom of the value chain are the manufacturers of devices such as thermostats and fridges but the real value is not expected to be at this level. It is further up the value chain where the data is captured and analysed that the real profits are probably to be found. Google did not spend $3.2 billion buying smart thermostat manufacturer, Nest, because it saw huge profits in the hardware. Google's underlying business model is in having access to and control of massive quantities of data. Nest, for example, makes money by doing deals with energy companies to give them a degree of control over the thermostats. This allows the energy providers to micro-manage the temperatures of their customers' houses resulting in cost savings for the customers and more efficient utilisation of the grid and power generation.

Questions over who controls the data generated by the IoT and what is then done with it will be central to who makes the biggest profits. It is no coincidence that IBM, Cisco, Microsoft, Google and others are investing billions of pounds in platforms and infrastructure to manage these data flows and make sense of it in a way that allows new services to be created, many of which have not yet been thought of. Imagine going back to 1994 as the first web browsers were being launched and the Internet was just starting to take a hold as a communications and information-sharing platform. Who would have imagined that twenty years later it would be so embedded in our private and work lives? I suspect we are at about the same stage with the IoT.

How Good is your Content Quality?

Martin White, Managing Director, Intranet Focus Ltd

martin.white@intranetfocus.com

I wrote on this important theme in issue 3-4 of eLucidate last year but I would like to revisit the issues I raised in order to articulate and expand on a much wider range of topics.

1. Content quality management

Having not worked for a law firm for the last fifteen years I am now working for two of the world's leading firms at the same time. As you might expect there are both similarities and differences, and both are quite fascinating projects. However there is a common problem that both firms face, and that is that the content quality of their intranet content is below the level that would be appropriate for any firm, especially so in a major law firm. If you look at the individual content items on the intranets it might be quite hard to spot low quality content. However it becomes painfully obvious in search results, where tens if not hundreds of items are supposedly listed in decreasing order of relevance.

The other firm, which is in the process of a migration to SharePoint 2013, has set a target of only migrating high-quality content and then maintaining this level of high quality in the future. That is a very laudable objective but it begs the question about how content quality can be measured. Some years ago a multi-national engineering company commissioned me to develop a set of content quality guidelines and I have been using them, occasionally somewhat modified, ever since.

2. Four key principles

In no particular order there are four key principles of information quality.

a. Purpose

Every piece of information on an intranet should be there for a purpose. Before publishing, the content owner (who may not be the publisher) should consider the potential value of the content they are publishing and write it in a way that will help an individual employee or a group of employees to use and share the content.

b. Discovery

Intranet users will find the information they need by browsing through the site navigation, by carrying out a search or by setting up a profile to alert them to information as soon as it is published. The role of a content publisher is not just to create information but also to ensure that the people who will benefit from it can find it.

c. Trust
Every piece of information on the intranet will be used in some way to make a decision that could affect the reputation of the firm. Every user of information has to be able to trust it implicitly or if they have any doubts they are able to check with the content owner.

d. Ownership
It is the responsibility of all employees to ensure that content that they own is maintained in a way that it meets the purpose for which it was published, that it is findable and that staff can trust that it is accurate and valid.

3. Standards and guidelines

A standard is an absolute. Either an element of content meets the standard and can be published, or it does not meet the standard and cannot be published. A standard must be carefully defined so that there can be no dispute that the standard has, or has not been met. A guideline is a recommendation. It can be ignored and there is no sanction that can be applied.

As an example, a title such as:

Presentation at the 2014 Berlin workshop – New opportunities at sea…

…might meet the standard for a title but does not take account of a guideline on titles which might suggest that in fact a better title might be:

Expansion of marine contract capabilities in Singapore in 2015-2016

The standard title is not incorrect but neither is it at all useful as a means of helping a user decide on the potential value of the content item.

Successful content governance models have:

- A minimum number of standards which are rigorously enforced
- A set of guidelines which are illustrated with examples of good and poor practice
- Certified training procedures that ensure that publishers have a justification and a reward for using guidelines

As far as possible standards should be device-independent but guidelines should take into account the differences between desktop, tablet and smartphone devices.

4. A minimum acceptable quality standard

From the principles above there are three minimum standards for content

- It has a title
- It has a date of publication
- It has a content owner

A good quality title is important because we use titles as a filter, on a page, in a navigational list or a search results page. Search software also is biased towards a title. Perhaps counter-intuitively a title should not contain a version number. This is because only the latest version should be on an intranet! However there will need to be both a standard and a guideline. The standard states that a title must be present. An associated guideline sets out some characteristics of a good title. Individual publishers may have different views on what constitutes a good title, which is why the standard only relates to its presence or otherwise, as these cannot be disputed.

If you want to see an example of what happens when title management is missing just undertake a search of the Ofcom website.

The date should indicate the date on which the content owner passed the content as being fit for publication. In effect before this date the content did not exist. It is not the date of publication or the date of migration. In due course there should be a review date but for now the content owner certifies that a) the content is valid and definitive and b) the content owner will ensure that the content is updated and republished should it become invalid. At the same time the earlier version should be removed from the active server. For an example of an unusual approach to date management run a search on Syngenta and you will see that all the results carry the date of the last working day, a result of a server management issue.

Unlike web pages the name of the content owner is itself a very important indication of quality. The name signifies that the content owner is putting their reputation on the line for the quality of the content. The user, if they wish, is able to check the credentials of the owner and also contact the owner for additional information. If the owner is not with the firm then it is the role of the manager of the owner to appoint a new owner. No item of content can be owned collectively by a department or by 'Admin'.

To strengthen the definition the owner must have a corporate email, a corporate internal telephone number and/or be listed in the staff directory.

With this information a user is able to assess the content in three steps

1. Does the title suggest that this is at least close to the information I am looking for?
2. Does the date indicate that this is reasonably current information, something that I can also verify with the owner if needed?
3. Does the owner of the content have the authority (in expertise terms) to publish this content, which the user can always verify through the people directory?

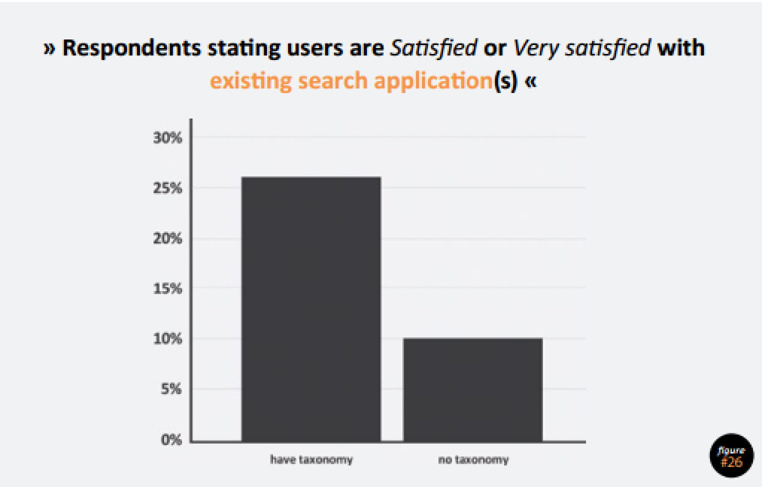
5. Content quality guidelines

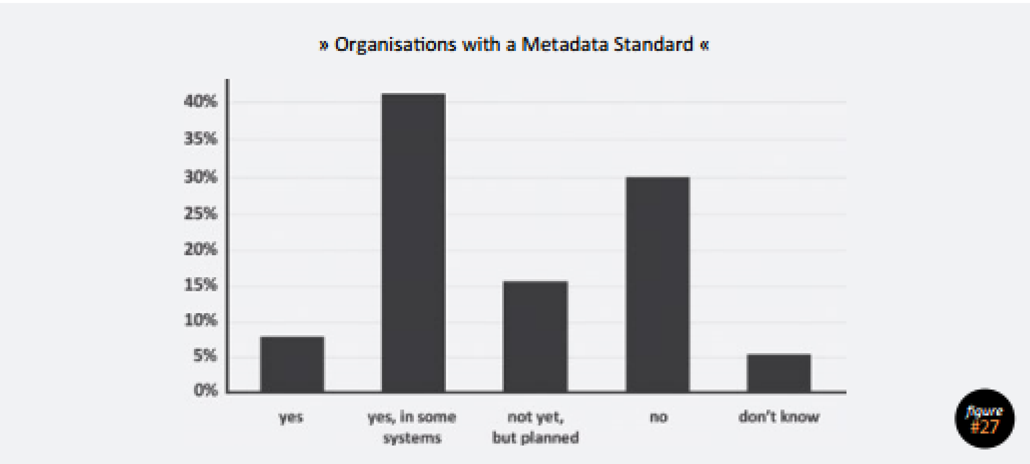
In the table below are brief descriptions of parameters that could define content.

Topic	Summary of scope
Audience	Style and language-level should be appropriate to the intended readership
Authors	Defining the authorship of content
Copyright	The copyright owner of 3rd party content should be identified
Date formats	Dates should be unambiguous
File types	Guidance on the applicability of web page, Word, pdf etc. formats
Images	Guidance on the use of images
Language	Guidance on 'corporate' English
Links	Ownership of content also entails ownership of links, and links should also be managed with care
Metadata	Metadata schema descriptions. This is a massive topic in its own right, especially when taxonomy management is considered.
Mobile content	Content likely to be used extensively on mobile devices should be written in an appropriate format
Names	How employee names should be presented
Owners	The owner of the content may not be the author
Page-level identity	Any individual pages of content should be able to be uniquely identified from metadata on each page
Protective marking	Security management for content
Record	When and how should content be a declared record
Related content	Ensuring that related content is identified so that content can be placed in context
Re-use	Where content is re-used the origin of the content should be declared
Review	The review period of content
Scanning and structure	Presenting content, especially on a web page so that it can be scanned in an F-pattern
Style sheet	Extent of linkage to corporate brand guidelines
Titles	Ensuring that titles are consistent and informative
Translated content	Linking to original versions of translated content
Validity	Any validity issues – i.e. for use in US only
Version control	Consistent version control designations

The challenge with deciding how comprehensive to make a set of guidelines is that there has to be balance between meeting these guidelines and not putting too great a workload on content publishers who are almost certainly carrying out the publishing work in their spare time. Another aspect to consider is how different categories of content (e.g. videos and PowerPoint presentations) may need to variants on these guidelines. It is difficult to be definitive about which guidelines are important, and which may need to be expressed as both guidelines and standards.

The benefits on fundability are shown in the results from the Findwise Enterprise Search and Findability Survey 2014. The charts below show a significant improvement in search performance where metadata is managed well and again where there is a taxonomy.





6. Managing content curation

The process of content curation (which includes publishing, revising, reviewing and deletion) is far more complex that most organisations comprehend. There is a superb paper by Stephen Dale entitled Content Curation: The Future of Relevance in *Business Information Review*, 2014, Vol. 31(4) 199-205.

In the abstract the author comments:

"We're all creating content, as originators or commentators, which is then shared and re-shared many times over. The resulting cascade of information requires new content organization and consumption techniques, and the disciplines, competencies and skills of content curation are now critical. This article argues that effective content curation requires real-time technology and tools used by knowledge domain experts who can interpret and add insight to content."

This paper is essential reading for any web or intranet content manager.

7. Training content publishers and owners

Content publishers need to be trained in how to contribute high-quality content, and I have found it useful to set up at least a two-level certification covering 'routine' content contribution (uploading Office, HTML and PDF files) and 'advanced' contribution for (e.g.) images and videos. However it is not enough to train the publishers. Content owners also need to realise that they have a crucial role to play in understanding the complexities of content contribution and being able to work in partnership with publishers.

8. Staff directories

The quality of the information in staff directories can be very variable. In the case of one of the law firms only the lawyers in the firm have profiles. To find information on 2,300 business support staff requires a search through LinkedIn. To me it seems that rarely is any thought given to the information that should be presented in a staff directory. In the case of the other law firm the same information is given in the profiles on the website as in the internal directory even though this information is used by a different group of people for a different purpose.

Even agreeing on how a name should be presented is not an easy problem. Take a look at this briefing paper from Basis Technologies (Site registration required).

For further reading see:

- Creating Intranet Content
- Organising Digital Information for Others
- SharePoint Content Authoring Quick Guide
- Quality-Biased Ranking of Web Documents
- Dublin Core Metadata Initiative
- Website Product Management

If you would like more details of the content quality guidelines mentioned above please email me at *martin.white@intranetfocus.com*.

Post-it Notes for Web Content

Michael Upshall, Consult MU

michael@consultmu.co.uk

Why do we use Post-it notes? To remind ourselves of something. We like them because they are so flexible – you can scribble notes and reminders on them and they're easy to find, reorder and change. Today Post-its are so ubiquitous that you have to think hard to remember what life was like before they were invented.

In this article, I will be introducing and describing three similar tools for linking or sharing content: Declara, Stackly and Pocket (there are others, to be sure, but these three looked the most interesting from my trawl). All three could all be described as Post-it notes for the Web: you can make notes about websites you have seen, then reorder your notes and come back and find them again. All three tools were created for academic and educational publishing. The nature of academic publishing is to share what you have discovered or created, and so tools that enable users to join A with B can provide valuable productivity benefits. Of course, discovering content of value and wanting to share it with others is universal across all sectors. All three use recommendation engines – tools that automatically suggest content that is linked in some way to pages you have selected before. Two were launched as recently as this year.

Declara

Declara is the newest of the three services, launched on 7th April. It combines an online playlist and a recommendation engine, and states proudly that its role is to "make it easy to discover, share and organise knowledge." Users can tag pages and share recommendations. After creating a collection of web links – to documents, content, websites or people – users can search for related content using the Declara search engine. The Declara app places a small icon on your browser bar, and clicking on the icon adds the current page to your collection.

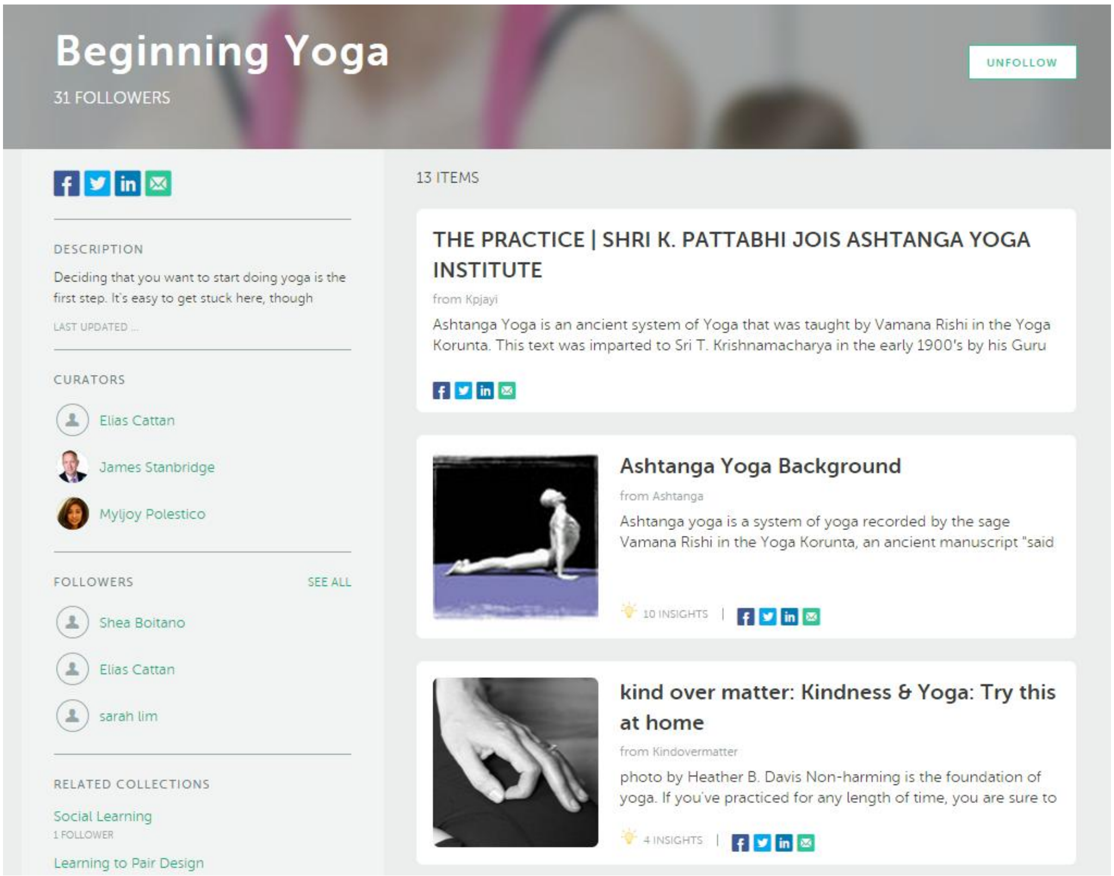


Figure 1 Example of a Declara Collection

Having browsed through the collections, you can then opt to follow any of them (see Figure 1). When these collections are updated, Declara sends you an update notification.

Collections can be open (public) or private. Open collections are, as you would imagine, shared with everyone, while private collections can be opened to as many users as you would like. Individual content items can be annotated (creating what Declara calls "insights") and those can be shared with other people.

It might be surprising, based on the above, that Declara calls itself not a link-management tool but a "predictive learning company", which analyses user behaviour to predict recommendations. It makes use of a "Cognitive Graph" that identifies what you might like to learn, based on what you already know (which includes social media). It appears to be a text-mining recommendation engine with some social media features added.

Stackly

Stackly is another recent service (launched in March 2015), developed by HighWire, the US company that provides hosting tools for many journal publishers. As you would imagine from a company that provides tools for academic users, it is presented as an academic tool, but there is no reason why it should be restricted to academic use.

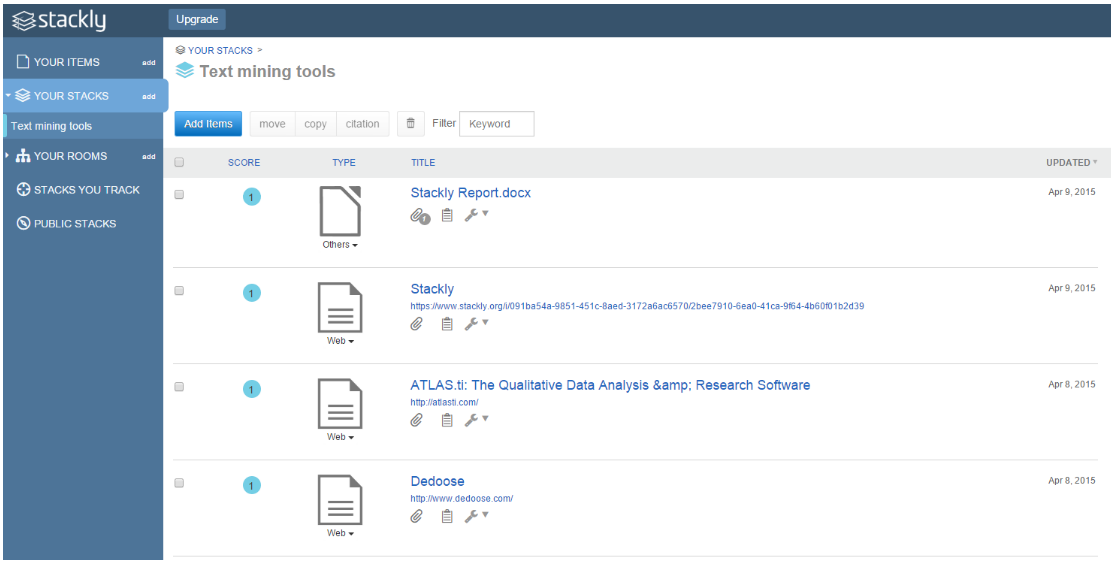


Figure 2 *A Public stack*

Like Declara, Stackly has public and private collections, although Stackly calls collections "stacks". Public stacks are free, while private stacks are payable. Stacks can be grouped into "rooms", which are collections of stacks (I fail to understand the analogy here, although I suppose many rooms are full of stacks of things). The service is for personal use, but Stackly provides a paid model for teams to collaborate with each other. Clearly, HighWire believe it has some relevance to journal publishers, since it provides useful tools for them to disseminate information about their journal, and certainly some HighWire journals have a Stackly icon on their journal home page. However, Stackly seems to have more in common with the two other services here than with journal promotions – such a link seems almost peripheral.

Pocket

Pocket is the granddaddy of these three, a service for storing and sharing content (it was launched as "Read It Later", which gives some idea of what it does). It has a similar look and feel to the other two apps, although Pocket presents itself as a more general-purpose tool than primarily for academic use. Its strapline is "When you find something you want to view later, put it into Pocket." This seems to sum up Pocket quite well. However, Pocket has the added feature of not requiring an online connection to view the saved content – it remains in your Pocket collection. As might be expected from a service founded as long ago as 2007, Pocket has more users than Stackly and Declara put together – they claim 17 million registered users. An integral feature of Pocket appears to be that it has apps for mobile devices, and certainly making a note of something on your mobile or tablet seems like a good motivation for using Pocket.

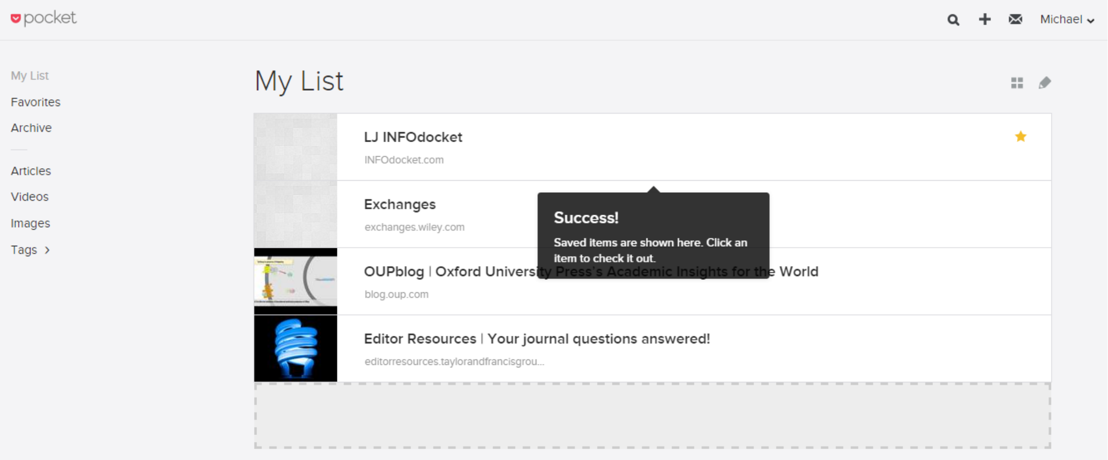


Figure 3 Pocket Personal list

Like the other two services, Pocket has a free basic service and a paid premium offering. It provides two benefits for premium members:

- Content can be saved to a permanent library – hence it will still be available even if the web page is altered. This could be a brave attempt to make a feature out of a limitation, since many people would want to keep the web page in its current, latest version – something that would be more difficult to achieve with an offline store
- Pocket suggest automatic metadata to catalogue your content, although as a user you can edit and change these tags

Conclusions

What is the appeal of all these tools? They are all simple to use; they have all gained the endorsement of attracting substantial external funding, which suggests that investors think they are a good idea. The idea seems to be to circulate links, in the way that Twitter sends "signposts" of websites and events, for example.

They all resemble Twitter in that they can simply be signposts to other content, but at the same time, all of them are an attempt to get beyond Twitter, by providing more graphical indications of the content they are describing.

However, they are not formal citation tools like Endnote or Mendeley. Although each service provides tools for linking content within each service, that is, from one user-based collection to another, most users will probably not want to restrict their searches to what other users of Declara, Stackly or Pocket have found - they will want to keep searching the Web. Hence although these tools are useful ways of noting things you like, none of them has sufficient appeal for a user to want to use them as a means of communicating with others for more than an occasional share of a specific list. I can't imagine any of these services becoming a genuine community, in other words, unlike Facebook or Twitter.

The arrival of three rival services within a few years suggests that Post-it notes for the Web meet a genuine need. Unfortunately, the chances are that either all three will fail to achieve substantial use, or that (more likely), one will become dominant, and the others fade away. None of the three has the convenience of the real Post-it note - if only that could translate to the screen.

Book Review: Knowledge Management in the Pharmaceutical Industry: Enhancing Research, Development & Manufacturing Performance

Goodman, Elisabeth and Riddell, John
Gower Publishing 2014.
978-1-4094-5335-2
ISBN 978-1-4094-5336-9 (ebook)

Reviewed by John Wickenden, FCLIP
Retired, ex. Eli Lilly & Co. Ltd

The Pharmaceutical Industry is intensely information and knowledge rich across all of the significant disciplines it embraces so is an ideal sector to demonstrate the value of Knowledge Management (KM). The authors have extensive experience with the UK's biggest pharmaceutical company GlaxoSmithKline (GSK) and its predecessor SmithKline Beecham. As part of their research they interviewed twenty-seven leading pharmaceutical KM practitioners and used their experiences to detail various practical applications of KM. This gives the book real insight into the value of KM across the industry.

The pharmaceutical value chain has gone through radical change in recent years and the book reviews in detail how KM can be applied to catalyse further change and improvements in order to speed up the process of bringing effective new products to market. Each of the pharmaceutical development cycles are reviewed, including Research and Development, Manufacturing, Functional and Commercial areas. These later areas are also relevant to other industries.

The book starts by detailing the different ways of approaching and defining Knowledge Management to bring added value to information in the organisation, so that it moves from being just information to being usable and effective knowledge ('wisdom'). It then goes into the importance of linking people to content, while the following chapters look at its application in different functional areas of the pharmaceutical industry. The final two chapters look at strategy (supporting and sustaining) and the future of KM.

I recommend this well written and easy to read book to everyone interested in Knowledge Management, as it gives numerous practical examples of how to undertake KM in organisations using various techniques, (e.g. Communities of Practice). It will encourage readers to look for opportunities of applying KM in their own organisations.

Sample Chapter (Ch.3 'Realising Pharmaceutical Value')

Additional Reviews

This book demonstrates the critical importance of knowledge management and data sharing to translate the new vision of drug development into concrete actions. This is a timely endeavor since more than ever therapeutic advances depend on integrative analysis of big data by scientists acquainted to the principles of collective intelligence.
Michel Goldman, Executive Director, Innovative Medicines Initiative (IMI)

This isn't just a book describing the theory of knowledge management, but rather an illustration of how it can be applied to the real, challenging world of the pharmaceutical industry. For those attempting to improve ways of working collaboratively in pharma, look here for some practical and pre-tested ideas, even if a formal KM strategy is already in place.
Alison Zartarian, AstraZeneca

With little published in this area, this book provides valuable, concrete evidence of the value of Knowledge Management (KM) to Pharma operations. Key KM principles are exemplified by a blend of case material and anecdote - easy to read and tempting to adopt. The content should stimulate readers to seek the KM opportunities in their own organisations - in Pharma and beyond.
Sandra Ward, Principal Consultant, Beaworthy Consulting

Goodman and Riddell have extensive experience in supporting pharmaceutical development through information management and bring their expertise to the text. …the authors not only impart their combined knowledge but also collect and document the experiences of many interviewees with relevant industry backgrounds. …it does give a clear overview of the theory of knowledge management and provides the reader with the key elements needed to create a culture that supports it. …detailed and descriptive, with diagrams to aid understanding where necessary. …Those wanting to implement or improve knowledge management within their own companies can learn from the successes and failures of others through this text. This book will be of most interest to those managing research and development or manufacturing processes. However, the principles discussed could be applied to other areas of the pharmaceutical business.
The Pharmaceutical Journal, vol. 293, no. 7835 2014

Book Review: Preserving Complex Digital Objects

Delve, Janet and Anderson, David (eds)
Preserving Complex Digital Objects
Facet: 2014
ISBN: 978-1-85604 -958-0

Reviewed by Margaret Katny, BBC Archives margaret.katny@bbc.co.uk

This collection comprises contributions from 33 authors and is based on the results of Jisc-funded Preservation of Complex Objects (POCOS) symposia, which took place in the UK in 2011 and 2012. The symposia brought together international experts on complex digital objects and aimed to review research and practice to date, point to solutions and suggest areas where more work was needed.

The last five years or so of the digital revolution have produced a proliferation of technologies. Many of them are ephemeral and involve social networking, cloud computing, online gaming or virtual art. Digital objects, simple and complex, surround us. Simple digital objects such as documents, images or e-books are fairly easy to make safe and accessible for the future generations. This is considerably more difficult with complex digital objects. For the purpose of this book, the types of complex objects have been narrowed down to simulations/visualisations, digital/software art and gaming/virtual environments. These by no means cover all types of complex objects but represent a selection of most distinct categories.

The three types of complex digital objects are also very different in nature. Simulations and visualisations tend to be more generic and produce outputs in many different fields such as archaeology, film, architecture or engineering. We are used to simulations and visualisations in these fields and indeed expect them in exhibitions, galleries or presentations. Software or digital art presents its own problems for curation and preservation. There is often no distinction between the master art piece and its clones so it may hard to decide which art piece should be retained and preserved, even before the method of preservation can be considered. In the case of games, the main issue is intellectual property rights as many games are orphans.

All complex digital objects share overarching characteristics – they are technologically highly advanced and analysing and preserving them involves multiple levels of difficulties. Interestingly but perhaps not surprisingly, some are also hybrid, part digital and part physical artefacts. Needless to say, this does not make their preservation easier.
The book raises questions about why and what to preserve. In the case of complex digital objects, the answer to these questions is in itself complex. The authors explore to what extent the process of selecting for preservation may interfere with the process of creation and delve into the meaning of preservation in the context of a virtual world. After all, how can we preserve interaction? The collection looks at how different memory institutions attempt to preserve complex digital objects and suggests which approaches may bring

authentic and long-term results. The chapter on case studies illustrates challenges faced by digital preservation in the areas of art, animation and archaeology.

The authors do not shy away from offering practical advice. The chapter on approaches, practice and tools focuses on technical know-how and includes guidelines for the preservation of complex objects. Techniques and methods are examined, as are metadata issues, which arise in the preservation process. As is often the case, digital preservation involves legal issues and this book offers an interesting summary on how the law views preservation activities.

This collection is a valuable contribution to the digital preservation literature. Although somewhat academic in parts, it is sufficiently practical to be of use to practitioners, particularly those working in memory institutions and faced with issues of digital preservation of complex digital objects.

Online Resource Update

Joy Cadwallader, Aberystwyth University (Aberystwyth Online User Group).

Please send your submissions for the next edition to:
jrc@aber.ac.uk

British Library / Heritage Lottery Fund

In a press release the British Library have announced their plans to digitise, "500,000 rare, unique and at-risk sound recordings from its own archive and other key collections around the country," following a successful Heritage Lottery Fund (HLF) bid for over nine and a half million pounds. The sound recordings will be rescued from a variety of fragile formats including wax cylinders and acetate discs, and the content will be made freely available online. British Library sound holdings include authors reading their own work (Tennyson, Plath and Joyce), Radio Luxembourg broadcasts, slang dialects and accents from the 1950's Survey of English Dialects, and sounds of wildlife. The HLF money will also support the development of a network of regional centres of archival excellence and a school and community outreach programme to raise awareness of sound archives.

Elsevier

More conflict between the academic community and Elsevier, this time over Elsevier's new hosting and sharing policy announced on April 30th. The Confederation of Open Access Repositories (COAR), supported by 23 co-signing organisations including Research Libraries UK (RLUK), have asked Elsevier to reconsider the policy because it, "represents a significant obstacle to the dissemination and use of research knowledge, and creates unnecessary barriers for Elsevier published authors in complying with funders' open access policies". In their statement they oppose over-long embargos for some journals (up to 2 years), over-restrictive licenses on articles uploaded to open access (OA) repositories, and the planned retrospective application of the policy. Elsevier's response on May 21st was followed up by COAR on May 28th with specific recommendations to free authors to, "make their 'author's accepted manuscript' openly available upon acceptance through an OA repository or other open access platform", to choose the type of open license they want and to not be dictated to regarding individual sharing of journal articles which they describe as, "a scholarly norm". Watch this space!

Jisc / Spare Rib / British Library

It may have been all over the traditional press but I had to include the full run of the feminist magazine Spare Rib 1972-1993 going online, thanks to the British Library collaborating with the original Spare Rib community and Jisc. In the days before the Internet and social media became the place to campaign, there were posters, pamphlets and magazines and here is Spare Rib freely available on the Jisc Journal Archives site. It is going to be a fascinating read; the Jisc press release mentions, "big-name contributors including Betty Friedan, Germaine Greer, Margaret Drabble and Alice Walker, but alongside these were the voices of ordinary women telling their own stories", and a sample of articles about offensive advertising, sex lives, FGM and body hair – plus ça change. Spare Rib worker Sue Sullivan is quoted, "The digitised magazines will be a wonderful resource for younger historians and feminist activists, researchers and all the women (and men) who wonder what their mothers, aunts, grannies and older friends got up to all those years ago."

National Library of Scotland / ProQuest

UK House of Lords papers from the nineteenth century are to be digitised and made available online in a collaboration between the National Library of Scotland (NLS) which holds one of the rare surviving sets of approx. 3,000 volumes and ProQuest whose online resource catalogue already includes the House the Commons parliamentary papers from the 18th century onwards. The NLS press release announces that NLS members will have free online access to the House of Lords content while ProQuest will offer online access on a commercial basis. Scotland's National Librarian Dr John Scally notes that, "More British Prime Ministers served in the Lords in the 19th century than in the House of Commons", and researchers will be able to explore the changes made to by the Lords to bills before they were passed and the bills that were rejected.

ProQuest

In more news from ProQuest, they have taken over Coutts Information Services (which includes MyiLibrary and OASIS) from the Ingram Content Group. InfoDocket notes that it was only in February this year that EBSCO took over YBP Library Services, two big players in the information market building on their academic book/e-book supply profiles.

UK Supreme Court

The UK Supreme Court has launched a trial on-demand online archive of past hearings. Their press release points out the educational value to lawyers and law students as well as the potential to reduce transcription costs. Footage of court proceedings will be uploaded on the next working day and, "Once judgment is delivered, footage of the Justices' summary in court will also be published alongside the full judgment text and press summary". The on-demand service is scheduled to be reviewed in March 2016. The Supreme Court has had live streaming from its 3 courtrooms since October 2014, during which time 15,000 people have used it each month.

UKeiG, Information Management & Intranets

Dion Lindsay

dion@dionlindsayconsulting.com

UKeiG has been busy developing its support of electronic information management in the UK, and we are very happy to be associated with IM and Intranets events coming up in the Summer and Autumn. There are double benefits: not only do we further the interests of professional information management by sponsoring these events, but UKeiG members get substantial discounts from many of those we are involved in.

UKeiG membership is free if you are a member of CILIP, and £39 if you aren't – so it's a clear financial benefit to join if you're going!

Conferences

Internet Librarian International 2015 is at the Olympia Conference Centre in London on 20th and 21st October and as a "supporting organisation" we can offer 25% discounts to UKeiG members. As regulars will know ILI is all about the exchange of experience and ideas and this year the uppermost concerns are about new technologies, services and business models, and how the Internet librarians can continue to provide constant innovation.

Enterprise Search Europe has combined with ILI this year – same location, same dates. Key topics will include how to make the business case for investment in search at the enterprise level and as media partners UKeiG can again offer 25% discount to members. Further information on both events is published in this issue of Elucidate.

We are particularly proud to be a main sponsor of **Intranet Now**; an annual one-day conference that those who went last year will know was a dramatically successful event, with lots of learning and energy, and a very popular un-conference feature in the afternoon. This year's edition is in **London on 13th October**, and is already booking well. Tickets are very good value (there are still some tickets available at £95 at the time of going to press) and prices, booking and agenda are all at http://intranetnow.co.uk/

Information Manager of the Year Award 2015
We have initiated this award as part of our commitment to raising the profile of the information profession in the UK, particularly in the e-information arena. It will be presented at this year's Intranet Now conference to the person who has made the greatest contribution to information management or is the year's best exemplar of an information manager. It comes as an impressive glass plaque and, naturally includes a free ticket to this year's Intranet Now conference. Closing date for nominations is 24th July, and the process and forms are available on the web site.

"Simplicity must be the Mantra of the Digital Workplace"

The organisers of Intranet Now held a "Simplify your Intranet" workshop on May 12th, led by Gerry McGovern which was well attended and well tweeted at *#IntranetNow*.

Here's a blog item by Gerry which he posted shortly before the workshop, and which he has given us permission to reproduce:

Employees are reaching a breaking point. They are exhausted and overwhelmed, alienated and disengaged. They are cynical and distrustful.

"The elephant in offices all around the world is that people are running on empty," Tony Schwartz wrote for The New York Times in March 2014. "If you are expected to work 60 or 70 hours a week, or to stay connected in the evenings and on the weekends, or you can't take at least four weeks of vacation a year, or you don't have reasonable flexibility about when and where you work, then your company can't be a great place to work."

The intranet / digital workplace could contribute in some small way to reducing employee workload. It could save time by making things easier to find and do. And yet most intranets are still an appalling mess, a time-wasting black hole. They are so far away from their potential.

I've just had a chat with Martin Edenström, who writes for a popular Swedish blog on intranets. He spoke about how he felt that a lot of employees had simply given up on the intranet.

Why are so many employee systems so awful? Why is enterprise search a joke in most organizations? Why is it so hard to sign up for a training course, book a meeting room, or get the most up-to-date product details?

Because management doesn't care about employee time. If they can't get their work done in 40 hours, then let them stay 50, 60, 70 hours is the thinking. We must change that thinking because it is ruinous to everyone's long term success.

"The majority of corporate infrastructures, tools, budgeting, resourcing and reporting relationships are still corporate-centered -- designed to make it simpler for the company to succeed, but usually more complex for each individual," Bill Jensen wrote for Huffington Post in March 2015.

"The future of simplicity is far more workforce-centered," he continues. "In a knowledge and service economy, this approach to simplicity is the ultimate competitive advantage. We have the tools. We have the knowhow. What's missing is the leadership will and commitment that simplicity for each individual is crucial to organizational success."

The public Web, according to Jensen, is shining a big, glaring torchlight on internal systems. "The standard against which the workforce will evaluate all corporate IT and work process experiences is each individual's phone," he states. "And the standard against which the workforce will evaluate company culture, managers and leadership is each individual's tribe, community."

If you work with an intranet you must make simplicity your number one priority. Without simplicity the digital workplace is a digital quagmire, a digital dump, a digital propaganda rag.

Over the years I have asked many employees what their definition of a perfect intranet was. By far the best definition was this: "A perfect intranet is a survivor's guide to a shitty week."

Employees are not looking for miracles. They just want to find stuff quickly. They want to be able to trust that what they find is accurate and up-to-date. They want to be able to do stuff as quickly and easily as possible.

Discount for UKeiG members at Internet Librarian International 2015 & Enterprise Search Europe 2015

Jason Farradane & Tony Kent Strix Awards to be presented

UKeiG is once again working with these two important conferences which this year will take place side by side at London's Olympia Conference Centre on 20th-21st October 2015. As part of this co-operation, the Tony Kent Strix Award will be presented at Enterprise Search Europe, and the Jason Farradane Award will be presented at Internet Librarian International.

As a supporting organisation, UKeiG members are entitled to a 25% discount on the full registration fee for either conference.

Internet Librarian International 2015 (ILI) is all about the exchange of ideas, knowledge and experience, exploring the big questions that challenge libraries and information professionals. The 2015 theme is 'Dynamic disruption: transforming the library' and the conference features seven tracks dedicated to service and technology trends, including new strategies for libraries and new roles for information professionals; new technology and new user experiences; understanding usage and users; and search, discovery and visibility.

Running alongside ILI is **Enterprise Search Europe 2015**, chaired by Martin White. The conference promotes effective enterprise search, which is critical in supporting decision-making, managing risk and dealing with increasing volumes of information. Coverage includes how to make a business case for search investment; practical advice on search applications, selection and rollout; the opportunities of search, big data and content analytics; search on mobile devices; federated search solutions and lessons learned from search projects including critical success factors for search implementation. The emphasis is on case studies with speakers on hand to share their search problems and solutions. More details and the programme have been published on the conference web site.

enterprise
search
europe
2015

Further details - and links to book your place and claim your 25% discount - can be found on the UKeiG web page or via www.internet-librarian.com/2015/ or www.enterprisesearcheurope.com/2015/.

About eLucidate

eLucidate is the journal of the UK Electronic Information Group. It is usually published four times each year, in March, June, September and December. Its aim is to keep members up to date with developments in the digital information environment, as they affect professionals. The journal is provided free to UKeiG members.

Notes for contributors

eLucidate welcomes articles or ideas for articles in the areas covered by the journal. UKeiG is always on the lookout for feature writers, reviewers both for books and for meetings, as well as respondents to articles. Sadly, we don't pay contributors, but contributors retain copyright of their articles and can republish their articles elsewhere.

If you are writing for eLucidate, please follow these simple guidelines:

About the members

Our membership comprises information professionals involved in the dissemination and/or delivery of digital content and services. Our membership base is two-thirds academic, one-third commercial, as well as some public libraries. A key benefit of the group is that meetings and forums provide "crossover" insight from one area to another: members see it as a way of keeping up to date in areas outside their core business. Few other organisations provide this kind of cross-sectoral awareness. The focus of the group is the UK electronic information sector, but issues of digital provision are of course global. The most popular training courses we run are on search tools – Google and others, ebooks and how to deal with them. Other popular strands include Intranets, content management and ebooks.

Technical level

Although members rate themselves highly for technical awareness, they are typically users rather than creators of technology. Articles should not assume understanding of technical terms without explanation.

Length of article

Feature articles should be in the region of 1500-2500 words. Each article should be prefaced by a short summary (around 50 words) that can be used when displaying on public search engines an outline of the article, and to display on the non-member section of the website.

What to write

A key aspect of UKeiG is that it provides insight from one area to another – members see it as a way of keeping up to date in areas outside their core expertise. Because the membership is disparate, ranging from pharmaceutical information professionals to public librarians, you should not assume readers are as familiar as you in the subject area.

The most valuable viewpoint you can give is that of an end user. UKeiG is not a place for theoretical debate, but a forum where peers can share their practical experiences and understanding. So, if it worked for you, tell others. If it didn't, tell others why not.

How to submit

Please e-mail your articles to gary.horrocks@gmail.com Articles should be delivered in Word format. Images are welcome – they may be in gif or jpeg formats.

Rights

By submitting an article to eLucidate, authors grant UKeiG the non-exclusive right to publish the material in any format in perpetuity. However, authors retain full rights to their content and remain the copyright owner.

About you
Please provide a 10-20 word biographical summary about yourself to appear at the end of the article.

Editorial process
Your article will be copy-edited for spelling and for sense. If there are major changes to the article we may return it to you for your comments and approval, but most articles require only light corrections before appearing in eLucidate, and do not need a further review by the author.

Brief for book reviews
Book reviews are typically 600-1000 words. Because UKeiG is independent of any publisher, we are not obliged to have favourable reviews. If you think a book is poor, then by all means explain why. Members and non-members alike are welcome to suggest books for review or to submit reviews.