9/4/2021

Digital Skills for Research Postgraduates in the Humanities and Social Science

Individual Portfolio

David Cass
STUDENT NO: 120225753
WORD COUNT (EXCLUDING BIBLIOGRAPHY AND DIGITAL RESOURCE): 2482

Introduction

Attempts to define digital humanities have brought their own contradictions to the field. A focus on the tools employed belies the type of discoveries that would be impossible without such tools and introspective debate on whether digital humanities should be classified as a true component of the larger field could obfuscate some of the elemental reasons for such scholarship in the first place. This paper explores such questions by first considering whether digital humanists really need to know how to code, then exploring the various beneficial qualities of digital tools with the example of data narratives, before concluding by providing a related practical example of digital humanities scholarship in the form of a network graph.

PART I – The Importance of Coding; Seeing the Wood for the Trees

digital HUMANITIES or DIGITAL humanities?

If we take it that Ramsay's main concern when answering the question of whether we, as digital humanists, need "to know how to code" is the action of moving from "reading and critiquing to building and making", his answer in the affirmative should not even seem contentious (Ramsay, 2013b; a). Coding, in such light, is merely the means to the end, the process by which tools can be employed for the goal of better scholarship in "traditional" fields of humanities. If, however, the concern relates more to the disruption caused by the use of digital tools in the study of these traditional areas of humanities, and more particularly the updated theory-practice divide (Fitzpatrick, 2012), the question is more nuanced.

Coding may of course be seen as a form of the interpretation on which much humanities scholarship is based. At its simplest, for example, it is not necessarily about computer languages or complex data arrays; marking up text is a form of interpretation and assigning meaning, and such "humanistic interpretation" has always had a technological dimension (Underwood, 2015). The image of an "avalanche of printed numbers" in the 19th century has been used in discussing Foucault (Hacking, 2015) and until the 21st century these arrays of data may have looked just like conventional or "analogue" representations (Robertson and Travaglia, 2015). Is the nub of the question then, as Fitzpatrick asks, the exploration of the

differences that digital approaches can make to our studies, both within and across disciplines?

Doing Digital

Unsworth noted the need to distinguish a tool from the various uses that can be made of it, arguing that "in some form, the semantic web is our future, and it will require formal representations of the human record" (Unsworth, 2013). If the structured data of the ideal form of the semantic web is truly part of our future, however, is it always possible to separate the tool from the use made from it? Are the annotations in the William Blake archive to be filed under tool, use or humanist endeavour? If the annotations and their study are categorised as worthwhile humanist endeavour, does the definition survive transfer from paper to the digital realm? Conversely, would the scholarship produced through the Text Encoding Initiative really be redundant in a world free of digital? Hitchcock picks apart the disconnect between the study of analogue materials using digital tools, including digitisation and keyword searches, while remaining in thrall to "older taxonomies of knowledge", arguing that people can engage with sources in novel ways to create "maps of meaning" that "would not undermine traditional forms of scholarship" but rather give them "more intellectual purchase" (Hitchcock, 2013). In the same way, it can be argued that digital culture and practices cut across traditional opposites and "unsettle distinctions", leading to technology and practice becoming more entangled (Says, 2014). Perhaps it is not the idea of needing to code that is unsettling but its portent.

Just Digital Humanities?

What does the bridging of a digital divide look like in practice? We can take the example of a study of women in Irish politics when it was difficult to assemble data on Oireachtas Members precisely because of a lack of available data (Carty, 1980). Bigo et al. argue that data have become a social and political actor because of the way they change relationships (Bigo, Isin and Ruppert, 2019) but in the humanities, along with other disciplines, not only is the definition of "data" problematic, but so is the relation between this "data" and objects of study (Schöch, 2013).

In the example, we can reduce the data to a basic form for the sake of comparison. Table 1 in the Carty example details the number of women who took seats in seven Dáileanna (Figure 1) but the fundamental numbers are incorrect in several instances (Figure 2),

according to the information from the Houses of the Oireachtas open data portal at *api.oireachtas.ie*.

Table 1 WOMEN IN THE DÁIL			
	# Seats	# Women	% Women
1927(2)	153	1	0.6
1938	138	3	2.2
1948	147	4	2.7
1957	147	4	2.7
1965	144	3	2.1
1973	144	4	2.8
1977	148	6	4.0

Figure 1: Reproduction of table from Carty (1980)

Dáil	Term Start	First ..	Last Name
6th Dáil	11/10/1927	Margaret	Collins-O'Driscoll
10th Dáil	30/06/1938	Bridget Mary	Redmond
		Briget Mary	Rice
		Mary	Reynolds
13th Dáil	18/02/1948	Bridget Mary	Redmond
		Briget Mary	Rice
		Honor Mary	Crowley
		Mary	Reynolds
		Mary Bridget	Ryan
16th Dáil	20/03/1957	Brigid	Hogan-O'Higgins
		Celia	Lynch
		Honor Mary	Crowley
		Mary	Reynolds
		Mary Bridget	Ryan
18th Dáil	21/04/1965	Brigid	Hogan-O'Higgins
		Celia	Lynch
		Eileen	Desmond
		Honor Mary	Crowley
		Joan	Burke
20th Dáil	14/03/1973	Brigid	Hogan-O'Higgins
		Celia	Lynch
		Eileen	Desmond
		Joan	Burke
21st Dáil	05/07/1977	Eileen	Desmond
			Lemass
		Joan	Burke
		Kathleen	Ahern
		Máire	Geoghegan-Quinn
		Síle	de Valera
	07/11/1979	Myra	Barry

Figure 2: Current data from Oireachtas website

Data on the Houses of the Oireachtas application programming interface are structured, containing information on Members that includes identity, constituency base and date of term served, for example, and because of this highly structured form of text encoding, data on Members can also be easily linked to other data, including contributions to debates (Figure 3).

containsDebate	showAs	parentDebateSect...	speakerCount	speechCount	memberCode	role	speaker showAs	speaker uri	bill
True	DEPARTMENT OF DEFENCE.	null	18	51	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	CEISTEANNA (QUESTIONS).	null	14	44	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	IRISHWOMEN AND THE FRANCHISE.	null	21	148	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	MOTION RE FRANCHISE BILL.	null	2	4	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	DISCUSSION ON MOTION RE FRANCHISE BILL RESUMED.	null	11	25	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	RECEPTION OF DEPUTATIONS.	null	18	55	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	DEBATE ON TREATY	null	49	245	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	DEBATE ON TREATY	null	26	143	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	Prelude	null	50	244	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	Prelude	null	42	283	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	THE NEW MINISTRY.	null	26	43	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	PUBLICITY DEPARTMENT REPORT: DISCUSSION	null	13	15	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	MOTION RE ALTERATION OF CONSTITUTION	null	13	29	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null
True	ELECTION OF DEPUTY SPEAKER.	null	3	3	Katherine-O'Callagh...	null	Katherine O'Callaghan	https://data.oireacht...	null

Figure 3: Example of metadata available from api.oireachtas.ie, viewed through Tableau Desktop

There is therefore a notable difference in these examples not only between the availability of data at the time but in how the data, once obtained, can be used for further study. Data on the number of female Members of Dáil Éireann no doubt have existed since the First

Dáil, not least through the roll of Members, but they may have been difficult to procure or collate even up until their digitisation. In contrast, such data are now available in open formats such as the JSON from which this example was produced, and this in turn facilitates further data analysis, interpretation and making of meaning.

Coding as Collaboration?

If Ramsay's thoughts on whether we need to know how to code speak more to building things than to knowledge of command line interfaces, it follows that his point about the dangers of overprecise definition at the expense of "engaging with disciplinary questions" is perhaps more salient (Ramsay, 2013a). He and others have spoken to the cloudy definitions of digital humanities and the fuzzy edges may not deserve their share of the debate. The core qualities of humanities scholarship should certainly survive and perhaps even thrive with these new modes (Fitzpatrick, 2012). Moreover, neither collaboration nor a kindred spirit for scholarship have coding skills as a prerequisite but both can be enhanced with digital tools produced by coding skills; if the response to Ramsay is it does not matter if we do not know how to code as long as we know somebody who does (O'Sullivan, Jakacki and Galvin, 2015), it may be beneficial to consider this in more depth in Part II with a critical analysis of the argument that data journalism "is the new punk" (Rogers, 2014).

PART II – Data-Led Research and Democratising the Humanities

Array over Disarray?

In his short article, Rogers compares data narratives – or data journalism – to punk in the importance of its influence and encouragement to take part (Rogers, 2014). He specifically mentions the "DIY ethos and a shake-up of the old established order", perhaps not far from the ideas outlined by Fitzpatrick. Crucially, however, the aim of data journalism is to get a story from the data, which is again analogous to the engagement with disciplinary questions so valued by Ramsay. The most important element of the Rogers argument however, may be the question of the "democratisation of data"; if we remember the good punk bands, Rogers notes, their work could only have come about because "of kids experimenting and sounding awful". Moreover, Rogers notes that data does not just belong to experts – or perhaps even scholars – and in a follow-up piece argues that data, processing, visualisation and output can lead to a virtuous circle to a point where "creativity can run free" (Rogers, 2021).

Building and Democratising Data

Ramsay speaks about the importance of building and making and O'Sullivan et al. conclude by focusing on the importance of collaboration in digital humanities (Ramsay, 2013a; O'Sullivan, Jakacki and Galvin, 2015). In a subsequent piece Rogers, in making a similar argument about the spirit of collaboration, highlights the work by data journalists at *La Nación* in Argentina as an example of producing data-led narratives (Rogers, 2021). The process at this newspaper is an interesting case study in how digital humanities might deal with the conundrum of dealing with "smarter big data" or "bigger smart data" (Schöch, 2013). *La Nación* has used open source platforms to collaborate with readers, educational institutions and other actors in types of "hackathons", enabling the journalists to gather specific and structured data from tens of thousands of documents and audio files, with the work done by volunteers because they "believe a project matters"(Coelho, 2021).

The work at *La Nación* highlights the power of bringing people into the process, facilitated by digital tools and open digital platforms. Rogers speaks about "the great democratisation of data" and the importance of the information, noting that people "are willing to forgive a lack of perfection; they are much less forgiving for those who get the facts wrong". Does this democratisation in the making of data really make the information more reliable?

Doing Data Differently?

Despite the proliferation of data, there may be pitfalls among the benefits in using information without applying critical thought. If structural forces lend themselves to systemic data-driven narratives, there is a need to question whether it is possible that such narratives can inadvertently perpetuate injustice while at the same time looking to dismantle it (D'Ignazio and Klein, 2020). Building on the matrix of domination concept (Collins, 2002), D'Ignazio notes the complicated nature of power of which data users should be aware even while seeking to dismantle structural prejudicial forces. If such bias extends well beyond the limits of academic disciplines (How Medicine Discriminates Against Non-White People and Women, 2021), is it possible to square the circle of democratising the "tsunami" of data that Rogers mentions while limiting or eliminating the effect of structural prejudice?

D'Ignazio advocates the taking of a feminist approach to data narratives (D'Ignazio, 2021) in trying to uncover hidden biases and the example of women in the Irish parliament mentioned in Part I may be a useful case study, specifically because of the paucity of women represented since the foundation of the State.

If data not only capture but colonise minds, bodies and spaces (Bigo, Isin and Ruppert, 2019, p13) , it would be a judicious practice to view all data with this in mind. Carter's examination of women in the Oireachtas should be considered through this lens as much as it is seen with the consideration that data at the time were lacking in quantity; in the same way, the data and his findings would have to be seen with the caveat of both political studies and practice being dominated by men at the time, and therefore partially the reason Carter's findings appear mostly to be of comparison with male Dáil Deputies.

In contrast, more recent studies of female Dáil Deputies approach the similar subject matter to Carter's essay in a different way. Rather than seeing women's participation in politics through the lens of hegemonic patriarchy, such studies can consider the same core data through the type of feminist approach outlined by D'Ignazio to formulate more nuanced conclusions, including the gendered obstacles faced by early women Dáil Deputies, their own agency in representation and activism and how this can lead to a reappraisal of women in politics, particularly in the early years of an independent Ireland (McGing, 2020).

I Got This Guitar and I Learned How To Make It Talk

Echoing Ramsay, Rogers answers the question of whether we need to be coder or developer but unlike Ramsay, his answer is in the negative. To form a punk band, he notes, we just needed three chords, and he equates those chords to the many available tools that can be used to make meaning from data narratives.

If we again take the example of examining women representatives in the Irish Parliament, it is possible to consider simple data, such as the number of women and the number of speeches made, but to get deeper insight it is necessary to look at more data, which brings with it the pitfalls inherent in bigger or messier data outlined by Schöch and the bias concerns noted by D'Ignazio. In doing so we might consider text or network analysis as "chords". Narratives based on unstructured data, such as the text of parliamentary speeches, are much less common than those based on structured datasets, but they can reveal new insights (Maseda, 2021).

In the same way, network graphing could provide a different perspective when distant reading by providing different contexts. Although the Rogers argument that a viewer needs to only learn how to read a chunk of such types of visualisation to get the entire story is somewhat reductive, it is correct for him to say that simplicity lends itself to quick understanding while complexity can prevent people reaching a conclusion.

PART III – Seeing the Wood. And the Trees.

Networking Members of the Oireachtas

To this end, using datasets available at the Oireachtas open data portal, *api.oireachtas.ie*, I have produced a network graph of Dáil Deputies elected since 1919, including name, constituency and political party. I have cleaned superfluous data and changed the null values under the "gender" column to the value appropriate to the Deputy in question. The result is an interactive network graph of Deputies, demonstrating relationship to constituency, gender and political party.

There are some shortcomings in the network graph. For example, where a Deputy changed constituency or political party, only one instance of the constituency or party in question is recorded so as not to have duplicates of the Deputies in question or a double count of gender representation. Nonetheless, it should be a useful resource to continue the type of

feminist studies of women representatives in Dáil Éireann discussed in Parts I and II. The graph is available to view at https://bubcass.github.io/DH/dail/ and the Gephi .gml file is available at:

https://uccireland-my.sharepoint.com/:u:/g/personal/120225753_umail_ucc_ie/ER9E8AG4_oFPocCtEgr2wjQBWKMAb-BNgr5Wz1BFdhVPZw?e=M0h1fE

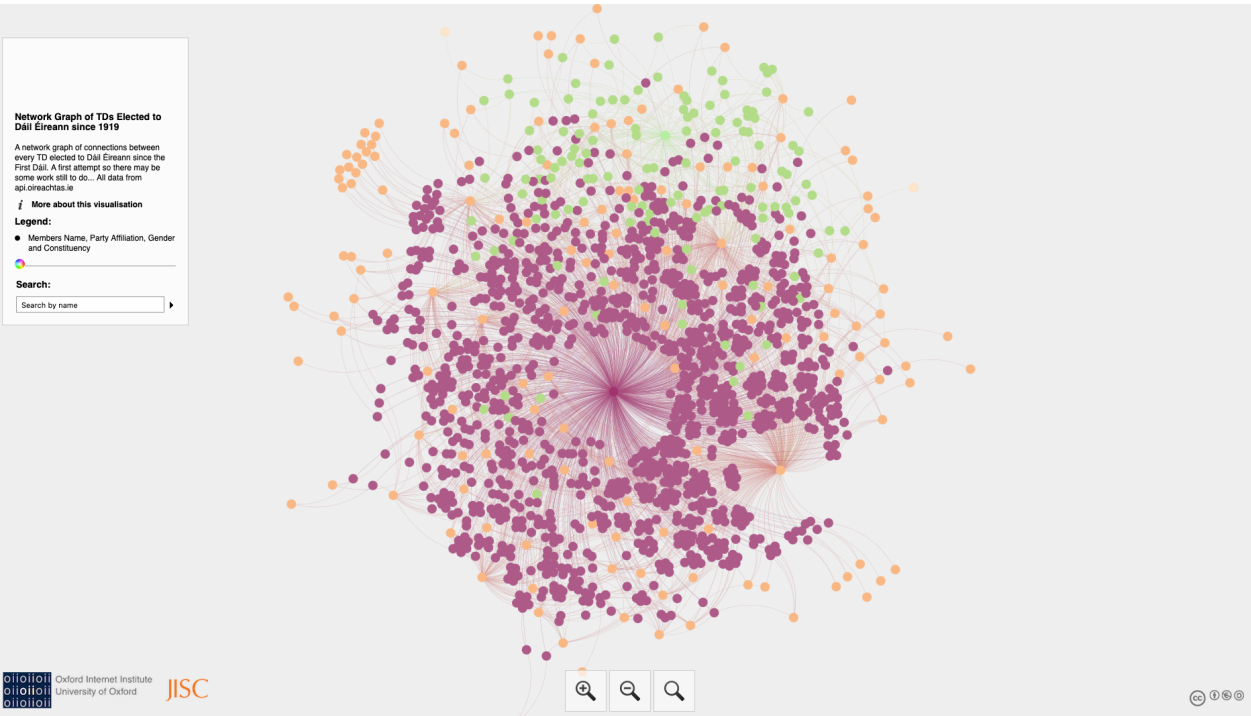

Figure 4: Detail of Network Graph of Dáil Members, from gephi

Conclusion

Rogers and Ramsay, on a superficial level, answer in opposite ways the question of whether we should know how to code, with Ramsay answering in the positive and Rogers in the negative. The more important consideration, however, may be an element of digital humanities on which they agree; it is that the work – either humanities scholarship or the "punk band" ethic of data journalism - is about building something to make meaning, and a network graph may be one chord of many in trying to make that meaning.

Bibliography

Anon 2021. How Medicine Discriminates Against Non-White People and Women. *The Economist*. [online] 8 Apr. Available at: <https://www.economist.com/science-and-technology/2021/04/08/how-medicine-discriminates-against-non-white-people-and-women> [Accessed 8 Apr. 2021].

Bigo, D., Isin, E.F. and Ruppert, E.S. eds., 2019. *Data Politics: Worlds, Subjects, Rights*. Routledge studies in international political sociology. London ; New York: Routledge, Taylor & Francis Group.

Carty, R.K., 1980. Women in Irish Politics. *The Canadian Journal of Irish Studies*, 6(1), pp.90–104. https://doi.org/10.2307/25512494.

Coelho, F., 2021. Making Data With Readers at La Nación. In: L. Bounegru and J. Gray, eds. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press.pp.91–95.

Collins, P.H., 2002. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge.

D'Ignazio, C., 2021. Data Journalism: What's Feminism Got to Do with I.T.? In: L. Bounegru and J. Gray, eds. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press.pp.103–108.

D'Ignazio, C. and Klein, L., 2020. 1. The Power Chapter. In: *Data Feminism*. [online] PubPub.pp.21–47. Available at: <https://data-feminism.mitpress.mit.edu/pub/vi8obxh7/release/3> [Accessed 8 Apr. 2021].

Fitzpatrick, K., 2012. The Humanities, Done Digitally. In: M.K. Gold, ed. *Debates in the Digital Humanities*. U of Minnesota Press.pp.12–15.

Hacking, I., 2015. Biopower and the Avalanche of Printed Numbers. In: V.W. Cisney and N. Morar, eds. *Biopower: Foucault and Beyond*. University of Chicago Press.

Hitchcock, T., 2013. Confronting the Digital. *Cultural and Social History*, 10(1), pp.9–23. https://doi.org/10.2752/147800413X13515292098070.

Maseda, B., 2021. Text as Data: Finding Stories in Text Collections. In: L. Bounegru and J. Gray, eds. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press.pp.116–123.

McGing, C., 2020. 'Women of character': Women's Political Representation in Dáil Éireann in Revolutionary and Post-Revolutionary Ireland. (11), p.23.

O'Sullivan, J., Jakacki, D. and Galvin, M., 2015. Programming in the Digital Humanities. *Digital Scholarship in the Humanities*, 30(suppl_1), pp.i142–i147. https://doi.org/10.1093/llc/fqv042.

Ramsay, S., 2013a. On Building. In: D.E. Vanhoutte, M. Terras and J. Nyhan, eds. *Defining Digital Humanities: A Reader*. Ashgate Publishing, Ltd.pp.243–245.

Ramsay, S., 2013b. Who's In and Who's Out. In: D.E. Vanhoutte, M. Terras and J. Nyhan, eds. *Defining Digital Humanities: A Reader*. Ashgate Publishing, Ltd.pp.239–241.

Robertson, H. and Travaglia, J., 2015. Big Data Problems We Face Today Can be Traced to the Social Ordering Practices of the 19th century. [Online resource] *Impact of Social Sciences Blog*. Available at: <http://blogs.lse.ac.uk/impactofsocialsciences> [Accessed 8 Apr. 2021].

Rogers, S., 2014. Data Journalism is the New Punk. *British Journalism Review*, 25(2), pp.31–34. https://doi.org/10.1177/0956474814538181.

Rogers, S., 2021. From The Guardian to Google News Lab: A Decade of Working in Data Journalism. In: L. Bounegru and J. Gray, eds. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press.pp.279–286.

Says, T., 2014. Noortje Marres: Technology and culture are becoming more and more entangled. *Impact of Social Sciences*. Available at: <https://blogs.lse.ac.uk/impactofsocialsciences/2014/12/03/philosophy-of-data-science-noortje-marres/> [Accessed 8 Apr. 2021].

Schöch, C., 2013. Big? Smart? Clean? Messy? Data in the Humanities. *Journal of Digital Humanities*, 2(3), pp.2–13.

Underwood, T., 2015. Seven Ways Humanists are Using Computers to Understand Text. *The Stone and the Shell*. Available at: <https://tedunderwood.com/2015/06/04/seven-ways-humanists-are-using-computers-to-understand-text/> [Accessed 8 Apr. 2021].

Unsworth, J., 2013. What is Humanities Computing and What is Not. In: D.E. Vanhoutte, M. Terras and J. Nyhan, eds. *Defining Digital Humanities: A Reader*. Ashgate Publishing, Ltd.pp.35–48.