

**ABD
BVD**

ASSOCIATION BELGE
DE DOCUMENTATION

BELGISCHE VERENIGING
VOOR DOCUMENTATIE

Bladen voor **DOCUMENTATIE**
Cahiers de la **DOCUMENTATION**

Trimestriel | Driemaandelijks | Septembre | September

Numéro spécial
Speciaal nummer

Inforum 2015

The I&D Puzzle :
new pieces to solve it





Bladen voor DOCUMENTATIE

Cahiers de la DOCUMENTATION

Trimestriel | Driemaandelijks | Septembre | September

Rédactrice en chef
Hoofdredactrice
 Catherine Gérard

Ont participé à ce numéro
Werkten mee aan dit nummer

Christopher Boon
 Benoit Collet
 Sara Decoster
 Guy Delsaut
 Catherine Gérard
 Paul Heyvaert
 Arnaud Seeuws
 Marc Van Den Bergh

Mise en page
Opmaak
 Stéphanie Fort

Conception de la couverture
Coverontwerp
 Image Plus

Image de couverture
Afbeelding cover
 Photo Pixabay, libre de droit

Impression
Druk
 Ciaco

Pour tout renseignement sur les *Cahiers de la documentation*
 ou pour soumettre un article :
 Voor alle inlichtingen over de *Bladen voor documentatie*
 of om een artikel voor te stellen:

cahiers-bladen@abd-bvd.net

Sommaire Inhoudstafel

70ème année - 2016 - n° 3

70ste jaargang - 2016 - nr 3

▪ Éditorial – Woord vooraf Catherine Gérard	3
▪ Digital preservation strategy and preservation planning at the British Library Michael Day	5
▪ Information is not for storage, information is for action Het wijzigende karakter van informatiebeheer in een web 2.0-kenniswerkomgeving Filip Callewaert	14
▪ Informatie opsporen, uitgaande van een beeld Paul Nieuwenhuysen	25
▪ Le chemin d'évolution du métier de la veille Résultats d'une enquête menée lors de la journée Inforum, à Bruxelles, le 4 juin 2015 Béatrice Foenix-Riou	35
▪ Notes de lecture - Boekbesprekingen	42
▪ Nouvelles parutions - Nieuwe publicaties	44

Les sommaires depuis 1947 et les articles des numéros 1999/1 à 2015/2
sont disponibles à l'adresse :

<http://www.abd-bvd.be/cahiers.php>

De inhoudtafels sinds 1947 en de artikels van de nummers 1999/1 tot 2015/2
zijn beschikbaar op:

<http://www.abd-bvd.be/bladen.php>

Publié par
Association Belge de Documentation, asbl
c/o Bibliothèque royale de Belgique
Boulevard de l'Empereur 4
1000 Bruxelles
Belgique

Les articles n'engagent que leurs auteurs
De inhoud van de artikels valt onder de verantwoordelijkheid van de auteurs

Uitgegeven door
Belgische Vereniging voor Documentatie, vzw
p/a Koninklijke Bibliotheek van België
Keizerslaan 4
1000 Brussel
België

ÉDITORIAL

WORD VOORAF

par / door

Catherine GÉRARD

Rédactrice en chef / Hoofdredacteur

traduit par / vertaald door

Marc VAN DEN BERGH

Quinze mois déjà se sont écoulés depuis notre Inforum 2015, intitulé *The I&D Puzzle: new pieces to solve it*. Mieux vaut tard que jamais ! Après deux numéros spéciaux très denses (2015/3 *L'audiovisuel sous les projecteurs* et 2016/2 *Année européenne du patrimoine industriel et technique*), vous trouverez, dans ce nouveau numéro spécial des *Cahiers de la Documentation*, les articles rédigés par quatre de nos cinq orateurs. Garantie du multilinguisme de l'ABD-BVD, l'équipe des *Cahiers* vous propose ainsi quatre articles, deux en néerlandais, un en français, et un en anglais, afin de conserver une trace écrite de tout ce que vous aurez pu découvrir et apprendre lors de cette journée du 4 juin 2015.

Le numéro spécial Inforum 2016 *In information we trust!!!* paraîtra quant à lui dans la première moitié de l'année 2017. Nous vous rappelons que vous pouvez à tout moment consulter et télécharger gratuitement les présentations de nos différents orateurs depuis 1996, via l'onglet «Inforum» sur notre site web¹. Vous étiez plus de 150 participants pour cette édition 2015 de l'Inforum, nous tenons à vous en remercier une fois encore. En espérant vous voir au moins aussi nombreux pour l'édition 2017 !

N'oublions pas non plus la tenue d'un *Doc'Moment* spécial, le 29 septembre après-midi, au Mundaneum, le «Google de papier», à Mons. Une visite guidée des lieux sera complétée par un exposé sur la visualisation des données. Le 17 novembre, nous organisons un autre atelier avec les sponsors de l'Inforum 2016, *Ebsco* et *LM*. Bonne rentrée à tous, et au plaisir de vous revoir lors de l'un de nos événements !

De tijd vliegt: er verstreken ondertussen vijftien maanden sinds Inforum 2015 *The I&D Puzzle: new pieces to solve it* doorging. Maar beter laat dan nooit. Na twee volumineuze speciaal nummers (2015/3 *Audiovisueel in de spotlight* gezet en 2016/2 in het kader van het *Europees Jaar voor het Industrieel en Technisch Erfgoed*) vindt u in dit speciaal nummer van *Bladen voor Documentatie* de referaten van vier van de vijf sprekers. De artikels staan tevens garant voor de meertaligheid van ABD-BVD: twee in het Nederlands, een in het Frans en een in het Engels. Op deze manier is het mogelijk om ook na de studiedag er teksten op na te slaan over wat u tijdens deze 4 juni 2015 heeft kunnen ontdekken en bijleren.

Het speciaal nummer over Inforum 2016 *In information we trust!!!* wordt op zijn beurt gepland in het voorjaar van 2017. We herinneren er u aan dat u de presentaties van de Infora die sinds 1996 georganiseerd werden, steeds gratis kan bekijken en downloaden vanaf de ABD-BVD website². We danken u om met meer dan 150 actief deel te hebben genomen aan Inforum 2015. We hopen dat u voor Inforum 2017 geen verstek laat gaan en in minstens dezelfde getale aanwezig zult zijn.

Laten we ook het volgende Doc'Moment niet vergeten. Op 29 september hebben we een geleid bezoek van het Mundaneum (Bergen) dat omschreven wordt als de voorloper van Google en van het elektronisch opzoeken. We krijgen er een toelichting over datavisualisatie. Op 17 november organiseren we opnieuw een Atelier met de sponsors van Inforum 2016, *Ebsco* en *LM*. We wensen iedereen een frisse start toe na de vakantie en hopen u snel weer te zien op een van onze evenementen.

Notes

1. Association Belge de Documentation. *Inforum*. <<http://www.abd-bvd.be/activites-evenements/inforum/>> (consulté le 2 août 2016).
2. Belgische Vereniging voor Documentatie. *Inforum*. <<http://www.abd-bvd.be/nl/activiteiten-en-evenementen/inforum/>> (geraadpleegd op 2 augustus 2016).

DIGITAL PRESERVATION STRATEGY AND PRESERVATION PLANNING AT THE BRITISH LIBRARY

Michael DAY

Digital Preservation Manager, The British Library

■ The British Library is the custodian of diverse and rapidly-growing collections of digital content, both born-digital and digitised. Its recently published vision document, *Living Knowledge*, has confirmed that custodianship is the Library's "first and core purpose", the one on which all of its other purposes depend. The strategy also specifically highlighted the need for the Library to address the challenges of the long-term preservation of digital content, in particular for the born-digital content received through legal deposit, which includes the periodic (at least annual) capture of the entire UK Web domain. This paper explores some of the activities being undertaken by the Library's digital preservation team as it seeks to implement a preservation planning capacity, focusing on the initial capture of preservation requirements through collection profiling and a programme of file format assessments.

■ La British Library est dépositaire de collections de contenus numériques, diverses et en rapide expansion, tant numérisées que sous forme numérique d'origine. Elle a récemment publié un document à vision stratégique, intitulé "Living Knowledge" (connaissance vivante), confirmant que cette mission de conservation est la mission première et fondamentale de la Bibliothèque, celle sur laquelle reposent toutes les autres missions. Cette stratégie a également spécifiquement mis en évidence la nécessité pour la Bibliothèque de prendre en considération les défis posés par la conservation à long terme des contenus numériques, en particulier les contenus numériques d'origine reçus au travers du dépôt légal, qui inclut la capture périodique (au minimum annuelle) de l'ensemble du domaine web du Royaume-Uni. Le présent article explore quelques-unes des activités entreprises par l'équipe de la Bibliothèque chargée de la préservation digitale, dans sa recherche pour implémenter une capacité de planification en vue de cette préservation, en mettant l'accent sur la saisie initiale des exigences de préservation au travers du profilage des collections et sur un programme d'évaluation des formats de fichiers.

■ British Library is de hoeder van diverse en snel groeiende verzamelingen van digitale content, van zowel born-digital als gedigitaliseerde documenten. Zijn onlangs verschenen visiedocument "Living Knowledge" bevestigt dat de bewaarfunctie van de bibliotheek primordiaal is en de core business blijft waarvan de andere doelstellingen afgeleid worden. De strategie beklemtoont de noodzaak van British Library om een bewaring op lange termijn van digitale stukken uit te werken, vooral dan wat betreft de born-digital content afkomstig uit het wettelijk depot, dit inclusief het periodiek (ten minste jaarlijks) vastleggen van het gehele Britse webdomein. Het artikel onderzoekt een deel van de activiteiten uitgevoerd door het digital preservation team in zijn streven om voor deze bewaring een capaciteitsplanning te realiseren via het focussen op de initiële bewaringseisen dit via een collectieprofilering en een programma van bestandsformaatevaluaties.

The British Library is the custodian of ever-increasing amounts of digital content, both born-digital and digitised. As a national library, it is the Library's responsibility to preserve this content and make it available for current and future users. The preservation of digital content, is not straightforward, however, in that it typically requires actions and interventions throughout content lifecycles, often far earlier and far more frequently than would be typical with physical collections¹. Above all, choices need to be made about which preservation strategies to adopt, which are currently mostly variations upon emulation or content migration². Regardless of preservation strategy, however, decisions need to be made about target formats or operating environments, as well as on the tools and software required to

manage digital collections over the long-term, e.g. for format identification and validation, migration, integrity checking, preservation assessment, etc. The process of making and recording these decisions is broadly known as preservation planning³.

Preservation planning depends upon the production of preservation plans, understood by Becker, et al. as action plans for "preserving a specific set of objects for a given purpose".⁴ A typical preservation plan would need to take into account, for example, organisational contexts and technical environments, as well as the preservation objectives and requirements of particular collections. It would also need to justify any decisions taken and specify in more detail the series of activities needed to preserve the collection (the latter is often known

as a preservation action plan)⁵. The importance of preservation planning for digital preservation is indicated by its inclusion in emerging standards for repository audit and certification, e.g. ISO 16363:2012⁶.

The British Library is currently in the process of developing its preservation planning capacity. This is being done both pragmatically and incrementally. After providing an overview of the Library and its digital collections, this paper will introduce some of the initial steps undertaken by the Library's digital preservation team in implementing a preservation planning capacity, focusing on the initial capture of preservation requirements through collection profiling and a programme of file format assessments.

The British Library

The British Library is the UK's national library. It was established on 1 July 1973 and defined in legislation as "a national centre for reference, study and bibliographical and other information services, in relation both to scientific and technological matters and to the humanities."⁷ The Library inherited much of its national library role – including its status as a legal deposit library – from the British Museum, whose printed book and manuscript collections provided the core of the Library's initial holdings. These have been supplemented since by the incorporation of other collections, including the India Office Library and Records and the National Sound Archive. The British Library is now one of the largest research libraries in the world, with significant collections of printed works, manuscripts, maps, sound recordings, postage stamps, and much else. The Library is the custodian of a great many unique items, including two of the four surviving copies of the 1215 Magna Carta manuscript, which was the focus of a major exhibition held at the Library in 2015⁸. The British Library is based in two main locations. Firstly, there is a purpose-built library building at St Pancras in central London – now part of a fast-growing "Knowledge Quarter" in the capital. This building was designed by Colin St John Wilson and built between 1982 and 1997; it was awarded grade I listed building status by Historic England in August 2015. Secondly, there is a site at Boston Spa in Yorkshire, which provides the hub for the Library's document supply services, as well as having significant physical storage capacity, including a dedicated Newspaper Storage Building that was opened in 2013⁹. Living Knowledge

In January 2015, the British Library published a new vision document to cover the development

of the Library up to its 50th anniversary in 2023. Living Knowledge: The British Library, 2015-2023 outlined six main purposes of the British Library¹⁰:

- Custodianship – We build, curate and preserve the UK's national collection of published, written and digital content;
- Research – We support and stimulate research of all kinds;
- Business – We help businesses to innovate and grow;
- Culture – We engage everyone with memorable cultural experiences;
- Learning – We inspire young people and learners of all ages;
- International – We work with partners around the world to advance knowledge and mutual understanding.

Living Knowledge reaffirmed that *custodianship* was the British Library's "*first and core purpose, the one on which all the others depend.*" In undertaking custodianship, it noted the Library's dependence on a wide range of specialist skills, covering collection management (e.g. content processing, metadata, conservation and preservation) as well as scholarly and curatorial expertise. The document also highlighted a few specific custodianship priorities that would need to be dealt with over the next few years. One of these was to address the preservation of audio and recorded music collections, where there was a concern about the longevity of the physical formats on which sound recordings were kept. This is already being addressed by the Save our Sounds project, initiated by a grant of £9.5 million from the UK Heritage Lottery Fund, announced in May 2015. Another custodianship priority included in Living Knowledge outlined the need to address the long-term preservation of born-digital content received through legal deposit. It described the recent extension of legal deposit to non-print content, including the periodic capture of the entire UK Web domain, as "the heart of our single greatest endeavour in digital custodianship".

Begun in April 2013, we are at the earliest stages of a journey that is set, over time, to transform what it means to be a national memory institution. Our challenges for the next phase of the project are to develop the scale and accessibility of the collection, and to ensure a proper level of investment in its future storage and preservation. (p. 13)

Living Knowledge acknowledges the extent to which the British Library is already a digital library.

Digital collections

The British Library's collections contain both born-digital and digitised content.

The growth in born-digital content to some extent simply reflects changes in the publishing industry, but, as Living Knowledge has noted, this has been amplified since April 2013 by the extension of UK legal deposit provisions to non-print content, including the Web. As a legal deposit library, the British Library inherited from the British Museum the right to receive copies of certain types of printed content published in the UK, including books, newspapers, printed music and maps (a right that ultimately dates back to the seventeenth century). Primary legislation supporting the extension of legal deposit to non-print content in the UK was passed in 2003. After a decade of planning and negotiation, the Legal Deposit Libraries (Non-Print Works) Regulations 2013 came into force on the 6th April 2013¹³. This, for the first time, enabled the British Library and the other UK copyright libraries to claim certain classes of non-print content under legal deposit provisions and to make it available to on-site users. In the British Library, this has led to the development of specialised ingest workflows to receive and process electronic journals and books, as well as a massive scaling-up of the Library's Web archiving activities to include the periodic (at least annual) capture of the entire UK Web Domain. The first of these domain crawls (running from April to June 2013) captured approximately 31 TB of compressed data¹⁵. Other born-digital content received by the Library includes geospatial datasets received from the Ordnance Survey and the Ordnance Survey of Northern Ireland. These include, from 2007, annual snapshots of Ordnance Survey MasterMap™ data, which is based on the Geography Markup Language (GML).

At the same time, the British Library has also spent a great deal of time exploring how to make elements of its existing collections available to wider audiences through digitisation and other digital scholarship initiatives. The Library's involvement in digitisation programmes started in the 1990s with some relatively small-scale projects undertaken as part of its Initiatives for Access programme. For example, the Electronic Beowulf project produced a facsimile edition of the Early English poem based on high-resolution images of the Library's version of the manuscript, linked to textual transcriptions made over several centuries and other critical apparatus¹⁸. This was initially published on CD-ROM in 2000, although the most recent version, Electronic Beowulf 4.0, is now a freely-available online resource. Since the 1990s,

the Library has scaled-up its digitisation activities considerably. The Library continues to digitise selected manuscripts and it makes many of them freely available through its Digitised Manuscripts website or (selectively) through online resources like Discovering Literature. The Library also undertakes much digitisation activity as part of funded grant programmes or in partnership with the commercial sector. For example, the Library has used grant funding to digitise many sound recordings and, wherever possible, it makes the results available through the British Library Sounds webpage; an activity that will be scaled up as part of the Save our Sounds project. Also, in the past five years, as part of a broader European collaboration linked to the centenary of the outbreak of the First World War, a range of different content types (monographs, photographs, archives, musical scores, etc.) were digitised and made available through the Europeana 1914-1918 portal. The Library's partnerships with the commercial sector (mainly secondary publishers) has also allowed it to experiment with large-scale digitisation, current examples being a partnership with Findmypast on the British Newspaper Archive and the Library's participation in the Google Books program.

The end result is that the British Library is the custodian of diverse and rapidly growing collections of digital content, both born-digital and digitised.

Digital preservation at the British Library

Given its responsibility to preserve this content and to make it available to current and future users, the British Library has, over the past few decades, made significant progress in responding to the organisational and infrastructural challenges of digital preservation.

Digital repository

A key part of the British Library's infrastructure is its digital repository, known as the Digital Library System (DLS). Prior to the extension of legal deposit to non-print content in 2013, the Library and the other UK Legal Deposit Libraries invested heavily in developing scalable solutions to the acquisition, storage and management of very large amounts of digital content. The repository has been described as a "single location to ingest, store, preserve, manage, discover and provide controlled access to digital content assets."¹⁹ While designed as an integrated storage system, it has been implemented in a highly distributed way, with content replicated in four geographically separated storage nodes

(based in London, Boston Spa, Edinburgh and Aberystwyth), with additional access gateways at two of the other legal deposit libraries (Oxford and Cambridge).

Some practical features of the British Library's repository have been described in an APARSEN project deliverable²⁰. Ingest usually takes place at one of the British Library's sites, with different ingest streams defined for different types of content, e.g. e-journals, digitized newspapers, or Web archive content. All objects have a signature file, which includes a hash value and timestamp, and content is automatically replicated on all four of the storage nodes after ingest. The system assumes that some bit-loss is inevitable when storing extremely large amounts of data. The repository has, therefore, been designed to be self-checking and self-healing; there are periodic integrity checks, and if an object is found to be damaged, it is replaced by a good copy from another node. The repository is designed to be scalable and vendor-independent, using commodity hardware that can be added to as required.

Digital preservation strategy

The British Library's Digital Preservation Strategy, 2013-2016 starts with the assumption that it is the Library's responsibility to preserve and make available content to current and future users. The strategy itself, which was approved in March 2013, outlined four main strategic priorities, i.e. to²¹:

- Ensure our digital repository can store and preserve our collections for the long term;
- Manage the risks and challenges associated with digital preservation throughout the digital collection content lifecycle;
- Embed digital sustainability as an organisational principle for digital library planning and development;
- Benefit from collaboration with other national and international institutions on digital preservation initiatives.

Fundamental to the implementation of this strategy is a sound understanding of the Library's digital collections and the preservation challenges that underpin them. The gathering of preservation requirements provides a baseline for the Library's development of a preservation planning capacity.

Preservation planning

Preservation planning is an essential part of any successful digital preservation strategy because it reflects the need for decisions to be made about appropriate preservation approaches, tools and formats at multiple points within content lifecycles.

Preservation planning potentially covers a very wide range of activities, but the starting point – as with so much else in digital preservation – has usually been the Reference Model for an Open Archival Information System (OAIS), which has also been approved as an international standard (ISO 14721:2012)²². The OAIS Model defines a functional entity for Preservation Planning, that would provide "the services and functions for monitoring the environment of the OAIS [i.e. an archive, broadly defined], providing recommendations and preservation plans to ensure that the information stored in the OAIS remains accessible to, and understandable by, the Designated Community [i.e. the expected users] over the Long Term, even if the original computing environment becomes obsolete."²³ The model also provides some specific examples of what functions might be required²⁴:

Preservation Planning functions include evaluating the contents of the Archive and periodically recommending archival information updates, recommending the migration of current Archive holdings, developing recommendations for Archive standards and policies, providing periodic risk analysis reports, and monitoring changes in the technology environment and in the Designated Community's service requirements and Knowledge Base. [...] Preservation Planning also develops detailed Migration plans, software prototypes and test plans to enable implementation of Administration migration goals. (p. 1-14)

Following the OAIS Reference Model, the concept of preservation planning has also started to be embedded in the emerging standards being developed to support repository audit and certification. For example, Section 4.3 in the ISO 16363:2012 Audit and Certification of Trustworthy Digital Repositories standard includes four separate requirements relating to preservation planning²⁵:

- 4.3.1 The repository shall have documented preservation strategies relevant to its holdings;
- 4.3.2 The repository shall have mechanisms in place for monitoring its preservation environment;

- 4.3.3 The repository shall have mechanisms to change its preservation plans as a result of its monitoring activities;
- 4.3.4 The repository shall provide evidence of the effectiveness of its preservation activities. (p. 4-16 ff.)

Perhaps the most comprehensive attempt so far to develop a structured approach to preservation planning has been the Plato decision-support tool, initially developed as part of the European Union-funded Planets (Preservation and Long-term Access through Networked Services) project²⁶. Plato provides a methodology and a software tool to support the systematic capture of preservation requirements from various stakeholders and then to match these to potential preservation strategies for further analysis. The result of this is a recommendation that can form the basis of a preservation plan, which contains information on contexts as well as recording the evidence base underpinning the decision. Several attempts have been made to integrate Plato with other digital preservation tools. For example, researchers from the KeepIt and Planets projects integrated Plato and other digital preservation tools with the ePrints repository software, creating plugins to ePrints that would support the development of preservation workflows, including the generation of preservation plans and action plans²⁷. Also, the European Union-funded SCAPE (Scalable Preservation Environments) project also explored how best to integrate Plato with a range of other digital preservation tools and services²⁸. In terms SCAPE, it was intended that the resulting preservation plans would be able to document collections, their institutional context, and all of the decision-making process that led to the selection of a particular preservation action. It would also contain a preservation action plan that would contain all of the information necessary to apply the preservation action as well as an executable action plan that could be deployed through a workflow management system like Taverna²⁹.

Preservation planning at the British Library

The British Library is still very much in the process of developing its preservation planning capacity. However, its digital preservation team has begun the process of developing a deeper understanding of the Library's collections and their preservation requirements. Two outcomes of this approach have been the development of collection profiles and a programme of file format assessments.

Collection profiling

Collection profiling was an attempt to work with curators and other content specialists to develop descriptive profiles of the British Library's high-level digital collection areas, with the aim of capturing key knowledge about the collections and their specific preservation requirements³⁰.

The British Library's collection profiling activity drew its inspiration from other content profiling activities, largely those based on a structured dialogue with curators and other content specialists. Particularly influential was the work of the National Library of Australia (NLA) on identifying 'preservation intent' for their digital collections³¹. As part of their approach to preservation planning, digital preservation specialists at the NLA consulted with content specialists (collection managers, curators) in order to develop 'plain-language' statements about "which collection materials, and which copies of collection materials, need to remain accessible for an extended period, and which ones can be discarded when no longer in use or when access to them becomes troublesome." Content specialists at the NLA were also "asked to make broad statements clarifying what 'accessible' means by stating the priority elements that need to be re-presented in any future access for each kind of digital object type in their collections." This both became a means of ensuring that curators and other collection specialists take responsibility for deciding what will happen to collections and is essential for preservation planning. Webb, et al. write that "without it, we are left floundering between assumptions that every characteristic of every digital item has to be maintained forever (almost certainly an impossible expectation) and assumptions that it is good enough to store data safely and let future users worry about how to access it (almost certainly an inadequate response)." Capturing elements of what the NLA refers to as preservation intent seemed vital for the success of the British Library's collection profiling activity.

The digital preservation team also looked at a number of other content-based profile initiatives, in particular the Digital Content Reviews (DCR) for Life Cycle Management developed by MIT Libraries and the Data Curation Profiles developed by Purdue University Libraries. Purdue's Data Curation Profiles are a means of capturing basic information about research datasets in order to support their curation and reuse. The profile provided a comprehensive framework (an interview structure) that could be used to gather information about datasets and their potential

re-use. Once completed, profiles would then be able to help guide decision-making about the management of datasets, as well as inform those providing research data management services of any specific requirements³². While the Purdue team's framework was probably too focused on one particular type of content to be useful for the British Library's immediate purposes, the general approach clearly demonstrated the benefits of using content profiles to support lifecycle management. The MIT Libraries Digital Content Reviews for Life Cycle Management took a similar lifecycle management view to the Purdue team, but – more in common with the emerging British Library profiles – they were primarily intended to help capture information about the implications of collecting certain types of digital content³³.

The development of collection profiles at the British Library was broken down into a number of smaller stages. The initial tasks were to identify the British Library's high-level digital collection areas and to develop an initial template to capture the required information³⁴. Identifying high-level collections was not easy, as there was no agreed list of digital collection types held by the Library. Those lists that did exist - e.g. those provided by the catalogue or website - often included, for reasons of practicality, content types at several different levels of granularity. The initial list of high-level collection types included some that were firmly based on resource type (e.g. sound, multimedia), others that were multi-faceted but based on particular content streams (e.g. Web archives); and others that followed more traditional categorizations of library collections, updated for the digital era (e.g. journals, books). Experience with developing the profiles led us to take a slightly more pragmatic approach, largely because we were not trying to produce a definitive taxonomy of all digital collection types held by the Library, but simply to be able to identify collections at a sufficient (and logical) level of granularity in order to get started on the development of content profiles.

Developing the profiles themselves was also an excellent opportunity to develop conversations with curators and other content specialists on identifying the specific preservation requirements of the British Library's digital collections. As a contribution to the development of a more comprehensive preservation planning function, this has had a number of benefits:

- The massive scale of content held by the British Library means that profiling can be undertaken at collection level, hopefully simplifying the identification of preservation

requirements, and the tools necessary to facilitate these;

- Collection profiling opens a forum in which all collection stakeholders (the people who make decisions at different lifecycle stages) can express the challenges faced by specific content types. This will help with the development of a shared understanding of digital preservation requirements from both curatorial and technical perspectives;
- Corporate understanding of the collections held by the British Library is enriched through the sharing of information about collections between the different departments which make collection management decisions.

The British Library now has a set of collection profiles that can inform the further development of its preservation planning capability and provide a baseline for future requirements gathering.

File format assessments

Another strand of work led by the digital preservation team has been based on the assessment of file formats³⁵. File formats are problematic in digital preservation for a variety of reasons, but it was felt that some of the criteria used to assess file formats were not always consistent with practical experience. It has been noted that working with formats often reveal significant challenges that are not always apparent in assessments made from mainly theoretical perspectives³⁶.

When developing a methodology for completing file format assessments for selected formats used by the British Library, the digital preservation team was keen to base its assessments on evidential principles. In particular, the team was determined to move away from numerical scoring based approaches or simple measures based on the availability or completeness of documentation. In addition, any recommendation to choose a preservation format different from that in which the content was received needed to be supported by strong empirical evidence of the risks and benefits involved.

The file format assessments produced by the British Library aim simply to provide evidence-based recommendations around the use of that format. Any potential risks in adopting the format are identified, noting any possible mitigation. Format assessments have been produced so far for a number of formats of interest to the Library, including: TIFF (Tagged Image File Format), JPEG 2000 (Joint Photographic Experts Group), PDF (Portable Document Format), XML (Extensible

Markup Language), EPUB, JATS (Journal Archiving and Interchange Tag Suite), NTF (National Transfer Format), and WAVE (Waveform Audio File Format). Public versions of the assessments are available from the Digital Preservation Coalition wiki³⁷. The results from these assessments will, in turn, feed into the evidence base that is required to support successful preservation planning.

As a national library, the British Library has a responsibility to preserve its ever growing digital collections and to make them available to current and future users. The extension of UK legal deposit provisions to non-print works in April 2013 represented a major milestone in the Library's acceptance of its role as a digital library. Over the past decade or so, the Library has made significant progress in responding to the organisational and infrastructural challenges of digital preservation, not least in the establishment of infrastructures

to support the processing and storage of digital content as well as in the development of a digital preservation strategy. The next steps for the British Library will be to develop a comprehensive preservation planning capacity. The Library's digital preservation team have made a start on this by developing a deeper understanding of the collections and their curatorial and technical context.

Michael Day

The British Library
96 Euston Rd
London NW1 2DB
United Kingdom
<http://www.bl.uk>
michael.day@bl.uk

February 2016

Notes

1. British Library. *Digital Preservation Strategy, 2013-2016* [on line] <<http://www.bl.uk/aboutus/stratpolprog/collectioncare/digitalpreservation/strategy/dpstrategy.html>>. The British Library, 2013 (consulted on 17 February 2016).
2. Borghoff, Uwe M.; Rödiger, Peter; Scheffczyk, Jan; Schmitz, Lothar. *Long-term preservation of digital documents*. Springer-Verlag, 2006.
3. CCSDS 650.0-M-2. Reference Model for an Open Archival Information System (OAIS) [on line]. CCSDS Recommended Practice, 2012, p. 1-14. <<http://public.ccsds.org/publications/archive/650x0m2.pdf>> (consulted on 17 February 2016).
4. Becker, Christoph; Kulovits, Hannes; Guttenbrunner, Mark; Strodl, Stephan; Rauber, Andreas; Hofman, Hans; Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. *International Journal on Digital Libraries* [on line], 2009, vol. 10, p. 133-157. <<http://link.springer.com/article/10.1007%2Fs00799-009-0057-1>> (consulted on 17 February 2016).
5. See note 4.
6. ISO 16363:2012. Space data and information transfer systems – Audit and certification of trustworthy digital repositories. International Organization for Standardization, 2012.
7. *British Library Act 1972* [on line]. Her Majesty's Stationery Office, 1972 <<http://www.legislation.gov.uk/ukpga/1972/54/contents>> (consulted on 17 February 2016).
8. Breay, Claire; Harrison, Julian (eds.) *Magna Carta: law, liberty, legacy*. The British Library, 2015.
9. Novotny, Deborah. The changing face of storage at the British Library [on line]. In *Proceedings of the 78th IFLA General Conference and Assembly*. Helsinki, Finland, 11-17 August 2012 <<http://conference.ifla.org/past-wlic/2012/102-novotny-en.pdf>> (consulted on 17 February 2016).
10. British Library. *Living knowledge: The British Library, 2015-2023* [on line]. The British Library, 2015 <<http://www.bl.uk/aboutus/foi/pubsch/pubscheme3/living-knowledge-2015-2023.pdf>> (consulted on 17 February 2016).
11. See note 10.
12. Brazier, Caroline. born.digital@british.library: the opportunities and challenges of implementing a digital collection development strategy [on line]. In *Proceedings of the IFLA World Library and Information Congress*, Singapore, 17-23 August 2013 <<http://library.ifla.org/222/1/198-brazier-en.pdf>> (consulted on 17 February 2016).
13. S.I. 2013 No. 777. *Legal Deposit Libraries (Non-Print Works) Regulations 2013* [on line]. The Stationery Office, 2013 <<http://www.legislation.gov.uk/uksi/2013/777/contents>> (consulted on 17 February 2016).

14. Gibby, Richard; Brazier, Caroline. Observations on the development of non-print legal deposit in the UK. *Library Review* [on line], 2012, vol. 61, no. 5, p. 362-377. <<http://www.emeraldinsight.com/doi/full/10.1108/00242531211280487>> (consulted on 17 February 2016).
15. Webster, Peter. *Crawling the UK web domain* [on line]. UK Web Archive blog, 16 September 2013 <<http://britishlibrary.typepad.co.uk/webarchive/2013/09/domaincrawl.html>> (consulted on 17 February 2016).
16. Fleet, Christopher; Kowal, Kimberly C. Planning the migration to OS MasterMap in the UK Legal Deposit Libraries. *Cartographic Journal* [on line], 2005, vol. 42, no. 3, p. 214-221. <<http://www.tandfonline.com/doi/abs/10.1179/000870405X77147>> (consulted on 17 February 2016).
17. Carpenter, Leona; Shaw, Simon; Prescott, Andrew (eds.) *Towards the digital library: the British Library's Initiatives for Action programme*. The British Library, 1998.
18. Prescott, Andrew. The *Electronic Beowulf* and digital restoration. *Literary and Linguistic Computing*, [on line]. 1997, vol. 12, no. 3, p. 185-196. <http://www.uky.edu/~kiernan/eBeo_archives/articles90s/ajp-llc.htm> (consulted on 17 February 2016).
19. Fleming, Patrick. The British Library Newspaper Strategy: developing collaboration with publishers to digitise back runs and to ingest born digital newspapers. In: Walravens, Hartmut (ed.) *Newspapers: legal deposit and research in the digital era*. De Gruyter Saur, 2011, p. 21-30.
20. APARSEN. *Storage solutions summary of inputs* [on line]. APARSEN Deliverable D23.1, 2013 <<http://www.alliancepermanentaccess.org/index.php/about-aparsen/aparsen-deliverables/>> (consulted on 17 February 2016).
21. British Library. *Digital Preservation Strategy, 2013-2016* [on line]. The British Library, 2013 <<http://www.bl.uk/aboutus/stratpolprog/collectioncare/digitalpreservation/strategy/dpstrategy.html>> (consulted on 17 February 2016).
22. ISO 14721:2012. Space data and information transfer systems – Open archival information system (OAIS) – Reference model. International Organization for Standardization, 2012.
23. CCSDS 650.0-M-2. Reference Model for an Open Archival Information System (OAIS) [on line]. CCSDS Recommended Practice, 2012 <<http://public.ccsds.org/publications/archive/650x0m2.pdf>> (consulted on 17 February 2016).
24. See note 21.
25. CCSDS 652.0-M-1. Audit and certification of trustworthy digital repositories [on line]. CCSDS Recommended Practice, 2011 <<http://public.ccsds.org/publications/archive/652x0m1.pdf>> (consulted on 17 February 2016).
26. Becker, Christoph; Kulovits, Hannes; Rauber, Andreas; Hofman, Hans. Plato: a service oriented decision support system for preservation planning [on line]. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, Pittsburgh, PA, USA, 16-20 June, 2008. ACM Press, 2008, p. 367-370. <<http://dl.acm.org/citation.cfm?doid=1378889.1378954>> (consulted on 17 February 2016).
27. Hitchcock, Steve; Tarrant, David; Carr, Les; Kulovits, Hannes; Rauber, Andreas. Connecting preservation planning and Plato with digital repository interfaces [on line]. In *Proceedings of iPRES 2010, 7th International Conference on Preservation of Digital Objects*, Vienna, Austria, 19-24 September 2010 <<http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/tarrant-65.pdf>> (consulted on 17 February 2016).
28. May, Peter; Wilson, Carl. *Technical architecture report*, v2 [on line]. SCAPE Deliverable D2.3, 2014 <<http://www.scape-project.eu/deliverable/d2-3-technical-architecture-report-v2>> (consulted on 17 February 2016).
29. Wolstencroft, Katherine; Haines, Robert; Fellows, Donal; Williams, Alan; Withers, David; Owen, Stuart; Soiland-Reyes, Stian; Dunlop, Ian; Nenadic, Aleksandra; Fisher, Paul; Bhagat, Jiten; Belhajjame, Khalid; Bacall, Finn; Hardisty, Alex; Nieva de la Hilda, Abraham; Balcazar Vargas, Maria P.; Sufi, Shoaib; Goble, Carole. The Taverna workflow suite: designing and executing workflows of Web Services on the desktop, web or in the cloud. *Nucleic Acids Research* [on line]. 2003, vol. 41, no. W1, p. W557-W561. <<http://nar.oxfordjournals.org/content/41/W1/W557>> (consulted on 17 February 2016).
30. Day, Michael; MacDonald, Ann; Kimura, Akiko; Pennock, Maureen. Identifying digital preservation requirements: Digital preservation strategy and collection profiling at the British Library [on line]. In *iPRES 2014: Proceedings of the 11th International Conference on Preservation of Digital Objects 2014*. Melbourne, Australia, 6-10 October 2014 <https://phaidra.univie.ac.at/detail_object/o:378119> (consulted on 17 February 2016).
31. Webb, Colin; Pearson, David; Koerbin, Paul. 'Oh, you wanted us to preserve that?!' Statements of preservation intent for the National Library of Australia's digital collections. *D-Lib Magazine* [on line]. 2013, vol. 19, nos. 1/2. <<http://www.dlib.org/dlib/january13/webb/01webb.html>> (consulted on 17 February 2016).
32. Witt, Michael; Carlson, Jacob; Brandt, D. Scott; Cragin, Melissa H.; Constructing Data Curation Profiles. *International Journal of Digital Curation* [on line]. 2009, vol. 4, no. 3, p. 93-103. <<http://ijdc.net/index.php/ijdc/article/view/137>> (consulted on 17 February 2016).
33. MIT Libraries. *Digital Content Management Infrastructure Improvement: FY13 strategic objective*. MIT Libraries, 2013.

34. Day, Michael; MacDonald, Ann; Kimura, Akiko; Pennock, Maureen. Implementing digital preservation strategy: developing content collection profiles at the British Library [on line]. In *Proceedings of JCDL '14: Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*, London, UK, 8-12 September 2014 <<http://dl.acm.org/citation.cfm?id=2740773>> (consulted on 17 February 2016).
35. Pennock, Maureen; Wheatley, Paul; May, Peter. Sustainability assessments at the British Library: Formats, frameworks, & findings [on line]. In *iPRES 2014: Proceedings of the 11th International Conference on Preservation of Digital Objects 2014*. Melbourne, Australia, 6-10 October 2014 <https://phaidra.univie.ac.at/detail_object/o:378110> (consulted on 17 February 2016).
36. Van der Knijff, Johan. *Assessing file format risks: searching for Bigfoot?* [on line]. Open Planets Foundation blog, 30 September 2013 <<http://openpreservation.org/blog/2013/09/30/assessing-file-format-risks-searching-bigfoot/>> (consulted on 17 February 2016).
37. Digital Preservation Coalition. *File formats assessments* [on line]. <http://wiki.dpconline.org/index.php?title=File_Formats_Assessments> (consulted on 17 February 2016).

INFORMATION IS NOT FOR STORAGE, INFORMATION IS FOR ACTION

Het wijzigende karakter van informatiebeheer in een web 2.0-kenniswerkomgeving

Filip CALLEWAERT

Verantwoordelijke Informatiebeheer, Havenbedrijf Antwerpen¹

■ In dit artikel beschrijft Filip Callewaert hoe en waarom hij met zijn team hun praktijk van informatiebeheer herontwierp in functie van nieuwe, 21ste eeuwse noden: *"information is not for storage, information is for action"*. Dit wordt gelinkt aan een lean framework, en staat helemaal in het licht van de creatie van waardetoevoegend kenniswerk. Webgebaseerd werken, en vooral de sociale, collaboratieve 2.0-dimensie ervan, maakt het nu mogelijk dat het eigenlijke kenniswerk als proces gefaciliteerd wordt, waar we in het verleden onze focus enkel hadden op het product van kenniswerk: afgewerkte documenten en bestanden. Het 2.0-web gooit het gesloten karakter van kenniswerk open, omdat het de conversatie als probleemoplossende techniek een fundamentele plaats geeft, en bewaart. Echter, software is nooit de oplossing; oplossingen worden altijd "ontworpen" in samenspel met de context, en liggen in se in de holistische kijk op processen, tools en mensen - met hun vaardigheden, gewoontes en culturen.

■ Dans cet article, Filip Callewaert décrit comment et pourquoi son équipe et lui ont réinventé leur pratique de la gestion de l'information, en fonction des nouveaux besoins du 21^e siècle : *"information is not for storage, information is for action"*. Ceci est lié à un mode de gestion "lean", allégé, au plus juste. Le but est de créer un travail de connaissance à valeur ajoutée. Le travail basé sur les technologies du web, et surtout sur sa dimension 2.0 sociale et collaborative, permet maintenant de faciliter le vrai travail de la connaissance en tant que processus, là où dans le passé nous nous focalisions uniquement sur le produit du travail de la connaissance : les documents et les fichiers finalisés. Le Web 2.0 renverse les barrières fermées du travail de la connaissance, car il attribue une place fondamentale à la conversation, considérée comme une technique de résolution de problèmes, et qui sera dès lors conservée. Cependant, le logiciel n'est jamais la solution ; les solutions sont toujours "conçues" en interaction avec le contexte : au fond, elles résident dans la vision holistique sur les processus, outils et personnes, avec leurs compétences, habitudes et cultures.

Ik was verheugd uitgenodigd te zijn als spreker op het congres ABD-BVD Inforum in 2015. Het was een druk jaar want elke maand sprak ik wel ergens in binnen- of buitenland een groep geïnteresseerden toe. Een trigger was zeker de awards die wij met het Havenbedrijf wonnen in 2014 in Washington voor een informatiebeheerconcept dat we hadden uitgewerkt. Het lean denkkader was ons uitgangspunt geweest: wat betekent het om vanuit het domein informatiebeheer waardecreërend te zijn? En we waren hierbij bereid zaken radicaal te herdenken. Het concept had een grotere impact dan we dachten; we bleken ver van het traditionele informatiebeheer geëvolueerd, en het concept en de praktijk bleken verrassend goed aan te sluiten bij de gekende uitdagingen om 21ste-eeuwse teams van kenniswerkers te organiseren om de bedrijfsdoelen te realiseren. Het is de eerste keer dat ik dit ook in een artikel uitschrijf - ik wou het al een hele tijd heel graag doen, maar de optredens waren even prioritair.

Omdat we er met het team in het Havenbedrijf ook al vaak hebben over nagedacht, wil ik even stilstaan bij "het recept" waarvan wij denken dat het het verschil uitmaakt. Ik schets de bredere context voor dit nieuw informatiebeheer, en moet zeker de evoluties van digitaal en het web hierin schatplichtig zijn. De krachten van het sociale Web 2.0 in onze bedrijven radicaal binnenbrengen is zeker een belangrijke differentiator. Maar daarmee doe je het niet; technologie alleen creëert niet het succes, het gaat ook over het meekrijgen van de gebruikers, de kenniswerkers, en hun leidinggevenden, waartoe zij nieuwe competenties moeten ontwikkelen en oude gewoontes afzweren. En een derde differentiator is het alom vergeten 'proces': hoe gaan we nu concreet te werk, als team, met informatie, om concrete bedrijfsproblemen tot een goed einde te brengen. En hier vond ik het belangrijk om naar de roots van kenniswerk te gaan.

Wie het recept zelf zou willen toepassen, raad ik aan om goed de basisprincipes van het geheel te

doorgronden, en de specificiteit van de differentiërende factoren te doorzien, om het dan "in het klein" (best in het eigen team eerst) op te zetten, om de praktische werking ervan te leren kennen, en de afstemming op de eigen bedrijfscontext zo goed mogelijk vorm te geven. Als het in je team werkt, kan je naar het volgende team, en dan alsmaar verder, en je schrijft een bedrijfsverhaal dat niet meer eindigt, waarvan je vaak zal merken dat je meehelpt het bedrijf écht anders te doen werken, heel wat meer afgestemd op de noden van deze nieuwe tijden.

Lean informatiebeheer vraagt een focus op efficiënter kenniswerk

Één van de zaken die mij uitermate boeit is hoe wij samen prutsen met informatie om de zaken gedaan te krijgen. Ik denk aan ons e-mailverkeer, aan ons "samenwerken op teksten", aan onze manier van inefficiënt vergaderen (tot slapens toe), aan onze gebruiksonvriendelijke BPM-systemen (Business Process Management), aan onze informatie-stress tijdens het bijeenprokkelen van informatiefragmenten om toch weer zicht op de zaak te krijgen, en aan de zoveelste formaat voor onze todo-lijst terwijl de vorige verre van afgewerkt is.

Het was een lean-cursus aan de Antwerp Management School die mij een heel interessant denk- en werkkader bood om informatie-werk efficiënter te bedenken. Lean gaat niet zozeer over efficiëntiedrang. Het gaat in de eerste plaats om de betrachting om de volle energie te investeren in zaken die er écht toe doen, zaken die een meerwaarde voor de klant realiseren. Als dat de focus is, ga je inderdaad ook aandacht hebben voor efficiëntie, en effectiviteit - maar dat zijn gevolgen, geen doelen. Als ik mij afvraag wat de waardecreatie van informatiebeheer moet zijn, dan link ik dat aan mijn persoonlijk professioneel credo, dat ik haal bij Peter Drucker, een ooit visionair man als het over "managen" ging. Hij zei in de vorige eeuw:

"The most important, and indeed the truly unique, contribution of management in the 20th century was the fifty-fold increase in the productivity of the manual worker in manufacturing. The most important contribution management needs to make in the 21st century is similarly to increase the productivity of knowledge work and the knowledge worker²".

Omdat de kenniswerker maar één grondstof heeft om zijn werk te doen, m.n. informatie, geloof ik inderdaad dat het daar vandaag en morgen in

informatiebeheer om gaat: het efficiënter maken van het kenniswerk.

Kenniswerk is eigenaardig informatiewerk

Kenniswerk is niet-routinematig werk, met één enkele grondstof: informatie. Het doel: bedrijfsuitdagingen en -problemen aanpakken en tot een goed einde brengen. Het is probleemoplossend werk en is daarom per definitie niet-routinematig en creatief.

In het kenniswerkproces doet informatie zich voor in twee gedaantes. Enerzijds gebruiken we bestaande informatie die je moet bijeenprokkelen - een essentiële fase voor probleemoplossing. Ik doe een beroep op diverse *repositories* in onze bedrijfssystemen en daarbuiten, op het geheugen van diverse collega's en betrokkenen, en op mijn eigen geheugen, om bestaande, voor dit probleem relevante informatie terug te vinden.

Anderzijds maken we in een probleemoplossingsproces heel wat nieuwe, tijdelijke, onaffe informatie - werkgeheugeninformatie - die we nodig hebben om onze zinnen op orde te krijgen, om de zaken te snappen. In deze context is de dialoog met andere betrokkenen (en jezelf) een noodzakelijk middel om met taal (en informatie) het proces van begrijpend oplossen vorm te geven. Dat onderdeel van kenniswerk is chaotisch en dynamisch.

Aandacht voor het "dialogische" (dynamische) kenniswerkproces hebben we traditioneel helemaal verwaarloosd, omdat we dat amper te managen achtten. Daar zit dus een belangrijk win-gebied. Traditioneel focussen we het liefst op de afgewerkte "producten" van kenniswerk: de beslissing, de visietekst, de analyse, de synthese, het besluit, het rapport, de beoordeling. Dat is makkelijk te vertalen als: aandacht voor documentbeheer en recordbeheer. De voorbije generatie kennismanagement was van dat soort. Dit heeft nooit gewerkt.

Het eigenlijke, dynamische kenniswerk, zo maakten we onszelf wijs, speelt zich af in de "black boxes", de breinen van onze medewerkers. We kunnen daar niet in kijken, en we kunnen dat dus niet beheren. Het makkelijkste was dan ook de resultaatgerichte "command & control"-stijl à la tegen maandag moet dat voorstel klaar zijn. We hadden wel zo wat procesjes in mekaar geknutseld om dat dialogische karakter toch wat vastere vorm te geven - het "in team nadenken en probleemoplossend communiceren". We vonden bijvoorbeeld een genre "vergaderen" uit, waarbij

we op hetzelfde moment op dezelfde plaats samenkwamen om het probleem te bespreken en te behandelen. Nog een genre dat we hebben uitgevonden om het kenniswerk gedaan te krijgen: projectbeheer, dat in principe ook volkomen abstractie maakt van het dialogische karakter van de probleemoplossing en enkel op (tussen)resultaten focust. Recentere afgeleide vormen zoals Scrum³ hebben al meer respect voor co-creatie en dialoog.

We laten het over aan het individu om het voor hem meest efficiënte middel in te zetten: "dialogische communicatie" via e-mail, telefoon, gesprek en meeting. We zien momenteel een explosie van de zogenaamde "productivity apps": software (heel vaak laag geprijsd of gratis, bv. *Endnote*, *OneNote* (is straatoud maar wordt nu populair), *Slack*, *WhatsApp*, ...) die helpt om je virtuele communicatiestromen wat fatsoenlijker te laten verlopen. Ze zijn inderdaad vaak een alternatief voor e-mail, maar creëren jammer genoeg vaak alweer een extra silo in het kenniswerklandschap van je team en organisatie. Op zich lossen ze geen e-mailprobleem op - ze maken de inbox minder vol door een ander vat ernaast te zetten. Het zijn maar suboptimale winsten, volgens mij, zolang de versnippering blijft.

Maar we beginnen hoe dan ook interessante veranderingen te zien, vooral nu we middelen krijgen om dat dialogische toch te vatten, te expliciteren, te bewaren, te ordenen en te herordenen IN het kenniswerk zelf, niet naast het werk. We gaan "luidop werken" in onze organisaties, op bijv. virtuele, collaboratieve, sociale platformen, die al wat meer features bieden dan de versnipperde productivity apps. En plots ligt het daar voor het grijpen: het probleemoplossingsproces, geëxpliciteerd in dialogen tussen medewerkers, waarbij af en toe wat informatie stolt, maar stap per stap evolueert naar een finaal product - de output.

Informatiebeheer vandaag

En hier zit voor mij het immense potentieel om kenniswerk efficiënter te maken, zoals Drucker het ons als 21ste-eeuwse opdracht meegeeft: een focus creëren, weliswaar in samenhang met record-denken, op het dynamische, betekeniswordingsproces. "*Information is for action, not for storage*"⁴. Technologie zoals sociale en collaboratieve platformen helpen ons om meer vat te krijgen op een dieperliggend niveau van kenniswerk, net zoals de microscoop en de telescoop ooit het begrip van de ons omringende wereld sterk hebben beïnvloed. Dit gaf dan weer aanleiding tot het ontstaan van nieuwe disciplines, en die leidden tot nieuwe technologieën, die op hun beurt weer mee bepalen

hoe onze dagelijkse wereld er vandaag voor ons uitziet. Esko Kilpi titelde in november 2015 een van zijn blogs: *Conversations and narratives are the new documents*⁵. Inderdaad, zij zijn de motor van de probleemoplossing, zij zijn de context, zij zijn de ware kennis. Kennis is informatie in actie. Dat wordt ook het motto voor het postgraduaat Kennismanagement dat in 2016 opstart aan de Arteveldehogeschool in Gent, en waarbij ik me gelukkig prijs eraan te mogen meewerken.

De toekomst is hier al

Informatiebeheer moet dus gaan om het efficiënter maken van kenniswerk, en ik vermoed dat we hier net nog niet begonnen zijn. In optimistische buien zeg ik: we zijn pas gestart - immers, de toekomst is nabij, want de ingrediënten zijn er al. William Gibson, die ik enkel ken van een inspirerend citaat, schreef ooit: "*The future is already here; it's only unevenly distributed*"⁶. Dat is heel mooi gezegd en geschreven. Het is ook zo: nieuwigheden komen nooit als een komeet uit de lucht vallen, ze bestaan al ergens. De toekomst ontwikkelt zich op basis van wat is, en wat wij in zwaar of licht innovatieve buien - meestal op heel bescheiden wijze - daaraan voort breien, steek per steek, en we laten er veel vallen. Het is een trager proces dan we vermoeden. Heel wat innovaties, uitvindingen, prototypes halen eigenlijk nooit het grote forum - vele sterven een stille dood na misschien eerst gehypet geworden geweest te zijn. Een decennium later, met mogelijk licht andere technologieën (misschien gewoonweg gebruiksvriendelijker en sneller), en wat kleine aanpassingen, kan mogelijk toch een groot verschil gemaakt worden, zodat het ding wel aanslaat en doorbreekt. Daarenboven zijn heel wat innovaties in principe "maar" knappe assemblages van allemaal bestaande zaken. De officiële uitvinder of innovator is altijd schatplichtig aan een resem onbekende collega's en voorgangers die in het centrum of in de marge van hetzelfde project werkten. Veel wetenschappers schrijven zich dood aan artikels en publicaties waarmee zij nooit beroemd zullen worden, maar toch zijn zij vaak een schakel in een wetenschappelijk proces dat op termijn van betekenis is. De academische wereld en pionierbedrijven als Google en Apple zijn goede voorbeelden van contexten waar de toekomst nu al bestaat. Het zal alleen nog wat duren voor die toekomst ook in onze straat zichtbaar is. Als ik in juni 2014 in Washington een dag kan rondlopen met de Google Glasses van de bevriende Chuck, en er een immens potentieel van inzie voor het havenbedrijf waar ik werk, dat ook beheerder is van heel wat geografische data, dan kan ik mijn fantasie niet de baas in toepassing van "augmented reality" voor brandweer,

ingenieurs, aannemers, stuurliu enzovoort. Als ik echter terug thuis ben (zonder bril weliswaar want op het voorstel er een mee te brengen hebben ze nooit geantwoord), dus thuis, weer dichterbij de IT-directeur die me het jaar ervoor nog aanmaande "*dat ik er hier niet moest aan denken om mobiele apparaten zoals iPads te introduceren*", beseft ik het maar al te goed: *de toekomst is er al, maar nog niet hier*. Doch we moeten moedig blijven en desondanks het verleden dat hier nu ook nog aanwezig is, er toch blijven in geloven. Om innovatief te zijn, moet je soms maar rond je kijken, want de ingrediënten zijn aanwezig. Wat creativiteit, de nodige moed om radicaal te zijn, en veel daadkracht, ontdoen innovatie van zijn huidig-romantische invulling, en maken er iets doenbaars van.

Kijk even rond je: belangrijke (r)evoluties en disrupties die zich momenteel voordoen in onze bedrijven en organisaties zijn er in veel gevallen gekomen door de grillige krachten van 4 hoofdbestanddelen: digitaal, het web 1.0, het web 2.0 en het web 3.0. Het zijn essentiële bouwstenen om informatie en informatiebeheer vandaag beter te begrijpen. Ik besluit echter niet in te gaan op Web 3.0 in de scope van dit artikel - het "intelligente web", met in dit licht immer geciteerde Watson van IBM⁷. 3.0 zit nu in een marketinghype, volgens mij, met nog veel lucht, en het levert nog niet die toepasbare intelligente technologieën die ik in mijn informatiebeheer vandaag zou wensen. Mail me als je daarover in dialoog wil gaan. In ons concept gaat het nu vooral nog over de kracht van het Web 2.0 - dat krijgt hier daarom het meeste aandacht.

Digitaal

Digitale informatie gaat over bits en bytes, die het maken, kopiëren, distribueren, transporteren van informatie snel en goedkoop maken. Het is een evolutie die we sinds de jaren 80 sterk hebben zien toenemen in onze bedrijven (en daarbuiten natuurlijk ook), voornamelijk door de introductie van de computers (- let wel: de eerste computer bestond al voor de Tweede Wereldoorlog!). Gordon Moore, mede-oprichter van Intel, "voorspelde" in 1965 dat hij de snelheid van zgn. chips elke 12 maanden zou zien verdubbelen, tegen dezelfde kost. Hij had gelijk. Dat levert snellere computers op, maar ook miniaturisering van digitale toestellen. Het is natuurlijk deze processorkracht die van onze GSM een smartphone heeft kunnen maken, die touch gebruiksvriendelijk heeft gemaakt, en die nu de opkomst mogelijk maakt van de vele sensoren die het "internet of things" voeden - digitaal zit overal. Netwerken zijn op

zich natuurlijk ook erg geëvolueerd: sneller, meer bandbreedte, maar ook... draadlozer.

Ook in onze bedrijven zijn wij in veel gevallen ook nog vooral blijven focussen op de kracht van digitaal, zonder daar de revolutionaire krachten van "het web" in te betrekken. We werken nog veelal in mappen, met bestanden, met ingescande documenten, met business process management applicaties, databanken, krachtige grafische applicaties, krachtige rekenprogramma's enzovoort. Daar biedt digitaal ons zijn meerwaarde. Het doet ons bepaalde zaken veel sneller doen, dan we ze in het verleden konden doen. We kunnen in elk geval die computers niet meer wegdenken uit onze praktijk.

Web 1.0

Digitalisering en de genetwerkte computers liggen natuurlijk aan de basis van het wereldwijde web. Zonder digitaal, geen internet, en geen web. Het was in 1993 dat het WWW "open" werd gesteld voor het grote publiek. We konden met zijn allen gaan surfen - nog niet vanuit de luie zetel weliswaar, want onze volumineuze pc hing nog vast met den draad aan een 28.8 Kb/s modem. Wij waren in hoofdzaak informatie-consumenten in het Web 1.0; wij consulteerden websites.

De essentie van Web 1.0 is de gehyperlinkte content in een open systeem. Het web verbindt content, dingen (het *Internet of Things* - dat de al aanwezige big data nog bigger maakt) en mensen (d.m.v. de content die zij produceren en d.m.v. hun *wearables*). Dit gebeurt in een open systeem dat "anywhere anytime" tot onze beschikking staat, zeker sinds "mobiel" zo enorm gebruiksvriendelijk is geworden. Het open karakter ervan is essentieel. Je geniet er niet enkel van een kritische massa content, data en mensen als nooit tevoren, maar je laat ook het toeval een belangrijke rol spelen in het vinden en gevonden worden. Geen serendipiteit zonder dit open karakter.

Het web heeft ten slotte ook geen gepredefinieerde structuur (zoals b.v. een databank dat wel heeft). Dit is een belangrijke karakteristiek. Er is geen keurslijf, er is enkel de mogelijkheid content te maken en te publiceren, en die te linken met andere content. De structuur op het web zelf is emergent van aard: ze "ontstaat" - letterlijk op te vatten. Ze krijgt vorm door het gebruik ervan. Volgens nood en behoefte schikken, linken en herschikken we content.

Web 2.0

Vanaf ongeveer 2004 veranderde het Web 2.0 het WWW. Waar het Web 1.0 een vrije "statische" html-aangelegenheid was, ontstond met de komst van nieuwe scripttalen en technologieën de mogelijkheid tot veel meer interactiviteit op het web. Content opladen werd heel makkelijk en verliep snel, ook door een toegenomen performantie van het netwerk (steeds betere modems, betere verbindingen, snellere verwerking van informatie in snellere processors, ...). Zo werd het een realiteit dat de vele consumenten-gebruikers stilletjes aan ook op een vrij makkelijke manier producenten van "user-generated" content werden.

Traditioneel proclameert *Time Magazine* op het einde van het jaar een 'persoonlijkheid van het voorbije jaar'. In 2006 werd dat "YOU"⁸: jij. "Ja, jij. Jij controleert het Informatietijdperk. Welkom in jouw wereld." – stond (in het Engels weliswaar) te lezen op de voorpagina. Want inderdaad, iedereen met een simpele internetaansluiting was auteur geworden, en kon zijn stem laten horen. Een aantal "sterke" verhalen werden hoe dan ook gehypet, van slecht-behandelde klanten die op sociale media het bedrijf ten schande brachten, wat viraal ging, en het bedrijf publiek tot verontschuldigen dwong. En zo zijn we in mini-versies allemaal "likers" geworden, post'ers van onze ervaringen en opinies. Genereerde het Web 1.0 al een overvloed aan informatie, het Web 2.0 deed dat exponentieel toenemen, omdat snelle feedbackloops ongelimiteerd mogelijk werden.

Met 2.0 werd het web vanaf ongeveer 2004 stelselmatig een steeds groter platform waar jij en ik elkaar kunnen vinden, waar we het virtueel gezelliger maken of elkaar de duivel aan doen, met elkaar in gesprek gaan, en samen nieuwe content produceren. "*Conversations and narratives are the new documents*"⁹, schrijft Esko Kilpi in zijn blog in 2015; deze virtuele interdynamiek creëert zeker ook een nieuwe dimensie in ons kenniswerk. Het concept van het web als platform is belangrijk: een soort van "podium" waar mensen samenkomen en dingen doen. *Communication, social, collaboration* zijn hier de wachtwoorden. Verwar het niet met e-mail (dat eigenlijk meer op sms'en lijkt). E-mailen is niets anders en niets meer dan een vorm van a-synchroon digitaal communiceren, dat weliswaar gebruik maakt van de internetinfrastructuur, maar dit slechts als transportmiddel gebruikt. Je stuurt informatie van punt a naar punt b. Met het web 2.0 publiceer je content OP het web, en anderen komen naar dat punt om daar op jouw content te reageren, of om er nieuwe content te posten. De al aanwezige kenmerken van het

web 1.0 worden volkomen gerespecteerd, vooral dat "open" karakter van het platform: iedereen kan deelnemen en reageren. En zo ontstonden in het begin fora, blogsites, wiki's, en wat later *Facebook*, en microblogsites zoals *Twitter*.

De overvloed aan informatie lijkt met de vele en makkelijke 2.0-feedbackloops onhoudbaar, en we moeten zeker wennen aan deze hyper-interdynamiek. Maar 2.0-platformen moeten het hebben van sociale dynamiek. Dat is met cafés ook zo. Een café zonder volk nodigt niet uit om binnen te gaan. Op sociale platformen heb je twee soorten "volk" nodig: enerzijds de producenten: zij die content posten, en anderzijds de consumenten: zij die content lezen. Eén persoon kan makkelijk in beide rollen meespelen: wie posts "likt", wordt ook producent. Nieuwe Facebookgebruikers starten met het "consumeren" van content: ze lezen wat andere posten. Dan gaan ze al eens durven "liken", om in het beste geval ook al zelf te durven posten. Als ze echter nooit feedback krijgen op hun posts, gaan ze er mee ophouden, en worden ze weer louter consument. *Community management* is daarom erg belangrijk op dergelijke platformen om deelnemers aan te zetten actief te participeren. Er zijn cases bekend van bedrijven die met succes sociale dynamiek hebben gefaket, om nieuwe bezoekers het gevoel te geven dat zij al in een halfvol café binnenkwamen.

Wat doen we met het web 2.0 in onze bedrijven?

Digitaal, Web 1.0, 2.0 en 3.0 hebben elk afzonderlijk hun eigen evolutielijnen, maar die lijnen spelen ook op elkaar in. En ze zijn op zich elk erg beloftevol. Maar als ik terugdenk aan wat kenniswerk is, aan het dialogische karakter ervan, het spel tussen ongestructureerdheid en gestructureerdheid, het belang van gelinkte informatie en daartussen en daarboven de betekeniscreërende dialogen en feedbackloops, dan wil ik voor de kenniswerkers van mijn organisatie op een authentieke manier inzetten op de krachten van het Web 2.0.

In 2006 en later schrijft de M.I.T.-onderzoeker Andrew McAfee een aantal blogs waarin hij de term "*Enterprise 2.0*" lanceert – in 2009 verscheen zijn gelijknamige boek¹⁰. Daarin behandelt hij de kracht van het Web 2.0 binnen de bedrijfscontext: hoe de inzet van wiki's, fora, blogs, microblogging als interactieve webtoepassingen de bedrijfsvoering grondig kan beïnvloeden.

Ik ga voor onze doeleinden vooral in op de krachten van het Web 2.0 voor de *interne* organisatie van onze bedrijfsvoering. Het Web 2.0 heeft zijn

krachten tot nu toe vooral bewezen in de externe relaties van bedrijven, naar de klanten toe, wat in 2.0-termen betekent: *met de klanten*. In eerste instantie was het het marketingdepartement dat zich als trojaans paard op de content-platformen begaf, om daar "gunstige" content over het bedrijf te stimuleren. In een tweede fase hebben nieuwe bedrijven het Web 2.0 als puur platform ontdekt. Daar gebeuren nu de meeste disrupties. Dergelijke nieuwkomers hebben soms maar één product: "een web-platform" waarop ze waarde-producenten en waarde-consumenten samenbrengen en interconnecteren. AirBnB heeft een miljoenenomzet in de business van vakantieverblijven, zonder ook maar zelf één kamer te bezitten; idem voor Uber en taxi's. Zij zorgen dat uit te wisselen diensten zo geruisloos en efficiënt mogelijk verlopen, ze creëren een erg "lean" proces: de klant en de aanbieder kunnen volledig op de waardecreatie focussen; AirBnB biedt de nodige ondersteunende processen (reservatiesoftware, de betaling, consumentenfeedback,...). Klant en leverancier laten hun stem horen op deze platformen over hun al of niet tevredenheid van elkaars aandeel.

Kunnen Web 2.0-krachten ook de interne organisatie van onze bedrijven mee vorm geven, om management en medewerkers tot een hechte gemeenschap van kennisproducenten en -consumenten om te toveren? Omdat het om virtuele omgevingen gaat, kwamen de eerste succesverhalen van multinationals, die er met dergelijke technologie uiteraard beter in slaagden om tijds- en ruimtelijke grenzen te overstijgen. Medewerkers creëren content (bv. hun gebruikersprofiel, of blogjes over hun projecten) en laten zich op die manier opmerken door collega's aan de andere kant van het bedrijf, waarvan zij het bestaan niet eens afwisten. Omdat dit "open" stromen van informatie zijn, gaat serendipiteit hier ook een positieve rol spelen: *toevallig* stoten op interessante informatie gebeurt in dergelijke omgevingen natuurlijk veel meer. Dit onderdeel van de 2.0-Enterprise vinden we vooral terug op congressen over (en in content van leveranciers van) het bedrijfsintranet, dat al lang de bedoeling had, maar er niet in slaagde, om enerzijds de medewerkers onderling, en anderzijds medewerkers en management "dichter" bij elkaar te brengen via een communicatieplatform (het "teamgevoel", "de visie en missie", "betrokkenheid", "engagement", "alignment", ...). Een stap verder dan dat intranet gaat dan over gefocuste "collaboration" (het begrip "collaboratie" wordt na lange tijd in onze contreien weer positief), waar samen content gemaakt wordt in functie van de oplossing van een of ander bedrijfsprobleem. Voor zover het asynchrone communicatie betreft die

"open" is en webgebaseerd, laten we het onder de 2.0-koepel vallen.

Nu, desondanks de vele gepretendeerde succesverhalen over het Intranet 2.0 (en nu meer en meer ook de evolutie ervan richting "digitale werkplaats"), vraagt McAfee in een blog in 2014¹¹ zich mogelijk terecht af hoe het komt dat de doorbraak van 2.0 binnen bedrijven zo lang op zich laat wachten. *"The news from Facebook, along with rapid take-up of new tools like Slack, the continued success and growth of Salesforce's Chatter and Yammer (now part of Microsoft), and evidence of a comeback at Jive, indicates that the business world might finally be coming around to Web-style communication and collaboration tools. Why did it take so long? I can think of a few reasons. It's hard to get the tools right —useful and simple software is viciously hard to make. Old habits die hard, and old managers die (or at least leave the workforce) slowly"*. En terwijl McAfee deze blog schrijft tuimelt de beurskoers van Jive al een tijdje naar beneden, en ontslaat Microsoft een kernteam dat rond Yammer werkt... En ik moet ook toegeven dat nog geen enkele van mijn tweets, LinkedIn- of Facebook posts viraal is gegaan. Hooguit een 20-tal "likes". En eigenlijk is dat evenmin het geval in mijn kennissenkring en bij hen die ik volg... Het blijft precies toch wel allemaal een beetje naar hype ruiken. En ja, er zijn al heel wat Yammer-ontgoochelingen genoteerd ook - breder te interpreteren als ontgoochelingen met *enterprise social networks*. Moeten we het daarom opgeven en er niet meer in geloven?

Het hoort er blijkbaar bij om in ons domein van het huidige informatiebeheer te leren omgaan met hype-mist: hoe deze te identificeren, en hype-kaf van koren te onderscheiden, om het kind niet met het badwater te moeten weggooien, want dat mogen we inderdaad niet doen. Digitaal, web 1.0 en web 2.0 hebben wel degelijk al hun krachten sterk bewezen in bepaalde domeinen, en meer zelfs: deze krachten zetten bedrijfsdomeinen werkelijk op hun kop – ze verstoren de markten grondig, oude modellen gaan failliet, de druk om mee te zijn is groter dan ooit.

Laten we duidelijk zijn: voor mij is Enterprise 2.0 in geen geval een hype. Ik zie en voel dagelijks van heel dichtbij de krachten ervan in mijn eigen team. 3.0 laat ik nog even voor wat het is. Maar 2.0 werkt echt. En we hebben al veel nagedacht waarom het werkt en hoe het nog beter kan. Is er dan iets als een "succesrecept"? Dat zit volgens ons in deze combinatie: een *radicale* inzet op Web 2.0-tools, ondersteund door mensen met de nodige competenties, die samenwerken met informatie in

een proces dat we "*adaptive case management*" noemen. En omdat het een vorm van patisserie is, moet je alles goed afwegen.

Web 2.0-tools

Terug naar McAfee. Hij haalt twee interessante aspecten aan voor het traag doorbreken van Enterprise 2.0, maar hij vergeet een derde. Hij ziet voor het uitblijven van de doorbraak van Enterprise 2.0 ten eerste al een oorzaak in de "tools". Hij heeft gelijk: veel van die 2.0-tools waren er al in 2003 (fora, blogs, wikis), maar moesten beter worden, sneller, gefocuster (bv. in de vorm van een app), 'mooier' (wat betekent: ook voor niet-IT'ers), gebruiksvriendelijker, en mogelijk ook "ingebed" in een ruimer ecosysteem van "kenniswerken" - ze stonden te veel "naast" het werk.

In ons havenbedrijf is de 2.0-tool Confluence van het bedrijf Atlassian gekozen, na wat experimenten en vergelijkende testen. Het heeft een wiki-motor; die zorgt op zich al dat alle Web 1.0 en 2.0-benodigdheden aanwezig zijn. Dan is er nog een blogfunctionaliteit die dienst doet als interactief ad valvas bord van teams. Er zitten taak- en workflowmogelijkheden in, en er zijn nog meer geavanceerde mogelijkheden met bv. JIRA (een zustersoftware van Confluence) om content te linken met meer gestructureerde informatie, bv. met e-loketten, projecten e.d. De wiki-motor doet ons in termen van open, gehyperlinkte content denken en handelen, omdat we er niet enkel snelheid, maar ook betrokkenheid en serendipiteit mee willen realiseren. Het sloeg aan. Het traditionele Enterprise Content Management systeem (dat in termen van "bestanden" en "mappen" dacht), leed er sterk onder. Confluence-gebruikers maken geen Word-bestanden meer, ze mailen niet meer, ze vergaderen veel minder; ze maken web-content en delen die al of niet met gerichte instructies met teamgenoten, die er op hun beurt hun ding mee doen. Ze ervaren dat informatie-zen tot de mogelijkheden behoort, i.p.v. informatiestress. De sterkte is dat deze kenniswerkers op één "virtuele"plaats, met en rond informatie (product + proces) bedrijfsproblemen kunnen oplossen - geen versnippering van informatie in diverse silo's. Is de tool misschien "het geheime ingrediënt"?

Ik heb het op diverse conferenties kunnen ervaren: als er over tools gesproken wordt, belanden we al snel in bizarre discussies. Het lijkt ten eerste al ongeoorloofd om te veel over "de tool" te spreken. Hier spelen de change managers hun rol, en zij duiden op het grotere belang van het verander- en adoptieproces bij gebruikers en organisatie i.p.v. de focus op "de tool" zelf. Ze hebben gelijk en geen

gelijk. De tool is weliswaar een hulpmiddel, maar zonder die tool kan ik de efficiëntie en effectiviteit (en nog veel meer) niet realiseren die ik met wel realiseer. Het kiezen van "de juiste tool" is voor mij essentieel: die bepaalt niet enkel technisch wat je allemaal zal kunnen en niet kunnen, maar ook vanuit het perspectief van gebruiksvriendelijkheid zal die meer zijn dan een detail in het verhaal. Een geslaagde gebruikerservaring betekent dat je gebruikers er al of niet bewust toe aanzet dingen te doen die ze moeten doen.

In het kader van de 2.0-krachten die ik ambieer in werking te stellen wil ik met mijn tool mensen op een platform brengen, waar zij met en rond informatie kunnen samen werken naar een oplossing voor een organisatie-uitdaging toe. Er *ontstaat* een open netwerk van alle relevante en valabele informatie, enerzijds referentiemateriaal voor onze probleemoplossing, anderzijds conversaties daarover en er rond. De betrokkenen moeten daartoe zelf de vrijheid hebben om volgens de noden van het oplossingsproces de nodige structuur aan te brengen in de zich ontwikkelende content-lijnen, richting oplossing. Afgaand op deze "vereisten" kan ik al heel wat tools uitsluiten als ongeschikt. Zo zijn tools als Yammer en Slack maar half geschikt voor wat we ambiëren, want we moeten ook nog te veel andere tools extra inschakelen om het werk consistent gedaan te krijgen - dat is het probleem van al die productivity apps. Sharepoint of MSO365 is voor mij ook (nog) niet geschikt omdat het nog te veel *het bestand* centraal plaatst, i.p.v. simpelweg *gehyperlinkte "content"*. Geloof me, op content werken in plaats van op (Word-) bestanden werken is een wereld van verschil.

Competenties, gewoontes, cultuur en bazen

Tools kunnen je positief helpen in het creëren van "*nudges*" (duwtjes) om gebruikers toch maar die kleine maar belangrijke dingen anders te gaan doen. Maar het blijft een belangrijk change management proces om mensen aan te zetten om écht anders te werken - het impliceert meestal ook "anders gaan denken". McAfee heeft het terecht over het breken van "*old habits*". Zonder voldoende aandacht voor cultuur, competenties, gewoontes, stort je Enterprise 2.0-verhaal in mekaar of blijft het een leuk projectje in de marge dat waarschijnlijk een stille dood sterft. Op meta-niveau moet je hoe dan ook inzetten op verandermanagement: je wil de "*minds*" van de betrokkenen mee krijgen in dit belangrijke veranderverhaal, en daartoe is luisteren, praten en begeleiden van groot belang. Het veranderen zelf krijgt de nodige aandacht.

Maar "Web 2.0" gaan werken vraagt eigenlijk ook totaal nieuwe competenties van kenniswerkers - je gaat bv. heel anders met elkaar samenwerken (geen e-mail meer, veel minder vergaderen, informatie ga je zien als een grondstof voor probleemoplossing). Zelf hebben we in onze 2.0-projecten aan den lijve ondervonden dat je makkelijk de aandacht voor deze nieuwe, nodige competenties negeert, hoogstwaarschijnlijk omdat het WWW een evidentie is geworden in onze wereld - maar niet op en in het werk! Enkel al de introductie van de hyperlink doet het lineaire werken en denken van de medewerker wankelen. Hoe hij de hyperlink ten volle benut in een werkomgeving van kenniswerkers, moet getoond, ervaren en geoefend worden. Luidop werken ("*working out loud*") is nog zo'n competentie die we mogelijk al te evident achten. Je doet kenniswerkers, die gewoon zijn om erg individueel te werken, plots al hun werk "in het openbaar" doen (weliswaar relatief openbaar: minstens openbaar voor het hele betrokken team, in het beste geval openbaar voor het hele bedrijf, én de betrokken klanten én de betrokken leveranciers). Dat vraagt een mentale shift bij mensen. Zo hebben wij een twintigtal nieuwe competenties geïdentificeerd, waarvan wij nu weten dat deze extra aandacht moeten krijgen in onze nieuwe manier van werken, anders wordt onze opzet geen succes.

Web 2.0-werken is in elk geval geen cultuur-onschuldige manier van werken. Omdat je het gedrag van kenniswerkers beïnvloedt en wijzigt, wijzig je ook het gedrag van teams en organisaties. Het Web 2.0-werken sluit verrassend sterk aan bij een totaal nieuw organisatieparadigma dat vandaag heel erg het Tayloriaanse organisatie-model uitdaagt. Het Web 2.0-werken maakt hoe dan ook je organisatie "platter", want mensen communiceren met elkaar, niet via elkaar. Het maakt het kenniswerk door de vele, transparante en snelle feedbackmogelijkheden ook veel sneller, en hyperdynamisch - daar kunnen bureaucratieën niet mee om. Dezelfde transparantie creëert betrokkenheid en respect: de chef gaat niet meer naar zijn baas pochen over *zijn* realisaties, neen, in de Web 2.0-omgeving zie je de veelheid van betrokkenen én hun daadwerkelijke input. Door de geconnecteerdheid van content en mensen, zie je ook veel makkelijker het grotere geheel, dus is Web 2.0-werken een vorm van zinvoller werken, van holistischer denken en doen. De emergente structuren geven ook een gevoel van vrijheid, noodzakelijk voor die creatieve kenniswerkers die willen proberen, en testen, en zien of het werkt. Maar de belangrijkste win die het Web 2.0-werken volgens mij brengt is de mogelijkheid tot authentieke en open dialoog - ook al is het web-gebaseerd en virtueel en niet

face-to-face... Het brengt wel degelijk mensen samen, en verrassend genoeg, door de openheid brengt het altijd méér mensen samen dan we oorspronkelijk hadden gedacht. En dat is voor een organisatie een supergevoel.

Er is mogelijk toch een minpunt in dit veranderingsproces. McAfee haalt een interessant aspect aan dat erg determinerend is voor bedrijfsculturen: de managers - "*old managers die slowly*"¹², zegt hij. Het is waar: we zitten met generaties managers die op Tayloriaanse wijze veel successen hebben geboekt, en dus niet geneigd zijn van deze methodieken af te stappen. Het is niet altijd een bewuste keuze - eigenlijk zijn ze vaak gewoonweg blind voor het nieuwe paradigma - ze begrijpen er de eerste letter niet van. "*Leading by example*" helpt ons in dit geval averechts: het werkt vaak tegen het 2.0-gebeuren. Dat is voer voor een ander artikel.

Processen: adaptive case management

De meest vergeten factor om 2.0 tot een succes te brengen (McAfee vergeet die ook), is volgens mij aandacht voor "het proces". Als ik gadesla wat bedrijven met de 2.0-tools doen (al of niet goed ondersteund door de nodige change management energie), dan is dat toch altijd vrij mager. Het is meestal een vorm van "interne communicatie": ergens een kanaal waar je berichten kan posten, die anderen al of niet toevallig kunnen opmerken, om er dan hun opmerkingen te posten. Of, er is een initiatief om werknemers te laten bloggen, of hun online intranetprofiel "rijker" aan te vullen zodat men makkelijker in huis aanwezige competenties kan ontdekken en vinden. Of ze doen medewerkers "tweeten" over hun werk dat ze momenteel aan het doen zijn. Of het gaat over de opzet van communities of practice. Of teams worden makkelijker ondersteund in teamcommunicatie rond 1 of ander opduikend probleem. Het zijn hoegenaamd allemaal winsten die je met 2.0 kan realiseren, maar het blijven nogal "flarden" die gebeuren in extra, parallelle kanalen en applicaties - waarmee ik uitmaak dat de 2.0-krachten niet geïntegreerd zijn geraakt IN het werk, maar enkel stukken ervan ondersteunen. We hebben nu vooral een praktijk van kenniswerk 0.7 met flarden 1.0 (beetje opzoeken op het web) en flarden 2.0 (beetje Yammeren). Kenniswerk moet helemaal 2.0 worden. Maar dan hebben we ook een goed proces nodig dat dat mogelijk maakt.

We hebben in onze bedrijven processen gedefinieerd om zaken te bestellen, om onze assets te beheren en te onderhouden, om (elkaar) jaarlijks te evalueren, om dienstverplaatsingen vergoed te krijgen. Die zaken zijn in functie van efficiëntie en

risicobeheer duidelijk gestructureerd geworden zodat iedereen weet/kan weten wie wat wanneer hoe doet om de zaken juist te doen. Business Process Management Softwares zijn een "poka yoke" middel dat vermijdt dat we fouten maken: het zijn harde structuren die totaal geen flexibiliteit toelaten, maar die bv. de bestelling wel tot een goed einde brengen.

Maar al bij al hebben we geen gepast proces dat ons reguliere kenniswerk het nodige respect betuigt, respect voor de dynamiek, het dialogische karakter, en de ambivalentie (proces/product) ervan (zie boven). We proberen wel processen en systemen op te zetten voor het beheer van de output van kenniswerk, documenten. En we probeerden dat ook te kunnen voor kennis als zodanig - kennisbeheersystemen, maar die bleven grotendeels een grote ontgoocheling. Rapporten uit BPM- en Business Intelligence (BI) systemen kunnen ons wel interessante input aanleveren, maar ze lossen geen problemen op. Het gaat steeds om transactionele processen waarbij we informatie gebruiken i.f.v. beslissingen - het gaat veel minder om het maken van nieuwe informatie zelf (het kenniswerkproces).

Kenniswerk (als werkproces) blijft een enorm grote vlek die onbeheerd blijft - de actie, zo menen we dan maar, blijkt zich vooral in de "black boxes" van individuen af te spelen (- "ssst, ze denken na"). Dan lijkt er voor bedrijven en managers niet veel over te blijven dan de best mogelijke context te creëren voor kenniswerkers zodat deze kenniswerk-individuen zich optimaal kunnen laten gaan. We ontwerpen aangename activity-based werkplaatsen (deze horen weliswaar onder de "tools", niet de processen), zodat er ruimte is voor concentratie, maar ook samenwerking, ontspanning en creatieve buien.

Procesmatig bestaan er toch pogingen om het harder te gaan spelen. We doen pogingen onze concentratie beter te beheren (lees: "19 tips voor een beter time management"; lees ook: Hoe multitasking onze hersenen pijnigt), we proberen doelstellingen te beheren (lees: Why so many projects fail), we proberen onze communicatie beter vorm te geven (lees: 148 tips voor beter e-mailen), en we proberen ook het teamproces rond probleemoplossingen te verbeteren (lees: 8 killer tips om beter te vergaderen; lees: Het slaagmodel voor team-brainstorming; lees: Visual management voor de nieuwe leider; en lees: Scrum voor beginners). Maar dat is allemaal fragmentarisch, vaak elk ook in één aparte app gepropt, naast elkaar, niet in elkaar. Het is alsof je voor een bestelcyclus in je bedrijf een software zou hebben om de

leverancier voor je product in te selecteren, een andere applicatie om eenheidsprijs van je product in te verwerken, en nog een andere applicatie om de hoeveelheid in te brengen.

De problematiek rond het vatten van generiek kenniswerk als proces, heeft te maken met de grote ongestructureerdheid van kenniswerk. Vandaar ook de versnipperde aandacht op deelaspecten ervan.

Deels uit de Business Process Management wereld groeide een nieuwe discipline, Adaptive Case Management (ACM), vanuit het besef dat heel wat kenniswerk helemaal niet voorspelbaar is, en dus niet in vaste processen gegoten kan worden. Adaptive Case Management (of dynamic/advanced/intelligent case management) gaat uit van "een case/een dossier", en daar gebeurt het. Praktisch: de identificatie van een "niet-routinematig of uniek probleem" in je bedrijf of team leidt in de praktijk tot het openen van "een case". Een nieuw probleem? Open een nieuw dossier. Cases zijn zo erg resultaatgefocus.

In de case komt alle informatie die van belang is voor het probleemoplossingsproces, en als dit in een 2.0-omgeving gebeurt, verloopt ook alle communicatie over het probleemoplossingsproces IN het dossier. In de 2.0-omgeving ontstaat er ook een groot bewustzijn omtrent het ordenen van informatie, nl. om via taal klaarheid in het denken te scheppen. Het verloop en de structuur van een case is daarom emergent: "actoren" lossen naar best vermoeden en vermogen deelproblemen op; ze communiceren met elkaar, en voegen nieuwe informatie toe. Die bepaalt de volgende stap. We brengen structuur in de te-weinig-structuur. Hier verkennen we de labiele grens tussen orde en chaos. De wisselwerking tussen denk- en communicatieproces, en de bestaande en gegenereerde deel-producten is de motor van de case. Een case is niet zoals een BPM-proces rond workflow gecentreerd, maar het is informatiegecentreerd. Het is duidelijk dat informatie beheren niet iets is wat je doet "naast het werk" (even mijn desktop opruimen en bestandjes klasseren), neen, kenniswerk betekent essentieel informatie beheren, dus je doet het IN het werk, maar je beseft het met moeite.

Na een tijd bevindt zich heel wat in de case - ruwe en verwerkte data uit BPM en BI-systemen, records uit ECM-omgevingen (Enterprise Content Management), flarden communicatie, taken, stukjes workflow, maar ook de 'deelnemers' zitten "in" de case... Daardoor wordt het een erg rijk contextueel gegeven, dat veel meer dan

BPM-omgevingen en andere, gewone informatie-omgevingen (bv. EIM - Enterprise Information Management, ECM) aanleiding geeft tot een holistische kijk op het probleemoplossingsproces. Door de verstrengeling van het oplossingsproces met de (deel)producten organiseer je ook een belangrijk metaniveau in het dossier zelf. Het is het proces van "connecting the dots", maar je brengt in de case eerst "the dots" samen. 2.0 staat dan eerder in voor het "connecting"-gedeelte. Dat denkproces wordt geëxpliciteerd in deze 2.0 omgeving, omdat alle actoren a.h.w. "luidop aan het werk zijn" in deze dynamische dossieromgeving. Dit levert een schat aan informatie op omdat "eindelijk" de kenniswerker massa's van zijn "tacit knowledge" kan delen. Daar droomden we al 40 jaar van.

Ik stel voor dat je de vorige alinea's nog minstens één keer herleest. Echter, nooit vergeten: "ACM is something you do, not something you buy" (Nathaniel Palmer¹³).

Besluit

Je merkt hoe verstrengeld het proces en de tools raken. Voor vele onderdelen van het dynamische dossierwerken kan ik niet zonder de 2.0-omgeving. En het succes van ACM in een 2.0-omgeving is zeker afhankelijk van de juiste competenties bij de actoren in de case.

2.0 kan heel wat winst realiseren binnen onze bedrijven, zeker voor de interne organisatie van "het werk". Zelfs met de fragmentarische inzet ervan in bv. enterprise social networks brengen

we de medewerkers dichter bij elkaar om sneller de zaken gedaan te krijgen. Echter, als je de stap durft zetten om AL het kenniswerk binnen je organisatie op 2.0-platformen vorm te geven, dan realiseer je een exponentiële winst, en die jackpot wil je toch niet laten liggen? Het betekent écht anders gaan team-werken om organisatieproblemen te lijf te gaan. Het 2.0-kader is een authentiek kader voor kenniswerkers, maar emergent dat het is, verrast het jouw organisatie elke dag met nieuwe, onverwachte, gerealiseerde "adjacent possible" stappen.

Niet voor niets is (tools, people, processes) de heilige drie-eenheid van mijn kenniswerkomgeving. Dit trio levert mij impact én duurzaamheid op. Ik eer hen elke dag. Zij vormen samen het succes van deze praktijk, die voor mij heel wat meer is dan informatiebeheer. Het is de basis voor een nieuw soort organisatie waar meer en meer de kenniswerker en het kenniswerk op authentieke wijze centraal staat. Het Web 2.0 is de meest natuurlijke partner in deze dynamische informatie-omgeving. En we zien vol verwachting uit naar een rijper Web 3.0, want we geloven zeker dat dit de dynamische dossiervorming nog zal verrijken.

Filip Callewaert

Havenbedrijf Antwerpen
Entrepotkaai 1
2000 Antwerpen
filipcallewaert2@gmail.com

Februari 2016

Noten

De case van het Havenbedrijf over ACM was in 2015 gepubliceerd in het boek door Palmer, Nathaniel. Thriving on Adaptability. Best Practices for Knowledge Workers. Future Strategies Inc., 2015. ISBN 978-0986321405. Wie verder in dialoog wil treden over de inhoud van dit artikel, wordt van harte uitgenodigd dat te doen: contacteer de auteur via filipcallewaert2@gmail.com

1. Drucker, Peter F. Knowledge-Worker Productivity: The Biggest Challenge. California Management Review, winter 1999, vol. 41, n° 2, p. 79-94.
2. Scrum is een flexibele manier om (software)producten te maken. Er wordt gewerkt in multidisciplinaire teams die in korte sprints, met een vaste lengte van 1 tot 4 weken, werkende (software) producten opleveren». Scrum (softwareontwikkelmethode). Wikipedia [on line], 27 mai 2016 (geraadpleegd op 10 juni 2016). <[https://nl.wikipedia.org/wiki/Scrum_\(softwareontwikkelmethode\)](https://nl.wikipedia.org/wiki/Scrum_(softwareontwikkelmethode))>.
3. Cfr noot 2
4. Kilpi, Esko. Conversations and narratives are the new documents. WTF? What's The Future of Work? [on line], 26 november 2015 (geraadpleegd op 24 februari 2016). <<https://medium.com/the-wtf-economy/conversations-and-narratives-are-the-new-documents-256ebcd3d37f#.drit537g4>>.
5. Gibson, William. The Science in Science Fiction. Talk of the Nation, NPR [on line], 30 november 1999 (geraadpleegd op 24 februari 2016). <<http://www.npr.org/templates/story/story.php?storyId=1067220>>.
6. IBM Watson: The platform for cognitive business. IBM [on line]. <<http://www.ibm.com/watson/>> (geraadpleegd op 24 februari 2016).
7. Grossman, Lev. You — Yes, You — Are TIME's Person of the Year, Time Magazine [on line], 25 december 2006

(geraadpleegd op 24 februari 2016). <<http://content.time.com/time/magazine/article/0,9171,1570810,00.html>>.

8. Cfr noot 5
9. McAfee, Andrew. Enterprise 2.0 - New Collaborative Tools for Your Organization's Toughest Challenges. HBR, 2009. ISBN 978-1422125878
10. McAfee, Andrew. Enterprise 2.0, Finally?, Andrew McAfee [on line], november 2014 (geraadpleegd op 24 februari 2016). <<http://andrewmcafee.org/2014/11/andrew-mcafee-enterprise-2-0-finally/>>.
11. Cfr noot 11
12. Palmer, Nathaniel. BPM and Case Management Global Summit Hosts Global Awards for Case Management Excellence, BPM [on line], 11 februari 2014 (geraadpleegd op 24 februari 2016). <<http://bpm.com/bpm-today/headlines/598-bpm-and-case-management-global-summit-hosts-global-awards-for-case-management-excellence>>.

INFORMATIE OPSPOREN, UITGAANDE VAN EEN BEELD

Paul NIEUWENHUYSEN

Hoogleraar, Universiteitsbibliotheek, Vrije Universiteit Brussel (VUB)

■ Deze bijdrage geeft een kort overzicht van een lopend onderzoek en evaluatie van de methode om informatie op te sporen waarbij de gebruiker start met een beeld. Dit is een relatief nieuwe techniek waarbij een zoekvraag niet bestaat uit woorden maar uit een beeld dat is gecodeerd in een computerbestand. Verschillende tests zijn uitgevoerd. De resultaten tonen aan dat zoeken met een beeld is geëvolueerd tot een krachtige, bijkomende methode om tegemoet te komen aan informatienoden die moeilijk aan te pakken zijn met andere, meer klassieke methodes. Verschillende types van toepassingen worden beschreven.

■ Cette contribution fournit un bref aperçu d'une recherche en cours et de l'évaluation d'une méthode de recherche d'information où l'utilisateur entame sa recherche au départ d'une image. Il s'agit d'une technique relativement nouvelle où la requête ne se compose pas de mots, mais d'une image codée dans un fichier informatique. Différents tests sont effectués. Les résultats démontrent que la recherche sur base d'image est devenue une méthode complémentaire puissante, permettant de rencontrer des besoins en information difficiles à satisfaire à l'aide d'autres méthodes plus classiques. Différents types d'application sont décrits.

Dit artikel is een verkorte, vertaalde, geactualiseerde versie van een hoofdstuk gepubliceerd in een gedrukt boek¹.

Inleiding / Context / Achtergrond

Twee evoluties vormen de context van dit artikel:

Toenemend aantal digitale beelden

Het aantal beschikbare beelden neemt snel toe, samen met de afname van de kosten en technische problemen, die verband houden met

- digitalisering van hard-copy beelden,
- digitale camera's en fotografie,
- publicatie/verspreiding van beelden via het web en zelfs sociale interactie in verband met beelden.

Beelden zoeken en zoeken met beelden

Om aan verschillende informatiebehoeften te voldoen, zijn verschillende zoekmethoden op het web beschikbaar gekomen en worden deze voortdurend verbeterd. Verscheidene van deze methoden zijn niet beperkt tot tekst, maar afbeeldingen zijn er op één of andere manier bij betrokken. De overkoepelende term "image searching" wordt vaak gebruikt, hoewel de methoden sterk uiteenlopen qua kenmerken, doelen, kracht en beperkingen.

De toepassing van "image searching" is soms aantrekkelijk, efficiënt, productief en interessant, om verschillende redenen:

- "Image searching" laat toe relevante afbeeldingen te vinden, als een specifiek beeld nodig is.
- De zoekdienst haalt beelden op en toont die als miniatuurtjes, zogenaamde "thumbnail images", zodat de beoordeling van elk resultaat in een lijst met resultaten minder tijd kost dan met de meer klassieke, normale weergave van de resultaten die hoofdzakelijk bestaan uit tekstfragmenten.
- Wanneer u een beeld vindt dat is gerelateerd aan uw vraag en dat ook relevant is, dan biedt het zoekstelsel normaliter ook een link naar de webpagina plus de site die dat beeld bevat; dit alles kan relevant zijn in de context van uw informatienood.

We kunnen zoeken naar afbeeldingen via een geschikte zoekmachine door een vraag te zenden die uit tekst bestaat, net als in meer klassiek zoeken naar documenten. Om op deze manier te zoeken door het web zijn verschillende systemen beschikbaar, zelfs gratis; voorbeelden zijn de specifieke zoekdiensten die deel uitmaken van *Bing*, *Yahoo* en *Google*².

Deze systemen zijn erg populair geworden; bijvoorbeeld, om afbeeldingen te vinden steunen gebruikers in een universiteit sterk op *Google Image Search*³ en in China op de Chinese zoekmachine *Baidu*, evenals op *Google*⁴. Deze zoeksystemen werken voornamelijk op basis:

- van de bestandsnaam en andere metadata toegevoegd aan het beeldbestand,
- van de tekst rond de afbeelding en

- van de tekst in hyperlinks naar de afbeelding, die zich bevinden in hetzelfde document, in andere documenten van de website of in een andere website.

Dergelijke zoeksystemen hebben last van de klassieke problemen bij het opsporen van informatie, zoals synoniemie en polysemie. Erger nog:

- De beschrijvingen van het beeld (metadata) ontbreken of zijn schaars.
- De verbanden tussen het beeld en de beschrijvende woorden en termen zijn meestal vaag.
- Deze beschrijvingen worden bij de creatie van een document meestal niet gecontroleerd en beperkt met een thesaurus of ontologie.

Later dan de techniek van zoeken op basis van een tekstuele vraag is een techniek ontwikkeld op basis van de inhoud van een beeld, nl. de zogenaamde "content-based image retrieval" of afgekort "CBIR". In deze benadering worden de beelden geïndexeerd door een computersysteem op basis van de visuele inhoud, zoals kleuren, texturen en vormen, met de bedoeling om zo een aantal bijkomende nuttige mogelijkheden voor het zoeken aan te bieden.

Talrijke publicaties handelen over deze methode^{6 7 8 9 10 11 12 13 14} en een kort overzicht is gratis beschikbaar voor iedereen op *Wikipedia*¹⁵. Sommige van deze systemen laten gebruikers toe om het zoeken te beginnen met een vraag die bestaat uit een beeldbestand dat naar de zoekmachine wordt gezonden. Algemene namen voor deze methode zijn in het Engels:

- search(ing) by example
- reverse image lookup = RIL
- reverse visual search(ing)
- reverse image search(ing)
- backwards image search(ing)
- search(ing) by image
- inside search(ing)

Het populaire bedrijf *Google* biedt o.a. een zoekdienst van dit type en noemt dit "reverse image search". De auteur van dit artikel vindt de term "search by image" het meest eenvoudig en duidelijk.

Doel van dit onderzoek

Het kader van deze bijdrage is een lopend onderzoek en beoordeling van de diensten die ons in staat stellen om te zoeken naar informatie via het web met methodes waarbij beelden betrokken zijn. Sommige systemen zijn voor iedereen gratis beschikbaar. Meer in het bijzonder: deze bijdrage geeft een overzicht van onze beoordeling van de

systemen die ons toelaten om te zoeken, niet door het indienen van een vraag in de vorm van tekst, maar door de meer recente en minder bekende methode waarbij zelfs de vraag bestaat uit een afbeelding, uiteraard gecodeerd in een computerbestand. Een dergelijke "search by image" kan zoekresultaten leveren van uiteenlopende types. Dit wordt geïllustreerd in Fig. 1 en wordt verder in meer detail beschreven.

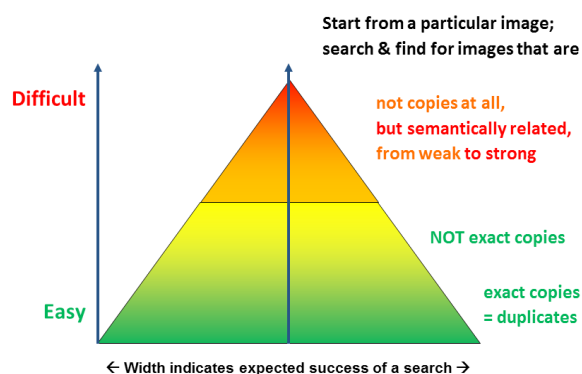


Fig. 1: "Search by image" kan verschillende soorten zoekresultaten opleveren.

Beelden die ten minste enkele elementen van het bronbeeld bevatten. Hierin kunnen wij weer diverse types onderscheiden, als volgt:

- exacte kopieën of duplicaten van het bronbeeld;
- bijna exacte kopieën/duplicaten, maar toch verschillend van het bronbeeld qua aantal beeldpunten (pixels) en/of kleuren en/of mate van datacompressie die oorzaak is van verlies van gegevens uit het beeld;
- andere beelden die nog slechts bestaan uit een deel van het bronbeeld;
- duidelijk andere beelden, maar die toch wel gebaseerd zijn op het bronbeeld, in de zin dat ze het bronbeeld omvatten, maar ook andere afbeeldingen en/of teksten, alle samengesteld, gemengd tot één resulterend nieuw beeld;
- bewerkte/gewijzigde versies van het bronbeeld;
- uiteraard ook combinaties van bovenstaande variaties.

De grenzen tussen de hierboven genoemde types zijn niet scherp; de types overlappen. Daarom beschouwen we deze samen, hier en ook in de Fig. 1, als een categorie.

Beelden die geen elementen van het bronbeeld bevatten, maar die een verband hebben met inhoud/betekenis/onderwerp/semantiek van het bronbeeld; in andere woorden, beelden die semantisch gerelateerd

zijn aan het bronbeeld. Deze relatie kan variëren in kracht:

- De semantische relatie kan zwak of oppervlakkig of breed zijn. Bijvoorbeeld elke bepaalde boomsoort behoort tot de planten, of bijvoorbeeld een bepaald, specifiek, individueel masker dat behoort tot het type gemaakt en gebruikt door een bepaalde etnische groep in Afrika in de 20ste eeuw behoort uiteraard tot de grote groep van voorwerpen die men "maskers" noemt.
- De relatie kan zeer sterk, scherp, specifiek zijn.

Het hierboven geschetste spectrum van mogelijke overeenkomsten tussen beelden bereidt ons voor op de volgende probleemstellingen en onderzoeksvragen:

- Welke systemen/diensten bieden gratis "search by image" via het web?
- Waarin verschillen deze diensten opmerkelijk voor een gebruiker in de praktijk?
- Welke van die zoekdiensten presteert het best?
- In welke mate is een zoekdienst in staat om exacte duplicaten te vinden die aanwezig zijn op het web?
- Hoe effectief kan het zoekstelsel beelden vinden op het web, die geen exacte duplicaten/kopieën zijn, maar die toch wel elementen gemeen hebben met een bepaald bronbeeld?
- We kunnen "search by image" gebruiken om beelden te vinden die visueel gelijkaardig zijn aan het bronbeeld, in de zin dat ze dezelfde kleuren, texturen of vormen bevatten; maar kan deze methode ook beelden leveren die 'semantisch' gelijkaardig zijn aan het bronbeeld?

Methoden die worden gebruikt in dit onderzoek

In elke "search by image" op het web wordt een bronbeeld van de computerschijf gezonden naar de zoekdienst. De bronbeelden die worden gebruikt in de tests zijn vooral foto's die gemaakt zijn en gepubliceerd zijn door de auteur van dit artikel

- in jpeg/jpg-formaat, omdat dit nu een klassiek, populair, goed bekend, standaardformaat is,
- op de centrale web server computer van een universiteit,
- al sinds een aantal jaren,
- als afbeeldingen opgenomen in een web pagina die gewoon bestaat uit HTML-opgemaakte tekst op de klassieke, normale manier, met behulp van hyperlinks naar de beeldbestanden die op de pagina verschijnen

en die zich bevinden op dezelfde web server; elk bronbeeld bevindt zich dus niet op het web op een manier die het opsporen moeilijker maakt, namelijk b.v. als deel van een container-documentbestand in een van de populaire bestandsformaten zoals een versie van Word of PDF formaten.

Om te kunnen werken op een gebied waarmee de auteur vertrouwd is en om een soort van opsporen van informatie te onderzoeken wat nuttig kan zijn in digitale geesteswetenschappen ("digital humanities"), gebruiken we als bronbeelden voornamelijk eenvoudige, documentaire foto's van statische, sculpturale, traditionele, etnische (kunst)voorwerpen.

We vonden dat "search by image" bij Google niet werkt met grote bronbeelden. Daarom gebruikten we in sommige gevallen een verkleinde versie.

"Search by image" is relatief nieuw en er is geen standaard beschikbaar als beeldcollectie en voor de methode om de prestaties van bepaalde systemen te meten en te vergelijken^{16 17}. Voor elke vraag inspecteren we de zoekresultaten geleverd door elke zoekdienst en noteren we onze waarnemingen in een tabel. De prestaties van een zoekstelsel worden vooral gemeten/geëvalueerd door te kijken naar de hoogst gerangschikte resultaten en door het tellen van het aantal relevante resultaten. Dit komt overeen met het meten van de precisie; dat is een methode die op grote schaal is aanvaard als praktisch en nuttig¹⁸ en die nu de meest populaire keuze vormt, volgens een recent overzicht van interactief zoeken met beelden¹⁹. Daarom beschouwen wij dit als een praktische, realistische en redelijke aanpak, ook al houden we in het achterhoofd dat het concept van "relevantie" ingewikkeld is²⁰.

Bevindingen / Resultaten

Beschikbare zoekdiensten voor "search by image"

Verschillende diensten zijn gratis beschikbaar om te zoeken in het web met "search by image"²¹.

Het stelsel beschikbaar bij *Tineye* is beschikbaar sinds 2008 en is relatief vaak vernoemd in publicaties. "Unfortunately, very little is known about the actual systems"²².

Meer recent, sinds 2011, wordt "search by image" ook aangeboden door het grote en populaire bedrijf Google; het is kort en oppervlakkig beschreven in een gids voor gebruikers²³. Dit stelsel

is geïntegreerd met de algoritmen en de enorme database van webdocumenten, die al ontwikkeld zijn voor de meer klassieke *Google* zoekmachines.

Een paar add-ons/extensies voor de internet-browser software Firefox zijn gratis beschikbaar, waarmee de gebruiker een afbeelding op een webpagina kan aangeven in de browser, om dan snel en gemakkelijk deze afbeelding als een zoekvraag te zenden aan enkele diensten die "search by image" mogelijk maken. Zulk een add-on voor *Google* "search by image" wordt ook beschikbaar gesteld door *Google*.

Verschillen tussen de diensten voor "search by image"

Na de uitrol van *Google* "search by image", zijn een paar snelle en zeer beperkte testen online gemeld:

- Een zeer kort verslag vermeldt een proef met 10 zoekopdrachten, waaruit bleek in alle gevallen dat de oudere *TinEye* niet zo goed presteerde als de meer recente dienst aangeboden door *Google*²⁴.
- Enkele zoekopdrachten gaven bij *TinEye* en *Google* gelijkaardige resultaten; in één geval, wanneer het bronbeeld een foto was van een stadsgezicht, mislukte *TinEye*, terwijl *Google* andere, gelijkaardige beelden leverde²⁵.

Recenter hebben we meer volwassen versies van beide diensten vergeleken, in meer detail en op een meer systematische manier²⁶. We hebben 10 bronbeelden gebruikt waarvan een duplicaat, een exacte kopie aanwezig is op het web; elk is als een zoekvraag gezonden aan zowel *TinEye* als *Google*. Van de 10 corresponderende duplicaatbeelden op het web onthulde *TinEye* er slechts 3, en *Google* wel 7. Verder hebben we onderzocht hoe goed de zoekdiensten beelden leveren, die zijn afgeleid van de 10 originele foto's en die op een andere website zijn gepubliceerd: *TinEye* onthulde geen enkele afgeleide afbeelding, terwijl *Google* ten minste één afgeleid beeld onthulde bij 7 bronbeelden. De verschillen tussen de zoekdiensten zijn dus aanzienlijk.

Prestaties van de zoekdiensten

De bovengenoemde tests tonen dat *Google* beter presteert dan *TinEye*. Daarom werden de volgende tests uitgevoerd met behulp van *Google*. Enkele variaties van "search by image" zijn beschikbaar bij *Google* "To search using an image, go to images.google.com and just put your picture in the search box. There are many ways to do this. You can click the camera icon in the search box and upload a photo from your computer or paste the URL of an

image from the web. You can also drag and drop pictures from webpages or your computer into the search box. To search images on the web even faster with just one click, you can download the Chrome or Firefox extensions."²⁷

Vinden van kopieën van een beeld

In elke test gebruiken we een bronbeeld waarvan we weten dat een kopie aanwezig is op een statische, stabiele, openbare website, al sinds vele maanden of jaren. Bij 25 vragen onthulde *Google* "search by image" 15 van die duplicaten op het web²⁸. Deze techniek blijkt dus succesvol, effectief, efficiënt, maar niet perfect, niet volledig.

Vinden van gewijzigde versies van het bronbeeld

We hebben ook onderzocht hoe effectief het zoekstelsel beelden op het web kan vinden die geen duplicaten (exacte kopieën) zijn van een bepaald bronbeeld, maar die toch gemeenschappelijke elementen hebben met het bronbeeld²⁹.

Hoger zijn bevindingen vermeld die we hier kort moeten herhalen in dit verband: de *TinEye* zoekdienst onthulde geen beelden die afgeleid zijn uit onze oorspronkelijke bronbeelden, terwijl *Google* wel een groot aantal beelden onthulde, die elementen gemeen hebben met onze originele bronbeelden.

In een verdere reeks van tests werden 16 vragen uitgevoerd met *Google*, elk uiteraard met één bronbeeld. Hier richten we onze aandacht dus niet meer op het opsporen van het duplicaat van het bronbeeld wat bestaat op de website van de auteur. In 12 van deze testen onthulde *Google* ten minste één ander beeld op een andere website, wat gebaseerd is op het bronbeeld. In de volgende paragrafen behandelen we meer specifieke vragen:

- Hoe effectief kan het zoekstelsel beelden vinden, die door een manipulatie/wijziging/verandering van het bronbeeld een verschil in resolutie of in het aantal beeldelementen vertonen? Verschillende tests onthulden een beeld afgeleid van het bronbeeld, beschikbaar op verschillende websites, zelfs wanneer dat beeld geen exacte kopie is, met een verschillend aantal beeldpunten (pixels) of met gewijzigde kleuren of met een combinatie van beide.
- Hoe effectief kan het stelsel voor "search by image" andere afbeeldingen op het web onthullen, die een versie van het bronbeeld bevatten, niet alleen dus, maar samen met

andere toegevoegde beelden en/of letters en cijfers, die niet voorkomen in het bronbeeld; met andere woorden, in hoeverre kan "search by image" beelden vinden die het bronbeeld bevatten als een onderdeel, als slechts één van de meerdere beeldcomponenten? De zoekdienst bleek zelfs bronbeelden terug te vinden die functioneerden als een onderdeel in nieuwere beelden, naast andere afbeeldingen of letters en cijfers. De zoekdienst onthulde ook afbeeldingen die onder andere een fragment van het bronbeeld bevatten, zelfs als dat fragment gewijzigd was, b.v. waarbij de originele kleuren omgezet waren in grijstinten.

Hieronder volgen een aantal concrete voorbeelden van dergelijke onthullingen met "search by image"; deze zijn opmerkelijk in de zin dat ze de kracht van de techniek tonen; daarenboven zijn het voorbeelden van het kopiëren, veranderen en gebruiken van beelden, zonder medeweten van de auteur/maker³⁰:

- Een foto van een traditioneel Kuba Bwom masker van Congo in Afrika is gewijzigd, is geplaatst op een achtergrond die een oude kaart van Afrika toont en het resulterende beeld wordt gebruikt als onderdeel van een jaarkalender. Dit wordt geïllustreerd in Fig. 2.
- Een foto in kleuren van een masker uit Afrika is gebruikt in een poster zonder kleuren maar slechts in grijsniveaus.
- Een foto van een kleine, houten sculptuur uit Afrika is gepubliceerd op een andere website en blijkbaar ook op de omslag van een boek; niet het geheel is overgenomen, maar slechts een fragment; daarenboven is tekst toegevoegd.
- Een foto van een Afrikaanse sculptuur is gebruikt als beeld op de cover van gepubliceerde hedendaagse muziek; als we die cover nader bekijken, dan blijkt een tweede foto van dezelfde originele website ook gebruikt te zijn in hetzelfde nieuwe beeld.



Fig. 2: Illustratie van Google "search by image" startend met het bronbeeld (links) toont deze zoekdienst o.a. dat dit beeld is gebruikt als een component in een publicatie op het web (rechts).

- Een webpagina met foto's is geheel vertaald en is op een andere website gepubliceerd. Een meer klassieke zoekvraag met tekst zou dit niet eenvoudig onthullen, want de teksten zijn vertaald.
- Een foto van de berg Kilimanjaro in Afrika, gemaakt vanuit een vliegtuig, is gekopieerd en gepubliceerd op de populaire, grote Flickr foto website; daar is de foto al bekeken door duizenden gebruikers en heeft ze talrijke reacties ontvangen.

Vinden van semantisch gelijkaardige beelden

Stel dat we meer informatie willen vinden over een object wat we dus al summier kennen of waarvan we ten minste toch al een beeld hebben, maar waarover we wel meer willen te weten komen. Dan zou een nuttige toepassing van "search by image" erin bestaan dat we starten met een foto of zelfs met een meer grafische afbeelding van dat gekend object en dat de zoekactie andere beelden zou leveren van objecten en teksten over dat of andere objecten die niet alleen visueel lijken op het bronbeeld, maar die zelfs inhoudelijk/semantisch gelijkaardig zijn aan het object wat ons interesseert. Dit geldt zeker in het bijzonder wanneer de gebruiker de algemene aard van het object in het beeld wel kent, maar niet op de hoogte is van een meer specifieke benaming, zodat de gebruiker de specifieke behoefte aan informatie niet kan uitdrukken door het formuleren van een specifieke, gerichte tekstuele vraag in een meer klassieke zoekmachine.

We hebben praktische testen uitgevoerd door het zenden van beeldbestanden naar Google "search by image". Hier moeten we ons realiseren dat het zoekstelsel de beeldinhoud uiteraard zal analyseren, maar dat het stelsel ook rekening kan houden met de naam van het beeldbestand. Deze bestandsnaam kan aanvullende, nuttige informatie geven aan de zoekmachine, als deze naam gerelateerd is aan de betekenis, de inhoud van het bronbeeld. In dat geval zou de interpretatie van de zoekresultaten en de evaluatie van de kracht van de techniek moeilijker zijn. Daarom hebben we bewust vermeden dat elke zoekactie die bestaat uit een beeldbestand niet alleen via het beeld maar ook via de bestandsnaam aan de zoekmachine een idee zou geven over de betekenis van de inhoud/onderwerp van het beeld. Meer concreet: de naam van elk bestand, die significant/zinvol was, is vervangen door een naam die neutraal, niet zinvol is, zoals "x.jpg".

Een reeks van 20 zoekacties diende als eerste test in 2013. De resultaten waren beelden die

gelijkaardige kleuren, vormen en compositie bevatten, maar geen beelden van andere objecten die semantisch verband houden met de bronbeelden. De resultaten toonden zelfs geen semantisch gerelateerde beelden in die gevallen waarbij veel inhoudelijk verwante afbeeldingen toch wel toegankelijk zijn op het web, zoals we weten uit eerdere, andere, verschillende manieren om het web te verkennen³¹.

Deze teleurstellende, onproductieve resultaten zijn niet verrassend, want het ontwikkelen van een effectief systeem voor een algemene semantische "search by image" doorheen een grote en diverse, heterogene verzameling van beelden is heel uitdagend, zo niet onmogelijk. Het is immers moeilijk of onmogelijk voor een computersysteem een verband te ontdekken tussen

- enerzijds één of meerdere van de vele mogelijke inhoudelijke componenten of concepten die kunnen worden gezien in het bronbeeld door de gebruiker van het zoekstelsel, met
- anderzijds de meer visuele eigenschappen van het bronbeeld en van miljarden andere beelden die ontleed werden en opgeslagen werden door het computersysteem, zoals de kwantificeerbare eigenschappen van kleur, textuur, vorm en de ruimtelijke verdeling of regio's, die automatisch kunnen worden gedetecteerd in de beperkte verzameling van beeldelementen waaruit elk digitale beeld is opgebouwd en die dan worden geïndexeerd in de database van de zoekmachine.

In het algemeen zijn de verbanden tussen enerzijds die concepten en anderzijds die beeldeigenschappen zwak of zelfs afwezig^{32 33 34}. Dit komt tot uiting in de slechte prestaties van de zoeksystemen, met een lage recall en precisie qua semantische zoekresultaten; dit geldt met name in het geval van grote verzamelingen afbeeldingen met een diverse inhoud.

Met andere woorden:

- Enerzijds: beelden die gelijkaardig zijn qua semantiek/inhoud/betekenis kunnen toch erg uiteenlopen qua kleur, vorm, textuur... Dit veroorzaakt een lage "recall" bij de zoekresultaten.
- Anderzijds: beelden die erg uiteenlopen qua semantiek/inhoud/betekenis, kunnen wel visueel gelijkaardig zijn. Dit veroorzaakt een lage precisie bij de zoekresultaten.

Dit probleem is een vorm van het meer algemene probleem bij opsporen van informatie, wat bekend staat als de "semantic gap"^{35 36 37 38}.

Meer recent, in 2014, hebben we meer testen uitgevoerd, waaronder een herhaling van de hoger vermelde. Onverwacht leverden een aantal testgevallen toen wel meer semantisch relevante zoekresultaten. Sommige resultaten gaven een correcte beschrijving van het bronbeeld en in sommige gevallen was die beschrijving niet slechts correct hoewel oppervlakkig, maar zelfs correct en zeer specifiek; dan leverden de zoekresultaten ook semantisch verwante beelden³⁹. Dit geeft aan dat de prestaties van "search by image" beter worden qua vinden van beelden die semantisch gelijkaardig zijn aan het bronbeeld.

De gevallen die dergelijke vruchtbare resultaten leveren starten hoofdzakelijk met een bronbeeld wat ook aanwezig is op internet/web in de vorm van meerdere kopieën, met sommige zelfs in een context die significante tekst bevat⁴⁰.

Dit klopt met:

- de opmerkingen van de producent van het systeem, Google dus, zoals geciteerd⁴¹: *"The feature works best for images of things that are quite well documented, such as often-photographed tourist attractions or images that are seen frequently online."*,
- de constatering dat zoekopdrachten elk met als bronbeeld een beroemde bezienswaardigheid of oriëntatiepunt ook een goed resultaat leverden, d.w.z. met een correcte beschrijving en met andere foto's van dezelfde oriëntatiepunten⁴².

Dit alles klopt ook met het feit dat het Google zoekstelsel niet louter en alleen werkt met beelden. Inderdaad Google "search by image" exploiteert niet alleen een database die puur bestaat uit afbeeldingen, maar kan ook gebruik maken van de teksten in verband met de afbeeldingen op webpagina's en zelfs in links naar de beelden.

*"Search by Image returns the best results for images that have related content already on the web, so you're more likely to get relevant results for distinctive landmarks or paintings than you will for more unique photos like your toddler's latest finger painting... The technology behind Search by Image analyses your image to find its most distinctive points, lines and textures and creates a mathematical model. We match that model against billions of images in our index, and page analysis helps us derive a best guess text description of your image."*⁴³.

*"Search by image looks for similar content on the web, so unique or never-before-seen images won't work well"*⁴⁴.

Na het zoekproces biedt de zoekdienst dan idealiter:

- een correcte en specifieke beschrijving van het bronbeeld,
- semantisch gelijkaardige beelden en
- de webpagina waarin deze optreden.

Deze complexe procedure om informatie op te sporen is een recente en geslaagde demonstratie van de algemene opvatting dat het benutten van zowel tekst als afbeeldingen in een enkele, hybride zoekactie vruchtbaarder kan zijn dan het benutten van slechts één soort zoekproces^{45 46 47 48 49 50 51 52}. *"...a text retrieval and an image retrieval system can simply complement each other."*⁵³.

Toepassingen

Mogelijke toepassingen van "search by image" worden samengevat in het volgende:

Vinden van kopieën van uw beeld

U kunt "search by image" toepassen om te onderzoeken of een bepaald beeld wat u hebt gemaakt ook gebruikt wordt op een andere website. Zelfs gewijzigde versies kunnen worden getraceerd. Waarschijnlijk is de kopie gepubliceerd zonder toestemming en zelfs zonder medeweten van de auteur of van de oorspronkelijke uitgever. Dit kan interessant zijn op verschillende manieren:

- Inbreuken op het auteursrecht kunnen worden ontdekt. Dit is heel belangrijk voor professionele fotografen.
- Iets positiever gezien en geformuleerd: Curators of eigenaars van een verzameling objecten kunnen een indruk krijgen van de impact en het hergebruik op het web van foto's van de fysieke objecten in hun collectie, en dit op wereldwijde schaal. Fotografen of kunstenaars kunnen de impact en hergebruik op het web evalueren van beelden die ze hebben gecreëerd. Nota bene: uw bronbeeld kan de basis vormen van een ander beeld op het web, maar daar stopt het niet, want dat digitaal, virtueel beeld op web kan ook gebruikt zijn in de meer tastbare realiteit, b.v. als illustratie die deel uitmaakt van een gedrukte publicatie; hoger hebben we daarvan enkele concrete voorbeelden getoond.

Vinden van andere versies van een interessant beeld

U kunt starten met een bronbeeld wat u kent en interessant vindt, maar wat u niet zelf gemaakt hebt en wat misschien niet de originele versie

is. Dan kan "search by image" nuttig zijn voor het opsporen van:

- een meer geschikte versie van dat beeld; bijvoorbeeld een versie die dichter staat bij het oorspronkelijke beeld, met een hogere resolutie of kwaliteit of integriteit,
- de maker/auteur of de persoon die auteursrecht kan claimen voor dat beeld,
- een kopie of een andere versie van dat beeld plus ook de locatie op een website die u dan kan voorzien van meer informatie over de afbeelding.

Vinden van visueel gelijkaardige beelden

Met "search by image" kunt u visueel vergelijkbare beelden vinden (qua kleuren, vormen, texturen...). Dit kan nuttig zijn, zelfs al zijn deze vergelijkbare beelden in de meeste gevallen semantisch heel verschillend van het bronbeeld.

Vinden van semantisch gelijkaardige beelden

Het volgende is waarschijnlijk meer belangrijk en interessant: als er van uw bronbeeld ook kopieën voorkomen op het web, dan kan "search by image" zelfs resultaten leveren waaronder:

- rechtstreeks een juiste en informatieve beschrijving in woorden van dat bronbeeld,
- evenals andere beelden die semantisch gelijkaardig zijn,
- plus links naar webpagina's die meer informatie kunnen verschaffen.

Discussie/Conclusies/Aanbevelingen

"Search by image" aangeboden door Google is niet star en stabiel, maar evolueert, zoals hoger weergegeven in dit artikel. Ook het publiek forum⁵⁴ aangeboden door Google over hun zoekdiensten bevat meerdere berichten waaruit blijkt dat de dienst voor "search by image" zich soms vreemd gedraagt. Dit bemoeilijkt de reproductie van zoekresultaten en tests van de prestaties.

Om het web doorzoekbaar te maken, werden zoekmachines ontwikkeld die webpagina's oogsten (d.w.z. vinden en kopiëren naar de maker van de zoekmachine) en die deze gegevens gebruiken om hun doorzoekbare databases te maken. Deze webpagina's bestaan uit html (tekst) bestanden en stand-alone image-bestanden in verband met deze html-bestanden, plus tal van andere soorten bestanden. Op zijn minst sinds 2015 haalt Google ook zelfs beelden uit geoogste bestanden in PDF formaat, zodat ook deze kunnen worden geïndexeerd en gebruikt in de zoekdiensten die worden aangeboden door Google⁵⁵. Dit heeft de dekking

van "search by image" dus uitgebreid en dit soort zoekacties dus nog krachtiger en aantrekkelijker gemaakt.

Uit de hierboven beschreven bevindingen concluderen we dat "search by image" evolueert naar een krachtige, extra methode om informatie op te sporen, gebaseerd op het benutten van het snel toenemende aantal beelden op het web plus de bijbehorende teksten.

Een aanbeveling: *"...libraries should consider providing additional information literacy courses in the areas of image information seeking and visual literacy."*⁵⁶. Deze aanbeveling werd opgesteld na een onderzoek van de manier waarop studenten zoeken naar beelden, vooral met de

meer klassieke methode, nl. door een zoekvraag te formuleren die bestaat uit tekst. De meer recente "search by image" vormt nu duidelijk een bijkomende, nuttige methode; die verdient het dus om gedemonstreerd en uitgelegd te worden aan potentiële gebruikers.

Paul Nieuwenhuysen

Universiteitsbibliotheek
Vrije Universiteit Brussel
Pleinlaan 2
1050 Brussel
paul.nieuwenhuysen@vub.ac.be
<http://www.vub.ac.be/BIBLIO/>

Februari 2016

Noten

1. Nieuwenhuysen, Paul. Search by image through the Internet: an additional method to find information. In Kataria, Sanjay et al (ed.) Transforming Libraries and Librarianship. KBD Publication, 2015, p. 179-194.
2. Nieuwenhuysen, Paul. WWW image searching delivers high precision and no misinformation: reality or ideal? Issues in Informing Science and Information Technology [on line], 2010 (geraadpleegd op 24 februari 2016), vol. 7, p. 109-131. <<http://iisit.org/Vol7/IISITv7p109-131Nieuwenhuysen770.pdf>>.
3. Kandiuk, Mary; Lupton, Aaron, Digital images in teaching and learning at York University: Are the libraries meeting the needs of faculty members in fine arts? Evidence Based Library and Information Practice [on line], 2012 (geraadpleegd op 9 augustus 2016), vol. 7.2, p. 20-48. <<https://ejournals.library.ualberta.ca/index.php/EBLIP/article/viewFile/12412/14059>>.
4. Kun Huang; Kelly, Diane, The daily image information needs and seeking behavior of Chinese undergraduate students. College & Research Libraries [on line], mei 2013 (geraadpleegd op 9 augustus 2016), vol. 74, n° 3, p. 243-261. <<http://crl.acrl.org/content/74/3/243.full.pdf+html>>.
5. WenMing Cao; Ning Liu; QiCong Kong; Hao Feng, Content-based image retrieval using high-dimensional information geometry. Science China Information Sciences [on line], juli 2014 (geraadpleegd op 9 augustus 2016), vol. 57, n° 7, p. 1-11. <<http://link.springer.com/article/10.1007/s11432-014-5086-8#enumeration>>.
6. Dinakaran B.; Annapurna, J.; Aswani Kumar, Ch., Interactive image retrieval using text and image content. Cybernetics and information technologies [on line], 2010 (geraadpleegd op 24 februari 2016), vol. 10, n° 3, p. 20-30. <http://www.cit.iit.bas.bg/CIT_2010/v10-3/20-30.pdf>.
7. Enser, Peter G.B.; Sandom, Christine J.; Hare, Jonathon S.; Lewis, Paul H., Facing the reality of semantic image retrieval. Journal of Documentation [on line], 2007 (geraadpleegd op 24 februari 2016), vol. 63, n° 4, p. 465-481. <<http://www.emeraldinsight.com/doi/full/10.1108/00220410710758977>>.
8. Ying Liu; Dengsheng Zhang; Guojun Lu; Wei-Ying Ma, A survey of content-based image retrieval with high-level semantics. Pattern Recognition [on line], januari 2007 (geraadpleegd op 24 februari 2016), vol. 40, n° 1, p. 262-282. <<http://www.sciencedirect.com/science/article/pii/S0031320306002184>>.
9. Omhover, Jean-François; Detyniecki, Marcin, Combining text and image retrieval Proceedings of the EUROFUSE Workshop on Data and Knowledge Engineering [on line], 2004 (geraadpleegd op 24 februari 2016), p. 388-398. <http://www.poleia.lip6.fr/~marcin/web/wp-content/papercite-data/pdf/Omhover_EUROFUSE_2004.pdf>.
10. Patil, Anita S.; Patil, Neelamma K.; Gejji, V.P., Content based image retrieval – a review Ird India [on line], 2013 (geraadpleegd op 9 augustus 2016), vol. 1, n° 1, p. 38-43. <http://www.irdindia.in/journal_ijeecs/pdf/vol1_iss1/7.pdf>.
11. Rahimzadeh, R.; Farzan, A.; Fanid Fathabad, Y., A survey on semantic content based image retrieval and CBIR systems. International Journal on Technical and Physical Problems of Engineering (IJTPE) [on line], maart 2014 (geraadpleegd op 9 augustus 2016), vol. 6, n° 1, p. 64-69. <<http://www.iotpe.com/IJTPE/IJTPE-2014/IJTPE-Issue18-Vol6-No1-Mar2014/9-IJTPE-Issue18-Vol6-No1-Mar2014-pp64-69.pdf>>.
12. Felci Rajam, I.; Valli, S., A survey on content based image retrieval. Life Science Journal [on line], 2013 (geraadpleegd op 9 augustus 2016), vol. 10, n° 2, p. 2475-2487. <http://www.lifesciencesite.com/ljsj/life1002/343_B01072life1002_2475_2487.pdf>.

13. Rehman, Mehwish; Iqbal, Muhammad; Sharif, Muhammad; Raza, Mudassar, Content based image retrieval: survey. World Applied Sciences Journal [on line], 2012 (geraadpleegd op 9 augustus 2016), vol. 19, n° 3, p. 404-412. <<http://idosi.org/wasj/wasj19%283%2912/17.pdf>>.
14. Weyand, Tobias; Deselaers, Thomas, Combining content-based image retrieval with textual information retrieval. [online], oktober 2005 (geraadpleegd op 9 augustus 2016) <http://thomas.deselaers.de/teaching/files/weyand_sa.pdf>.
15. Content-based image retrieval. Wikipedia [on line], 2 augustus 2016 (geraadpleegd op 9 augustus 2016). <http://en.wikipedia.org/wiki/Content-based_image_retrieval>.
16. Cfr noot 8
17. Thomee, Bart; Lew, Michael S., Interactive search in image retrieval: a survey. International Journal of Multimedia Information Retrieval [on line], juli 2012 (geraadpleegd op 9 augustus 2016), vol. 1, n°2, p. 71-86. <<http://link.springer.com/article/10.1007/s13735-012-0014-4>>.
18. Cfr noot 8
19. Cfr noot 17
20. Wang, Peiling. Information behavior and seeking. In Ruthven, I. and Kelly, D. (ed.) Interactive information seeking, behavior and retrieval. Facet Publishing, 2011, p. 15-41.
21. Nieuwenhuysen, Paul, Search by Image through the WWW: an additional tool for information retrieval. [on line] In international conference on Asia-Pacific Library and Information Education and Practices = A-LIEP 2013. Issues and challenges of the information professions in the digital age. Pullman Khon Kaen Raja Orchid Hotel, in Khon Kaen City, Isan, Thailand, 10-12 juli 2013 (geraadpleegd op 24 februari 2016). <http://www.vub.ac.be/BIBLIO/nieuwenhuysen/presentations/2013-07-aliep-khon-kaen/ALIEP2013_eProceeding.pdf>.
22. Stewénius, Henrik; Gunderson, Steinar H.; Pilet, Julien, Size matters: exhaustive geometric verification for image retrieval. In Fitzgibbon et al. (ed.) ECCV 2012, part II. LNCS 7573, 2012, p. 674-687.
23. Google Inside Search, Google [on line]. <<http://www.google.com/insidesearch/features/images/searchbyimage.html>> (geraadpleegd op 9 augustus 2016).
24. Chapman, Stephen. Google rolls out reverse image search: RIP TinEye. ZDNet [on line], 22 juni 2011 (geraadpleegd op 9 augustus 2016). <<http://www.zdnet.com/article/google-rolls-out-reverse-image-search-rip-tineye/>>.
25. McGee, Matt. Up close with Google's Search By Image: Hits, Misses & More. Search Engine Land [on line], 17 juni 2011 (geraadpleegd op 9 augustus 2016). <<http://searchengineland.com/up-close-with-google-search-by-image-82313>>.
26. Cfr noot 21
27. Cfr noot 23
28. Cfr noot 21
29. Cfr noot 21
30. Cfr noot 21
31. Cfr noot 21
32. Datta R.; Joshi, D.; Wang, J.Z., Image retrieval: ideas, influences, and trends in the new age. ACM Computing Surveys, 2008, vol. 40, n° 2, p. 1-60.
33. Cfr noot 7
34. Cfr noot 8
35. Cfr noot 8
36. Cfr noot 32
37. Cfr noot 7
38. Tsai, Chih-Fong; Lin, Wei-Chao, Scenery image retrieval by meta-feature representation. Online Information Review [on line], 2012 (geraadpleegd op 9 augustus 2016), vol. 36, n° 4, p. 517-533. <<http://www.emeraldinsight.com/doi/full/10.1108/14684521211254040>>.
39. Nieuwenhuysen, Paul. Search by image through the Internet: applications and limitations [on line]. In Libraries in the Transition Era: New Space –New Services – New Experience. Seventh Shanghai International Library Forum. Shanghai Library, China, 9-11 juli 2014 (geraadpleegd op 9 augustus 2016). <<http://www.vub.ac.be/BIBLIO/nieuwenhuysen/presentations/>>.
40. Cfr noot 39

41. Parker, Pamela. Google voice search & search by image comes to desktops. Search Engine Land [on line], 14 juni 2011 (geraadpleegd op 9 augustus 2016). <<http://searchengineland.com/google-voice-search-search-by-image-comes-to-desktops-81633>>.
42. Cfr noot 25
43. Wright, Johanna. Search by text, voice, or image. Google Inside Search [on line], 14 juni 2011 (geraadpleegd op 9 augustus 2016). <<http://insidesearch.blogspot.be/2011/06/search-by-text-voice-or-image.html>>.
44. Cfr noot 25
45. Cfr noot 6
46. Cfr noot 7
47. Gkoufas Yiannis; Morou, Anna; Kalamoukis, Theodore, Combining textual and visual information for image retrieval in the medical domain. The Open Medical Informatics Journal [on line], juli 2011 (geraadpleegd op 9 augustus 2016), vol. 5, p. 50-57. <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3178904/>>.
48. Goel Nidhi; Sehgal, Priti, A refined hybrid image retrieval system using text and color. International Journal of Computer Science Issues [on line], juli 2012 (geraadpleegd op 9 augustus 2016), vol. 4, n° 9, p. 48-56. <<http://www.ijcsi.org/papers/IJCSI-9-4-1-48-56.pdf>>.
49. Cfr noot 8
50. Mistry, Yogita; Ingole, D.T., Survey on content based image retrieval systems. International Journal of Innovative Research in Computer and Communication Engineering [on line], october 2013 (geraadpleegd op 9 augustus 2016), vol. 1, n° 8, p.1827-1836. <http://www.ijrcce.com/upload/2013/october/34_Survey.pdf>.
51. Cfr noot 9
52. Cfr noot 14
53. Cfr noot 9
54. Welcome to the Google Search Forum! Google [on line]. <<https://productforums.google.com/forum/#!forum/websearch>> (geraadpleegd op 24 februari 2016).
55. Chitu, Alex. Google Indexes Images from PDF Files. Google Operating System - Unofficial news and tips about Google [on line], 9 augustus 2015 (geraadpleegd op 24 februari 2016). <<http://googlesystem.blogspot.be/2015/08/google-indexes-images-from-pdf-files.html>>.
56. Cfr noot 4

LE CHEMIN D'ÉVOLUTION DU MÉTIER DE LA VEILLE

Résultats d'une enquête menée lors de la journée Inforum, à Bruxelles, le 4 juin 2015

Béatrice FOENIX-RIOU

Fondatrice de BFR Consultants, auteure de l'ouvrage "Recherche éveillée sur Internet : mode d'emploi" et du site Recherche éveillée

■ Dans le cadre de la journée annuelle Inforum 2015, organisée par l'Association Belge de Documentation (ABD-BVD) sur le thème *The I&D Puzzle: new pieces to solve it*, nous avons été sollicitée pour intervenir sur le sujet de "La veille sur les réseaux sociaux : méthodes, outils et astuces". Pour compléter notre intervention, nous avons mis en place un dispositif de veille éphémère¹, permettant de capter d'autres pratiques que la nôtre, dans un esprit d'exploration. Ce dispositif comprenait en particulier un questionnaire, qui interrogeait les participants à la journée sur leurs usages des réseaux sociaux et leurs pratiques de veille. Cet article propose les résultats de l'analyse quantitative et qualitative des réponses. En particulier, nous analysons l'évolution du métier de la Veille par l'usage d'une boussole dynamique.

■ Het door de Belgische Vereniging voor Documentatie (ABD-BVD) georganiseerde Inforum 2015 ging door onder de titel *The I&D Puzzle: new pieces to solve it*. Ons werd gevraagd om in te gaan op het "Attenderen op sociale netwerken: methoden, instrumenten en tips". Om onze toelichting te documenteren werd voor die dag een ad hoc excerperingstool¹ ontwikkeld om mogelijk andere instrumentaria op te sporen. Er werd vertrokken van een vragenlijst om na te gaan of de deelnemers van de studiedag sociale media gebruiken als attenderingsinstrument. Het artikel geeft een kwantitatieve en kwalitatieve analyse van de antwoorden. In het bijzonder wordt gekeken hoe de manier van attenderen evolueert.

Le contexte

Le dispositif de veille éphémère mis en place lors de la Journée Inforum 2015 avait pour objectif d'étudier les usages et les pratiques de veille sur les réseaux sociaux², sous différents angles :

- des méthodes, outils et astuces (support PDF de notre intervention),
- des pratiques (exposées lors de notre intervention),
- du partage en direct avec les participants grâce à un live-tweet pour la journée, avec une consolidation des échanges a posteriori dans un document *storify*³,
- et de l'exploration, par le biais d'un questionnaire remis aux participants et les interrogeant sur leurs pratiques de veille, aujourd'hui et dans l'idéal.

Les questions étaient de deux types :

- des questions fermées visant à mieux cerner l'usage que font les répondants des réseaux sociaux, avec un questionnement large ("*Utilisez-vous déjà les réseaux sociaux ? oui/non*"...) ou ciblé ("*Utilisez-vous les réseaux sociaux*" : - *comme source pour la veille (oui/non)*, - *comme canal de diffusion (oui/non)*, - *comme espace de partage/discussions (oui/*

non), - *dans un but de "réseautage" (oui/non) ?*) ;

- des questions ouvertes, permettant d'identifier concrètement les pratiques de veille des répondants aujourd'hui, les améliorations qu'ils voient et les pratiques qu'ils imaginent dans les années à venir.

Nous proposons dans cet article une analyse des 64 réponses aux questionnaires, sous deux angles :

- l'analyse quantitative nous donne des chiffres illustrant l'usage des réseaux sociaux par les participants, que nous avons comparés avec ceux qui existent sur les utilisateurs de réseaux sociaux aux États-Unis et dans le monde ;
- l'analyse qualitative éclaire la face cachée de l'impact des réseaux sociaux sur les pratiques des veilleurs, et fait émerger la dynamique de changement du métier de la veille.

Analyse quantitative

Les participants sont actifs sur les réseaux sociaux dans une grande proportion (70%)

70% des répondants – qu'ils exercent ou non une activité de veille – ont répondu "oui" à la question

"Utilisez-vous déjà les réseaux sociaux dans le cadre de vos missions ou fonctions ?".

77% des participants ont également répondu "oui" à la question "Exercez-vous une activité de veille ?".

Ces réponses étaient attendues, puisque l'on interrogeait les participants à une journée spécialisée organisée par l'Association Belge de Documentation.

Pour autant, le recouvrement entre ces deux pratiques est partiel.

Les participants qui ne pratiquent ni la veille, ni les réseaux sociaux, sont très rares (5%) et a contrario, un participant sur deux "seulement" (51%) est à la fois veilleur et actif sur les réseaux.

Plus précisément, 33% des veilleurs présents ne sont pas actifs sur les réseaux sociaux.

Corollaire : 67% des veilleurs présents le sont.

Ces pourcentages sont à mettre en regard de ceux que l'on retrouve dans des études généralistes :

- l'enquête annuelle menée par PewResearchCenter⁴ sur le profil des utilisateurs de médias sociaux, aux États-Unis. La dernière indique que 76% des internautes adultes américains utilisent au moins un réseau social. Ce pourcentage est de 65% si l'on prend comme base l'ensemble des Américains adultes ;
- l'enquête de Global Web Index⁵, menée chaque trimestre auprès d'un panel de 40.000 internautes dans 32 pays. D'après le rapport du 1^{er} trimestre 2015, 7 internautes sur 10 sont actifs sur les médias sociaux, le pourcentage atteignant 80% dans certains pays (Argentine, Mexique, Turquie).

La proportion des participants actifs sur les réseaux sociaux est donc comparable à celle des internautes dans le monde (Fig. 1).

LinkedIn, Facebook, Twitter : le trio de tête des réseaux sociaux utilisés

Les réponses à la question "Quels réseaux sociaux utilisez-vous ?" montrent un panel d'outils relativement large (plus de vingt), parmi lesquels on compte le réseau social scientifique *ResearchGate*, le réseau social d'entreprise *Yammer*, la plateforme de curation *Scoop.it*, le site de bookmarking social

Delicious, le réseau professionnel *Viadeo*, les sites de partage de photos *Pinterest* et *Flickr*, etc. Sans grande surprise, trois réseaux sortent très nettement du lot : *LinkedIn* (67%), *Facebook* (62%) et *Twitter* (47%) (Fig. 2).

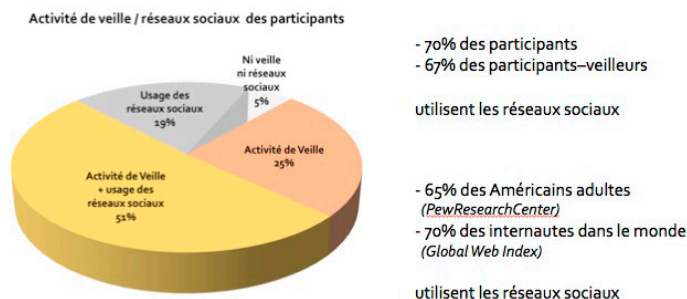


Fig. 1 : L'utilisation des réseaux sociaux par les participants.

Des réseaux sociaux utilisés sur l'ensemble des activités de la veille, depuis le sourcing jusqu'au partage et la diffusion

Nous avons souhaité connaître les pratiques qu'avaient les participants des réseaux sociaux. Nous nous sommes inspiré pour les questions des quatre grands usages autour desquels s'organise, d'après Frédéric Cavazza⁶, l'écosystème des médias sociaux – à savoir la publication, le réseautage, le partage et la discussion – et avons ajouté une question spécifique au sourcing, eu égard à la population de veilleurs.

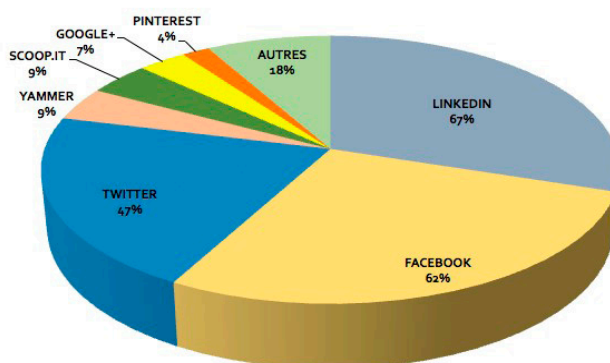


Fig. 2 : Réseaux sociaux utilisés par les participants.

Trois usages se partagent les faveurs des veilleurs de métier, à égalité : pour 64% des veilleurs, les réseaux sociaux sont utilisés :

- comme source pour la veille
- comme canal de diffusion
- et comme espace de partage et de discussion.

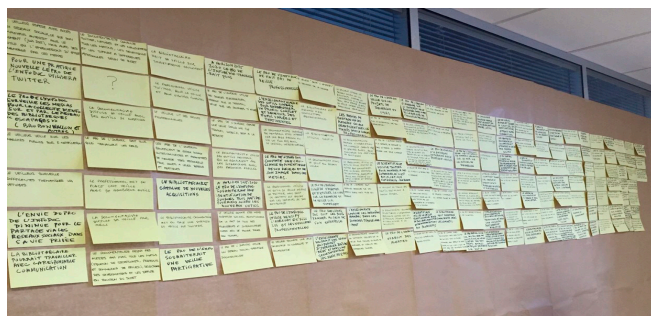


Fig. 3 : Mur de post-it rassemblant les 157 idées des questionnaires.

Le réseautage en revanche est une pratique légèrement moindre (58%).

Si l'on prend en considération l'ensemble des répondants actifs sur les réseaux sociaux (veilleurs ou non), la préférence d'usage varie :

- l'usage prioritaire (67%) est le partage et la discussion,
- viennent ensuite des pratiques de diffusion et de réseautage (62%),
- et enfin de sourcing (53%).

A noter que veilleur de métier ou non, le sourcing est un usage à part entière sur les réseaux sociaux.

Il est également notable que 60% des sondés ont une utilisation des réseaux sociaux dans le cadre de leur vie personnelle, en complément de leurs pratiques professionnelles.

Analyse qualitative

L'analyse quantitative des réponses valide les participants comme représentatifs des professionnels de l'information et en particulier des veilleurs. Elle nous donne une illustration de leurs usages des réseaux sociaux.

Mais quid des évolutions des pratiques des veilleurs en tant que telles, et plus largement des changements induits pour le métier de la veille ?

C'est ce que nous éclairons avec l'analyse qualitative des réponses aux trois questions :

- Citez une ou deux pratiques significatives de votre veille aujourd'hui ?
- Si vous deviez mettre en place une pratique nouvelle aujourd'hui, quelle serait-elle ?
- Quelle pratique de veille imaginez-vous avoir dans cinq ou dix ans ?

Nous avons retranscrit sur des post-it l'ensemble des idées contenues dans les réponses et avons obtenu 157 "idées", que nous avons rassemblées sur un "mur de post-it" (Fig.3).

Nous avons alors sélectionné les idées qui nous semblaient les plus représentatives, pour répondre à la question "Quelles pratiques clés des professionnels de l'information liées à la veille ?". À l'issue de plusieurs phases de sélection, 38 idées ont été retenues.

Nous avons analysé la sélection selon la méthode ABCDE par Elisabeth Brigant⁷ (Apprendre par la Boussole Consciente des Dynamiques Énergiques), en bénéficiant de son accompagnement. Cette méthode permet d'éclairer la face cachée de tout projet, toute situation, à travers une Boussole, grille de lecture des dynamiques sous-jacentes aux actions vécues ou souhaitées.

En synthèse, le Professionnel de l'information est dans un cercle vertueux commençant à germer : la "Découverte" amenée par la pratique des réseaux sociaux fait émerger une "Compréhension" renouvelée de la Veille.

Nous avons représenté dans un schéma 3D le développement du métier de la veille par l'usage d'une boussole dynamique.

Découverte

Découvrir, au sens : explorer, prendre connaissance, mettre en évidence.

Que découvre le Professionnel de l'information en pratiquant avec ces nouveaux outils numériques ? (Fig. 4)

Le Professionnel de l'information prend conscience que la nature de l'information a évolué, tant en quantité qu'en durée de vie, voire en qualité/fiabilité. Cette perception peut être vécue avec angoisse :

- Le professionnel de l'information a très peur de se noyer, un jour, dans la profusion – tsunami d'information.
- Le professionnel de l'information pense que l'information devient de plus en plus courte, éphémère et souvent inutile.

Se sentant pris dans un tourbillon, il se représente difficilement l'avenir pour lui et son métier :

- Le professionnel de l'information s'interroge sur l'intérêt de se projeter dans 10 ans, vu la rapidité de l'évolution des outils.
- A l'horizon 2015-2020, le professionnel de l'information souhaite se recycler.
- Pour autant, le Professionnel de l'information intègre les médias sociaux dans ses pratiques

et les utilise dans toutes les phases de sa veille, transposant les process jusque-là adoptés.

Ainsi :

Il diversifie les sources de veille grâce aux nouveaux outils :

- Le professionnel de l'information veille sur les sites web des institutions de normalisation des documents et aussi de façon hasardeuse sur *Scoop.it*.
- Le veilleur intègre les vidéos, podcasts et tous types d'événements retranscrits via audio / images.

Il organise ses canaux de veille :

- Le service communication fait une veille dans la presse sur la société et sur certaines thématiques.
- L'étudiant repère les sources à surveiller qui émergent automatiquement.

- Le professionnel de l'information utilise des requêtes automatiques dans diverses banques de données.
- Le professionnel de l'information s'abonne à des newsletters.
- Le professionnel de l'information utilise des alertes sur des sujets qui l'intéressent.

Il utilise en particulier Twitter comme canal de veille :

- Le professionnel utilise un outil de veille inspiré de Twitter, mais dont le contenu est stocké sur les serveurs de son entreprise.
- Le professionnel utilise Twitter pour la veille et pour d'autres finalités.

Il structure les résultats de sa veille :

- Le documentaliste utilise un tableau de bord qui rassemble les résultats de veille et qui pourrait être paramétré selon les différents besoins des utilisateurs.
- L'enseignante voudrait une veille qui permet d'intégrer rapidement et dans une seule interface les informations reçues.

Recherche éveillée – La Boussole du métier de la veille : La Découverte (méthode ABCDE par Elisabeth Brigant)

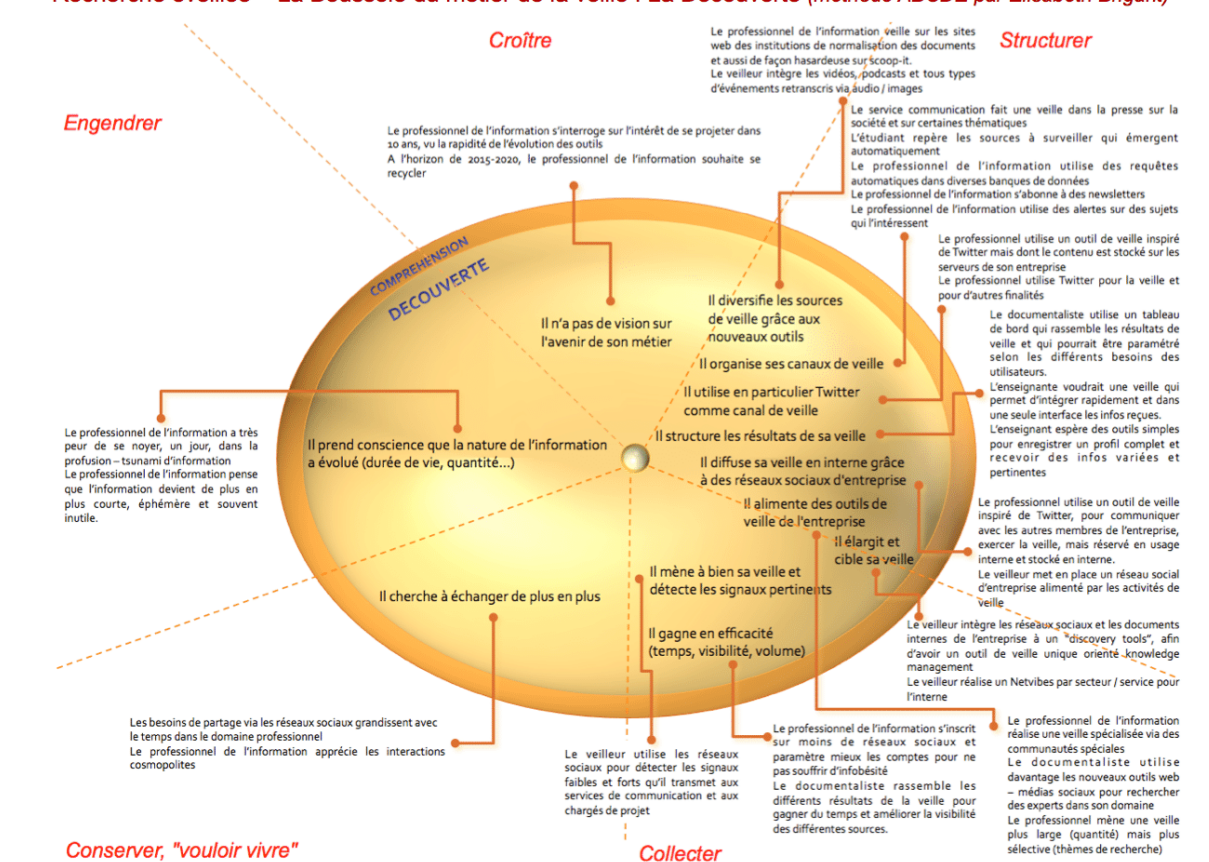


Fig. 4 : La Boussole du métier de la Veille / méthode ABCDE par Elisabeth Brigant – La Découverte.

- L'enseignant espère des outils simples pour enregistrer un profil complet et recevoir des informations variées et pertinentes.

Il diffuse sa veille en interne grâce à des réseaux sociaux d'entreprise :

- Le professionnel utilise un outil de veille inspiré de *Twitter*, pour communiquer avec les autres membres de l'entreprise, exercer la veille, mais réservé en usage interne et stocké en interne.
- Le veilleur met en place un réseau social d'entreprise alimenté par les activités de veille. Il alimente enfin des outils de veille de l'entreprise :
- Le veilleur intègre les réseaux sociaux et les documents internes de l'entreprise à un "discovery tools", afin d'avoir un outil de veille unique orienté knowledge management.

- Le veilleur réalise un Netvibes par secteur / service pour l'interne.

Avec ces nouvelles pratiques, il élargit et cible sa veille :

- Le *professionnel de l'information* réalise une veille spécialisée via des communautés spéciales.
- Le documentaliste utilise davantage les nouveaux outils web – médias sociaux pour rechercher des experts dans son domaine.
- Le professionnel mène une veille plus large (quantité), mais plus sélective (thèmes de recherche).

À l'usage, avec ces nouveaux outils, le veilleur gagne en efficacité :

- Le professionnel de l'information s'inscrit sur moins de réseaux sociaux et paramètre mieux les comptes pour ne pas souffrir d'infobésité.

Recherche éveillée – La Boussole du métier de la veille : La Découverte (méthode ABCDE par Elisabeth Brigant)

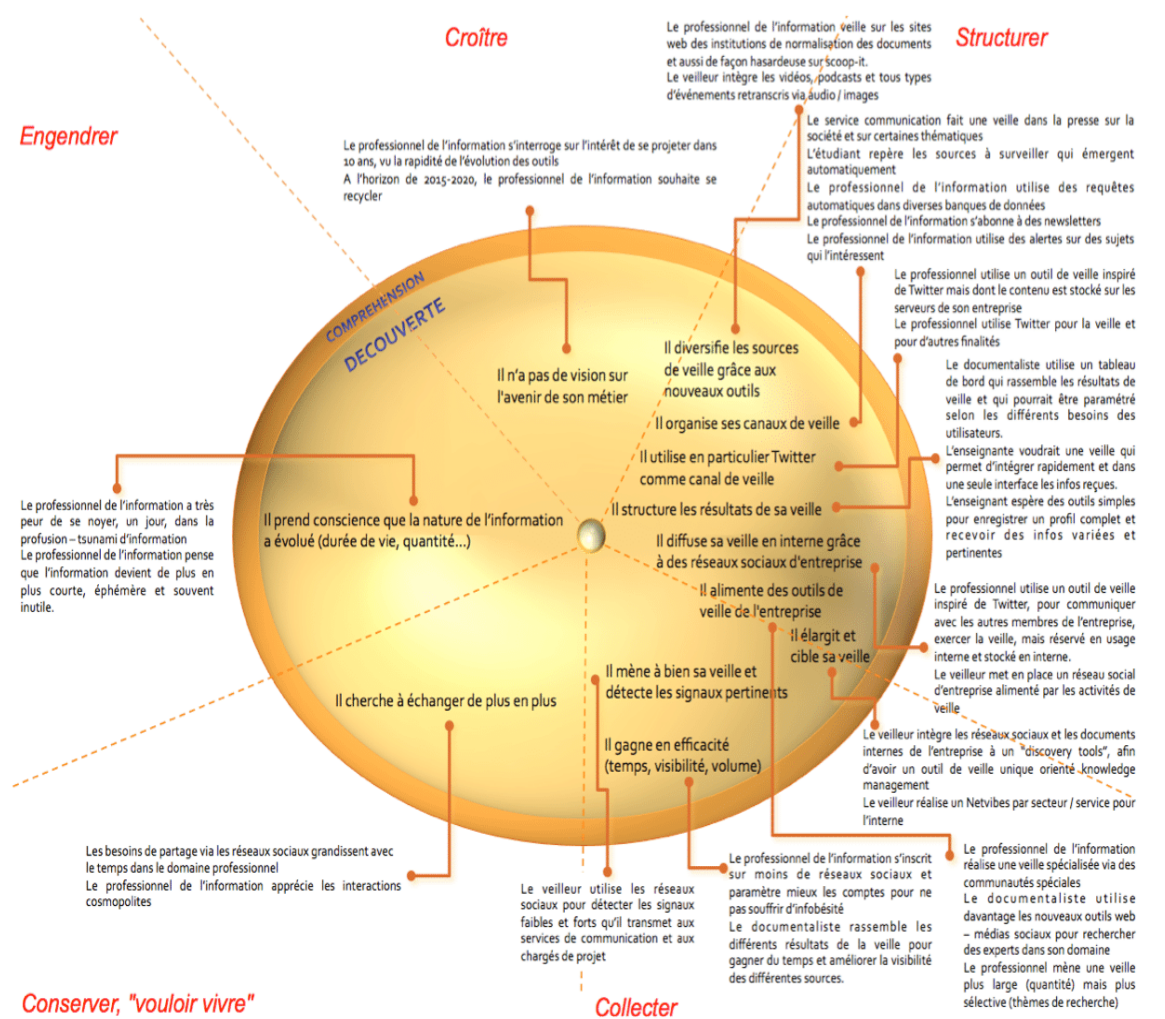


Fig. 5 : La Boussole du métier de la Veille / méthode ABCDE par Elisabeth Brigant – La Compréhension.

- Le documentaliste rassemble les différents résultats de la veille pour gagner du temps et améliorer la visibilité des différentes sources.

Il mène à bien sa veille, et détecte notamment les signaux pertinents pour son entreprise :

- Le veilleur utilise les réseaux sociaux pour détecter les signaux faibles et forts qu'il transmet aux services de communication et aux chargés de projet.

À l'usage également, il prend goût aux échanges facilités par ces nouveaux outils, en particulier à l'échelle internationale. Ces partages lui deviennent nécessaires :

- Les besoins de partage via les réseaux sociaux grandissent avec le temps dans le domaine professionnel.
- Le professionnel de l'information apprécie les interactions cosmopolites.

Ainsi, le Professionnel de l'information prend conscience que des fondamentaux de sa profession sont remis en cause, par exemple la durée de vie de l'information. Il perçoit l'incertitude de l'avenir pour sa profession et pour lui-même. Pour autant, il adapte son savoir-faire à ces nouveaux outils et en tire des bénéfices en termes de pertinence et d'efficacité. Et conquis par l'échange sur les réseaux sociaux, il se rend à l'évidence : ils lui sont devenus essentiels !

Et de la "Découverte" émerge, peu à peu et naturellement, la "Compréhension".

Compréhension

Comprendre, au sens : faculté de saisir intellectuellement, voir facilement les causes et les conséquences qui expliquent les faits, la réalité de quelque chose.

Que fait le Professionnel de l'information sur la voie de sa propre compréhension ? (Fig. 5)

Il veut mieux comprendre les nouvelles tendances accompagnant les réseaux sociaux et est prêt à s'investir pour cela :

- Le professionnel de l'information souhaite s'investir plus dans la compréhension des nouvelles tendances liées aux réseaux sociaux.

Il voit l'avenir dans la veille collaborative :

- Le veilleur souhaite une veille davantage co-construite et collaborative.
- Le professionnel de l'information souhaite une veille participative.

- À l'horizon 2015-2020, le professionnel de l'information participe à une veille collaborative pour alimenter une bibliothèque numérique avec une indexation par tag.

De façon volontaire, il développe le partage et l'échange au sein de son organisation :

- Le veilleur pratique le partage d'expériences et d'information en interne, au sein de son organisation.
- Le bibliothécaire veut davantage communiquer avec les usagers.
- Il connaît précisément les personnes pour lesquelles il œuvre :
- Le professionnel de l'information sait sur quoi travaillent les gens.

Au-delà des outils, le Professionnel de l'information voit la valeur ajoutée des rencontres pour sa veille :

- Le documentaliste utilise des outils nouveaux, qui ne remplacent pas les interactions avec des personnes réelles.
- Le professionnel de l'information s'informe via les médias, les blogs, le bouche à oreilles, pour connaître les personnalités et profils à suivre.
- Le professionnel de l'information utilise des sources électroniques, observe sur le terrain, veille sur les journées d'étude.

Le Professionnel de l'information développe sa compréhension des réseaux sociaux au-delà des outils, depuis l'évolution culturelle, sociale, voire historique dans laquelle ceux-ci s'inscrivent, jusqu'à sa vision de la veille collaborative pour demain.

C'est en pleine conscience qu'il oriente son savoir-faire vers le partage et l'échange. Il voit et comprend la valeur des liens qui se tissent autour de lui, avec les usagers, les personnes référentes... C'est son propre savoir-être qui s'en trouve peu à peu modifié.

Pour autant, des dimensions clés restent à éclairer : quel est son rôle en tant que Professionnel de l'information dans cette vision de la veille collaborative ? Selon quelles modalités, quelles pratiques, quelle organisation de travail ? Pour quel résultat, quel effet ? Pour lui ? Pour les usagers ?

Naturellement, la compréhension continue de se déployer. Et de la compréhension toujours plus aiguisée naîtra, tout aussi naturellement, l'action renouvelée.

Ainsi, prendre conscience de la face cachée de la réalité de la veille à l'heure du Digital, tant sur le plan des fondamentaux et de la représentation du métier, que du savoir-faire et des modalités, c'est fortifier l'action. Et au final, accélérer l'alignement du métier de la veille avec son environnement Humain, Organisationnel et Business.

Béatrice Foenix-Riou

*BFR Consultants
Recherche éveillée
bfoenix@bfr-consultants.com*

Février 2016

Notes

1. Foenix-Riou, Béatrice. Un dispositif éphémère pour explorer les pratiques de veille (Inforum 2015). 24 juin 2015. Blog de *Recherche éveillée* [en ligne]. <<http://www.recherche-eeveillee.com/2015/06/un-dispositif-ephemere-pour-explorer-les-pratiques-de-veille-inforum-2015.html>> (consulté le 19 janvier 2016).
2. Cet article étant centré sur les pratiques des utilisateurs, et non sur les fonctionnalités des services, nous utiliserons l'appellation communément usitée "réseaux sociaux" au sens large, pour désigner des services ayant des usages souvent plus diversifiés que le simple réseautage.
3. Inforum 2015 ABD-BVD. *Storify* [en ligne]. <<https://storify.com/BFoenix/inforum-abd-bvd>> (consulté le 29 janvier 2016).
4. PewResearchCenter. Social Media Usage: 2005-2015. 8 octobre 2015. *Pewinternet.org* [en ligne]. <<http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/>> (consulté le 29 janvier 2016).
5. Perrin, Andrew. Social Media Usage: 2005-2015. *PewResearchCenter* [en ligne], 8 octobre 2015 (consulté le 29 janvier 2016). <<http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/>>.
6. Mander, Jason ; McGrath, Felim. Social Media Engagement - Summary. *GWl Insight Report* [en ligne]. Q1 2015. *GlobalWebIndex*, 2015 (consulté le 29 janvier 2016). <<http://insight.globalwebindex.net/hs-fs/hub/304927/file-2615393475>>
7. Elisabeth Brigant. Profil LinkedIn [en ligne], 29 mai 2015 (consulté le 29 janvier 2016). <<https://www.linkedin.com/in/elisabeth-brigant-a2a1973a>>.

NOTES DE LECTURE

BOEKBESPREKINGEN

UTILISER WIKIPÉDIA COMME SOURCE D'INFORMATION FIABLE

Guy DELSAUT- Klog Éditions - 2016 - 182 p.
- ISBN 979-10-92272-12-3.

Tout ce que vous auriez voulu savoir...

Vous êtes aussi un utilisateur occasionnel, habituel, obsessionnel de l'encyclopédie libre sur le Net ? Ou alors vous pensez que c'est un pis-aller superficiel où chacun écrit n'importe quoi ? Vous vous interrogez sur son succès fulgurant, sur la fiabilité de cette ressource gratuite en ligne ? Notre confrère Guy Delsaut, a écrit un guide¹ court mais très complet sur son utilisation ; c'est un peu *"Tout ce que vous auriez voulu savoir sur Wikipédia sans jamais oser le demander"* !

Guy Delsaut avait déjà abordé le sujet en 2013 dans les *Cahiers de la Documentation*². Après 15 ans d'existence de l'encyclopédie libre, une étude plus complète s'imposait. C'est à double-titre, comme bibliothécaire-documentaliste et comme contributeur aux articles de l'encyclopédie en ligne, qu'il nous invite à en explorer les arcanes. Il nous prévient dès l'introduction : *"l'information présentée dans ce livre vaut surtout pour la version francophone de Wikipédia"*. Et il rappelle ces sages conseils : *"la vérité absolue n'existe pas"* et il faut toujours aller voir ailleurs !

Pour préparer le thème de la fiabilité, l'auteur nous explique l'historique de la culture encyclopédique et comment Wikipédia s'y est intégré, d'abord sous une première mouture payante, puis - assez vite - sous sa forme actuelle d'encyclopédie collaborative et libre d'accès. Il nous présente la structuration des pages (articles et autres) ; même si vous êtes déjà un utilisateur, vous apprenez : il est explicite, donne des conseils pratiques... La lecture pourrait s'arrêter là !

Dans la deuxième partie du livre, nous rentrons dans le vif du sujet. Sur l'origine des erreurs, accidentelles ou intentionnelles, l'auteur analyse les dysfonctionnements du système et comment

les structures communautaires mises en place les corrigent. Car tout un chacun peut facilement créer un article ou en modifier un (langage Wiki). Wikipédia, c'est à la fois une entreprise qui fonctionne bien, avec des filiales, comme *Wikidata*, *Wikimedia Commons* etc. Mais aussi une communauté d'utilisateurs, bénévoles répartis en "chapitres" linguistiques autonomes, qui se distribuent les tâches de gestion de l'encyclopédie³. Ainsi, début 2016, la version française cumulait à elle seule plus de 1.700.000 articles, ce qui est énorme en 15 ans ! Donc : centralisée et souple, verticale et horizontale. Les articles sont en licence *Creative Commons* (libres de droit). C'est ce mélange détonnant, à un moment T, qui a fait son succès.

Dans la troisième partie du livre, plutôt technique, l'auteur analyse le fonctionnement global et nous explique comment repérer les erreurs et vérifier les informations. C'est là que l'on peut apprécier une approche professionnelle de critique des médias et des sources en général. Il y a de nombreuses fonctionnalités, que la plupart des usagers ignorent et qui peuvent donner des indications sur l'élaboration des articles, leurs référencements, leurs historiques, etc. Précisons que si la plupart des contributeurs sont "d'honnêtes encyclopédistes", les enjeux sont tels que personnalités, entreprises et groupes divers tentent de manipuler des informations sensibles. Cependant, les auteurs ne sont pas si anonymes que l'on peut le croire et beaucoup de ces tentatives sont vite éventées ! Bien que contributeur et convaincu de l'avenir de l'encyclopédie, l'auteur critique Wikipédia, tout en reconnaissant une amélioration de la qualité des contenus.

Dans la quatrième partie du livre, intitulée *Wikipédia et la fiabilité, quel avenir ?*, l'auteur nous livre quelques conclusions issues de son retour d'expérience. Par exemple, il se demande avec pertinence si la réflexion sur les sources ne devrait pas évoluer plus vite, si la règle de toujours citer des sources secondaires (les autres auteurs), en priorité absolue sur la source primaire (l'auteur de l'article) est pertinente ; la créativité est importante, dans un système d'information horizontale, même si les vérifications restent valables ! Un petit regret à la lecture de ce stimulant ouvrage : l'auteur ne compare pas assez le champ francophone de Wikipédia avec les champs linguistiques voisins (néerlandophone, germanophone, et bien entendu, avec le champ anglophone) de l'encyclopédie collaborative. Peut-être cela pourrait-il constituer la matière d'un prochain ouvrage ?

Alain Reisenfeld

Notes

1. Editions Klog [en ligne]. <<http://www.editionsklog.com/product/utiliser-wikipedia-comme-source-d-information-fiable>> (consulté le 15 avril 2016).
2. Delsaut, Guy. Face-à-face entre Wikipédia et l'Universalis en ligne. Peut-on tenter une comparaison ? Cahiers de la Documentation - Bladen voor Documentatie [en ligne], 2013 (consulté le 15 avril 2016). 2013/2, p. 52-63. <http://www.abd-bvd.be/wp-content/uploads/2013-2_Delsaut.pdf>
3. Wikipédia [en ligne]. <https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Accueil_principal> (consulté le 15 avril 2016).

**NOUVELLES
PARUTIONS**

**NIEUWE
PUBLICATIES**



Éditions du Cercle de la
Librairie

<http://www.editionsducercle-de-lalibrairie.com/>

ÉDITER LA BANDE DESSINÉE

Benoît BERTHOU - Collection :Pratiques éditoriales - janvier 2016 - 123 p. - ISBN 978-2-7654-1485-8

Si le mode d'expression de la bande dessinée a fait l'objet de nombreuses études, il n'en est pas de même de ses publications. Se focalisant sur des pratiques éditoriales, cet article entend servir la compréhension des processus et des enjeux liés à la production du neuvième art. Ainsi la traduction de bande dessinée conduit à penser la rencontre entre des cultures de l'écrit fort différentes, tandis que la formule éditoriale du "classique" est sans cesse réinventée. Structurante, la pratique de publication en série constitue de même un dispositif suscitant l'ingéniosité d'éditeurs et d'auteurs largement professionnalisés. Et la commercialisation de cette production dans tous les canaux de vente représente encore l'une des singularités d'une bande dessinée qui, forte de multiples "déclinaisons" au sein d'autres médias, éclaire le monde du livre d'un jour nouveau.

PUBLIER LA LITTÉRATURE FRANÇAISE ET ÉTRANGÈRE

Joachim SCHNERF - Collection :Pratiques éditoriales - mars 2016 - 120 p. - ISBN 978-2-7654-1493-380730031

Pour qu'un texte puisse faire son entrée en littérature, pour qu'il soit proposé aux lecteurs sous la forme d'un livre soigné et exigeant, il doit être sélectionné puis soumis à un long processus éditorial qui lui ouvrira l'accès des librairies et de nos bibliothèques.

C'est ce parcours que l'ouvrage de Joachim Schnerf propose de retracer, une trajectoire qui démarre avec l'arrivée d'un texte jusqu'à son aboutissement littéraire, puis sa commercialisation. En arrière-plan, se dessine peu à peu la figure de l'éditeur de littérature, un professionnel habité par la passion et la rigueur qui se donne pour mission de choisir, éditer et promouvoir les textes qui. Il compte pour lui. Il peut travailler sur des ouvrages français ou bien en traduction, mettant en œuvre des savoir-faire parfois différents mais qu'il est intéressant d'observer pour leur complémentarité.

DIRIGER UNE BIBLIOTHÈQUE : UN NOUVEAU LEADERSHIP

Marielle de MIRIBEL, avec la participation de Brigitte IVANO, Christophe GRELET, Sandrine HAON et al. - Collection :Bibliothèques - mai 2016 - 448 p. - ISBN 978-2-7654-1497-81

Dans un environnement politique, administratif, économique, technique et professionnel en constante évolution, le métier de responsable de bibliothèque requiert de fixer des objectifs stratégiques pour son établissement, tout en ayant la confiance de ses collaborateurs pour les atteindre.



Éditions La Découverte

<http://www.editions-ladecouverte.fr/>

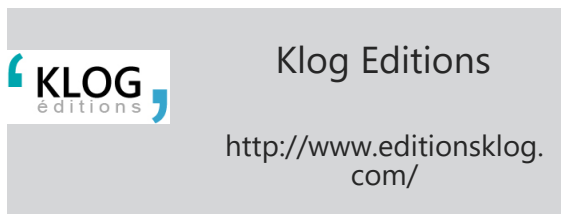
LES MOTEURS DE RECHERCHE

Guillaume SIRE - Collection : Repères n° 671 - mars 2016 - 128 p. - ISBN 978 2 7071 8495 5

Comment se repérer sur le Web ? Comment ne pas se perdre dans ce déluge d'informations ? Les moteurs de recherche sont à la fois des cartes, des boussoles et des topographes. Ils indiquent le nord et nous guident en nous disant : "Ce contenu est plus pertinent que celui-là". Ils répondent à nos questions les plus farfelues en quelques microsecondes. Ils indexent les informations accessibles et dressent pour nous des listes vertigineuses. Ils organisent la connaissance à l'échelle du monde. Oracles modernes, pythies algorithmiques.

L'objectif de cet ouvrage est de révéler ce qui se cache derrière l'interface en apparence extrêmement simple depuis laquelle nous avons accès à

tant d'informations. Le lecteur apprendra comment ont été inventés les moteurs de recherche. Il comprendra comment ils fonctionnent, comment ils sont financés, quel est leur statut juridique, les controverses dont leur activité est susceptible de faire l'objet, quels sont les principaux moteurs sur le marché, comment ils se distinguent les uns des autres et quelles sont les perspectives d'évolution du secteur.



CRÉER ET GÉRER UNE PHOTOTHÈQUE : ORGANISER SON FONDS D'IMAGES

Caroline LEFRANC - Juin 2016 - 200 p. -
ISBN 979-10-92272-16-1

Alors que la photographie existe depuis bientôt deux siècles, les collections photographiques – ou photothèques – restent encore largement méconnues du public et parfois même des professionnels de l'information eux-mêmes. Il en existe pourtant une grande diversité : photothèques publiques, photothèques d'entreprises, photothèques d'agences commerciales, collections privées de photographes, etc. Constitués au fil de l'histoire, les fonds photographiques l'ont d'abord été à des fins documentaires avant d'obtenir le statut de collection à part entière.

Mais si les collections photographiques prennent aujourd'hui une place de plus en plus importante dans les collections des bibliothèques, des archives et des musées, comment sont-elles constituées, organisées, valorisées ?

La création d'une photothèque nécessite des compétences spécifiques : outre une méthodologie documentaire solide, le photothécaire doit disposer d'une connaissance technique des images, argentiques comme numériques, être en mesure d'assurer leur conservation et leur diffusion, respecter les normes de traitement de ces fonds, résoudre les questions juridiques que soulève leur exploitation et répondre enfin, et surtout, aux besoins des différentes communautés d'utilisateurs.

L'ouvrage de Caroline Lefranc apporte un nouvel éclairage sur les différents aspects du métier de photothécaire. Le lecteur – professionnel de

l'information ou étudiant – y trouvera des informations d'ordre méthodologique et technique, de nombreux exemples commentés, un glossaire et différents modèles de contrats. Une bibliographie-webographie complète cet ouvrage résolument pratique et didactique.

LES 500 MOTS MÉTIERS - BIBLIOTHÈQUES, ARCHIVES, DOCUMENTATION, MUSÉES

Jean-Philippe ACCART, Clotilde VAISSAIRE-
AGARD - Juillet 2016 - 192 p. - ISBN
979-10-92272-11-6

"Les 500 mots métiers" répertorie les termes utiles aux métiers des bibliothèques, des musées, de la documentation et des archives. Il revisite les définitions existantes, les illustre d'exemples récents, de schémas et de liens, les situe dans le contexte professionnel et technologique que nous connaissons aujourd'hui.

Le lecteur y trouvera donc une sélection de mots métiers et notamment les termes d'actualité qui lui permettront de nourrir sa réflexion et d'enrichir ses connaissances. Cet ouvrage se révélera particulièrement utile pour des préparations aux différents concours de la fonction publique.

Pour chaque terme, une traduction en anglais et en allemand est proposée afin de répondre aux besoins des professionnels travaillant dans un environnement multilingue.

Écrire pour les Cahiers

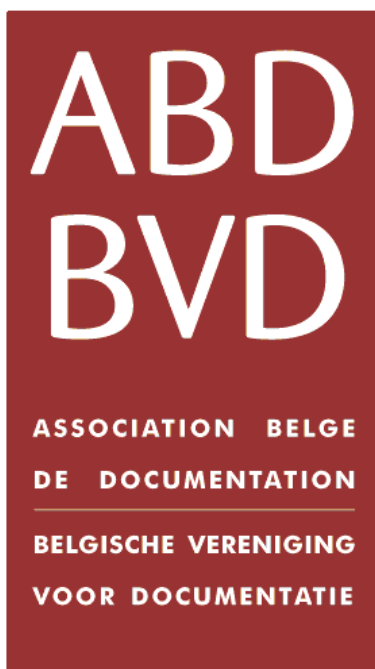
Les *Cahiers de la documentation* sont alimentés par leurs auteurs. Si vous souhaitez partager avec l'ensemble des membres de l'ABD votre expérience dans un domaine ou vos connaissances d'un sujet ou faire le compte rendu d'une conférence à laquelle vous avez assisté, n'hésitez pas à prendre contact avec le Comité de publication : <cahiers-bladen@abd-bvd.net>

Afin d'assurer une présentation cohérente de notre périodique, nous demandons aux auteurs de respecter les instructions aux auteurs disponibles sur <http://www.abd-bvd.be/wp-content/uploads/instr-aut_fr.pdf>

Schrijven voor de Bladen

Bladen voor Documentatie bestaat dankzij de auteurs. Indien u uw ervaringen binnen een domein of uw kennis van een bepaald onderwerp wilt delen met alle BVD-leden of een verslag wilt maken van een studiedag waaraan u heeft deelgenomen, aarzel dan niet om het Publicatiecomité te contacteren via <cahiers-bladen@abd-bvd.net>

Om een coherente presentatie van ons tijdschrift te verzekeren, vragen wij de auteurs de auteursaanbevelingen te respecteren : <http://www.abd-bvd.be/wp-content/uploads/instr-aut_nl.pdf>



asbl créée le 21 mars 1947
vzw opgericht op 21 maart 1947

Plus de 500 professionnels de
l'information et de la documentation

Meer dan 500 informatie- en
documentatiespecialisten

<http://www.abd-bvd.be>

Correspondance

c/o Bibliothèque royale de Belgique
Boulevard de l'Empereur 4
1000 Bruxelles
Belgique
abdbvd@abd-bvd.be

Briefwisseling

p/a Koninklijke Bibliotheek van België
Keizerslaan 4
1000 Brussel
België
abdbvd@abd-bvd.be

Tarif anciens numéros

Prix au numéro : 20 EUR
Prix par article : 10 EUR

Tarief vorige nummers

Prijs per nummer: 20 EUR
Prijs per artikel: 10 EUR

Commandes

tresorier-schatbewaarder@abd-bvd.net

Bestellingen

tresorier-schatbewaarder@abd-bvd.net



KBR•g

INFORUM 2016



In information we trust!!!

Thanks to the Royal Library for the infrastructure.

Thanks to our sponsors without their kindness this day would be impossible.

Thanks to our members for showing their interest.

