

## **Information quality assessment and source selection on the Internet for competitive intelligence: fieldwork research on 53 belgian executives**

**Jeremy Depauw**, Université Libre de Bruxelles, Belgium

### **Abstract**

The internet is seen as a challenging workspace for Corporate Communication and Information Management purposes. It enables a growing number of people to publish, share and relay information on any subject. The recent emergence of user-friendly content creation tools and networking facilities has increased that phenomenon and brought new settings to the informational landscape of organizations. We argue that the evolution of internet has created new needs in practices and theoretical understanding of information quality assessment and source selection.

Based on a survey of corporate information specialists in the private sector in Belgium, the study examines the perceived shift of information quality assessment criteria when dealing with online sources. This paper aims to present a conciliation of usual information evaluation criteria to five formats of information sources, considered as specific to what people call Web 2.0: weblogs, wikis, podcasts, file sharing platforms and social networks sites. 53 phone interviews were conducted in 2008.

Results confirms the existence of a perceived shift in the information quality assessment process. They suggests that it remains globally the same, but that it mostly needs to be adapted to better cope with the new reality of sources formats. Nuances to this observation are discussed in this paper.

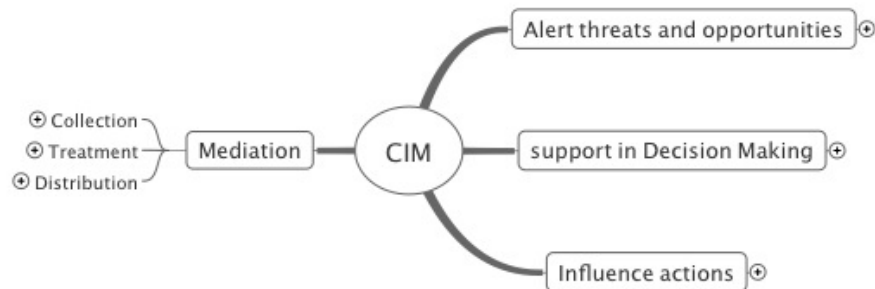
### **New settings, new practices?**

This research investigates how corporate information managers assess the quality of information and select online sources in the recently reconfigured online informational landscape. It is based on the thesis that the new tools and the new modalities of online content creation cause an uncertainty in information quality assessment that changes the criteria used to assess it and force information management professionals to reconsider their selection and validation processes. This paper fits into a larger research addressing three issues: corporate information management, the evolution of the online informational landscape - and especially the internet - and information quality assessment and source selection. This paper will focus specifically on the third aspect and present results and primary findings of the first phase of the fieldwork.

Assessing information quality and providing resources to improve sources selection has always been a subject of intense attention in the field of Information Sciences. On the practitioners as well as on the academics side, those issues remain crucial for all kinds of "audience". For years, information has been identified as a major asset that enables people or organizations to reduce their environment's uncertainty and help them optimize their decision process (Culnan, 1983; Hardy, 1982; O'Reilly, 1982). The

characteristics of the channels used to convey information have an impact on the information selection and evaluation processes. Its form and its content reflect methodological and editorial processes that the user must take into account. Is the information reviewed? Who is the initial source? For what purposes was it generated? These questions are neither new, nor even recent. For every major new means of communication, information evaluators need to adapt their analysis methods to cope with the new channel and the resulting new configuration. This can be observed in the recent evolution of Information and Communication Technologies (ICTs) and especially the content creation tools supported by the Internet as a global platform. Authors usually agree that the need to filter and select the most appropriate source and manage information requirements effectively is compounded due to the exponential rate of growth of literature via diverse media, resulting in information overload, and the lack of knowledge and skills on the part of managers to maximize the available resources (G. de Alwis, Majid, & Chaudhry, 2006; Gina de Alwis, 2001).

Corporate Information Management (IM) is defined here as all the activities that help companies to become aware of threats and opportunities, that help in the decision making process and/or in the influence action strategies. These activities are carried out through an information mediation that can be described as the recurrent cycle of information collection, treatment and distribution within the organization. Information managers act as the eyes and ears of the company and are a major asset to reduce the uncertainty of the ever more complex corporate context. Terms generally used for such activities include "environmental scanning" (Aguilar, 1967) or "competitive intelligence" (M. E. Porter & Millar, 1985; Michael E Porter, 1998). For the purpose of this research, we will be using the broader term of "Information Management" to encompass the wide range of terms and definitions that exist (especially in French). IM has been conceptualized based on a literature review (Author, 2006). The main advantage of proposing a tailor-made conceptualization is that it provides this research with a structured framework of IM activities.



**Figure 1:** Information Management conceptualization

The development of online content creation tools and social software has a significant impact on the corporate environment. For some time, the internet has been seen as a challenging workspace for Corporate Communication and Information Management purposes. In addition, there is a general agreement on the fact that the internet is a gateway to a wide range of direct or indirect sources of competitive information (Teo & Choo, 2001). In particular, it enables a growing number of people to publish, share and relay information (facts, opinions or contacts) on any subject.

"Web 2.0" is far from being a stable and reliable concept in academia. The phrase is used as such in this research, only after we come to terms with this otherwise known, but vague concept. Based on a literature review of definitions and descriptions of Web 2.0, a synthetic framework has been elaborated (Author, 2008). It covers most of the aspects and characteristics that are emphasized in discourses about Web 2.0, which is described as a set of tools, a set of practices and/or a set of tendencies. These would, in turn, be divided up in three categories: content management & communication, collaboration and community. Thus, it allows proposing a fast and comprehensive description of what is covered in this research by the term Web 2.0 and helps setting up the whole research methodology. The tools to be considered are weblogs, wikis, podcasts, file sharing platforms and social networking sites. Examples of typical practices that should be watched include all the tasks of information transformation allowed more than ever by Web2.0 formats (and their user-friendliness), such as information remix, retarget, content creation anytime and anywhere, comments, relay, user generated content, etc. As a set of trends and tendencies, Web 2.0 is usually held responsible for yet-another explosion of data; it is seen as dynamic, open, social, able to reduce the limit between information producer and consumer. It doesn't require a high level of technical skills to publish, so anyone is now able to publish. Again this framework may be represented visually (see Appendix 1).

### **Theorizing Information quality assessment and source selection**

Defining and conceptualizing information quality has proven unattainable, thus hindering the deployment of a method that would appropriately grasp the different dimensions of information quality and transpose the theoretical concepts into an operational design. Authors have generally offered two ways of approaching information quality. On the one hand, academic discussions usually frame the concept in a rather abstract way. On the other hand, professional writers, consultant and educationalists provide students, professionals or executives with guidelines and check-lists that offer an overview of usual criteria for information quality. Outside of literature, hands-on experience from the professionals on the field

deserves full attention as a primary source of insight about professionals' practices. Taken together these channels of knowledge about information quality cover a wide scope of aspects and settings that may lead a positive or a negative assessment of the piece of information under scrutiny. Our fieldwork aims to gather professionals' insight on the basis of concepts and dimensions found in literature.

In 2001, Alison Cook published a book entitled *A Guide to Finding Quality Information on the Internet: Selection and Evaluation Strategies*. It is based on a classic checklist structure. It is assumed that major criteria of information quality, such as authority or accuracy for example, cannot be assessed directly and objectively. In order to fill this gap, she broke down the major information quality criteria in a list of simple and short questions on limited aspects. Taken together, they are supposed to help the assessor building his opinion on the global criteria analyzed. Consider for example authority. It is a criterion that is always pointed out in literature, both academic and professional about information quality. As such, it is almost impossible to rate it because of the wide scope of aspects that this criteria covers. With guidelines such as Cooke's, the assessor is provided with questions such as: What is the reputation of the author? What is the reputation of any other organization involved in the production of the source? What is the reputation of the source? Are there any reviews discussing the source? What is the address of the site? Etc. By looking precisely to each question, Cooke invites readers to focus on a limited aspect of information quality. The assumption is that the overall quality level will be proportional to the rate of positive, or at least satisfying, answers. Under this point of view, information quality is not a stable and quantifiable parameter but rather a construct based on a methodical process.

If most of the criteria tend to be rational, assessing information quality is largely dependent on experience, context and needs of the assessor. Therefore, the key is not to find the perfect and objective evaluation process but the one that best fits the assessor's needs. In the case of this study, Cooke's checklist has been chosen among others. This is justified by the fact that Cooke invites to pay attention either on the content (accuracy, currency) of the information to assess or on its form (presentation, accessibility, comparison) and its production mode (objectives, authority, ease of use). By the time this study investigates the way in which different formats may or may not change the information quality assessment, this wide scope is anticipated as a necessity. In addition, these guidelines are designed initially to assess the quality of online sources. This is very valuable because it takes into account the content production mode specific to the internet at the time the book was written (i.e. 2001). Web 2.0 formats may then be confronted to criteria fitted for the internet as it was in 2001. It is assumed that even though new formats, new practices and new trends may be attributed to the evolution of the internet, there is a majority of sources that have not changed significantly regarding either their form or their production processes.

Having stated this, it is necessary to identify a way to confront and to compare "classical online sources" to Web 2.0 source formats. The core question shifts from "How to assess the quality of information coming from an online source" to "How to assess the quality of competitive information that comes from one of the Web 2.0 formats". Again, it is argued that there is a long history of strategies and habits developed by users, and especially corporate information specialists, to deal with online sources. The novelty of this study resides in the fact that it seeks to renew the understanding of those practices, within the evolving context of Web 2.0.

### **Investigating Information Quality Assessment and source selection**

The main thesis of this research is the following:

New trends, tools, practices and modalities of online content creation cause an uncertainty in information quality assessment that changes the perceived importance information quality criteria and force corporate information management professionals to reconsider their selection and validation processes.

This rather broad thesis is investigated through a fieldwork research divided up into two phases. Phase 1, aims to focus on the perceived shift of importance of information quality assessment process by corporate information managers. Phase 2 is designed to bring qualitative insights about this perceived shift. While the second phase answers the question why? And How?, phase 1 aims to find out if there is actually a perceived shift and understand what is changing in IM professionals practices. In this paper, we will elaborate on phase 1 only. What is studied is the relative importance attributed by professionals to theoretical criteria of information quality. There is no absolute figures or measures. The method proposed here still relies on a social phenomenon and consist more in a qualitative approach of the field than a real quantitative study with representative sample and generalization ambitions. By combining phase 1 with a second phase of qualitative investigation, this research seeks to grasp a global and rather satisfying overview of the perceived shift of information quality assessment process of corporate information managers confronted with the renewed online informational landscape.

#### *Recruiting respondents*

The whole fieldwork research took place in Belgium, from December 2007 to July 2008. It is designed in two phases. Ideally, one of the best ways to leverage understanding of professional practices should have been a direct observation of actors, during their daily tasks and/or a tracking system on professional's

computer. Unfortunately, this was not a possible option because of the strategic position those corporate information professionals have in their organization. They deal with highly strategic and therefore confidential missions, information and knowledge management. Therefore, it was necessary to design an indirect observation of information quality assessment process. The method is explained in the next section of this paper.

Corporate Information Management is not stabilized and organized as other typical organizational activities. Unlike units such as human resources, public relations, accounting or administrative tasks in general, it is very difficult to identify and map services, people and resources allocated specifically to IM activities. This makes trying to approach those professionals particularly problematic. Therefore, the recruitment phase required a preliminary investigation within companies in order to identify individuals whose work and position was most appropriate for our study.

The field investigated may be described as "any company from the private sector in Belgium". The starting assumption supported that the bigger the company, the higher the chances to find elaborated and organized information management units. Therefore, major companies listing and ranking in Belgium TRENDS TOP 100.000 (Trends-Tendances, 2006) was used, with particular attention to specific sectors of activities where IM is likely to be of importance. This was completed with a list of the 500 largest companies as listed in the reference businesses database BELFIRST (Bureau Van Dijk, 2006). Selected companies were contacted and where an Information manager could be identified, a request was sent for participating in the study. The acceptance rate was about one in ten.

Recruited respondents were explained the aim of the study and their profile and experience were checked individually to ensure reliability of their answers to the study. The researcher assured anonymity and confidentiality. No names of people or companies will be shared among respondents or in any dissemination document.

#### *Questionnaire in details*

The questionnaire of phase 1 is built on the criteria listed by Alison Cooke's book cited above (Cooke, 2001). The checklist is aimed at guiding user to assess the quality of information on online sources and to improve their sources selection process.

Phase 1 consisted in 53 one-hour phone interviews with recruited respondents. As introduction to the questionnaire, they were all given the following context:

In the normal course of your work - neither especially calm nor in the middle of a crisis - you are browsing the internet. After a click you arrive on a page that you don't know but that catches your attention. It is the first time you see that page and there is information on it that seems very important about your business or your competition. You need to treat it and to relay it in your company. But before doing so, you need to assess its quality and verify the source. The following list of criteria is made of potential guidelines to help you in this task.

Cooke lists 9 main criteria of information quality. In this research they are called "categories" and were given a short identifying code: aims and objectives (OBJ-1), coverage (COV-2), authority (AUT-3), accuracy (ACC-4), currency (CUR-5), accessibility (ACCESS-6), presentation (PRES-7), ease of use (EASE-8) and comparison (COMP-9). Each category is declined in a list of precise questions that taken together are supposed to help user to deal with those rather wide categories as explained above.

The questionnaire consisted then in crossing each criterion of the list with the 5 formats specific to Web 2.0, identified in the literature review and the framework previously built upon it: weblogs (BLOG), wikis (WIKI), podcasts (POD), file sharing platforms (FS) and social networks sites (SN). Data of the phone interview were collected in a matrix.

**Table 1: Questionnaire illustration**

	BLOG	WIKI	POD	FS	SN
OBJ-1					
Criterion_01					
Criterion_02					
Criterion_03					

To fill in the table, respondents were asked a two-step question. Firstly: what level of importance do you think the criterion xx has when assessing the quality of an online source? The answer was not coded but served as reference (classic online sources) to compare each of the 5 new formats. Hence the second question: Do you consider this criterion as more, less or equally important when it is applied to format xx? Answers were coded -1, 0 or 1 (respectively for less, equally or more important). This is an ordinal scale where modalities are mutually exclusive. It was not allowed to give no answer. Note that the data

collected here consists in assessing the relative importance of the criterion, and not its absolute importance. Therefore, the comparison between respondents is relevant. In addition, a criterion that has a very weak importance may or may not change in importance, exactly as for those that are crucial. By working with relative importance, the questionnaire avoid the major problem related to the fact that information quality assessment and source selection are very personal, based on experience and habits processes.

#### *Data structure and statistical approach*

Collected data take the following structure. The first is labelled **NAMES** and has a unique nominal value for each case, i.e. each of the 53 respondents. The 360 next variables have a unique ID and are taken directly from question for information quality assessment from Cooke's list. The ID is identified by three main elements: the format involved, the category of criteria involved and finally the question of the list. It is structured as: **FORMAT\_N°CAT\_N°LIST**. For example, **BLOG\_1\_01** points out to the first question of the of Cooke's list in the category "objectives", applied to the format "blog": Is there a clear statement of the aims and objectives of the source?

A secondary group of variables was computed from the count of modalities frequency. **FREQ\_1**, **FREQ\_2** and **FREQ\_3** offers a view on the distribution of answers "<", "=" and ">" of all respondents. It shows that distribution on the 360 variables. In addition, this was done on formats specifically as well. For example, **FREQ\_1\_BLOG**, **FREQ\_2\_BLOG** and **FREQ\_3\_BLOG** show the distribution of modalities on the 72 answers related to blog. To finalize the data, the same process was applied to categories of criteria. Here, **FREQ\_1\_OBJ**, **FREQ\_2\_OBJ** and **FREQ\_3\_OBJ** offer a view on the distribution of answers on the category "objectives". It is computed from all criteria related to objectives through the 5 formats. This group is needed to begin the statistical analysis of data, especially for the central tendency and dispersion. As ordinal variables, the 360 criteria of information quality taken from Cooke's list raise problems with using classic descriptive statistics. Each variable may take three modalities. In order to compare respondents and variable, data needed to be processed into new variables. Additional scale variable were computed from the 360 initial ones, taken from the questionnaire. In the data table, the modality "<" was coded 1, "=" was coded 2 and ">" was coded 3. New variables were created by simple summation. This provides the analysis with comparable scale variables that can be approached with a broader set of statistical tools to help compare variables with one another. For example, the category "objectives of the source" is declined in Cooke's list in three questions, i.e. three variables. Each respondent gave one, and only one, answer for each of them. Following this, the category "objectives" on the format "blog" (code:



**BLOG\_1**) may take values ranging from 3 to 9. With this value, each case is then comparable. This was computed for all of them, providing the analysis with 45 new scale variables. Based on the same logic, additional variables were computed. Five of them aim to give a score to the formats and are identified as **BLOG, WIKI, POD, FS** and **SN**. For example, BLOG is computed from variables  $BLOG_1 + BLOG_2 + \dots + BLOG_9$ . Again, 9 variables were computed to give a score to the categories of criteria: **OBJ, COV, AUT, ACC, CUR, ACCESS, PRES, EASE** and **COMP**. For example OBJ is computed from the sum of  $BLOG_1 + WIKI_1 + POD_1 + FS_1 + SN_1$ . A final variable was computed and is entitled **ALL\_**. It is computed from the sum of  $BLOG + WIKI + POD + FS + SN$ .

### Investigation structure

The analysis of such a large set of data would not fit into this paper, which will focus on the general observation of results that addresses a first level of the research question: is there a perceived shift in the information quality assessment process when corporate information managers deal with Web 2.0 source formats?

In this study, the answers given by respondents point out their perception of this shift by pointing out a change of importance on information quality criteria. So every "<" and ">" answers may be considered as indicators of shifting criteria on specific formats and every "=" answers is supposed to point out criteria that are not influenced by Web 2.0 formats. A first level of analysis will then consist in considering frequencies of answers on different levels of detail in order to highlight the perceived shift that is under investigation.

A second level of questioning will consist in having a look at the distribution of answers. Do formats differ in the way the shift is leveraged? Are categories of criteria differently affected by Web 2.0 formats? This level offers the opportunity to bring nuances in the way the perceived shift in information quality assessment is effectively pointed out by respondents' answers.

A third level seeks to highlight the variables that show highest scores. By calculating indexes, it is possible to determine the formats and criteria that best show the perceived shift in respondents' information quality assessment process.

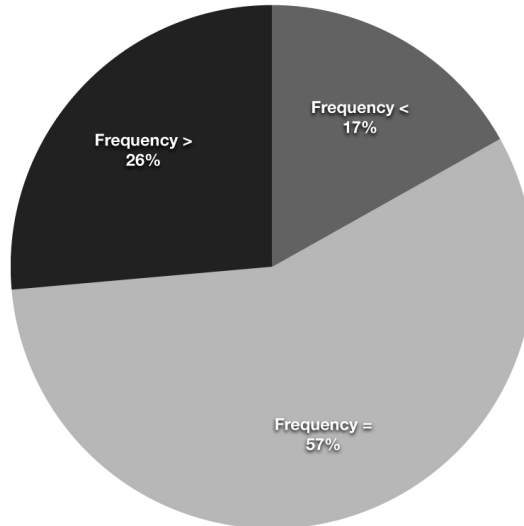
Finally, a fourth level of investigation aims at taking benefits from the quantitative data collected by trying to observe the extent to which variables may be associated (i.e. correlation among answers on formats or categories of criteria).

Short of providing an exhaustive analysis of the data, those levels of investigation aim to offer a first range of findings that are completed by an explanatory dimensions in the second phase of the research project.

**Analysis and findings**

*Dealing with frequencies*

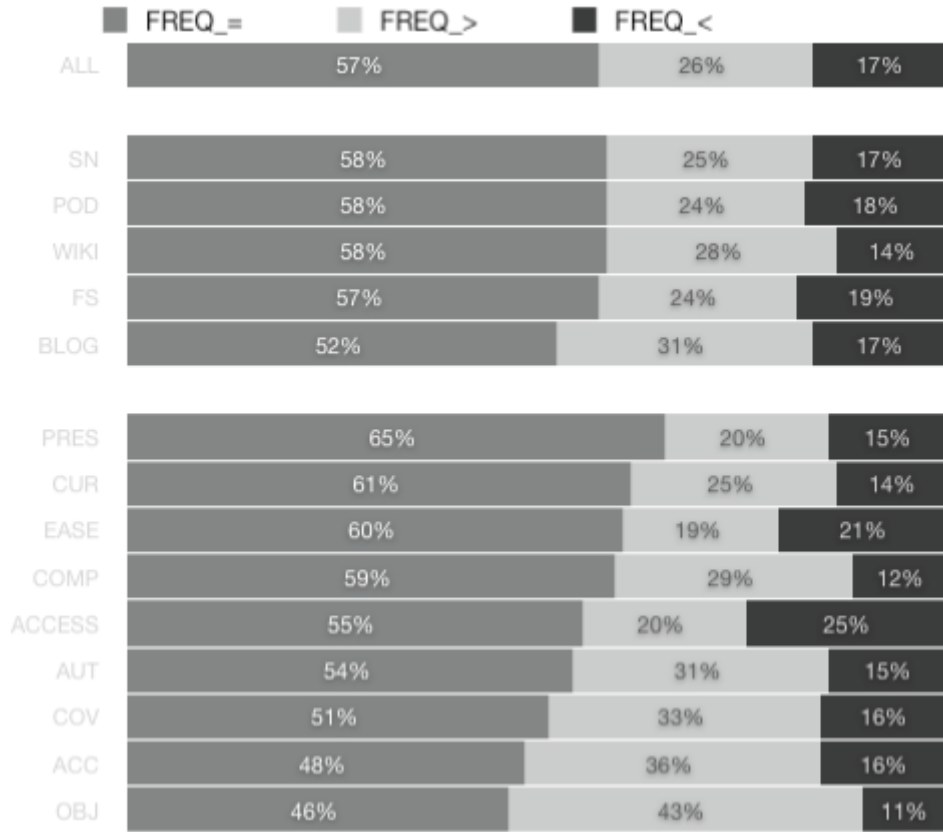
Distribution of frequencies shows that the most frequent answer is "=", meaning that the majority of criteria for information quality are judged as equally important. In other words, when an information, coming from an online source is assessed, the criteria investigated does not become more important, more crucial (nor less) when the information is found within a Web 2.0 format. For example, consider the criteria "Is the information up-to-date?", applied to the evaluation of an information coming from a blog. If the respondent answers "equally important", this means that he won't pay more nor less attention to this criteria to judge the quality of the information. Then the mode of phase 1 is "=" as it is from far the most frequent answer (Albarelo, Bourgeois, & Guyot, 2007).



**Figure 2:** Frequencies on 360 variables for all respondents

The pie chart shows that 57% of the criteria from the 53 respondents are equally important when assessing an information or a source, regardless the format of source it comes from. Back to the question of the perceived shift in information quality assessment process, the mode suggests that globally speaking, information quality assessment criteria are stable. This does not imply that there is no shift at all. Indeed, taken together, 43% (26% of ">" + 17% of "<") of the criteria change by growing or decreasing in importance when assessing an information or a source coming from Web 2.0 formats. It may be noticed additionally that the second most frequent answer is ">" - 26% of the criteria - suggesting that Web 2.0 format raise more questions, caution and challenges for respondents. This provides the research with clues that Web2.0 actually creates uncertainty in information validation process. Moreover, it is argued on the contrary that criteria that loose importance are no longer useful to assess the quality of a piece of information. As such they couldn't be a source of uncertainty for the evolving practices of interviewed corporate information professionals.

As such, this observation needs to be nuanced in order to leverage the second level of analysis. This can be found either through the distribution of answers by formats or by categories of criteria. Are there any significant differences in the distribution in formats? This was achieved by additional case summaries of the frequencies of answers by format and by categories that are gathered in the following bar chart.



**Figure 3:** Frequencies by formats and categories of criteria

The mode and the distribution of modalities in each format is similar for all formats (BLOG, WIKI, POD, FS, SN). Figure 3 shows a rather regular disposition of the 5 bars. Each of them represents a format, split in 3 modalities of answers in the interviews ("=", ">", "<"). The frequency of "=" answers remains the mode for formats as well. It is the most stable modality within formats as values are very close the 57% of the answers, already pointed out as the mode when all variables were considered together (ALL\_). If the amount of Only BLOG shows a weak loss of "=" answers. This is directly overtaken in ">" answers. This observation suggests that there seems to be no critical nuance brought by the distribution of modalities on formats variables. Furthermore, expect from tiny differences in the distribution of answers, the bars are quite well aligned. This suggests that when answering the questionnaire, the respondents have considered the formats more as a unique set than as five unique formats. This point is developed more in details

further. Anyway, the finding of a perceived shift remains as long as more than 40% of answers keep pointing to a perceived shift in information quality assessment process.

On the contrary, bars related to categories of criteria indicate a distribution of frequency that is shows more dissimilarity. It may be noticed that the answers on bars does not follow a regular circuit as they do when investigating Format as cause of shift in perceived information quality criteria. Frequency of "<" and ">" answers are drastically different, pointing out that category of criteria have a significant impact on the distribution of answers. The analysis hereafter will give more clues to understand how this is actually distributed.

#### *Central tendency, dispersion and comparable scores*

By computing scores, the analysis was provided with unique scale variables, instead of only three-folded frequencies and ordinal data, opening the way for the third level of analysis. As explained earlier, these scores, calculated from the sum of the virtual value of modalities 1, 2 and 3 in the data table, offer a comparative view on data. However, because the scale of scores between categories is not the same, as well as between categories and formats, indexes on a 100 base were calculated.

$$Index = \frac{mean}{maximum\ score} \times 100$$

As an example. OBJ=(34,92/45)\*100. The indexes do not figure a percentage or an absolute value, but they provide a robust basis for comparison. In this case, an index that is higher than another indicates that the variable raise more challenges in the information quality assessment process, considering that to be higher, it needed to gather more ">" answers during the interviews. However, those indexes are not to be confused with a real percentage such as those given in the pie chart of frequencies of answers. The higher the score is, the most challenging the format may be considered for the respondent.

**Table 2: Scores and indexes summary**

	<b>Scope</b>	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>Std. Dev</b>	<b>Index</b>
ALL_	405	489	894	<b>686,4</b>	66,803	<b>63,56</b>
BLOG	83	91	174	<b>135,55</b>	15,382	<b>62,75</b>
WIKI	68	112	180	<b>141,38</b>	12,472	<b>65,45</b>
POD	100	80	180	<b>136,87</b>	16,402	<b>63,37</b>
FS	98	82	180	<b>135,53</b>	17,271	<b>62,75</b>
SN	70	110	180	<b>137,08</b>	14,285	<b>63,46</b>
OBJ	21	24	45	<b>34,92</b>	5,106	<b>77,6</b>
COV	60	60	120	<b>86,79</b>	11,011	<b>72,33</b>
AUT	50	40	90	<b>64,96</b>	8,685	<b>72,18</b>
ACC	47	43	90	<b>66</b>	9,794	<b>73,33</b>
CUR	55	50	105	<b>73,94</b>	10,803	<b>70,42</b>
ACCESS	121	74	195	<b>126,77</b>	19,06	<b>65,01</b>
PRES	70	55	125	<b>92,11</b>	13,725	<b>68,23</b>
EASE	60	55	115	<b>89,25</b>	12,476	<b>66,11</b>
COMP	52	41	93	<b>69,75</b>	10,425	<b>42,27</b>

The indexes on formats variables confirm the previous observation of rather equal distribution of answers. This table shows also that WIKI is the format showing the highest score, while file sharing and blogs show the lowest. Social network produces the second highest score and podcast the third. A ranking may then be established, from the most challenging to the least challenging format in the context of information quality assessment. In decreasing order, they are: WIKI, SN, POD, BLOG, FS. As such, it is obvious that the difference in scores is thin. There is a smoothing effect that has to be taken into account. Here, standard deviation shows a relatively uniform dispersion of cases. This is consistent with one of the previous observation that shows a high stability of scores among formats (see above).

On the same logic, the indexes of categories of criteria variables are quite different, which was expected based on unequal distribution of answers (shown on bar chart). Here the indexes improve our capacity to compare the way categories are challenged when used to assess the quality of a piece of information coming from a Web 2.0 source formats. Two groups may then be identified. The first is made of the objectives of the source, accuracy, authority and coverage. Those categories seem to raise more questions

than the others when dealing with Web 2.0 formats. The second group, made of accessibility, ease of use and comparison have similar indexes, lower than those of the first group. Currency is a category that doesn't fit in any of the groups as it has an intermediate score. This suggests that even if it couldn't be concluded that the second group of categories lose importance in information quality assessment, it is basically handled more easily, or at least with less challenges than those of the first group. Back to the frequencies, distribution of answers shows actually a difference between the groups without the smoothing effect of the indices calculation.

**Table 3: Frequencies on categories of criteria**

	Frequency <	Frequency =	Frequency >
OBJ	11	46	43
COV	16	51	33
AUT	15	54	31
ACC	16	48	36
CUR	14	62	25
ACCESS	25	56	20
PRES	15	65	20
EASE	21	61	19
COMP	12	59	29

Note that in each group, the distribution of answers is similar. Currency has a high amount of "=" answers, which is more a characteristic of the second group. It also has a higher amount of ">" answers, which is more specific to the first group. This supports the initial idea of keeping it out of both groups. Helped by indexes and distribution of answers, it is argued that a kind of ranking of the most challenged category of criteria may be proposed. By far, respondents' answers point out objectives of the source (OBJ) as the most challenged category when applied to web 2.0 format, with 43% of ">" answers. It is followed by accuracy (36%) and coverage (33%). Authority follows very closely (31%). In the second group, the ranking is presentation, ease of use, accessibility and comparison. It may be noticed that even though accessibility has a higher indexes than comparison, it is also the category that gathers the highest amount of "<" answers (25%). This value means that a quarter or the criteria of accessibility, on every format, are considered as less important to assess the quality of information by respondents. On the same level of reading, ease of use follow the same line with 1/5 of the criteria that lose the critical power in the information quality assessment process.

Beyond the scores and indexes, the two groups cover two different dimensions of the information assessment. While the first group deals with aspects related to the information, the author and the content, the second is mostly related to the technical and structural ones. It is then argued that the perceived shift under investigation points to a growing concern by the respondents for aspects related to the content production. Who is the author? What are his purposes? What is his level of specialty? How accurate and trustworthy is the information he shares? Results of phase 1 suggest that such questions raise more concern when dealing with Web 2.0 formats.

### *Correlating variables*

Figure 3 (bar chart) and indexes pointed out that the distribution of answers on formats is relatively uniform. However, it is argued that blogs, wikis, podcasts, file sharing platforms and social network sites are rather different in essence. One would probably agree on the fact that reading a blog post and following updates on LinkedIn profile does not require exactly the same analysis process. Then it would be interesting to check if there is any association between formats. Pearson correlation has been computed on formats in order to point out those associations. For example, a positive correlation between blogs and wikis would indicate that a high score on blog tends to be associated with a high score in respondent's answers on wikis. This would support the idea that the two formats are considered on the same "level" in respondent's mind. Correlations table shows the following results:

**Table 4: Pearson correlation on formats and categories of criteria**

		<b>BLOG</b>	<b>WIKI</b>	<b>POD</b>	<b>FS</b>	<b>SN</b>
<b>BLOG</b>	Corrélacion de Pearson	1	,660(**)	,824(**)	,723(**)	,669(**)
	Sig. (bilatérale)		0,000	0,000	0,000	0,000
	N	53	53	53	53	53
<b>WIKI</b>	Corrélacion de Pearson	,660(**)	1	,617(**)	,630(**)	,663(**)
	Sig. (bilatérale)	0,000		0,000	0,000	0,000
	N	53	53	53	53	53
<b>POD</b>	Corrélacion de Pearson	,824(**)	,617(**)	1	,814(**)	,719(**)
	Sig. (bilatérale)	0,000	0,000		0,000	0,000
	N	53	53	53	53	53
<b>FS</b>	Corrélacion de Pearson	,723(**)	,630(**)	,814(**)	1	,798(**)
	Sig. (bilatérale)	0,000	0,000	0,000		0,000
	N	53	53	53	53	53



	Corrélation de Pearson	,669(**)	,663(**)	,719(**)	,798(**)	1
<b>SN</b>	Sig. (bilatérale)	0,000	0,000	0,000	0,000	
	N	53	53	53	53	53

Because variables show significant levels of correlation, we will consider the most significant correlations for interpretation. Blog is highly correlated with podcasts ( $r=0.824$ ), suggesting that if the score on blog is high for a respondent, the probability is great that the score of podcast will be high and close to it as well. The second mostly correlated format to blog is file sharing. Podcast is associated to file sharing as well, which is consistent to the link it has with blogs and but also with social networks. It is also the case with social network that file sharing is correlated when analyzing formats scores. As such, this does not allow to group format together as it was done for categories of criteria. Considering blog and podcast and file sharing and social network as related seems coherent. Further investigation will be undertaken to check the validity of this finding.

### Following up the study

This first set of findings bring interesting insights on the actual perceived shift of information quality assessment when dealing with Web 2.0 source formats. Again, it is important to keep in mind that this study does not intend to extrapolate its findings to all corporate information managers. Information quality assessment remains a very individual process. This is latent in this first of analysis through which interesting observations have been highlighted but with a limited explicative power. As explained earlier in this paper, the first phase of this fieldwork research is considered as starting point for further analysis. It reports on practices and perceptions of practices but does not give details about the reasons that motivate the rating of importance of criteria.

The second phase of the study intended to fill this gap. Fourteen in-depth semi-structured interviews were conducted among the fifty executives interviewed during phase 1. Respondents were presented the main findings of phase 1 and were invited to compare their own answers with that of the whole sample. The aim consisted in elaborating on them. Those semi-directed interviews were designed to understand why criteria change or not in respondent practices.

While the Web 2.0 framework (see appendix 1), elaborated in the context of this research project, points out the formats that needed to be included in phase 1, the practices and trends dimension suggested in the analysis are potential hypotheses that could explain the uncertainty that these formats create and the

changes in IM professionals' habits of information quality assessment and sources selection. Based on the Web 2.0 framework as reference repertoire, the interviews are analyzed by tracing the elements of the discussion that refers to one of the explicative hypothesis of the model. This process highlights the elements in discourses about Web 2.0 that really interfere with IM professionals' practices of information quality assessment. In addition, it points out those elements that are not relayed by professionals' discourses. The interviews could also bring new explanatory elements that are not relayed in Web 2.0 descriptions. To date, data collection is completed and undergoing analysis.

## References

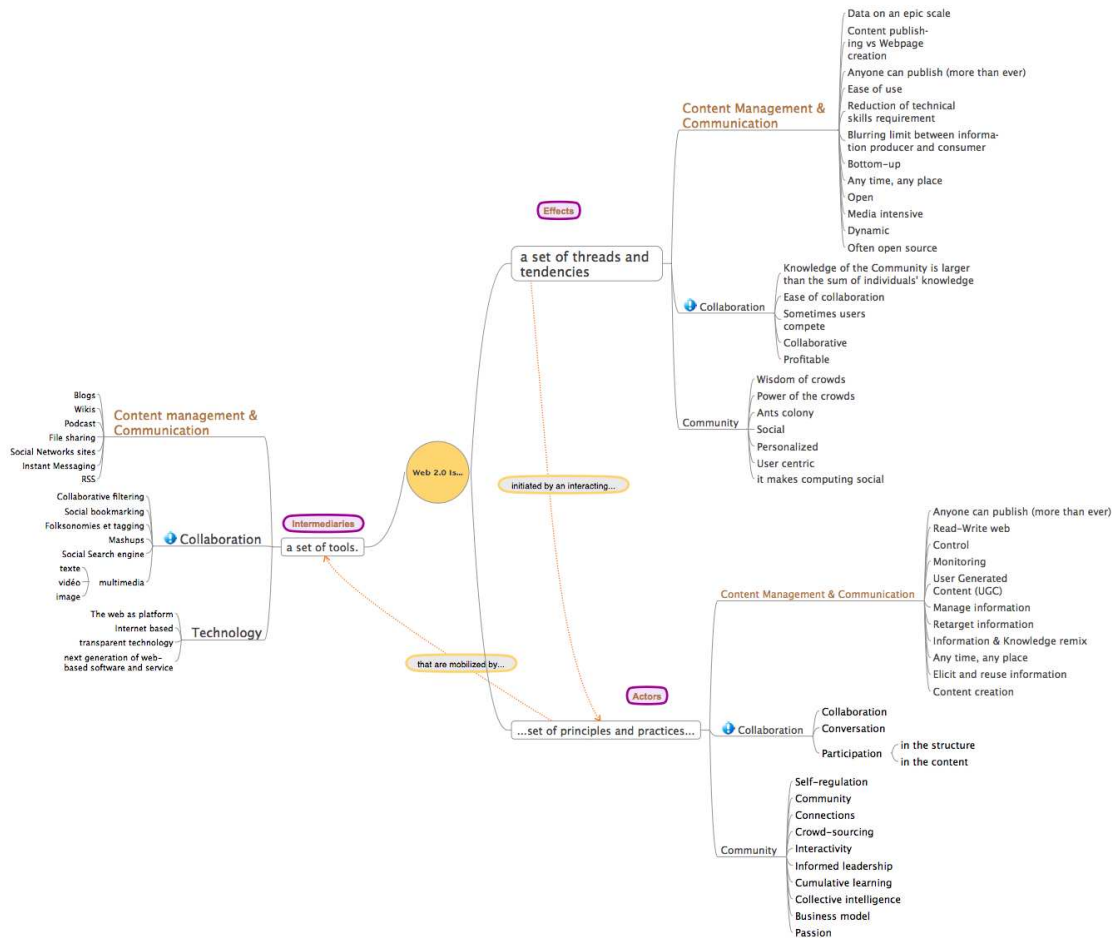
- Aguilar, F. J. (1967), *Scanning the Business Environment: Studies of the modern corporation*, New York, Macmillan
- Albarello, L., Bourgeois, E., Guyot, J. (2007), *Statistique descriptive: Méthodes en Sciences Humaines* (2 éd., p. 176), Bruxelles, De Boeck Université
- de Alwis, G., Majid, S., Chaudhry, A. S. (2006), "Transformation in managers' information seeking behaviour: a review of the literature", *in Journal of Information Science*, 32(4), 362
- de Alwis, G. (2001), "Information as a tool for management decision making: a case study of Singapore", *in Information Research*, 7(1), paper 114, <http://informationr.net/ir/7-1/paper114.html>
- Bureau Van Dijk (2006), Belfirst, Base de données, Bruxelles. <http://www.bibeco.ulb.ac.be/database/belfirst.htm>
- Cooke, A. (2001), *A Guide to Finding Quality Information on the Internet: Selection and Evaluation Strategies* (2 éd), London, Library Association Pub
- Culnan, M. J. (1983), "Environmental scanning: The effects of task complexity and source accessibility on information gathering behavior", *in Decision Sciences*, 14(2), 194-206
- Depauw, J. (2006), "La Gestion de l'Information des organisations: analyse de définitions et conceptualisation", *in Revue de l'Association belge de Documentation*, (2006/4), 23-29
- Depauw, J. (2008), "Web 2.0 under the Actor-Network Theory point-of-view: Conceptualization and definitions analysis", *in Proceedings of Politics: Web 2.0 - An international Conference*, New Political Communication Unit, Royal Holloway University of London, <http://newpolcom.rhul.ac.uk/politics-web-2-0-conference/>
- Hardy, A. P. (1982), "The Selection of Channels When Seeking Information: Cost/Benefit vs. Least-Effort", *in Information Processing and Management*, vol.18, n°6, p289-93
- Kolbitsch, J., Maurer, H. (2006), "The Transformation of the Web: How Emerging Communities Shape the Information we Consume", *in Journal of Universal Computer Science*, 12(2), 187-213
- O'Reilly, C. A. (1982), "Variations in Decision Makers' Use of Information Sources: The Impact of Quality and Accessibility of Information", *in The Academy of Management Journal*, 25(4), 756-771
- Porter, M. E., Millar, V. E. (1985), "How information gives you competitive advantage", *in Harvard Business Review*, 63(4), 149-160

Porter, M. E. (1998), *On Competition*, The Harvard business review book series, Boston, Harvard Business School Publishing

Teo, T. S. H., Choo, W. Y. (2001), Assessing the impact of using the Internet for competitive intelligence. *Information & Management*, Vol. 39 No.1, 67-83

Trends-Tendances (2006), *Trends TOP 100.000*, Roularta Media Group, <http://www.trendstop.be>

**Appendix**



**Figure 4:** Web 2.0 framework (Author, 2008)