SEARCH ENGINE OPTIMISATION IN UK NEWS PRODUCTION

Murray Dick

This paper represents an exploratory study into an emerging culture in UK online newsrooms – the practice of Search Engine Optimisation (SEO), and its impact on news production. Comprising a short-term participant observational case study at a national online news publisher, and a series of semi-structured, depth interviews with SEO professionals at three further UK media organisations, the author sets out to establish how SEO is operationalised in the newsroom, and what consequences these practices have for online news production. SEO practice is found to be varied, and application is not universal. Not all in the UK media are making the most of SEO, even though some publishers take a highly sophisticated approach. Efforts are constrained by time, resources and management support, as well as off-page technical issues. SEO policy is found, in some cases, to inform editorial policy, but there is resistance to the principal of SEO driving decision making. Several themes are established which call for further research.

Introduction

The UK media have been active online for around 15 years, but it is only in the last few years that newsrooms have seen SEO practice having a direct impact on journalistic work-flow. The incorporation of techniques designed to ensure high ranking in search engine results pages (SERPs) means that journalistic output is increasingly written not for the benefit of the audience per se, but for those popular search algorithms which determine 'relevance'. This raises a number of important questions with regard to traditional news values as applied in the selection, creation, presentation and distribution of news output. As competition for readership moves away from the news-stand and towards the browser, this research aims to trace how SEO, which seeks to drive organisational behaviour (Moran and Hunt, 2006) is informing UK journalism, and what consequences SEO is having for our news and wider culture.

Literature

This research overlaps by degrees with MacGregor's (2007) study concerning the influence of server tracking data in the newsroom. MacGregor sets out three strands in the literature relevant to his study; journalists' attitudes to their audience, market pressures in the web era, and news values. While the author acknowledges the relevance of these themes to the present study, he chooses instead to focus on scholarship in SEO, and an emerging 'rise of the clickstream' literature which addresses those areas where market pressures impact on journalistic values and routines.

SEO: an introduction

SEO is the blanket term for a range of practices whose end-purpose is to increase

organic search traffic referrals to web sites. It comprises four elements; keyword research, search engine indexing, on-page optimisation and off-page optimisation (Malaga, 2010). In theory therefore, the processes involved in SEO, from establishing appropriate keywords with which to frame (or inspire) stories, to developing relationships with external sources with whom to share links (link-bulding) can be significant to newsgathering, news production, and news distribution – the full online news production chain.

SEO is predicated on the assumption that people search with intent, rather than allowing serendipity to guide their interests, and that by tracking keyword search referrals, web publishers can gain an insight into the psychology of the online audience (Enge et al, 2010). The process offers a means of tracking the language and idioms used by search engine users, which can then be mirrored in the language and idioms used by web publishers in order to maximise search traffic.

Search engines are an important source of traffic to many news websites today. In May 2009, telegraph.co.uk received around a third of its traffic from search (Coles, May 2009), down from July 2008, when the figure was said to be around 50% (Charlton, 24/07/08), but nevertheless still significant.

High ranking in SERPs can have a significant impact on search-driven traffic. Around 62% of search engine users click on a result in the first page of results, and 90% within the first three pages (iProspect and Jupiter Research, 2006). A previous study of AOL server logs found that 42% of users tend to click on the first placed result in SERPs, while only 12% click on the second placed result, and 9% on the third. The first ten results receiving 90% of all click-through traffic, and the second page only just over 4% (Enge et al, 2010).

Ranking factors

Google is the pre-eminent search engine in the UK. In July 2010, Google domains accounted for more than 90% of the UK search market (Experian Hitwise Data Centre, 2010). Google claims to use over 200 'signals' in addition to their patented PageRank algorithm, in order to determine the ranking of web pages. But beyond providing general guidance and support on how to optimise for search, Google is evasive both on what these 'signals' are, and on how they are weighted in ranking. The lack of transparency around this issue stems from a fear of giving web publishers free reign to 'game' the search engine ranking system (Moran and Hunt, 2006).

In the absence of official guidance, the biennial SEOMoz search engine ranking factors survey, which aggregates the views of 72 of the world's leading SEO practitioners (SEOMoz, 2009), offers practice-based estimations. The key ranking factors in this study include:

- Keyword used in HTML title tags (online news headlines are normally placed in title tags)
- Anchor text used in inbound links: where external sites use rich, descriptive words when embedding hyper-links to a site, this site will experience a higher ranking.
- Global link authority of site: the authority (in ranking) of any given page in a site is influenced by the site's overall link authority (derived at via PageRank). The PageRank for a news domain is therefore a significant factor in the ranking of each of its individual stories.
- Age of site: sites which have developed their link authority over a longer period of time generally have a higher PageRank.
- Link popularity within the site's internal link structure: well-structured collections of links within websites are used by search engines to gain a sense of what publishers themselves consider their 'most important' content.

Relevant 'related links' sections, and well-structured topic (subject) pages contribute towards improved ranking.

- Topical relevance of inbound links: Reciprocal (and 'relevant') links between online sources who are concerned with the same or similar topics will boost the search ranking of both sources.
- Keyword use in body text: keywords and synonyms carried over from Title content into body text will boost ranking. Keyword use in file names, and in other fields (such as picture captions) are also ranking factors, albeit less significant.

The speed at which users navigate back to search results (the less time, the less 'relevant' the link clicked on is considered to be in terms of the search terms used) can be used in ranking. User data (or 'contextual relevance') including a surfer's location, search history, profile data (where applicable), and language preference are also weighted. There is disagreement on the degree to which these factors impact on relevance. Only the top three factors in this survey were agreed upon by over 60% of respondents, and indeed the degree of consensus about each of the other factors varies considerably.

Some factors have either declined in significance over the years, or are no longer considered significant at all. For example, keywords placed in the the HTML keywords tag no longer play any part in Google's ranking algorithm (Cutts, 2009), and keyword density, which used to be considered significant (Moran and Hunt, 2006), is no longer considered relevant.

Many of the UK's leading online news organisations now employ SEO specialists – The Telegraph have an SEO executive (Press Gazette, 2010), while the Daily Mail appointed an SEO manager in July 2008 (Oliver, 2008). SEO can be controversial – The Mirror's Associate Editor Matt Kelly has publicly criticised both its principles and practitioners (Andrews, 2009), and indeed, accusations have circulated concerning unethical (or 'black hat') SEO practices in the UK media (McGaffin, 2005; Cook, 2010).

On the other hand, the keyword research and trend-spotting elements of SEO can also be used to inspire original journalism (Broomhall, 2009), as well as offering insights into wider social trends (Tancer, 2009). For Richmond (2008), the rules of SEO should be as important as the 'Five Ws' and house style in the construction of online journalism, while for others it is considered more important yet (Niles, 2010). The online medium, Richmond argues, merits its own conventions, rather than re-packaged conventions carried over from print media. Basic on-page SEO techniques are in his analysis, beyond debate in ethical terms, and the rules of SEO are value-neutral (Richmond, 2008). But this belies the fact that 'relevance' in search is a quantitative assumption, weighted by a trademarked algorithm owned by a powerful corporate entity: Google.

The rise of the 'click-stream' in newsrooms

The interpretation of search data in newsrooms can be problematic. Outing (2005) demonstrates that traffic statistics do not exist in a vacuum, and that a surge in traffic to news sites can result from a link from a popular site, or a prominent ranking in a search index. Outing does not consider the influence SEO techniques can have in establishing such links, and in ranking.

Both Outing and MacGregor (2007) found that online news editors in their respective studies were exerting caution as to the use of (and conclusions drawn from) readership metrics. MacGregor shows that traditional news (and brand) values represent a counter-weight to commercial imperatives brought about by the increasing access to (and prominence of) click-stream data in newsrooms – the potential dangers of which were first articulated by Cohen (2002), in a re-consideration of McManus (1994). Nevertheless, while the 'nose for news' is ceding power to reason, ratings and reaction, otherwise termed

a 'culture of the click' (Anderson, 2008), in practice server data is not found to be unbalancing the relationship between an essentially passive audience, and the 'gatekeeping' editor.

Dissenting in part from this view, Lee-Wright (2009) has established that page hits influence story priority at BBC News Online, even though the lack of supporting data are acknowledged to skew assessment. Lee-Wright noted a disjuncture between "audience predilection and the News editors' perception of news priority" (Lee-Wright, in Fenton, Ed, 2009, p77), whereby 'hard' news stories continue to be prioritised over 'soft', more popular stories in terms of placement on topic pages.

This more nuanced interpretation is taken further by Currah (2009), whose research bears witness to a "growing tension between editorial values and knowledge of what will actually generate revenue" (Currah, 2009, p48), with the rise of what he terms 'the clickstream'. Currah's research gives voice to a central concern in SEO. While some feel search and traffic data can improve the quality of journalistic output by providing editors and journalists with a means of knowing what their readers want, others fear less popular public-interest stories will cede ground to "populist, click-friendly topics" (Currah, 2009, p48) in news-gathering and editorial decision making.

Currah maintains that the increasing visibility of click-stream data in the newsroom exerts pressure on journalists "to justify the value and relevance of news stories to the brand and the bottom line" (Currah, 2009, p78). As people increasingly go online for their news, the click-stream will "assume an even more important role in the shaping of the news agenda" (Currah, 2009, p88). Those media who opt to pursue the large volumes of traffic associated with soft, populist content, should be careful not to "erode the distinctiveness of the brand and its connection to a specific audience" (Currah, 2009, p89). Currah's analysis does not, however, accommodate the possibilities of diverging off- and on-line news brands.

Methodology

This research employs two methodologies; a short-term participant observation case study, and three series of semi-structured, depth interviews. All subjects have some grounding in the theoretical principles of SEO, though by varying degrees of technique and application. They are all SEO practitioners working in the UK news media, whose labours are exclusively online.

In total 10 hours of interactive observations were undertaken with BBC News Online SEO executive Martin Asser over three separate days: 02, 16 and 17 June 2010. These were followed by a post-observation interview conducted via email, sent on 27 July. During the period of observation, Martin provided several files and data, which allowed the author to check for bias in observed behaviour. The observation was designed to coincide with a discrete SEO campaign at BBC News Online – coverage and treatment of the group stages of the World Cup 2010.

Three semi-structured, depth interviews were undertaken with two SEO staff at The Guardian; Paul Roach (Technical Lead for SEO) and Chris Moran (SEO editorial executive). One telephone-based interview with Paul Roach took place on October 08, 2009, and two face-to-face interviews with Paul and Chris took place on November 20, 2009, and June 18, 2010. Two semi-structured interviews were undertaken with Northcliffe Digital Director Robert Hardie (on June 23 and July 07, 2010) with added input from Northcliffe's Head of Audience Response and Optimisation via email. A further semi-structured, depth interview was undertaken with an anonymous online journalist at an undisclosed national news site on July 01, 2010. In total, these interviews comprised 10 hours. Short-hand notes were taken during observations and interviews, and findings were summarised from all notes in around 11,000 words.

The subjects chosen for this research were not selected at random – the field of SEO professionals in UK news is small, and the author was guided primarily by access.

Questions asked during observation and interviews covered many areas, including:

- How long an SEO policy has been in place,
- How SEO policy is operationalised and applied
- What (if any) impact SEO has had on traditional journalistic work
- The extent to which SEO informs editorial
- How search traffic (and search terms) are measured and audited
- What (if any) resources are used to track 'buzz' online
- Whether or not stories have been inspired by search data
- Whether there is a linking and/or link-building strategy in place
- What (if any) are the consequences of SEO on news language, and on news values.

These research themes were developed iteratively, and framed according to subjects' experience of terminology and process. They were informed by the literature in SEO, as well as by the author's knowledge as a consultant to industry. A mixture of open- and closed-questions was used. The weighting of responses is broadly proportionate to the range of participants, but reflects individual concerns.

Participant observation case study findings

Background: BBC SEO policy, and the SEO Specialist

In April 2009, representatives from the BBC's Future Media and Technology Journalism (FM&T J) and Editorial Development (News) divisions worked on a trial SEO project for implementation across the World News team, focussing on the Middle East index page. Following a successful up-turn in traffic during this period, the chief implementer (Martin Asser - the subject for this research) was appointed SEO Specialist, on a short-term, rolling contract. Martin reports to the Head of Editorial Standards in BBC News, though his working relationship with those in FM&T J has been complicated by organisational change. Between October 2009 and March 2010, around 600 BBC journalists were trained in basic, on-page SEO theory and practice. Martin's job is, as he sees it, to monitor output and cascade best practice down to sub editors and journalists.

Day-to-day: keeping track of referrals, terms and competition

Martin uses various tools in his SEO work. The BBC has an account with Experian Hitwise, which provides details of performance on downstream clicks from Google and Google News (the two greatest sources of traffic to BBC News Online by a significant margin), against competitors (Guardian.co.uk and Telegraph.co.uk). Other competitors (such as the Independent.co.uk, theSun.co.uk, and DailyMail.co.uk) are also factored into analysis. Bespoke graphs are used for specific SEO projects – factoring in other competitors in niche fields.

Search referral data from an in-house system called Sage Analyst is used to inform Martin's keyword research. This data is neither provided in a user-friendly way, nor is it exhaustive for SEO needs. There is no breakdown of search terms by UK region (or indeed by country). The system does not appear to provide analysis of IP addresses, so it is not possible for Martin to monitor the proportion of visits that are new and returning. All BBC journalists have access to audience metrics via the "Livestats" service. A dashboard provides interactive heat maps of click-through for the top 10 pages on the BBC news site, which shows Martin how well headlines on index pages are performing in real-time.

Martin undertakes sporadic competitor analysis, tracking the development of a

particular story in terms of ranking performance in Google and Google News, but he does not have time to do this in a more systematic way. Tools for keyword research and monitoring search and social web trends are not generally used: "the philosophy is not to follow trends... it is to continue following our news agenda as we have always seen it, but to make sure we are availing ourselves of current SEO techniques in order to get the maximum amount of traffic from search for the agenda that we are following" (Martin Asser, personal communication, July 28 2010).

An instinctive knowledge of BBC news style drives SEO policy at the BBC, and preferred spellings and transliterations from BBC Monitoring take precedence. The BBC do not have a uniform, written SEO styleguide which journalists can consult. Martin will never alter headlines directly – this is editorially forbidden.

BBC News has no active presence on YouTube, though its videos are available via some other third party aggregators. Martin is fully aware that his SEO activity, in particular his ability to grow search referrals, is curtailed by this lack of presence, but the BBC has significant rights issues to consider. Martin has occasional meetings with BBC News' head of social media but they do not work closely. This compromises his ability to fully grasp the potential in SEO for planning, adapting to and developing story ideas around online 'buzz'. But having monitored search and social web trends in the past, Martin believes that many of the keyword topics which trend tend to be of little value to BBC News audiences.

SEO in practice

Martin spends most of his time working at the sub-editors desk in the BBC News hub. His day-to-day work on SEO is compromised by other project work which is considered more pressing. Throughout the observations, Martin was involved in dealing with problems with the latest release of the BBC News content management system, which he had been involved in for several weeks. Martin explained that this causes him frustration, because it means SEO is neglected. Martin does not attend editorial meetings, as he doesn't have the time, but he is well aware of any given day's news agenda.

When not working on particular SEO projects, Martin is concerned with fine-tuning headlines on stories linked 'above the fold' (prominently) on index (topical) pages. He checks that headlines are adequately keyword-rich and well-structured, and that links in 'related links' sections are up-to-date. This work is very much about influencing practice, rather than ranking.

Martin's project work takes a news diary approach. The pre-election budget in March, the general election in May, and the World Cup in June, all formed the basis for discrete SEO projects in the Business, Politics and Sports online teams respectively, throughout 2010. By choosing forward-planned stories around which to concentrate efforts, Martin is working around what could be natural increases in the volume of searchers seeking out BBC information, which could make it difficult to attribute search performance to SEO actions taken.

BBC News SEO policy aims are to increase the percentage of visits from search, and the percentage of total visits more generally. Without any conscious effort in SEO (prior to current arrangements), BBC News Online has achieved a PageRank matched only by the Guardian in UK media – reflecting the age of this site, and its wider online authority. Martin sees SEO as part of the public service approach the BBC must take. SEO is undertaken in the interests of the license fee payer, and in the wider public interest.

All of Martin's communications on SEO are by way of suggestion and promotion, though he does have the authority to communicate his concerns to senior news management. When Martin sends out emails about headlines, he will often send them to the senior journalist sub-editors group, carbon-copying in the author of the piece. Martin

has rarely experienced resistance or criticism, and being a journalist by trade, he is sensitive to his colleagues' working environment.

SEO and technical issues

The first sentence of any news story is automatically inserted into Meta 'description' tags by the in-house Content Management System (CMS), but the keywords tag is not used. Keywords are not used in naming conventions for the thumbnail pictures which accompany stories, even though the Google Images vertical is one of the top five referrers for BBC News content. Stories cannot be filed in the BBC News CMS without alternative descriptive text (in keyword form) for all images.

Since November 2009, BBC News has run dual headlines for every story – a short headline (31-33 characters) which features on index pages and Ceefax, and a longer headline (up to 55 characters) on story pages. This approach was taken in order to accommodate more keywords in headline fields. Index (short) headlines don't have to include keywords so long as they give what Martin calls a "whiff" (Martin Asser, observation, July 2, 2010) of the news. Through experience, Martin believes these headlines have little bearing on ranking. Some off-page SEO issues are felt to constrain Martin's efforts. The use of unique numbers rather than keywords in BBC News Online URLs (web addresses) is a legacy issue over which he has no control. Likewise, at the time of writing there is no site map (a file used by high-output publishers to improve SEO by automatically notifying search engines about new content), though this will soon change.

SEO project – World Cup 2010

Before the World Cup started, Martin had meetings with key editorial figures in BBC Sport to discuss SEO issues. A week into the World Cup, Martin put together a 'dashboard' of metrics on how the site would be monitored on a weekly basis throughout the campaign – made available to Sports staff. Martin used June 01 (10 days before the tournament started) as the benchmark for comparisons – he acknowledged that this is "arbitrary" (Martin Asser, observation, June 17, 2010) but felt he had to start somewhere.

After the first week of the tournament, Martin sent a round-up email to subs and other senior journalists outlining performance. The proportion of search (as a total of page views) peaked at 6.1%, averaging at around 4.5% throughout the week, around 2% higher than pre-tournament referral traffic. Martin was pleased that the second and third highest referring keywords on the June 14 were BBC "branded searches" (Asser, observation, June 17, 2010) – navigational searches including the term 'bbc' (these were 'bbc world cup' and 'BBC world cup 2010'). These navigational searches are of value, as they imply searchers are actively seeking out BBC News Online.

During the second day of observation, Martin sorted through the top 250 search terms, for the June 12 and 13. Those keywords yielding 30 searches or less were discounted, and the rest integrated into a pre-existing keyword list. Martin passed over those terms which have stopped trending since his monitoring began, so there is a danger he misses out terms whose popularity ebbs and flows over a longer period of time.

Martin's second weekly email to Sport news staff (sent on June 17) included a potential newsgathering issue:

'brazil fans – 221 searches on the 14th'
'brazil fans – 763 searches on the 15th'

Martin drew this issue to the attention of the Sports team, but in a very informal way – with good reason. A search of Google Images showed that this search term returns images of near-naked female Brazil fans. Martin does not feel it is his job to suggest

editorial treatments of data like this – he is happy just to supply the information, and let journalists and editors interpret it as they would any other source.

Kickers

Martin has found that inserting trending keywords into kickers (a print convention, whereby headlines are prefixed with a two- or three-word clause followed by a colon) across major on-going news stories can have a beneficial effect on traffic. He says he feels that this approach has been useful in instilling discipline in the process of headline writing. On June 14, the third most popular search term was 'BBC World Cup 2010'. The kicker 'World Cup 2010:' was applied to much World Cup coverage in BBC Sport, and proved to be popular across much of the UK's online media for several weeks (subsequent research showed some UK media had been using 'World Cup 2010:' in headlines since the first week of May 2010). Martin acknowledges the negative impact this approach has on SERPs for world cup content. He agrees that these Google News results are difficult to interpret – they are "a sea of information" (Asser, observation, June 17, 2010) – but he feels he is locked into an arms race. He sees that he has a public duty to compete, and that his competitors have a commercial need to compete, but can see no way around this problem, other than to vary their use from time to time, and so potentially lose traffic.

SEO and 'hard to find news'

On the third day of observation, Martin worked on a story about CCTV cameras in Birmingham that were being covered over with plastic bags to render them ineffective. Martin spent much time trying to re-write the original headline which started with the sub-optimal keyword 'Plastic bags'. It proved to be very difficult to re-write – all of the keyword elements were conceptually difficult. The cameras could be referred to as 'terror cameras', 'cctv cameras', 'surveillance cameras', or even 'spy cameras' – as was found in past BBC news output, and amongst the BBC's competitors.

Martin concluded that the headline was far from ideal, but also that it was a very difficult story to bring out in headline keywords – there were so many options it would have taken half an hour to consult keyword tools, and the competition, to check them all. Some stories, he acknowledged, are hard to optimise for, even relatively 'big' stories (this story was the second most prioritised story on the UK News index at the time).

SEO and the politics of news language

I ask Martin for his views on a story from late May 2010 concerning the murder of three prostitutes in West Yorkshire. I suggest that the prominent use of the keyword 'prostitute' in headlines might imply SEO values are impinging on the ethics of formal style across the media. Martin says he has no desire to be "prurient" (Asser, observation, June 17, 2010) in his pursuit of increased traffic. This remains essentially an editorial issue.

Internal linking, Link-building and linking out

Martin has input into the naming conventions used within the BBC News Online information architecture. During the third day of observation, the BBC news topic page concerning BP's Deepwater Horizon oil spill was called simply 'Oil and water' (albeit, the title for this page was 'Oil Disaster'). The visual heading for this page was inspired by the branding for a World Service radio series on the disaster by this name. However, accommodating a much wider pool of BBC multi-media content, the page subsequently reverted to the far more SEO-friendly title 'BP oil disaster'. Martin is not involved in any

formal link-building (establishing contact with other sites in order to share links, to boost ranking). On the BBC's general policy on linking out, Martin said that they are at "a crossroads" (Asser, observation, June 17, 2010). The public service brief informs a policy of increasing in-line (body text) and deep (direct) linking, but there has been some concern that this could lose the BBC traffic.

Interview findings

Background

The Guardian's in-house SEO team was formed in 2008, and comprises three members including; Paul Roach (Technical Lead for SEO, and SEO by trade) who is based in Technical, and Chris Moran (SEO Editorial Executive, and sub-editor), who is based in Editorial. At Northcliffe, an SEO specialist has been in place since 2006, working on the 'thisis' sites to Digital Director Robert Hardie. An undisclosed media group's SEO history is not shared (hereafter referred to as Y, and the SEO journalist subject X).

SEO culture in the news room

All Northcliffe and Guardian journalists are trained in SEO, but most SEO efforts are undertaken by Guardian sub-editors, Northcliffe Digital Editors and Y's sub editors respectively. There is no blanket SEO policy nor catch-all guide for writers at any of these organisations – and for Guardian and Northcliffe copy, optimisation is prioritised for those stories featuring on the home page, or prominently on topic (index) pages. At the Guardian, Reviews and Sport have particular SEO requirements.
Not everyone sees the importance of SEO to the same degree within these organisations. Guardian sub-editors prioritise headlines, standfirsts, and picture captions, though journalists select suitable (internal and external) related links. Body text is not a priority for SEO treatment at any of the three organisations.   Both Paul and Chris are wary of impacting the reading quality of Guardian copy by forcing (and overly repeating) keywords. Topical in-body links to internal keyword pages are automatically generated on all three sites. At the Guardian, the SEO team aim to look beyond isolated stories towards creating topically-consistent hyper-linked portmanteaus, to improve the site's wider optimisation. Topic (subject) pages are generated automatically at the Guardian, and at Northcliffe.
Northcliffe's general approach is to lead with real, intuitive language. The natural use of synonyms is encouraged, albeit top-ranking keywords will feature most prominently. Northcliffe Digital Publishers manually re-work automatically generated topics pages, as well as on-page SEO and link-building for key stories.
SEO is undertaken at Y by sub-editors, not all of whom are enthusiasts. There are no consistent rules on key-wording in HTML meta tags or picture captions.   A significant amount of news (including picture-led news) on Y is wires copy, which is not re-worked for SEO purposes. X estimates that one person will have around an hour to optimise the day's content – with a heavy emphasis on checking for errors in headlines.

Monitoring search traffic

The Guardian's SEO team have access to a Hitwise account, and all three organisations have in-house traffic referral monitoring consoles. At the Guardian, search traffic analysis is ephemeral, though monthly analysis is also carried out. Key metrics include "Unique users, page views, visits to various areas of the site from specific referrer types and level of engagement of each" (Paul Roach, interview, June 18, 2010).

Engagement is measured in several ways; from dwell time on site, through to sign-up, and sign-in to online community accounts, and even conversion to products bought online. Most literature conceives the end-goal of SEO to be a conversion or transaction (Kent, 2008, Thurow and Musica, 2009, Enge et al, 2010), but for the Guardian, the equivalent is a unique visit, followed by click-through to other pages, followed by a subsequent visit from this IP address to guardian.co.uk, via a navigational search (ie 'guardian football'). The difference in click-through between first and second place in search results is considered to be hugely significant, especially when it comes to the 'one-box' Google News vertical results found in Google search results for trending terms. The Guardian SEO team is therefore very proactive in its competitor analysis.

At Northcliffe, search traffic reports are analysed monthly, but there is a daily headline report too. Northcliffe use a combination of Hitwise and their own in-house Audience Dynamics research, data from which is shared with all Digital Publishers.

X and his team are aware of internal statistics throughout the day. Day-to-day search and third-party referral key words are not considered important – responding quickly to web trends and topics is not something Y has the resources, inclination or time to do. Neither is there time for competitor analysis. News carried in Y is published on the website at midnight. These are print stories without SEO fine-tuning which are retrospectively checked, optimised, placed and adjusted on the home page, or landing pages in the hours following publication.

Editorial policy and social web and search trends

Chris acts as SEO representative at Guardian digital editors meetings. At these meetings, and in his communications, he is very careful to only impart trending topics where they have (as he perceives) a resonance with Guardian readers. He is conscious that if he were to be seen to be pushing irrelevant topics, he may loose interest and hence influence within the organisation. This selective approach "leaves traffic on the table" (Chris Moran, interview, June 18, 2010) for competitors, but this is considered a price worth paying to maintain the Guardian brand. The SEO team's contribution to editorial comprises three strands: ensuring staff understand the culture of online news consumption, helping them optimise their content for search, and helping to identify opportunities in treatment, news-gathering and commissioning.

Chris distributes daily social and search trend updates at The Guardian using several freely available online tools. General web trends are considered useful not just to establish blank-spots in coverage in any given point in the news cycle, but also for helping determine when topics are ephemeral or long-lasting among the searching public.

There are no central web trend communications at Y. X dislikes the idea of being driven by trends, but he is aware of the pressure the promise of such traffic drives in industry. Y's current editor is adamant that heavy-trending stories which are "not right" for Y will not be published (X, interview, July, 2010).

SEO does influence editorial process at Northcliffe directly, but web trends are of minimal use in day-to-day work. Search and social web trends are considered to be too nationally-focused for a local news online publisher. There is no formal daily trends email sent round journalists at Northcliffe – but the Audience Response and Optimisation (ARO) team deal with this on a case-by-case basis, contacting particular digital publishers on specific topics.

SEO and news language

At The Guardian, SEO principles always defer to the in-house styleguide, and so Paul and Chris are not immediately concerned about the rise of jargon in keyword

selection. From time to time this is to the detriment of organic traffic. Traditional style conventions in headlines such as 'Formula 1' (as opposed to the more frequently searched 'F1') make optimising for SEO difficult.

Robert, as a former sub-editor, is concerned that search might "reduce the scope of language" (Robert Hardie, interview, July 07, 2010) in news. Any fears about clichés and jargon should be considered in the context of how readers use language, and how they refer to the world around them, rather than blindly following subbing advice from off-line. During the period of interviews, Northcliffe were starting to consider the incorporation of SEO into their styleguide, and Robert was aware that a potential conflict between SEO and conventional writing style is something he will have to confront.

X said he would not use a keyword he didn't think shared the values of Y just to maximise traffic. Use of the SEO-friendly term 'Crossbow cannibal' in referring to the alleged murderer Stephen Griffiths (who used this term to describe himself in Court, in May 2010), led to some editorial soul-searching at Y in terms of showing respect for the victim's families. They opted to run with it, though not from an expressly SEO-driven perspective – this was a decision shared by most UK media. X relies on his news instincts and editorial judgements (via the Y styleguide) when it comes to choosing keywords to research.

When copy-editing for online, X is wary of doing violence to big impact, visual front-page splashes which often use simple, intuitive, but striking headlines. He often ends up re-writing them completely, incorporating keyword-driven pull quotes, sub-headers, and other page furniture, into what he feels can be "a poor imitation" (X, interview, July, 2010). The trick, he suggests, is to "embody the editorially elegance" (X, interview, July, 2010) within such news stories online but this he concedes can be difficult.

Generating news ideas

None of the sources use trend data to inspire news stories or features regularly, and no examples were given of past use.

Kickers

Paul and Chris are opposed to the slavish use of kickers in pursuit of traffic, but acknowledge they have to do it sometimes. Kickers are "the easiest way to ensure the targeted keyword appears in the right order in every article... [but] ...it means that the news headlines are homogenised and you have no nuance within headlines or between publications" (Roach, interview, June 18, 2010). To obviate the problem, they try to "target search terms within the headline and page title in a way that makes sense, is in context and reads well" (Moran, interview, June 18, 2010). Living Stories, a Google experiment in creating interactive pages for big, rolling news events, was mentioned as an example of a possible design-led solution. At present it is felt Google News rewards those publishers who output lots of discrete stories and link well around a key-word relevant topic, rather than bringing ongoing stories around a topic into a rolling narrative on one page. Paul believes this is a crucial issue for the future of SEO in news.

X is aware that kickers are not ideal, and he is averse to the use of phrases in kickers. He feels search engines have a responsibility to ensure that when online publishers follow their terms and conditions (and guidance) to the letter of the law, that the end-user isn't presented with "a sea of near-indistinguishable headlines" (X, interview, July, 2010).

Robert is also aware of a problem with kickers but rather than consider this is an ethical dilemma for publishers, he believes it will be resolved by the market. The public will have to become more attuned to what they want, and so will change their behaviour, or indeed the search engines will have to work a path around this problem, or fail

competitively.

Calendar news stories

Paul suggests that SEO has the power to stretch the news agenda around big calendar stories. He is aware that there is a danger of media groups setting out to pre-prepare topic pages for news events months in advance of the actual event taking place, preparing this content more in the pursuit of ranking in future weeks and months, than in informing public awareness.

X regrets this approach, feeling that "a fair fight" (X, interview, July, 2010) on reporting, insight, and opinion at the actual time of the event in question, should be the primary influence on ranking. X believes this is contributing to information overload.

Hard-to-find news

Paul acknowledges that some stories are harder to find than others, but points out that the Guardian's internal key-wording taxonomy "means that an article will always be linked from several keyword pages as well as edited fronts" (Roach, interview, June 18, 2010). X also acknowledges some stories may be harder to find than others.

Linking

Paul says "Linking out to the relevant sources wherever possible is actively encouraged throughout guardian.co.uk" (Roach, interview, June 18, 2010), including competitor sites. The media "have a duty to link to relevant sources wherever appropriate", but they do not pass link authority on to commercial links or user generated content (Roach, interview, June 18, 2010). Robert sees external linking as an important way of harnessing the full power of the internet, though with similar caveats. When it comes to link-building, Northcliffe publish on average two to three stories per week, which are submitted to Fark.com (an aggregator of 'quirky' news). Northcliffe use a range of third party sites in order to build traffic, establish brand, and develop reciprocal relationships in a number of niche topical areas. Robert mentions a recent example of a local story concerning Cliff Richard, whereby one of his Digital Publishers spent around two hours researching and distributing this story around relevant fan sites and forums, to boost traffic to this page.

Analysis and conclusions

SEO is still very much in its infancy in the UK media, with some organisations still focusing on influencing culture, rather than applying policy. SEO activities are constrained by time, resources and support. Even when a specialist in place, SEO can defer to other tasks and projects. Some technical, rights-based and off-page SEO factors effectively tie the hands of SEO practitioners in terms of maximising traffic. In short, SEO is some way short of being institutionalised across the UK media.

The culture of SEO in these organisations is visible, but not dogmatic. This research found only one example of a media organisation where SEO directs editorial decision-making, for all others SEO is acknowledged to inform (by varying degrees), but never lead editorial, with significant resistance to this idea in some quarters. It is acknowledged that in yielding to editorial (and style) conventions these organisations are losing out on traffic. But this is broadly considered a price worth paying in order to preserve the brand, and maintain the news values of the profession.

Search traffic measurement tools are used across all organisations studied, but

some organisations have a clearer sense of how they are performing in SEO terms than others. Search and social web trends are not being routinely used to identify original areas for reporting, but they do inform news-gathering. Other SEO tools, including keyword research tools, and search or social web trend measuring tools, are not uniformly used in industry. On the contrary, language, style and ethical conventions take absolute precedence over the language and idioms people use to search.

The perceived bias towards 'soft' (including celebrity, entertainment and 'quirky') topics in search and social web trends is considered unfit for sharing with some journalists, and by extension with news audiences. Linking out is considered important amongst all the subjects I spoke to, but with caveats. The process of link-building is not a priority in national news organisations, but a local news publisher was proactive in attempting to build links with external sources towards driving traffic.

This research shows that approaches to SEO go beyond mere 'common sense'. Standards in journalistic writing reflect best practice and ethical concerns, and help publishers pull together a consistent news brand. But this research demonstrates that SEO is applying pressure to these standards not in the perceived interest of the reader or because of publication constraints (as once may have been the case), but in the interests of a third-party commercial arbiter in online distribution: Google.

A new conformity in the language and 'aboutness' of news online is emerging, mediated by Google. Google stands between those terms used by the reading public and journalist alike, implicitly regulating the 'market place of ideas' as it goes. Relevance as a quantitative measure is challenging what people feel qualitative 'relevance' to be. But for now at least, some of the UK's leading news brands are refusing to bow entirely to the demand for more and more search traffic.

This research uncovered various themes which merit further investigation.

The first relates to the use of trending keywords in kickers to prefix ongoing news headlines online. Kickers are an old media convention given new life and significance online. News organisations use this blunt tool because Google is perceived to reward their use, leading to an arms race whereby the surfing public suffers. There is a danger that news organisations will damage both their own brand, and the brands of leading search engines, if they do not reign in indiscriminate use. It has been suggested that were Google to reward news organisations who publish big news themes on one page (as opposed to a myriad of small, discrete stories), that this may obviate the problem. It is prescient that some in industry are calling for an end to the journalistic convention of the 'story' in online journalism (Marsh, in Miller Ed, 2009) at a time when such an approach might help lead to a solution to this problem.

Another theme relates to the increasingly early development of pages for big calender news events, designed with the intent of improving ranking nearer the time the news event takes place, as much as informing the public. This represents another potential arms race in news SEO. This approach has the potential to stretch the definition of 'news' to breaking point, while contributing yet further to the problem of information overload.

Thirdly, SEO professionals accept that some news stories are harder to find online than others by their very nature. Some stories do not lend themselves to simple keyword analysis, and without a firm editorial steer, these stories could be relegated in priority, in the same way that less visual stories are perceived (by some) to have less priority in broadcast news.

Present day SEO practices may be looked back upon in the decades ahead as a relic of early developments in search, as we move towards a 'semantic web'. But SEO will be with us for some time yet, and this research serves merely as a starting point for further study in this field.

References

Andrews, Robert (2009) "Daily Mirror's Matt Kelly: put SEO in its place", **The Guardian**, 2nd December:
http://www.guardian.co.uk/media/pda/2009/dec/02/mike-kelly-seo-journalism-world-newspaper-congress-keynote   accessed 17 December 2010

Anderson, Chris (2009) "Web Production, News Judgement, and Emerging Categories of Online Newswork in Metropolitan Journalism", presented at the Online Journalism Symposium, University of Texas, Austin, 18 April.

Brooker, Charlie (2008) "Online POKER marketing could spell the NAKED end of VIAGRA journalism as we LOHAN know it", *The Guardian*, 21 July:
http://www.guardian.co.uk/commentisfree/2008/jul/21/charliebrooker.pressandpublishing
                                     accessed 17 December 2010

Broomhall, Drew (2009) "Building content and traffic using search trends", presentation for the Association of Online Publishers SEO Forum, London, 25 March.

Charlton, Graham (2008) "Site review: Telegraph.co.uk revamp", *Econsultancy*, 24 July:
http://econsultancy.com/blog/2628-site-review-telegraph-co-uk-revamp
                                     accessed 17 December 2010

Cohen, Elisa (2002) "Online Journalism as Market-Driven Journalism", *Journal of Broadcasting & Electronic Media*, Volume 46, Issue 4 December, pp 532 – 548.

Coles, Malcolm (2009) "Find and Share Information on the Internet : Part 2", *Fumsi*, May:
http://web.fumsi.com/go/article/share/3907
                                     accessed 17 December 2010

Cutts, Matt (2009) "Google doesn't use the keywords meta tag in web search", *Matt Cutts Blog,* 21 September: http://www.mattcutts.com/blog/keywords-meta-tag-in-web-search/
                                     accessed 17 December 2010

Enge, Eric; Spencer, Stephen; Fishkin, Rand; and Stricchiola, Jessie (2010) *The Art of SEO: Mastering Search Engine Optimization*, Sebastapol: O"Reilly.

Cook, Mark (2010) "Express Group contacts SEOs to sell links", *Further.co.uk*, 22 April:
http://www.further.co.uk/blog/Express-Group-contacts-SEOs-to-sell-links-269
                                     accessed 17 December 2010

Lee-Wright, Peter (2010) "Culture Shock: New Media and Organizational Change in the BBC" in: Fenton, Natalie (Ed) *New Media, Old News: Journalism & Democracy in the Digital Age*, London: Sage.

Experian Hitwise Data Center (2010), "Top Sites & Engines":
http://www.hitwise.com/uk/datacentre/main/dashboard-7323.html
                                     accessed 17 December 2010

iProspect and Jupiter Research (2006) "Search Engine User Behaviour Study", April:
http://www.iprospect.com/premiumPDFs/WhitePaper_2006_SearchEngineUserBehavior.pdf
                                     accessed 17 December 2010

Kent, Peter (2008) *Search Engine Optimization for Dummies*, Indianapolis: Wiley.

MacGregor, Phil (2007) "Tracking the Online Audience", *Journalism Studies*, 8:2, pp 280-298.

Malaga, Ross, (2008) "Worst Practices in Search Engine Optimization", *Communications of the ACM*, December, Vol 51. No.2.

Marsh, Kevin (2009) "Death of the Story", in: Miller, Charles (Ed), *The Future of Journalism: Papers from a conference organised by the BBC College of Journalism in London*, 2008:
http://www.bbc.co.uk/blogs/theeditors/future_of_journalism.pdf

accessed 17 December 2010

McGaffin , Ken (2005) "The Financial Times Website and Hidden Links, *Linking Matters*, 11 June:: http://www.linkingmatters.com/2005/06/financial-times-website-and-hidden

accessed 17 December 2010

McManus, John (1994) *Market Driven Journalism: Let the Citizen Beware?*, London: Sage.

Moran, Mike, and Hunt, Bill (2006) *Search Engine Marketing* (2nd Edition), London: IBM Press.

Niles, Robert (2010) "Student journalists need to learn SEO more than they need AP style", *Online Journalism Review*, 21 April: http://www.ojr.org/ojr/people/robert/201004/1843/

accessed 17 December 2010

Oliver, Laura (2008) "Mail Online names Simon Schnieders as search strategy head", *Journalism.co.uk*, 11 July: http://www.journalism.co.uk/2/articles/531917.php

accessed 17 December 2010

Outing, Steve (2005) "Sophisticated Web Stats Give Editors Better Idea of Reader Interests", *Convenience Store News*, 26 July.

Richmond, Shane (2008) "How SEO is changing journalism", *British Journalism Review*, 19;51.

SEOMoZ Ranking Factors (2009): http://www.seomoz.org/article/search-ranking-factors#ranking-factors

accessed 17 December 2010

Press Gazette (2010) "Telegraph creates 'digital operations unit'", 20 April: http://www.pressgazette.co.uk/story.asp?storycode=45325

accessed 17 December 2010

Tancer, Bill (2009) Click: What We Do Online and Why It Matters. London: HarperCollins.

Thurow, Shari and Musica, Nick (2009) *When Search Meets Web Usability*, Berkeley: New Riders.