COMPUTATIONAL THINKING AND JOURNALISM EDUCATION

Radu Meza

(Babeș-Bolyai University)

This paper will review the most important aspects of computational thinking and try to emphasize the most relevant ones for the future of journalism education, by looking at scholarship in the field and assessing abilities associated with computational thinking in the context of journalistic work, from topic identification, to data collection and interpretation, to information presentation and content aggregation and distribution.

From Critical Thinking to Computational Thinking

As critical thinking was gaining momentum in all areas of education in the late 1980s and early 1990s, the concept was still vague and its role in education interpreted in different many ways. A consensus statement regarding critical thinking and the ideal qualities associated with the critical thinker is presented in Facione (1990: 3) as the "purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgement is based".

The ideal attributes of critical thinking as habits of mind in both their positive and negative aspects are summarized as follows (Facione, 1990; Facione *et al.*, 1995):

Positive	Negative
truth-seeking	intellectually dishonest
open-mindedness	Intolerant
analyticity	Inattentive
systematicity	Haphazard
critical thinking self-confidence	mistrustful of reason
inquisitiveness	Indifferent
maturity of judgment	Simplistic

Table 1. Characterological attributes associated with critical thinking

As the traditional media institutions and implicitly media education are undergoing major changes on the backdrop of shifting economic models in the context of the fast and unprecedented development of digital communication technologies, a new concept is being hailed as a solution that should be introduced in education as early on as possible: *computational thinking*.

Early scholarly work describes computational thinking as "a fundamental skill used by everyone in the world by the middle of the 21st Century" (Wing, 2006: 33). According to the same author, at the core of this concept, which seems a generalization of theoretical computer science, are two fundamentals – *abstraction and automation*. Essentially, computational thinking is first conceptualized as approaching problem-solving, system design or the study of human behaviour by relying on fundamental computer science concepts.

As other authors later approach computational thinking in education, the analytical aspect is also embedded in the concept. According to Lee *et al.* (2011), problem solving via computational thinking is defined by:

- *Abstraction* – generalizing from specific instances, stripping down problems to their essentials or capturing common characteristics or actions and using them to represent all other instances;
- *Automation* – a labour saving process whereby a computer executes repetitive tasks more efficiently than a human;
- *Analysis* – a reflective practice that refers to the validation of whether the abstractions made were correct.

Implementing the acquisition of computational thinking in education outside of computer science is a problem frequently tackled by progressive educators. A generally applicable model is the three-stage progression called *Use-Modify-Create* (Lee *et al.,* 2011).

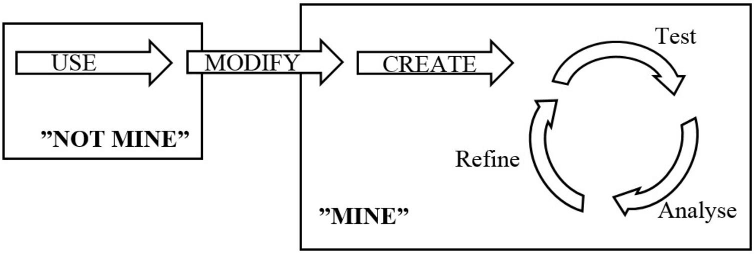

Figure 1. Use-Modify-Create Learning Progression – adapted from Lee *et al.* (2011).

This computational thinking-learning model is highly applicable to the use of free open-source applications and APIs. At the early levels, students would use a piece of software, then after gaining some understanding of the abstractions in use, they would be able to slightly modify the code to perform other automated tasks than originally designed, and eventually they would be able to apply iterative modifications (test, analyse and refine them) up to the point where the result is their own creation.

Journalism is moving in many different ways towards the new digital media, but journalism education has yet to acknowledge that the new media market and new media languages are different than the traditional media markets and languages. Future journalists might need to go further than just critical thinking skills that ensure their intellectual independence and foster creativity. Understanding the principles of the language of new media - *numerical representation, modularity, automation, variability and transcoding* (Manovich, 2001) and becoming educated in computational thinking are key to surviving and thriving in a global attention economy where the biggest players on the online advertising market rely heavily on automation to deliver precisely targeted, apparently innocuous messages at very low costs, thus undermining the news industry's traditional revenue streams.

Academic approaches to journalism education emphasized critical thinking through research, while professional approaches emphasize production skills. According to Reese (1999: 85) "journalism provides a valuable educational setting to explore the redefinition of scholarship as well as other major pedagogical reforms [...] such as collaborative learning and instructional application of information technology". However, "journalism practice need not cultivate intellectual autonomy, as indicated by the lack of self-critical insight exhibited by many of its practitioners, the formulaic recitation of quotes from experts and sources, media feeding frenzies, and pack journalism" (Reese, 1999: 86). Academic approaches to journalism are sometimes more likely to lead to innovation, especially in large universities, where the environment facilitates collaboration between academics and students across different fields, especially in the context of increased interest for computational methods in the social sciences and humanities. Professional environments tend to be more entrenched in traditional routines and rituals and (with some exceptions) are resistant to change, thus it's less likely to foster innovation, especially if such innovation requires interdisciplinary approaches.

Co-citation Networks and Interdisciplinarity

As academic research can open new paths in the subject matter of journalism education, especially when considering the new trends towards computational methods in all sciences (including social sciences and the humanities), it is interesting to look at who writes what in the area of practices associated with computational thinking and journalism.

One of the frequently used methods to look at relationships between ideas as they appear in academic research is studying co-citation patterns. According to Small (1973: 265-266), "if it can be assumed that frequently cited papers represent the key concepts, methods, or experiments in a field, then co-citation patterns can be used to map out in great detail the relationships between these key ideas". In this way, we might come to a somewhat objective way of modelling the intellectual structure of scientific fields and emerging areas of interest. Viewing such structures over several years might provide information that can lead to a better understanding of the development of ideas, concepts, or methods.

Co-citation analysis has always relied on graph theory for visualisation and analytical tools or measures as authors or papers are represented through nodes and citation relations through directed edges. Concepts such as the in-degree and out-degree of a node are useful in measuring the importance of authors or papers as they show how many authors/papers cited a particular author/paper and respectively how many papers/authors a certain author/paper references. The development of the World Wide Web over the past couple of decades has driven the need to better organize and rank web pages (which can be seen essentially as nodes in the web graph) to provide better, more relevant search results. One groundbreaking development in this field is Google's PageRank, which transformed the company's web search service into a near monopoly. As Google founders Sergey Brin and Lawrence Page admit "academic citation literature has been applied to the web, largely by counting citations or backlinks to a given page. This gives some approximation of a page's importance or quality. PageRank extends this idea by not counting links from all pages equally, and by normalizing by the number of links on a page" (Brin & Page, 1998: 2).

Co-citation and online social network research is making use of this measure to better rank authors and papers. For example, in Ding *et al.* (2009: 2232), the effectiveness in PageRank measures are tested in co-citation networks: "A page with a high PageRank means that there are many pages pointing to it, or that a page with high PageRank is pointing to it. Intuitively, pages that are highly cited are work browsing, and pages that are cited by the high PageRank pages are also worth reading".

Analogously, we can think about papers or authors. Authors or papers that are cited by many or are cited by authors or papers with a high PageRank must be important in a specific area.

However, the interdisciplinary nature of the emerging concept we are investigating requires more than just identifying works with high PageRanks. Relevant scholarship might be available in several different fields or the works cited by some papers might stray too far from our focus: the relationship between computational thinking and journalism. It is consequentially necessary to keep track of the field-specificity of the scholarly works and also using selection criteria that is not too rigid - some relevant contributions might not necessarily be published in databases like Web of Science.

For the purposes of identifying key themes that might be explored by computational thinking in the context of journalism education, co-citation network analysis will be used in conjunction with term/concept co-occurrence network analysis over recent scholarship relating to "computational journalism". This approach will allow us to assess what some of the most relevant recent topics are and how technical they are (and implicitly how accessible to journalism students) with respect to the context in which the scholarly work was published.

Literature Review

The quantitative turn in journalism is far from a new issue, but it is constantly developing new aspects. Journalism education and practice most of the times is quick to embrace technological developments. Three strands of journalism have been identified as being more prominent among the quantitative shifts in journalistic practice (Coddington, 2015): Computer-assisted reporting; Data journalism; and Computational journalism. However interconnected or even overlapping these concepts may be, this paper will mainly look at the newest of the three: computational journalism.

Computational journalism, or the application of computational thinking to the activities of journalism is a fairly new preoccupation of professionals and especially academics. Early in-depth treatments of the concept (Cohen, Hamilton & Turner, 2011) see it as a combination of two already familiar approaches – computer assisted reporting and the use of social science tools in journalism with the aim to enable journalists to explore increasingly large amounts of data to create stories through the use of algorithms and data in order to supplement the accountability function of journalism.

In Flew *et al.* (2012), the authors outline four key challenges in the online news environment that computational journalism may address:

1. The shift from mass audiences to niche interests the subsequent isolation of the journalism profession from the needs of its readers/viewers;
2. The decline in revenue for traditional news providers results in cost cuts, which means that fewer journalists have to report less news in fewer pages with a severe cost in terms of quality reporting standards;
3. The tediousness and high costs of original investigative journalism which could balance out the flood of tabloid sites relying on low cost newsfeed content;
4. The shift in business models – from large established organizational structures, to smaller, more flexible start-ups;

The same overview lists a series of techniques associated with computational journalism (Flew *et al.*, 2012):

- Statistical analysis,
- Regression analysis to predict changes, trends using connections in datasets,
- Correlation and matching,
- Visualisation, mashups and GIS (Geographic Information Systems),
- Parsing – or automated syntactic analysis using software tools,

- Personalisation – adapting open-source tools to one's own needs,
- Co-creation (crowd-sourcing, co-reporting and facilitating citizen journalism).

Aside from enthusiastic literature on the promises of computational journalism, some studies (Karlsen & Stavelin, 2014) have shown that practitioners in large newsrooms do not perceive computational journalism as something radically different and are sceptical about this new approach's ability to increase the efficiency of doing journalism. Still, as stated earlier, it is expected that professionals be more resistant to change (especially those that require the learning of new skills from a different field) than academics that are more used to continuous learning and interdisciplinary challenges.

Also, computational journalism could be seen as just another piece in the *Big Data* frenzy that has captivated academics, professionals and entrepreneurs alike. In a critical approach to this trend, Boyd and Crawford (2012: 662) define Big Data as: "a cultural, technological, and scholarly phenomenon that rests on the interplay of (1) technology – maximizing computation power and algorithmic accuracy to gather, analyse, link, and compare large datasets; (2) analysis – drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims; (3) mythology – the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity and accuracy".

The same article makes some fine key points about the phenomenon, its assumptions and biases, which could be also extended to the issue of computational journalism as it is associated with working with big data:

1. Big Data Changes the Definition of Knowledge
2. Claims to Objectivity and Accuracy are Misleading
3. Bigger Data are Not Always Better Data
4. Taken Out of Context, Big Data Loses its Meaning
5. Just Because it is Accessible Doesn't Make it Ethical
6. Limited Access to Big Data Creates New Digital Divides

Computational journalism is not the only concept to be proposed recently. Some scholars are working with more refined formulations such as *computational exploration in journalism* (Gynnild, 2014) to discuss specific practices that involve the journalistic co-creation of quantitative news projects with respect to three main pathways – the newsroom approach, the academic approach, and the entrepreneurial approach.

Some of the most extensive work on computational journalism and how this new concept is perceived especially by professionals in newsroom belongs to Erik Stavelin, who concludes his doctoral thesis on the subject by pointing out towards a gap that needs to be acknowledged between journalistic values and the reasons for using technology: "While computational journalism emerges from traditions of software-oriented news productions that to a large extent overlap as a merge of computer science and journalism, some distinctive features distinguish and define this field. Both internally in the newsroom and as journalistic output, computational journalism is defined to be a shift towards platforms, in creating spaces for finding, discussing and narrating stories. This can include the management of computable models, not merely collected sets of data. As a craft, creating software to solve journalistic problems, computational thinking becomes a key skill that defines both reasonable expectations and limitations, but also collaborations […] Programming journalists strive for higher journalistic capital, while newsrooms adapt by both embracing computational efforts as possibilities for journalistic reinvention and keeping a distance by labelling the work as technical" (Stavelin, 2014: 5-6).

As the literature regarding computational journalism is very recent and quite scarce, this paper will try to investigate the scientific literature connected to this topic in order to emphasize the diversity of contributors, approaches and the development speed with the goal of assessing the extent of the usefulness of integrating computational journalism in journalism education and the relevant areas of research associated with the concept.

Research Design

The main goal of this meta-analytical research is to explore key themes in the scholarship relating to the application of computational thinking in journalism. Identifying and categorizing research themes, their respective field and analysing the structure of relations between published researches can be seen as an intermediate step in defining the key points to be tackled by computational journalism courses in undergraduate or graduate Journalism School curricula.

Research Questions

RQ1: What are the main research themes relating to computational journalism?

RQ2: How interdisciplinary is the scholarship relating to computational journalism?

RQ3: What are the connections between the main themes relating to computational journalism?

Method

Data collection was done using the Google Scholar specialized search using the exact phrase “computational journalism”. A snowball sampling method was used to go through three levels of results provided by the search. For every first level result, we accessed the “Cited by” page (second level), we gathered all the articles citing the initial result. Then, for every result on every “Cited by” page, we accessed the “Cited by” page (third level). This non-probability sampling method is often used in social network analysis research (including co-citation research) when we are interested in the connections between actors. The total number of articles gathered initially was over 12,000. By eliminating duplicate records, we reached N1=7,747 scientific works. For an in-depth look at the most cited works we selected all scholarly works with a number of citations of 10 or above: N2=1,085 scientific works cited 10 or more times (according to Google Scholar). The records collected contain the publication title, the author(s), the year, the number of times it was cited and the database it is indexed in.

Tools

For constructing the co-citation graph and for data analysis NodeXL (Hansen, Shneiderman & Smith, 2010) was used along with other Excel add-ins used to clean-up the date. For term/concept extraction from the publication titles and for constructing the term/concept co-occurrence network we used KH Coder (Higuchi, 2001). For other descriptive visualisations of the dataset, we used Tableau Public.

Findings

We used NodeXL to construct the co-citation network in Figure 2. A clusterization algorithm was used to group the data. The groups with more nodes were then further analysed to reveal the most frequent terms and bi-gram syntagms (two term pairs) in the publication titles in each group. They are shown in the graph, in order of their frequency, with the number of occurrences for the most frequent term/syntagm.

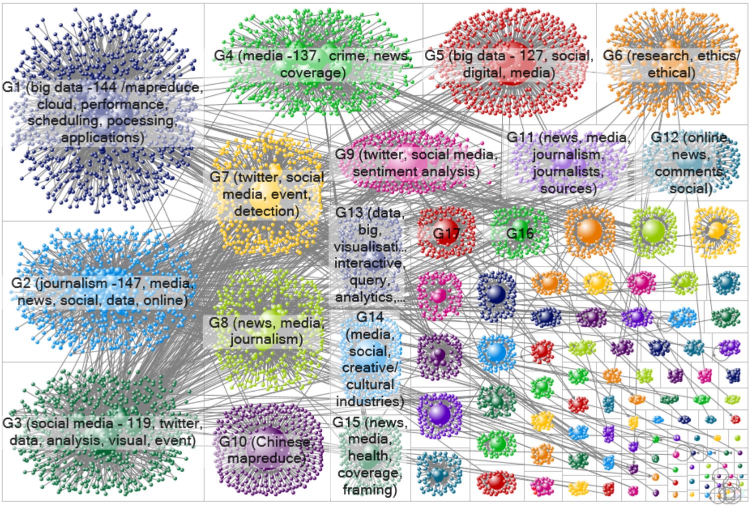

Figure 2. A visualisation of the co-citation network over the entire dataset (N1=7,747)

We can see several clusters that reveal key themes. *Big data* shows prominently in two clusters. Clusters that feature *mapreduce* (a framework used for processing large datasets) come from researchers in computer science and engineering. Also, big data research appears in conjunction with *visualisation, interactive* and *query*. There are three clusters that feature research relating to Twitter. The most frequent themes occurring in this context are *visual analysis*, *event detection* and *sentiment analysis* (the latter is also called *opinion mining* – using natural language processing tools to detect subjective information usually from messages in social media). It is also important to note, that there are also clusters that relate to *journalistic sources, comments and participation* and also, very important – *research ethics*. These results fit some of the challenges and key directions covered earlier in this paper. The sizes of the nodes are proportional to their PageRank score in the co-citation network.

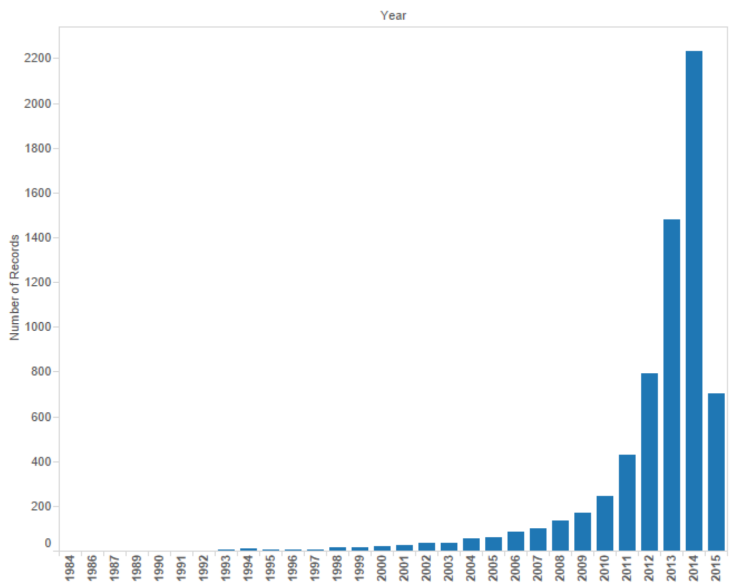

Figure 3. Descriptive statistics showing the number of published titles per year in the dataset (N1=7,747)

As Figure 3 shows, most of the collected records refer to publications published recently, demonstrating a steady increase in scholarly work published on related topics every year since 2010. Still, due to the sampling method used, a large part of this scholarship comes in fact from computer science and engineering. In order to try to differentiate between the more computer science oriented publications and the ones that apply computational thinking in relation to social sciences and humanities, or more specifically journalistic work, we looked at the databases indexing the publications.

Figure 4. Total records per databases / Total citations for publications per database

In Figure 4 we notice that computer science and engineering databases (ACM and IEEE) mainly indexing conference proceedings are quite prominent when looking at number of records. Big academic journal publishers like Taylor & Francis, Springer, Wiley and Elsevier are also quite visible. When looking at how cited the publications are per database, we notice an interesting shift. Books (as indexed by Google Books) are very much cited, while conference proceedings papers are cited less (IEEE). Besides the JSTOR database, other publishers also show up more prominently in the visualisation on the right (Cambridge University Press and Oxford University Press) as their publications are very often cited even less in total numbers, attesting to its prestige.

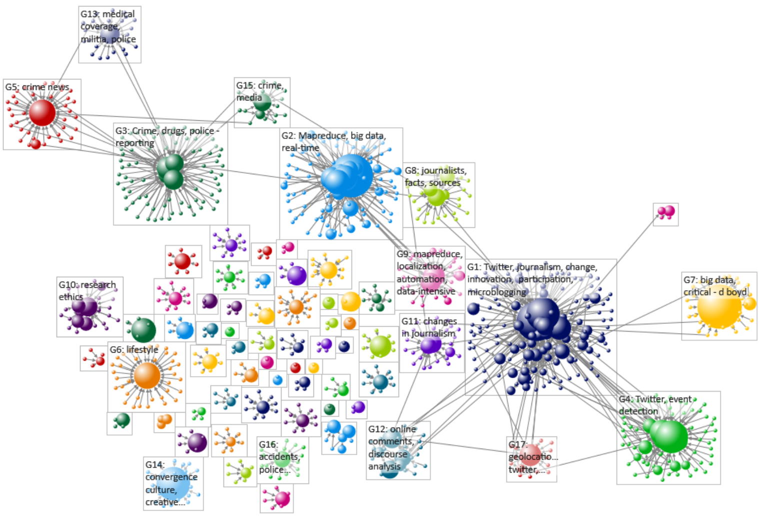

Figure 5. Co-citation network of most cited publications and key themes/group (N2=1,085)

Looking at the most cited publications and the key themes per cluster, we may notice the connections between different themes better. The *research ethics* cluster, although showing up in the context of *computational journalism* and *big data* research, does not appear to be connected to other areas. There is also a group of three clusters of publications relating to *crime, police, militia, drugs news/reporting* – probably relating to early attempts at mapping criminal/violent activity by use of public police records. It is fairly isolated from other scholarly work, which is more recent (as seen in Figure 6). The most visible cluster of research is G1, including topics such as *Twitter, journalism, change, innovation, and participation*. It is well connected to other clusters of research on Twitter (*event detection, geolocation), online comments, facts and journalistic sources and changes in journalism,* but also with critical meta-analytical approaches to big data research such as the ones produced by Microsoft researcher Danah Boyd. The more technical computer science clusters of scholarly work (G2 and G9 – featuring concepts such as *mapreduce, big data, automation, real-time, localization*) are well connected to each other and only loosely connected to the rest of the clusters.

Figure 6 shows a filtered view of the most cited publications by publishing year, focusing on only the most recent research. We can deduce that the focus on crime news and police reporting were probably of interest to scholars in the previous decades as this kind of public data was most likely made available by authorities. In both figures 5 and 6 the sizes of the nodes are proportional to their PageRank score in the co-citation network.

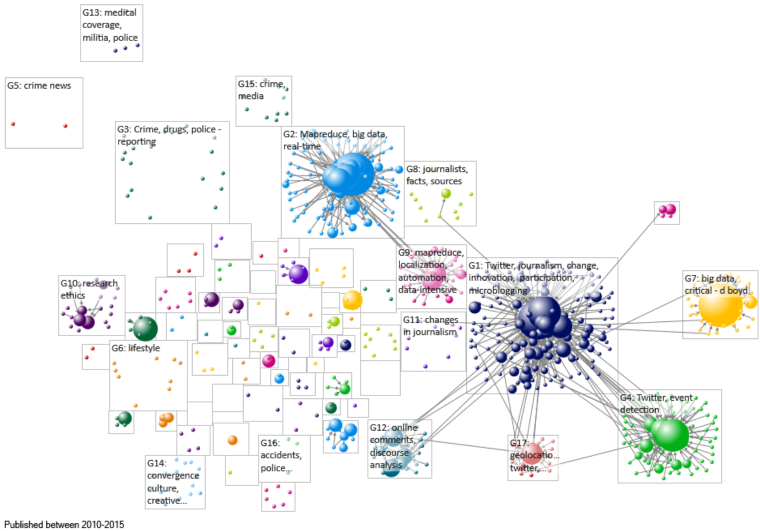

Figure 6. Co-citation network of most cited publications and key themes/group 2010-2015 (N3<1,085)

Figure 7 shows the highest scoring terms detected by KH Coder's Term Extract over the most cited titles. These results correspond to the themes identified by examining the co-citation clusters in figures 2, 5 and 6.

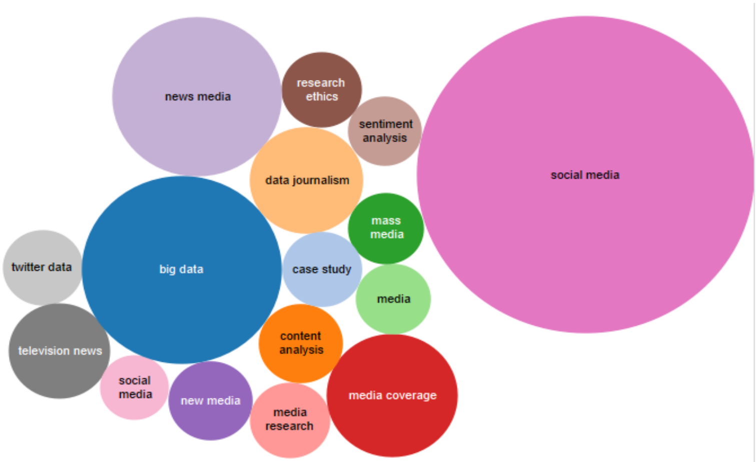

Figure 7. Most frequent concepts identified in titles (N2=1,085) using KH Coder

In order to look at the connections between these most frequent concepts found in titles of most cited publications, we used KH Coder to construct a co-occurrence network of terms/phrases. We coded both of the terms/phrases that had high scores, but also the names of the most prominent databases/publisher associated with the titles in order to establish connections between key themes and academic fields/research areas.

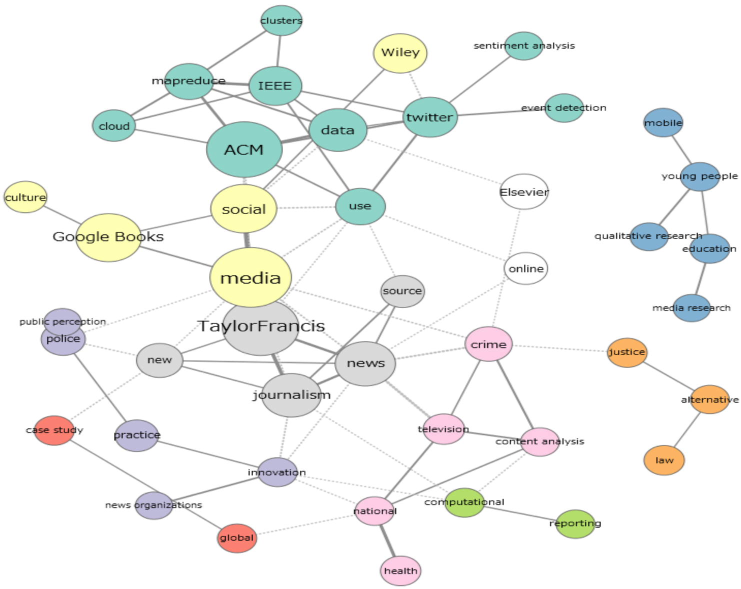

Figure 8. Co-occurrence network of frequent terms/concepts and databases

As we can see in Figure 8, the very highly cited publications on Google Books are mostly about social media and culture, the computer science and engineering databases that mainly index conference proceedings papers (IEEE and ACM) are related to big data research on Twitter – which is in turn associated most often with *sentiment analysis* and *event detection*. The topic of *education* shows up predictably in the context of research about *young people*, which is also connected to research on *mobile media*. There is considerable scholarship being published at Taylor & Francis on *journalism and news media* – as the publisher is heavily focused on behavioural, social sciences and humanities, but with a significant number of journals being published in interdisciplinary and emerging areas as well as some (although comparatively fewer) in computer science. The numbers of journals per field are available at www.tandfonline.com.

However, the most interesting aspect shown by the co-occurrence diagram is the relation between computational, reporting (in the same group), *innovation, journalism and content analysis.* Articles published about *computational journalism* often place it in the context of innovation in professional practice or link it to the social science research method of content analysis.

Discussion and Conclusion

One might speculate that the nature of the associations of *computational journalism*, with either *content analysis or innovation in newsroom practice,* depend on the context of the research, whether it is done by researchers in universities that emphasize professional skills oriented teaching or academic skills oriented teaching.

Also, we should note that this concept is not frequently associated with big data research being done mostly by computer scientists, information engineers or mixed teams. However, *computational journalism* is frequently associated with data journalism (and also *computer-assisted reporting*), being only the third concept in a line of concepts trying to connect journalism with the more technical aspects of social sciences, hence the connection with the big data research spectrum.

The key themes identified by this research mostly coincide with pathways proposed in the literature reviewed. However, our research has allowed a more in-depth perspective into recent trends in big data research that are still mostly explored by computer scientists or interdisciplinary research teams, but could very well be a part of the framework for a truly interdisciplinary computational journalism:

- *social media event detection* – using timestamps and geolocation data published on social media to detect events in real-time or gather multimedia associated to events;
- *social media sentiment analysis* – using natural language processing techniques to classify subjective information in social media messages;
- *(real-time) geo-location analyses/visualisations* – using timestamps, geolocation coordinates to create visualisations of multimedia from social media on maps.

As Boyd and Crawford (2012) pointed out, we should be mindful of the increasing gap created by differences in access to big data, ability to collect, process, analyse and make use of it when discussing computational approaches to journalism.

Also, the increasing number of scholarly works dealing with research ethics issues in relation to computational/big data approaches (collecting data from social media, respecting platform terms of service, truly anonymizing datasets collected from networked digital media) point out that we should not disregard journalism and education's strong need for continuing efforts into critical thinking that fosters creativity and independence in any field, but even more in the case of future journalists. As some others show (Stavelin, 2014), changes in journalistic practices might influence the core values associated with the occupation as journalism schools educate journalists to be perhaps future attention brokers.

To conclude, there is definitely an increasing interest towards integrating computational thinking into journalism education. However, the definitions of computational journalism depend on each researcher's understanding of computer science fundamentals. The concept of computational thinking in journalism practice is associated with earlier trends like computer-assisted reporting and data journalism and is mostly associated with use of certain pre-made tools (designed by programmers). Depending on the context, computational journalism might be framed as: (1) innovation that takes place in the newsroom by ways of automation tricks or hacks; (2) a deeper infusion of social science methodology into journalism education with emphasis on content analysis (which requires abstraction and analytical skills and now might be made more efficient with the help of automation); or (3) just one specific application/example of the larger computer science dominated area of big data research and analysis.

Consequentially, I suggest that a computational journalism course in a future journalism school curriculum would need to tackle all three perspectives and attempt to:

- foster creativity, innovation and entrepreneurship supported by critical thinking and the Use-Modify-Create model of teaching computational thinking;
- ground itself into the fundamentals of both social science content analysis methodology and computer science theory, but be mindful of research ethics when collecting/publishing

information; constantly align itself to new research being done tools being developed both in academia and the business sector by computer scientists as not to widen the gap and ensure the independence and self-reliance of future journalists (for example discuss sentiment analysis, event detection, geo-location analysis).

References

Boyd, D.; Crawford, K. (2012). "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon". *Information, communication & society*, 15(5): 662-679.

Brin, S.; Page, L. (1998). "The anatomy of a large-scale hypertextual Web search engine". *Computer networks and ISDN systems*, 30(1): 107-117.

Coddington, M. (2015). "Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting". *Digital Journalism*, 3(3): 331-348.

Cohen, S.; Hamilton, J. T.; Turner, F. (2011). "Computational journalism". *Communications of the ACM*, 54(10), 66-71.

Ding, Y.; Yan, E.; Frazho, A.; Caverlee, J. (2009). "PageRank for ranking authors in co-citation networks". *Journal of the American Society for Information Science and Technology*, 60(11): 2229-2243.

Facione, P. A. (1990). "Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations". Retrieved from: http://t.insightassessment.com/content/download/757/4801/file/Prof_Jdgmnt_%26_Dsp_CT_97_Frnch1999.pdf.

Facione, P. A.; Sánchez, C. A.; Facione, N. C.; Gainen, J. (1995). "The disposition toward critical thinking". *The Journal of General Education*, 44(1): 1-25.

Flew, T.; Spurgeon, C.; Daniel, A.; Swift, A. (2012). "The promise of computational journalism". *Journalism Practice*, 6(2): 157-171.

Gynnild, A. (2014). "Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets". *Journalism*, 15(6): 713-730.

Hansen, D.; Shneiderman, B.; Smith, M. A. (2010). *Analyzing social media networks with NodeXL*. Burlington, USA: Morgan Kaufmann.

Higuchi, K. (2001). *Kh coder. A free software for quantitative content analysis or text mining*. Available at: http://khc.sourceforge.net/en.

Karlsen, J.; Stavelin, E. (2014). "Computational journalism in Norwegian newsrooms". *Journalism Practice*, 8(1): 34-48.

Lee, I.; Martin, F.; Denner, J.; Coulter, B.; Allan, W.; Erickson, J.; Werner, L. (2011). "Computational thinking for youth in practice". *ACM Inroads*, 2(1): 32-37.

Manovich, L. (2001). *The Language of New Media*. London: The MIT Press.

Reese, S. D. (1999). "The Progressive Potential of Journalism Education Recasting the Academic versus Professional Debate". *The Harvard International Journal of Press/Politics*, 4(4): 70-94.

Small, H. (1973). "Co-citation in the scientific literature: A new measure of the relationship between two documents". *Journal of the American Society for information Science*, 24(4): 265-269.

Stavelin, E. (2014). Computational Journalism. When journalism meets programming. (Ph.D.), University of Bergen, Bergen. Retrieved from: https://bora.uib.no/bitstream/handle/1956/7926/dr-thesis-2013-Eirik-Stavelin.pdf?sequence=1

Wing, J. M. (2006). "Computational thinking". *Communications of the ACM*, 49(3): 33-35.