Dutch Journalism in the Digital Age

Periodismo Holandés en la Era Digital

Max Kemman
Junior Researcher (Erasmus University Rotterdam)

Martijn Kleppe
Post-doc researcher (Erasmus University Rotterdam)

Bob Nieman
Teacher & researcher (Hogeschool Rotterdam)

Henri Beunders
Professor (Erasmus University Rotterdam)

Fecha de recepción: 5 de mayo de 2013
Fecha de revisión: 6 de junio de 2013
Resumen

En un contexto de constante crecimiento de la oferta de fuentes en línea, la información para producir noticias parece estar a un clic de ratón. Pero, realmente ¿cómo están usando los periodistas holandeses las herramientas de investigación asistidas por ordenador? El artículo presenta un inventario de las distintas formas en las que los periodistas utilizan las fuentes y recursos digitales y explora las diferencias existentes entre expertos y novatos. Se aplica un enfoque metodológico que combina la realización de un estudio etnográfico con una encuesta. Los resultados muestran que los periodistas holandeses utilizan relativamente pocas herramientas digitales para encontrar información en línea. No obstante, aquellos que pueden ser considerados como expertos en el campo de la recuperación de la información utilizan una gama más amplia de motores de búsqueda y técnicas de recuperación, llegando más rápido al enfoque de su historia, y siendo mejores en la búsqueda de información relacionada con este punto de vista. Ello les permite dedicar más tiempo a escribir su noticias. Por su parte, los novatos son más dependientes de la información proporcionada por otros.

Palabras clave
Periodismo - Uso de internet - Sala de prensa - Tecnologías digitales - Producción de noticias - Observación - Estudio - Búsqueda de comportamientos - Google

Abstract

With an ever-growing supply of online sources, information to produce news stories seems to be one mouse click away. But in what way do Dutch journalists actually use computer-aided research tools? This article provides an inventory of the ways journalists use digital (re)sources and explores the differences between experts and novices. We applied a combined methodological approach by conducting an ethnographic study as well as a survey. Results show that Dutch journalists use relatively few digital tools to find online information. However, journalists who can be considered experts in the field of information retrieval use a wider range of search engines and techniques, arrive quicker at the angle to their story, and are better at finding
information related to this angle. This allows them to spend more time on writing their news story. Novices are more dependent on the information provided by others.

Key Words
Journalism - Internet use - Newsroom - Digital technologies - Newsproduction - Observation - Survey - Search behaviour - Google

1. Introduction

With an ever-growing supply of digital and online sources, information to produce news stories seems to be one mouse click away when using these computer-aided research tools. Several scholars point to the impact of digital technologies on the possibilities for citizens to participate in the journalistic production cycle (Bowman and Willis, 2003), the acceleration of convergent (Boczkowski, 2004; Deuze, 2004) or networked journalism (Heinrich, 2011; Jarvis, 2006) and its possibilities for the daily work situation of journalists (Pavlik, 2000; Deuze, 2007; van Dijk, 2004). Pleijter and Deuze (2003) for example coin the term 'editorial cybernetisation' to describe the increase of information in the newsroom, resulting in the reinforcement and acceleration of the journalistic skills needed. As a result, practical instruction books to do research online are being published widely (van Ess, 2010; Luckie 2012) and more recently we have seen the establishment of several associations that provide hands-on trainings both on international level such as the Global Investigative Journalism Network as well as national level such as the Dutch Flemish Association for Investigative Journalism (VVOJ) and the Association for Dutch Online Journalists (VOJN).

Survey studies have shown that the Internet is part of the common work practice for journalists (Middelberg and Ross 2001, 2005; Keel and Bernet, 2005; Wegner 2005). However, most of these studies formulated general conclusions based on the perceived behaviour of the respondents and failed to provide insight in the complex journalistic production process of which online research is only one part. Machill and Beiler (2009) tried to fill this gap by using a mixed method approach of both a survey as well as participatory research. In contrast with the studies that only used surveys, they showed German journalists indeed use online tools more
frequent, but for shorter periods than traditional research tools. They concluded that digital tools complement traditional research methods while the telephone remains the most important research tool. The same conclusions were drawn by Spyridou et al. (2013) for the Greek situation. By employing actor network theory, their results also showed that the Internet and other digital tools are seen as empowering journalists to do their jobs better, but journalists do not use its potential to create new products.

This kind of insight on the use of computer-aided tools by Dutch journalists is scarce. Hermans et al. (2009 & 2011) state the Internet is mainly used to follow the news, check facts and search for background information. However these conclusions were based on a larger inventory of Dutch journalistic practices in general. Both structured research on the use of computer-aided tools by Dutch journalists as well as ethnographic observations of the research process of journalists are scarce (Brants & Vasterman, 2010). This article aims to fill this gap by using the same approach of Machill and Beiler (2009) and answers the main research question: *In what way are digital (re)sources incorporated in the working process of Dutch journalists?* It gives more detailed information by answering the sub-research questions: 1) *To what extent are online databases and search engines used?* 2) *Which search techniques are applied?* 3) *Do we see differences in search strategies between experts and novices?*

Since we focus on the online information seeking behaviour of journalists, we use methods and insights from information science and library studies. Case (2007) states that information seeking is commonplace until time pressure makes it a concern and the need for new information is satisfied. Especially within the context of journalists, time and deadlines play a big role. An important step in the production process of an article is the angle or perspective of the story (Attfield & Dowell, 2003). Finding this angle or perspective of the story is reminiscent of the *formulation* stage in the Information Search Process (ISP) as formulated by Kuhlthau (2004). She analysed the information search behaviour of students when writing an academic paper and formulated six stages in her ISP-model: 1) *initiation*, becoming aware of a lack of knowledge, 2) *selection*, identifying a general area, topic or problem, 3) *exploration*, searching the broad domain, 4) *formulation*, forming a focused perspective, 5) *collection*, searching information relevant to the focused perspective and 6) *presentation*, in which the collected information
is presented in a unified form; in our case writing the news item. Although ISP focuses on a different group of information seekers, we claim it is also relevant for journalists when producing a news article since the steps are similar.

Comparing novice with expert information seeking, expertise was found to influence the Information Search Process. In previous research, the expert used more varied sources of information for complex tasks, in addition to sources also used as a novice. Moreover, the expert was found to have an increased interactivity with information sources (Kuhlthau, 1999). We identify two types of expertise; domain knowledge and information retrieval expertise (Wang, 2011). Domain knowledge refers to the knowledge a person has about the topic of his research, e.g. being an expert in Dutch politics. Information retrieval expertise refers to the experience a person has with using online search technology to find information. For example, both domain knowledge and search expertise influence the ability to select the right keywords (Hsieh-Yee, 1993; Wildemuth, 2004).

In order to research the use of digital sources by journalists, we thus first need to identify novice from expert journalists. We did this by focusing on their information retrieval expertise; in the rest of this paper we will thus refer to information retrieval expertise when we discuss expertise. In the discussion of domain knowledge we will explicitly refer to it as such. We did this by developing a survey in which we measured information retrieval self-efficacy, meaning:

“People’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances. It is concerned not with the skills one has but with judgments of what one can do with whatever skills one possesses” (Bandura, 1986)

Information retrieval self-efficacy thus refers to the judgment of one’s capability to search for information online. In previous research, Internet self-efficacy was found to discern between low and high self-efficacious students’ information searching strategies and learning in web-based learning tasks (Tsai & Tsai, 2003). Computer self-efficacy was found to influence individuals’ expectations of success, emotional reactions to computers and their actual computer use (Compeau, 1995). In this article we will further describe the influence of information retrieval self-efficacy on journalistic practices.
2. Method

Building upon the methodological approach of Machill and Beiler (2009) and similar research on the use of digital sources by academics (Kemman et al., 2013), we applied a combined methodological approach by conducting an ethnographic study as well as a survey. This approach allowed us not only to collect quantitative data (survey), but also gave us a better understanding of the actual working circumstances by observing the journalists while they were working on a story (ethnographic study).

We first observed the information search processes of 13 journalists of a Dutch daily newspaper working at different editorial departments. By using a semi-structured topic list, we observed each journalist individually during the time they were researching and writing an article for the daily newspaper, which took about four hours. We analysed how they used the Internet to complement their information sources and made notes on all their activities on the Internet. Afterwards each subject was interviewed using the same topic list we used during the observations.

We conducted the survey online from September to October 2012 amongst a broad group of Dutch journalists. We invited journalists both via email (response rate 12.5%) and also distributed the call to participate via social media and mailing lists. In total 298 respondents participated, working at newspapers, television programmes and magazines.

3. Results

In our discussion of the research questions, we combine results of the ethnographic study and the survey, with quotes from observed journalists. The results of the survey relevant to this paper are available online.

3.1. Demographics

We observed thirteen journalists in the ethnographic study. Of these, ten were male and three female. We selected journalists from a variety of editorial departments: two at World (WO), two at Business (BU), two at Sports (SP), two at Arts (AR), two at Opinion (OP) and three at Dutch News (DN). The abbreviations
are used to identify the quotes below.

Of the 298 respondents to the survey, 196 were male (65.8%) and 102 female (34.2%). Regarding age, five were in the age of 18-24 (1.7%), 54 between 25-34 (18.1%), 58 between 35-44 (19.5%), 89 between 45-54 (29.9%) and 92 were 55+ (30.9%). We are not able to make conclusions regarding the representativity of our sample, as there are no general statistics available about the background of Dutch journalists (Hermans, 2011).

3.2. Usage of online databases and search engines

The value of the Internet for journalistic research is generally agreed upon by journalists in our research:

"Without the Internet it would take much longer to search for information. Besides, where can certain information be found outside of the Internet? I wouldn’t know." (DN2)

On the other hand, they notice the Internet is no more than a tool to collect information to write an article:

"I cannot do without the Internet, but it remains raw material where you have to add something to make it into a product." (DN1)

To address the first sub-question “to what extent are online databases and search engines used”, we first consider the material that is digitally consumed by journalists. We asked for several types of digital material which participants rated on a 6-point Likert scale, where “I don’t know it” is lower on the scale than “never”, ranging to “very much”. We assume that when the score is regularly or higher it is part of the common journalistic practice. We found that only regular text (other news articles, stories, etc.) and still images such as photographs are used regularly or more often by a large majority of respondents. Other types of data such as scholarly publications, statistical data and multimedia are used less often. See Figure 1 for a comparison of the mean and mode of responses regarding the usage of data types.
We found in both the observations and the survey that journalists mainly use generic search tools when searching these data. We asked participants to rate a number of search engines and databases on a 6-point Likert scale, where “I don’t know it” is lower on the scale than “never”. We again assume that where the score is regularly or more it can be seen as part of common journalistic practices. We found almost all the respondents (97%) use the Google search engine often or very often. Furthermore, we found that Wikipedia is popular (60% use it often or more, 24% use it regularly) and Google Images (60% use it often or very often). However, other Google products are less commonly used. Only 13% of journalists use Google Scholar regularly or more. Instead, in the observations we found journalists searching for scholarly publications use the main Google search engine. For Google Alerts only 29% of respondents use it regularly or more, while this alerting service could be an interesting tool for journalists to keep up-to-date with recent developments or news. On the contrary, 13% did not know what Google Alerts was at all, one journalist noted in the observations:

“I do not know Google Alert. Are those sources reliable?” (SP1)

For a comparison of the mean and mode of responses regarding the usage of several search tools and websites, see Figure 2.

Figure 1: Mean and mode responses to “Which of the following digital data or sources do you use professionally?” ordered by mean score (N=298).
3.2.1. Trustworthiness

As can be seen from the above quote regarding Google Alert, an important factor in the use of a website is the (perceived) trustworthiness of this website. The most important reason to trust a website is a journalists’ previous use and experience through trial and error; in the survey 84.6% of respondents state this is a reason to consider a website trustworthy. This finding is confirmed in the observations, where journalists relied heavily on their bookmarks (see below for a further analysis); one journalist stated:

“ Websites I use are selected by earlier experience, proven trustworthiness and journalistic standards these websites impose upon themselves.” (WO1)

Furthermore, we asked respondents to rate their trust in search engines and databases on a 5-point Likert scale, ranging from “very little” to “very much”, with an additional “no opinion” option. From this, we found about 60% of the respondents indicate Google as a trustworthy source for finding information. Wikipedia, which is filled with user generated content, is agreed much or very much to be trustworthy by 36%, while 46% remains neutral. We found a significant, albeit small, correlation between trust in and usage of Wikipedia using Kendall’s tau $\tau = .265$, $p$(one-tailed)$<.001$. This correlation is even stronger for Twitter; while 48% have only little or very little trust in Twitter, people who use Twitter more often are also more likely to trust it more; we found a significant correlation using Kendall’s tau $\tau = .476$, $p$(one-tailed)$<.001$. 

Figure 2: Mean and mode responses to a subset (9/22) of “Which of the following search engines, websites or databases do you use?”, ordered by mean score (N=298).
From the observations, we found that Twitter and Wikipedia are mainly used to gather more background information and gain a better understanding of a topic, instead of functioning as direct input for an article, as stated by one journalist in the observations:

“I have a strong faith in Wikipedia. However, I use it mostly to inspire ideas for an angle and to confirm certain information.” (SP2)

3.3. Search techniques

Interaction with the search engines and databases discussed above is largely dependent of the search techniques a journalist employs. When having little information retrieval experience, the sheer size of information available can be a demotivating experience:

“You can search very focused, but there is so much information on the Internet.” (EC1)

With regards to the second sub-question “which search techniques are applied?” we look at functionality in search engines, databases or other websites that are employed by journalists to enhance their search process. We found in the observations that a common starting point is the list of bookmarks with links to websites and experts. Journalists with higher information retrieval expertise, create these bookmarks in their browser or with specialized start pages such as MyPIP. Information obtained from these bookmarks is also used to improve search keywords. As such, they influence the entire search process. Furthermore, in the survey we asked participants to rate several search functions on a 6-point Likert scale, where “I don’t know it” is lower on the scale than “never”, ranging to “very much”. We again assume that when a search function scores “regularly” or higher, it is part of the common search process. We found that keywords are the most important search ‘technique’ (see Figure 3) and the only one that is commonly used by a large majority of respondents. In line with Wildemuth (2004) we found that the more knowledge a journalist has about the topic he writes about, the better this journalist is able to formulate keywords in order to find the (right) information he or she needs.

“Combination of keywords ensures that what I am looking for, ends up high in the
search results. The rest of the results is not useful for me.” (W02)

In our observations we found that other advanced search techniques like Boolean operators or filters are only rarely used. In our survey we additionally found that 66% of the respondents use the advanced search page regularly or more. For Boolean operators we found that 44% never use these, while 31% use these regularly or more. We found a similar result for search filters; 34% never use these, while 29% use these regularly or more. For a comparison of the mean and mode of responses to the application of different search techniques, see Figure 3.

![Figure 3: Mean and mode responses to “While searching the web, which of the following options do you use?”, ordered by mean score (N=298).](image)

3.4. The influence of expertise

Regarding our third sub-question “do we see differences in search strategies between experts and novices?”, we measured information retrieval self-efficacy to analyse the influence of information retrieval expertise on journalists’ search strategies. We did so by letting users rate the questions below on a 5-point Likert scale from 0-4, ranging from “very little” to “very much”. The validity of these questions as an information retrieval self-efficacy scale was evaluated in other research (Kemman et al., 2013).

I’m confident that I know how...
1. ...to use filters to refine search results
2. ...to know which search engine would suit my search task best
3. …to appropriately use advanced search options
4. …to learn new functionality without a user guide
5. …to learn new functionality with a user guide
6. …to use Boolean operators
7. …to use Google’s search operators

By taking the average of these seven questions, we calculated the information retrieval self-efficacy scores for all the respondents. For a distribution of respondents, see figure 4. The average score was 2.33 (S.D.=.78). To compare between novice and expert respondents, we divided the respondents using the mean score, in line with Tsai & Tsai (2003). We thus define respondents with low self-efficacy scores below 2.33 as novices (N=150), and respondents with high self-efficacy scores equal to or above 2.33 as experts (N=148).

![Figure 4: Distribution of information retrieval self-efficacy, the line represents the mean where we distinguish novices from experts (N=298)](image)

To compare search strategies between these two groups, we performed a MANOVA with Pillai’s trace in which we compare the use of search functionality between the two groups. For this we found significant differences exist with F(8, 289)=7.877, p<0.001. From the between-subjects analysis, we found that for Boolean operators, the use of a thesaurus, advanced search options, visualizations,
filters and browsing of subject categories, the expert group scores significantly higher than the novice group, but not for the use of keywords and related terms.

We also performed a MANOVA with Pillai’s trace in which we compared the use of search engines and databases between the two groups. We selected a subset from the search engines and databases present in the survey, comparing those that we found interesting, see the list in figure 2. For this we found significant differences exist with \( F(9, 288)=3.777, p<0.001 \). From the between-subjects analysis, we found that experts score significantly higher compared to novices for Google Images, Google Alert, Google News and YouTube, but not for Google, Google Scholar, Wikipedia or Twitter.

In short, higher information retrieval expertise is related to a wider use of search engines and databases and a wider use of search techniques.

In the observations, we found that in the exploration phase, where journalists diverge their searches to find broad information, experts require less time to determine the angle of the article than novices. Experts use more advanced search strategies, as confirmed in the survey, and are sooner able to determine their angle of an article so the writing process can start earlier. Moreover, experts are faster at finding better information related to this angle, resulting in a more effective collection-phase. Additionally, domain expertise might lead to journalists accessing information directly, instead of searching via a search engine, providing another time advantage:

“More domain knowledge means better search strategies. Often, a search engine is then unnecessary for me.” (DN2)

As all journalists worked with the same deadline, the time advantage in determining the angle results in longer collection (where searches converge as they become increasingly focused) and presentation phases. However, the collection and presentation phases do not occur serially, but simultaneously. Upon deciding the angle to the story, information found in the exploration phase that is no longer relevant is dismissed. The effectiveness of the collection phase search is influenced by the strength of the angle; the more focused the angle, the better the journalist can focus the searches. As such, the search process is also influenced by domain expertise; knowledgeable journalists are better at finding the right keywords and quicker in selecting information. Novice journalists on the other hand fall back on other sources like press releases or articles from other news media:
“This is a clumsy website; I can find little information here. So, I go back to the press release.” (SP1)

This journalistic information retrieval process can be illustrated in the shape of a diamond, see figure 5.

![Figure 5: The search diamond: Journalistic information retrieval processes for novices (above) and experts (below), with identified phases as described by Kuhlthau (2004). Novices need more time for the exploration phase and thus have less time for the collection & presentation phase. Experts formulate an angle quicker, allowing them more time to write the article.](image)

4. Discussion

4.1. Using digital sources

After examining the results of our survey and ethnographic study and discussing the three sub-questions, we can return to the main research question: To what extent are digital sources incorporated in the working process of Dutch journalists? In general we see that the journalists who participated in our research use standard tools to find information on the Internet, mainly Google. More specific search tools like Google Alert or Google Scholar are used only by a small group of journalists. These results are in line with the research of Spyridou et al. (2013) for the Greek situation and Machill & Beiler (2009) for German journalists who also concluded online tools are only used as basic research tools. Moreover, we found
that higher information retrieval expertise is related to a wider use of search engines and databases and a wider use of search techniques, in line with Kuhlthau’s distinction between novice and expert information searchers (1999).

In our observations, we confirm that the search process of journalists is similar to Kuhlthau’s ISP model. Journalists started with a general idea for a news story (initiation & selection), followed by broad searching of information (exploration), after which an angle to the story is found (formulation).

After this angle has been found, journalists search more focused for relevant information (collection), and write their news article (presentation). One difference with Kuhlthau’s Information Search Process is that the collection and presentation phases are performed mostly simultaneously. Another difference is that we considered the initiation and selection phases as a single step, as this study focused on the search for information related to the final product, the article.

4.2. The benefit of digital sources for journalists

Pleijter and Deuze (2003) describe the increase of information in the newsroom. However, our observations show that this is mainly profitable for those who have higher information retrieval expertise and are capable of using advanced search strategies. A direct consequence of not being able to find the right information is dependency of others. Davies (2008) already pointed to the reuse of information by colleagues with sources that stem from other colleagues or news agencies, with or without mentioning the sources. In the Dutch study ‘Gevaarlijk spel’ (‘Dangerous game’) (Prenger et al., 2011), the relationship between journalists and public relation officers working for commercial organizations and the government has been further analysed. Our results substantiate the dependency on external sources when search techniques are not optimally used and there is insufficient time for efficient research.

Since search techniques to find information on the Internet and other digital sources have become such an important journalistic skill, our research shows training these skills deserves more attention in the education of current and future journalists. Moreover, we found that more domain knowledge led to more effective search techniques. As such, we underline the plea of Anderson, Bell & Shirkey (2012) for a further specialization of journalists in specific journalistic disciplines.
As we have shown with the search diamond, these skills will assist journalists to find better information in a quicker way, allowing journalists more time to focus their information and write their article.

Notes

[1] Corresponding author. Contact via kemman@eshcc.eur.nl.


Literature


Acknowledgements

We are grateful for the financial support from the EU FP7 project AXES – Access To Audiovisual Archives ICT-269980. Moreover, we thank the following people and organisations for their cooperation and input: Piet Bakker, Manouk van den Broek, Arjan Dasselaar, Henk van Ess, Harmen Groenhart, Wieneke Gunneweg, Eric Hennekam, Amanda Koopman, Alexander Pleijter, European Journalism Centre, NVJ, Vereniging van Wetenschapsjournalisten, VOJN.