City Research Online

City, University of London Institutional Repository

Collaborative learning: From CAR to data journalism and Hacks/Hackers

Data journalism in the UK grew out of computer-assisted reporting (CAR) in the USA. But how did it cross the Atlantic and also feed into journalism education in Britain? Jonathan Hewett traces the development and teaching of data journalism in the UK

We have come a long way since Tim Berners-Lee, the inventor of the world-wide web, declared 'data journalism is the future' in 2010 (Arthur 2010). By then the term was only beginning to become recognised outside specialist circles and 'data journalist' was a designated role in few newsrooms. Reporters and others were working with data, typically to source stories – notably in financial journalism. Some were familiar with spreadsheets and databases, of course, and a few with computer-assisted reporting (CAR), particularly if they had come across the latter in US investigative journalism. A similarly low level of recognition would probably have been evident at many university journalism departments in the UK.

A few years later, however, data journalism was appearing in job advertisements, conferences, university courses, discussions about the future of journalism and research projects. Its rapid and diverse evolution meant that data work was found among almost every strand of journalism, from local newspapers and hyperlocal bloggers to international news agencies, broadcast, investigative and specialist journalists and beyond.

This chapter explores both how this shift happened, including the US influence on data journalism's emergence in the UK, the role of freedom of information and the development of data-focused courses within journalism education. It draws on interviews with some of the key players and documentary research.

Tough times for investigative work

One factor underlying the relative lack of computer-assisted work in UK journalism at a time when it was well-established in the USA involved the challenges faced by investigative journalism, the main focus for CAR. 'Many would argue that the glory days of investigative journalism in the UK are well beyond us now,' Arjan Dasselaar suggested (2005: 221), attributing this to cuts in editorial budgets, fierce

competition and legal liabilities. David Leigh, then-investigations editor at the *Guardian*, bemoaned 'difficult and frustrating times for investigative journalism in Britain' (Meek 2005).

Dasselaar's survey noted also 'some distrust towards new methods of information gathering, such as the internet'. While the BBC's *Panorama* programme employed 'computer researchers, journalists seem to consider going out and talking to people as superior to using Google'. Others felt that 'information on the internet must be untrue, for otherwise it would have been picked up already' (ibid: 224) – a point well made, given recurring errors attributable to 'facts' taken from websites apparently without checking (Orlowski 2007).

Signs of change were already evident in 2005, as more journalists grasped the opportunities offered by the internet, particularly financial reporters accessing records: 'The advent of the internet has revolutionised this branch of journalism,' noted Dasselaar (op cit: 224). Even so, few could have anticipated the scale of a shift that led five years later to news organisations such as the *Guardian* dealing with huge volumes of data requiring detailed analysis. These included the Afghan war logs (92,201 rows of data), the Iraq war logs (391,000) and the US embassy cables (251,287) – all released through WikiLeaks (Rogers 2013a: 71).

Nerds plus words add up in the USA

Journalists on the other side of the Atlantic were far ahead of their UK counterparts. Already by 1999, a growing network of reporters, editors and journalism educators could look back on substantial developments in CAR over the previous ten years (Paul 1999). *When Nerds and Words Collide* reviewed progress since the 1989 creation of the National Institute for Advanced Reporting and the Missouri Institute for Computer-Assisted Reporting (which became the National Institute for CAR, NICAR). Education and training formed a recurring theme in that report: notably how to train practising journalists and integrate CAR into journalism education at universities.

The resistance to change among the latter has been documented extensively (Hewett 2016). It is linked to another long-standing debate on the stance of journalism education towards employers; some critics attribute a lack of innovation partly to it having been a 'handmaiden to industry, not its critic or visionary guide' (Dennis 1983: 3). In any case, it is hard to disagree with Folkerts that 'journalism education has, to a great degree, ignored the larger contours of the digital age' (2014: 63).

Even by the mid-2000s, the lack of UK networks such as NICAR in the USA for sharing ideas, skills and discussions was a pointed contrast with the situation not only across the Atlantic but also with many other European countries such as the Netherlands, Denmark, Sweden and Germany. Again, this may have reflected the state of investigative journalism in the UK which provided only 10 out of

450 delegates at the Global Investigative Journalism Conference in Amsterdam in 2005 (Meek 2005).

Some CAR training was taking place in the UK, however, particularly through the Centre for Investigative Journalism (CIJ). Established in 2003, it had close links with the USA – hardly surprising as it was founded by a journalist from the States. Gavin MacFadyen had worked as a producer of investigative documentaries on both sides of the Atlantic – including on *World in Action*, Granada Television's campaigning series – and became the CIJ's director. The CIJ's first summer school in July 2003 offered what may have been the first intensive training conference in the UK (albeit with participants from outside the UK, too) on investigative techniques for journalists. It included a half-day introductory CAR course, plus advanced classes, explaining: 'CAR is an increasingly important tool that enables journalists to add depth to their stories by accessing, making sense and presenting relevant government, financial, and social statistics' (CIJ 2003).

Transatlantic training triumphs

The influence of US journalism's experience in CAR was clear from the seven trainers who led the CIJ sessions. They all came from or had close links with NICAR and/or its parent body, Investigative Reporters and Editors (IRE), and included already experienced data practitioners such as Brant Houston, Aron Pilhofer and Jennifer LaFleur. The leading role played by NICAR can be inferred from the 40 to 50 seminars a year it ran over the period 1994-1999, and the estimated 12,000 journalists who attended its 300 conferences and seminars in its first ten years (Houston 1999: 7).

Houston became managing director of NICAR in 1993 (before it gained that name), having been database editor at the *Hartford Courant* and won awards for his investigative work. His book *Computer-Assisted Reporting: A Practical Guide*, into a third edition by 2003 (Houston 2003), became a key resource, drawing on 'what many of us had learned about training' and providing 'at least one road map for classes in newsrooms and journalism schools' (Houston 1999: 6).

Pilhofer was working as database editor at the Center for Public Integrity, in Washington DC, and later led the development of data and interactive journalism at *The New York Times*. LaFleur had worked as database editor at the *San Jose Mercury News* before becoming IRE's first director of training. She went on to be CAR editor at the *Dallas Morning News* and director of CAR at non-profit investigative newsroom *ProPublica*. As early as 1989, she had analysed the use of computers as part of her Master's degree at the University of Missouri School of Journalism (home to IRE and NICAR) (LaFleur 1999: 25).

The beginner's CAR workshop at the 2003 CIJ summer school focused on 'more effective searching techniques, resources and data on the internet, downloading data, and doing basic analyses using spreadsheet software such as Microsoft Excel'.

An 'intermediate track' involved moving 'from filtering and sorting data in Excel to calculating rates and ratios for news stories, cross-tabulating data and generating graphics' before 'showing how to select and filter information in a database manager and introducing users to summarising data effectively to find trends and story ideas' (CIJ 2003).

More complex techniques of data analysis were covered in an 'advanced track' with 'techniques of summarising data and using relational databases to compare different files of information to see connections that could lead to new stories', and used Microsoft Access. It also offered participants the opportunity to 'learn how to build their own databases when there is no electronic information or when governments or businesses refuse to release electronic data. Lastly, the track will include an overview of the new and increasing use of mapping (GIS) software to visualise the results of data analysis' (ibid).

Freedom of information, better journalism?

Such CAR training became a core feature of CIJ summer schools, along with another theme relevant to the emergence of data journalism: the use of freedom of information (FoI) legislation. Passed in 2000, the FoI Act came properly into force on 1 January 2005 and was fundamental to the development of data journalism in the UK. For the first time, journalists had a legal right to request information held by public bodies, which now had to respond (although not necessarily by providing the information requested). This was a huge advance on the preceding code of practice on access to government information, introduced in 1994 as an alternative to FoI legislation (Brooke 2004). Drawing on her background as a reporter in the USA, Heather Brooke wrote a pioneering book on using FoI in the UK, *Your Right to Know,* partly out of her 'frustration with the relationship between the citizen and the state in Britain, which was not as egalitarian as in America' (Brooke 2013). Primarily because a culture and legislative framework enabled access to information there, data and CAR techniques were becoming well-established among US investigative reporters in the 1990s, Brooke says:

> The main reason was that the datasets were available. A typical story was to get hold of a list of school bus drivers and cross-reference it with a list of sex offenders or other offences. They are both public records in America, with no privacy law, so you could find out whether any of the school bus drivers were paedophiles or had other convictions (ibid).

The appliance of science to reporting

In contrast to the UK – and its lack of public data – the US federal FoI Act had operated since 1967. The timing seems coincidental, but it was same year in which Philip Meyer and colleagues at the *Detroit Free Press* produced a Pulitzer Prize-winning investigation into the causes of riots in the city, often cited as the

beginning of CAR. Meyer saw the potential to apply social science techniques to journalism, having studied them the previous year while on a Nieman fellowship at Harvard. A course he took there also introduced him to computing in the form of an IBM 7090 mainframe machine (Meyer 1999: 4). In Detroit, Meyer deployed quantitative survey research, helped by two university professors, a team of 30 interviewers for field work – and a computer programmer. (This was not the first use of computing in support of journalism, however: in 1952, a Remington Rand Univac machine was used to help US television network CBS predict election results; see Chinoy 2010.)

Having demonstrated 'the application of the scientific method to the practice of journalism' (Meyer 1999: 4), he started to pass on his skills to other reporters in the same newspaper group as he developed his statistical and computing abilities. Published in 1973, his *Precision Journalism* – subtitled *A Reporter's Introduction to Social Science Methods* (Meyer 1973) – and its 1991 successor became landmarks in CAR. Meyer did not use public records in his investigative journalism until 1972 (Meyer 1999: 4), but his work converged with FoI in the use of computing tools and statistical techniques for analysis. As computers and software became more ubiquitous and easier for non-specialists to use, so US news organisations began to appoint dedicated database editors.

Grappling with hygienic spreadsheets

Interest in the UK appears to have remained very limited until more journalists began to realise the opportunities provided by FoI. From 2005, Heather Brooke trained hundreds of journalists in FoI techniques with the National Union of Journalists and elsewhere – and while she mentioned the value of obtaining information in the form of spreadsheets, she was not teaching analysis using them (Brooke 2013). She was also putting FoI through its paces herself as a freelance journalist – and learning CAR techniques to deal with the resulting data. One striking success published during the first year of FoI was an investigation with *The Times*. The resulting 'justice by postcode' story, published on the front page, revealed huge disparities in conviction rates around the UK (O'Neill, Gibb and Brooke 2005). It also highlighted the greater experience in data analysis that journalists elsewhere in Europe had developed.

'I talked to the FoI officer at the Crown Prosecution Service about when they switched to electronic data. I got three years' worth of data in Excel spreadsheets – which was great, but it was 42 different sets of records from different CPS areas,' Brooke recalls (2013; O'Neill and Brooke 2005). To help with the analysis, she turned to Tommy Kaas who had run CAR training sessions at CIJ summer schools. He had set up the Danish International Center for Analytical Reporting (DICAR) which evolved from the Association of CAR in Denmark, set up in 1997 with other CAR pioneers such as Nils Mulvad, who taught CAR at the Danish School of Journalism.

A tipping point for CAR and UK data journalism

The fourth CIJ summer school (2006) was a turning point for CAR and data training. CIJ director Gavin MacFadyen noted a 'surge in interest in computing. The rooms where those skills were being taught were packed and that's the first time that's happened. … The whole landscape has changed and journalists see the value of using electronic tools that we've taken for granted and don't really know much about' (Brooke 2006). CAR trainers at the summer school – as previously, mostly from the USA – were starting to use data about the UK, obtained from UK public bodies under UK FoI legislation, to demonstrate what could be done:

One of my very first FoI requests was for London councils' inspection reports on restaurant hygiene. Most of them were electronic datasets, which I didn't really know how to handle properly. It was around that time that Aron Pilhofer was over, and I gave him my restaurant inspection data. 'This is really great for teaching,' he said. 'I'm going to use this for our CAR classes in London.' He showed us different ways of analysing it using Excel – but also what the limits were and how you could switch over to Access and write SQL queries to drill down into the data and find out specifically which were the dirtiest restaurants in London. 'I think that was a real transitional point because it was teaching using real data from this country rather than America, and obtained from FoI' (Brooke 2013).

Another US CAR trainer, David Donald (then training director of IRE), was encouraged: 'I think you'll begin seeing many more in-depth investigative stories that will be based on using CAR,' he said (Brooke 2006). To support that growth and to share tips and ideas among interested journalists in the UK, a mailing list called BICAR emerged – inspired by the successful equivalents at DICAR in Denmark and NICAR in the USA. It was set up by Martin Stabe, then at journalism weekly *Press Gazette*, after post-CIJ summer school discussions in a pub with Brooke and freelance investigative journalist Stephen Grey. Alas, their enthusiasm seemed to outrun the wider interest among journalists for such a project, as Stabe recalls:

It never really amounted to much – there were almost no messages and it fizzled out quickly. There just wasn't the volume of material or people to make that viable, and I seemed to spend more time administering the server it ran on than actually having any content. I like to think that it was just ahead of its time (Stabe 2013).

Journalism education and training

Data-related work was also developing at City, University of London, where the CIJ was based and Brooke had become involved in teaching. I introduced FoI as part of the postgraduate Newspaper Journalism course at City in 2005; every student researched an FoI project to generate their own original story for publication. Some City students also worked with Brooke on other FoI/data projects, too. They

included Elena Egawhary (BBC *Newsnight* and *Panorama*), James Ball (later data editor at the *Guardian*) and Alex Wood (who worked as a data journalist at the BBC World Service).

FoI became important not only for the data it enabled journalists to access (for data journalism stories and projects); it also acted as a gateway for students – a valuable bridge from more conventional reporting to data journalism. Many characteristics of journalistic FoI work – from spreadsheets and the analysis of changes over time, to patterns and statistics – underlie data journalism, too. Arguably FoI serves as a useful introduction to computational thinking (Hewett 2014a), which underlies the pedagogy of some data journalism education – and the introduction of coding – in universities.

A stuffy computer room hosted a dozen or so participants for a key training event for UK data journalism in July 2007. Two hugely experienced trainers, Aron Pilhofer (then database editor at *The New York Times*) and David Donald (by then data editor at the Center for Public Integrity, in Washington) ran an intensive three-day 'training the trainers' programme at City. Subtitled 'How to teach computer-assisted reporting', the course – arranged through MacFadyen and the CIJ – aimed to 'show how CAR is successfully taught so that more CAR training can take place here and more home-grown, UK-based journalists can take advantage of these skills'. An outline of the course noted not only the importance of CAR – but also the lack of training in data-related skills and stories in the UK:

> Computer-assisted reporting (CAR) has led many reporting advances over the past 20 years in the United States, Europe and elsewhere. It's both a method to discover stories that otherwise would go unreported and a way of adding depth and context to existing stories. Historically, the United Kingdom has offered little training in these techniques for experienced journalists and novices alike (CIJ 2007).

What may have been the first such course in the UK – geared explicitly towards increasing the teaching of CAR and of working with data, and their future development – brought together a number of people who were making their mark in data journalism and its training or education, or were to do so subsequently, including:

- James Anslow – then a lecturer on BA Journalism at City; later developed modules with a more digital focus.

- Heather Brooke – FoI expert; went on to teach CAR and data at City, University of London.

- Elena Egawhary – then completing the Newspaper Journalism course at City; later became BBC researcher on *Panorama* and *Newsnight*, and data trainer.

- Stephen Grey – freelance investigative reporter; former editor of *The Sunday Times* Insight team; had investigated the CIA's secret 'extraordinary rendition' programme, analysing the details of more than 12,000 flights. Later he became special investigative correspondent at Reuters.

- Jonathan Hewett – then (and now) leading the MA Newspaper Journalism programme at City; later set up the Interactive Journalism MA with dedicated modules on data journalism.

- Mike Holderness – freelance journalist; ran training courses in online publishing and databases with the NUJ.

- Francis Irving – then a programmer with innovative NGO MySociety; main developer of its FoI site *WhatDoTheyKnow*, and later chief executive of *Scraperwiki* (which helps organisations to collect and analyse data).

- Adrian Monck – then head of the Journalism Department at City, encouraged the development of the MA Investigative Journalism. Later, managing director, head of communications and media, at the World Economic Forum.

- Cynthia O'Murchu – *Financial Times* reporter; later, investigative reporter on stories such as the data-intensive exposé of the EU Structural Funds (jointly with the Bureau of Investigative Journalism), and deputy interactive editor.

- Martin Stabe – then new media editor at *Press Gazette*; later head of interactive news at the *Financial Times*.

Also that summer, two university courses for postgraduates were preparing to welcome their first investigative journalism students. Both included the use of CAR techniques and elements – such as data-mining or scraping – of what one might now call data journalism, and were led by experienced investigative journalists (Waterhouse 2011; O'Neill 2011a). At the University of Strathclyde, in Glasgow, Eamonn O'Neill set up an MSc Investigative Journalism after studying the development of courses at universities in the USA. Investigative classes had been 'available on American campuses since at least the 1950s and possibly earlier' (O'Neill 2011b).

The Investigative Journalism MA at City, University of London, was run by Rosie Waterhouse, formerly with BBC *Newsnight* and *The Sunday Times*. She was able to build also on the teaching experience at City, including the CIJ and its director, Gavin MacFadyen; the investigations editor of the *Guardian*, David Leigh; David Lloyd, former head of news and current affairs at Channel Four, and Heather Brooke. While these two courses represented an important step in journalism education in UK universities, and were followed by a BA in Investigative

Journalism at the University of Lincoln, it would be misleading to suggest that investigative journalism had previously been absent from the curriculum.

At Sheffield University, for example, Mark Hanna had developed an investigative module for journalism undergraduates, the first to include a requirement to use FoI (Hanna 2008). At City, some courses included investigative research techniques and FoI, and Leigh and MacFadyen had already run a specialist investigative option. But such developments were relatively recent; O'Neill suggests that the UK 'did not offer investigative journalism classes until the mid-late 1990s' (2011b). An earlier Investigative Journalism MA, at Nottingham Trent University, had been launched in 1997 (Hanna 2000) but ran into difficulties Adams 2001).

More online, more open data

Journalism education was also reflecting the industry's shift to the web, which was another factor that enabled data journalism to take off. Some courses focused specifically on online, such as the Online Journalism MA started by Paul Bradshaw in 2009 at Birmingham City University. He had noted data's significance for journalists, and had been teaching students to use Yahoo Pipes to aggregate, filter, mash and map since 2006. By 2010 he was also teaching data journalism to established reporters and news organisations' trainees (Bradshaw 2013), and made an introduction to data part of the core MA journalism curriculum at City.

The timing probably helped the later courses to thrive. Online tools and web publishing were making new forms of storytelling possible. FoI was becoming better established and continuing to help journalists break stories. It was also at the heart of one of the biggest stories of the period in 2009, on MPs' expenses, even though the core material was ultimately leaked before it was due to be released (with redactions) under FoI legislation.

Although the MPs' expenses files were obtained electronically, it was not a database that the *Telegraph* obtained – it was a mass of PDF files. That meant the investigating journalists 'could mostly still operate like old-style reporters', says Brooke (2013), cross-referencing names, addresses and other details – even if spreadsheets were involved (Winnett and Rayner 2009: 220).

The open data movement was gathering pace, too, emphasising the importance of publishing data resulting from publicly-funded work, and in accessible formats. Technology journalists Charles Arthur and Michael Cross had kicked off a 'Free Our Data' campaign aimed at changing government policy (Arthur and Cross 2006). The *data.gov.uk* site eventually followed in January 2010, expanding to offer more than 9,000 datasets by October 2013. Spurred on also by the commercial possibilities, David Cameron made a series of commitments on open data after he became prime minister in May 2010 – complementing the momentum that FoI had provided for data journalism.

Dealing with data from FoI requests had been an essential part of James Ball's route towards data journalism in 2007-2008, when he worked with Brooke as a student on the pilot Investigative Journalism course at City:

> Lots of the early stuff I was doing was standardising FoI responses, getting them into a spreadsheet, and doing … everything the hard way actually, because I hadn't been taught lots of things that I would now do to make the process easier. … But this was before anyone was interested in data journalism (Ball 2013).

WikiLeaks – and a journalism MA goes interactive

In 2010, the year after MPs' expenses, and more significantly for data journalism, came the huge WikiLeaks releases of war logs from Iraq and Afghanistan, and of cables from US embassies around the world. 'It was a big deal for us – and it also made newsrooms see data people differently,' says Simon Rogers who had launched the *Guardian*'s Datablog the previous year (Rogers 2013b).

Working on the war logs and cables also proved formative for Rogers' successor as data editor, Ball, first at the Bureau of Investigative Journalism and then at WikiLeaks (Ball 2013). Ball was soon teaching his successors at City, first on the MA Investigative Journalism. Data experience from the course helped Conrad Quilty-Harper gain a job at the *Telegraph* in 2010, when he became interactive news editor. The next step in data journalism at City was the MA Interactive Journalism, which I set up in 2011. This included a separate module dedicated to data journalism, led by Ball and Bradshaw, with input from Rogers – at the same time as his team grappled with data from the riots that took place in London and other cities that August.

Out of the first 36 graduates, 11 went on to work as data journalists and a further eight in data-related roles (Hewett 2014b). These alumni went on to work at the BBC, *The Times* and *Sunday Times*, the *Guardian*, *Telegraph*, *Financial Times*, *Manchester Evening News*, *CityAM*, *Property Week*, investigative site *Exaro* and Trinity Mirror's data journalism team. From the start, the course also included a specialist module on strategic social media and online engagement – where data is also relevant, e.g. in analytics – and many alumni went on to work in this area as well as in more traditional reporting or editing roles.

Professional network for collaboration

More journalism courses were including data journalism in their curricula, too, notably at Birmingham City University, University of Central Lancashire, Cardiff and Goldsmiths, partly in response to the rise in prominence of data in the journalism industry (and more widely). Beyond universities, too, pioneering practitioners were getting together – often internationally -- to share emerging good practice, ideas for projects and to develop their own initiatives in Europe

(Lorenz 2010). Moreover, an international project produced the *Data Journalism Handbook* (Gray et al. 2012), supported by the European Journalism Centre and Open Knowledge Foundation. In addition, several massive open online courses (MOOCs) focused on data journalism (Howard 2013), and other organisations developed their own training courses.

Data teams at newspapers, the BBC and elsewhere were assembling or expanding their data teams, while other publishers looked to integrate elements of data journalism within their existing structures. In 2013, Trinity Mirror set up a central data journalism unit to work with its regional newspapers and the *Daily Mirror* – and it also piloted a fresh approach to data journalism with its standalone *ampp3d* project. With its rapid response to news events, tabloid-style approach, high-impact graphics and interactivity, produced for mobile access and social sharing, the site reached places (and audiences) that few data projects had tried to go. Although it closed after 18 months, *ampp3d*'s impact was felt widely (Bradshaw 2015) and Trinity Mirror's data journalism unit continued, later expanding to ten-strong (Southern 2016).

At a number of news organisations, developers and journalists were starting to get used to working together on specific projects or ongoing teams – but wider, more open networks for collaboration were playing an increasingly important role, too. In March 2010, some coders and journalists, including James Ball and Joanna Geary (then web development editor, business, at *The Times*), held the first of a series of 'Ruby in the Pub' meetings to explore areas of mutual interest – and to help the journalists to learn Ruby (Goodchild 2010). This soon evolved into Hacks/Hackers London, coordinated by Geary and launching in August 2010 as part of the international Hacks/Hackers network that was founded in San Francisco less than a year earlier to become the most prominent example of computing-journalism collaboration (Lewis and Usher 2014).

Taking a slightly wider focus than journalism and coding – to embrace media-related start-ups and entrepreneurial/tech initiatives – the monthly meetings of Hacks/Hackers London quickly outgrew the basement bar in which it had begun. Despite – or because of – moving to larger venues (often at tech/media companies' offices), with corporate sponsorship and online bookings, its meetings are consistently oversubscribed. Indeed, it has become one of the prime networking events for those working at the intersection of media and technology in London, attended by several hundred people, and sister groups elsewhere in the UK.

This left room for another group, Journocoders, to take a more explicit focus on learning technical skills for use in journalism. With a lower profile and smaller informal meetings, it began in London in December 2014 (Bentley 2014). The monthly sessions typically take a hands-on approach, with participants bringing laptops to work on, focused on sharing and learning journalistic uses of coding.

Within university journalism education the picture remained mixed, despite growing interest. As late as 2014, one journalism educator concluded that 'the route into data journalism is not an obvious one and a period of studying journalism at a UK university certainly doesn't seem to be part of that route' (Hannaford 2014). Most journalism schools 'don't get it', said the head of interactive news at the *Financial Times* (Tinworth, 2014). This may reflect the particular complexities of running data journalism courses, which include ensuring its currency; the 'million-dollar question' (McKerral 2013) of who will teach it; technical and statistical demands; and – in a market-oriented HE system – the need to attract students who may not be familiar with data journalism (Hewett 2016). Studies of data journalism education around the world have identified similar issues (e.g. Berret and Phillips 2016; Yang and Du 2016; Davies and Cullen 2016; Treadwell, Ross, Lee et al 2016).

Some British educators were starting to respond more fully, however, incorporating data journalism into textbooks for university courses, for example (e.g. Hill and Lashmar 2014; Holmes et al. 2013; Bradshaw and Rohumaa 2011)..

Data, coding and computational thinking

One key feature of their recent evolution has been the introduction of coding as a core element of journalism programmes. An additional data and coding module was introduced at City, University of London, in 2015 as part of the MA Interactive Journalism, for example; Cardiff launched an MSc in Computational Journalism in 2014; coding forms part of Goldsmiths' Digital Journalism MA/MSc and of the Data Journalism MA that starts at Birmingham City University in 2017/18.

How far journalism students – and established journalists – need to learn to code has been a point of discussion for some years, particularly since the emergence of newsroom developers or 'programmer-journalists' alongside data journalists (Royal 2010; Parasie and Dagiral 2012; Taylor 2009). Learning coding, in this context, means going far beyond mark-up and styling with HTML and CSS. Data journalism programmes in the UK have concentrated on developing relevant skills and understanding of Javascript, R and Python, and in some cases D3 and/or SQL.

Typically – and importantly – programmes tend to involve more than technical skills alone. The aim is also to enable students to develop computational thinking, for example (Wing 2006). This approach seeks to tackle problems by drawing on key concepts from computer science, such as automation and modularity. 'Most fundamentally it addresses the question: what is computable?' (ibid: 33). In many ways, computational thinking is about unpacking and understanding the processes of computing and applying critical thinking and reflective practice to these to design solutions – here, in the context of journalism (Meza 2016).

Data journalism has spread, evolved and diversified over the past decade. This 'quantitative turn' in journalism (Petre 2013) has prompted more research, including a typology that seeks to differentiate CAR, data journalism, and computational journalism – while recognising that this interests scholars more than practitioners (Coddington 2014). Regardless of such delineations, the increasing ubiquity and availability of data in many fields makes it hard to ignore – particularly for anyone concerned with the future of journalism. Data underlies many current media trends, from verification and automation to networked investigations and personalisation (Newman 2017), while data-driven processes appear essential for successful innovation in digital news (Küng 2015).

This rise of data makes a compelling case for journalism students to learn to handle data competently, at least at a basic level, and to understand its increasing strategic and range of uses. At a more advanced level, the position of data journalist remains largely a specialist role in the industry – and of limited appeal to many journalism students, it appears (Yang and Du 2016; Hewett 2016). But other positions in journalism now also rely on data as essential inputs to editorial decision-making – in social media and audience development roles, for example (Thurman 2016). The central place in journalism of data analysis (and coding) is suggested in their appearance among criteria for UK news organisations' journalism trainee schemes in recent years (Hannaford 2015).

The importance of including data work in university journalism programmes appears to be confirmed, in broad terms, by an international survey of data journalists (Heravi 2017). Initial findings indicate that many more (62 per cent) had a degree in journalism than one in a data-related or technical discipline (12 per cent). The majority were also relatively new to journalism, having 1-4 years' experience (41 per cent) or 5-9 years (26 per cent).

Where now – and where next?

While data journalism remains a specialist pursuit for most people in the media, its profile and authority are arguably higher than ever. It is becoming well established in most areas of journalism and can look to both thriving professional networks that involve developers as well as journalists, and a range of university journalism programmes offering data and coding courses to students in HE. In many ways, the demand for training and professional networking reflects the 'mainstreaming' of digital journalism and specific areas such as data (Tinworth 2016).

Back in 1999, Philip Meyer lamented that CAR was 'an embarrassing reminder' that journalism was 'the only profession in which computer users feel the need to call attention to themselves' (1999: 4). Nearly 20 years later, can we look forward to the day when journalists will no longer feel the need to call attention to themselves as users of data?

References

Adams, Catherine (2001) Inside story, *Guardian*, 13 March. Available online at http://www.theguardian.com/education/2001/mar/13/highereducation.uk, accessed on 30 October 2013

Arthur, Charles (2010) Analysing data is the future for journalists, says Tim Berners-Lee. *Guardian*, 22 November. Available online at http://www.guardian.co.uk/media/2010/nov/22/data-analysis-tim-berners-lee, accessed on 3 September 2015

Arthur, Charles and Cross, Michael (2006) Give us back our crown jewels, *Guardian*, 9 March. Available online at http://www.theguardian.com/technology/2006/mar/09/education.epublic, accessed on 30 October 2013

Ball, James (2013) Interview with Jonathan Hewett, 25 October 2013, London

Bentley, Elliot (2014) Journocoders inaugural [sic] meetup, *Meetup.com*. Available at https://www.meetup.com/Journocoders/events/219005707/

Berret, Charles and Phillips, Cheryl (2016) *Teaching Data and Computational Journalism.* Columbia Journalism School. Available online at https://journalism.columbia.edu/system/files/content/teaching_data_and_computational_journalism.pdf

Bradshaw, Paul and Rohumaa, Liisa (2011) *The Online Journalism Handbook: Skills to Survive and Thrive in the Digital Age*, Harlow: Longman

Bradshaw, Paul (2013) Interview with Jonathan Hewett, 14 October 2013, London

Bradshaw, Paul (2015) The legacy of Ampp3d, UsVsTh3m and Row Zed, 13 May. Available online at: http://onlinejournalismblog.com/2015/05/13/the-legacy-of-ampp3d-usvsth3m-and-row-zed/, accessed on 15 September 2015

Brooke, Heather (2004) *Your Right to Know*, London: Pluto Press

Brooke, Heather (2006) Investigate! *Journalist*, August/September, National Union of Journalists. Available online at http://heatherbrooke.org/2006/article-future-of-investigative-reporting/, accessed on 27 October 2013

Brooke, Heather (2013) Interview with Jonathan Hewett, 18 October 2013, London

Centre for Investigative Journalism (2003) Programme for CIJ Summer School, 18-20 July 2003, London

Centre for Investigative Journalism (2007) Programme for Training the Trainers: How to Teach Computer-Assisted Reporting, 23-25 July, London

Chinoy, Ira (2010) *Battle of the Brains: Election-Night Forecasting at the Dawn of the Computer Age*, Dissertation: University of Maryland. Available online at http://drum.lib.umd.edu/bitstream/handle/1903/10504/Chinoy_umd_0117E_11395.pdf;jsessionid=F52BCFE72351DABAE91D0418D470E70C?sequence=1

Coddington, Mark (2014) Clarifying journalism's quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting, *Digital Journalism*, Vol. 3, No. 3 pp 331–348

Dasselaar, Arjan (2005) United Kingdom, Van Eijk, Dick (ed.) *Investigative Journalism in Europe*, Amsterdam: Vereniging van Onderzoeksjournalisten pp 213-226

Davies, Kayt and Cullen, Trevor (2016) Data journalism classes in Australian universities: Educators describe progress to date, *Asia Pacific Media Educator*, Vol. 26, No. 2 pp 132-147

Dennis, Everette E. (1983) An exchange. Journalism education, *Planning for Curricular Change in Journalism Education*, Eugene: School of Journalism, University of Oregon

Folkerts, Jean (2014) *History of Journalism Education*: *Journalism & Communication Monographs*, Thousand Oaks, CA: SAGE Publications

Goodchild, Dave (2010) Developers and journalists forging common ground, *Journalism. co.uk*, 17 May. Available online at https://blogs.journalism.co.uk/2010/05/17/developers-and-journalists-forging-common-ground

Gray, Jonathan, Bounegru Liliana and Chambers, Lucy (eds) (2012) *The Data Journalism Handbook.* Sebastopol, CA: O'Reilly Media. Available online at http://datajournalismhandbook.org/1.0/en/

Hanna, Mark (2000) *British investigative journalism: Protecting the continuity of talent through changing times*. Paper presented to the International Association for Media and Communication Research, Singapore, 18 July

Hanna, Mark (2008) Universities as evangelists of the watchdog role: Teaching investigative journalism to undergraduates, De Burgh, Hugo (ed.) *Investigative Journalism*, Abingdon and New York: Routledge, third edition pp 157-173

Hannaford, Liz (2014) *Data Journalism: A Modern Day Bletchley Park*? Available online at http://www.lizhannaford.com/data-journalism/data-journalism-a-modern-day-bletchley-park/, accessed on 8 October 2014

Hannaford, Liz (2015) Computational journalism in the UK newsroom, *Journalism Education* Vol. 4, No. 1 pp 6-21

Heravi, Bahareh R. (2017) The state of data journalism globally, *Proceedings of the European Data and Computational Journalism Conference,* Dublin, 6-7 July

Hewett, Jonathan (2014a) *Engaging with Data: Reflections on Developing a Data Journalism Course.* Paper presented to Computation + Journalism symposium, 24-25 October, New York

Hewett, Jonathan (2014b) Growing data journalism in the UK, *American Journalism Review*, 13 November. Available online at http://ajr.org/2014/11/13/growing-data-journalism-uk/, accessed on 15 September 2015

Hewett, Jonathan (2016) Learning to teach data journalism: Innovation, influence and constraints, *Journalism*, Vol. 17, No. 1 pp 119-137

Hill, Steve and Lashmar, Paul (2014) *Online Journalism: The Essential Guide,* Los Angeles, CA: SAGE

Holmes, Timothy, Mottershead, Glyn and Hadwin, Sara (2013) *The 21st Century Journalism Handbook: Essential Skills for the Modern Journalist*, Harlow: Pearson

Houston, Brant (1999) Changes in attitudes, changes in latitudes, Paul, Nora (ed.) *When Nerds and Words Collide: Reflections on the Development of Computer Assisted Reporting*, St Petersburg, Florida: Poynter Institute for Media Studies pp 6-7

Houston, Brant (2003) *Computer-Assisted Reporting: A Practical Guide*, Boston: Bedford/ St Martin's, third edition

Howard, Alexander (2013) Data journalism MOOCs offer new options for distributed learning, *towcenter.org/blog*. Available online at http://towcenter.org/blog/data-journalism-moocs-offer-new-options-for-distributed-learning/

Küng, Lucy (2015) *Innovators in Digital News*, London/New York: I. B. Tauris

LaFleur, Jennifer (1999) Evangelizing for CAR, Paul, Nora (ed.) *When Nerds and Words Collide: Reflections on the Development of Computer Assisted Reporting*, St Petersburg, Florida: Poynter Institute for Media Studies pp 25-27

Lewis, Seth C. and Usher, Nikki (2014) Code, collaboration and the future of journalism, *Digital Journalism*, Vol. 2, No. 3 pp 383-393

Lorenz, Mirko (2010) *Data Driven Journalism: What is There to Learn*? Paper on the Data-Driven Journalism Roundtable, Amsterdam, 24 August

McKerral, Mac (2013) Steering data journalism into the curriculum. *Quill*, Vol. 101, No. 4 p. 24

Meek, Colin (2005) Analysis: Computer-assisted reporting leaves UK journalists in the slow lane, *Journalism.co.uk*, 5 October. Available online at http://www.journalism.co.uk/news/analysis-computer-assisted-reporting-leaves-uk-journalists-in-the-slow-lane/s2/a51543/, accessed on 25 October 2013

Meyer, Philip (1973) *Precision Journalism: A Reporter's Introduction to Social Science Methods,* Bloomington and London: Indiana University Press

Meyer, Philip (1991) *The New Precision Journalism*, Bloomington and Indianapolis: Indiana University Press

Meyer, Philip (1999) The future of CAR: Declare victory and get out, Paul, Nora (ed.) *When Nerds and Words Collide: Reflections on the Development of Computer Assisted Reporting*, St Petersburg, Florida: Poynter Institute for Media Studies pp 4-5

Meza, Radu (2016) Computational thinking and journalism education, Singla, Carles da Rocha, Irene and Ramon, Xavier (eds) *Shaping The Future Of News Media*, Integrated Journalism in Europe pp 179-206

Newman, Nic (2017) *Journalism, Media, and Technology Trends and Predictions 2017*, Oxford: Reuters Institute for the Study of Journalism

O'Neill, Eamonn (2011a) Digging deeper: Reflecting on the development and teaching of investigative journalism in a university setting in the United Kingdom, Mair, John and Keeble, Richard Lance (eds) *Investigative Journalism: Dead or Alive?* Bury St Edmunds: Abramis pp 291-307

O'Neill, Eamonn (2011b) Written evidence, *The Future of Investigative Journalism*, London: House of Lords Select Committee on Communications pp 425-429

O'Neill, Sean and Brooke, Heather (2005) Prosecutors in dock over disparity in convictions, *The Times*, 23 November p. 6

O'Neill, Sean, Gibb, Frances and Brooke, Heather (2005) Justice by postcode: The lottery revealed, *Times*, 23 November p. 1

Orlowski, Andrew (2007) Braindead obituarists hoaxed by Wikipedia: Only fools and journos, *Register*, 3 October. Available online at http://www.theregister.co.uk/2007/10/03/wikipedia_obituary_cut_and_paste/, accessed on 30 October 2013

Parasie, Sylvain and Dagiral, Eric (2013) Data-driven journalism and the public good: 'Computer-assisted reporters' and 'programmer-journalists' in Chicago, *New Media & Society*, Vol. 15, No. 6 pp 853-871

Paul, Nora (ed.) (1999) *When Nerds and Words Collide: Reflections on the Development of Computer Assisted Reporting*, St Petersburg, Florida: Poynter Institute for Media Studies

Petre, Caitlin (2013) *A quantitative turn in journalism?* Tow Center for Digital Journalism, October 30. Available online at http://towcenter.org/blog/a-quantitative-turn-in-journalism/

Rogers, Simon (2013a) *Facts are Sacred: The Power of Data*, London: Faber and Faber/Guardian Books

Rogers, Simon (2013b) Interview with Jonathan Hewett, 22 October 2013, San Francisco/London (via Skype)

Royal, Cindy (2010) The journalist as programmer: A case study of *The New York Times* interactive news technology department, *International Symposium on Online Journalism*. Available online at https://www.semanticscholar.org/paper/The-Journalist-as-Programmer-A-Case-Study-of-The-N-Royal/b77a51be50f449ef821e84c07848854a0cdc211f

Silver, Nate (2014) What the fox knows, *FiveThirtyEight*, 17 March. Available online at http://fivethirtyeight.com/features/what-the-fox-knows/, accessed on 15 September 2015.

Southern, Lucinda (2016) Inside Trinity Mirror's data-visualization unit, 8 December. Available online at: https://digiday.com/uk/trinity-mirror-data-visualization/

Stabe, Martin (2013) Email exchange with Jonathan Hewett, 30 October 2013, London

Taylor, Megan (2009) How computer-assisted reporters evolved into programmer/journalists, *MediaShift*, 7 August. Available online at http://mediashift.org/2009/08/how-computer-assisted-reporters-evolved-into-programmerjournalists219

Tinworth, Adam (2014) *Data Journalism:Buzzword or Baseline Skill?* Available online at http://www.onemanandhisblog.com/archives/2014/01/data_journalism_-_buzzword_or_baseline_s.html

Tinworth, Adam (2016) Hacks/hackers London: It's only going to get harder to get a ticket – and that's just how it should be. Available online at https://www.onemanandhisblog.com/archives/2016/03/hhldn-tickets-attendence-growth-mainstream-digital-journalism.html

Thurman, Neil, Cornia, Alessio and Kunert, Jessica (2016) *Journalists in the UK*, Oxford: Reuters Institute for the Study of Journalism

Treadwell, Greg, Ross, Tara, Lee, Allan, Lowenstein, Kelly, Jeff (2016) A numbers game: Two case studies in teaching data journalism, *Journalism & Mass Communication Educator*, Vol. 71, No. 3 pp 297-308

Waterhouse, Rosie (2011) Can you teach investigative journalism? Methods and sources, old and new, Mair, John and Keeble, Richard Lance (eds) *Investigative Journalism: Dead or Alive?*, Bury St Edmunds: Abramis pp 284-290

Wing, Jeanette M. (2006) Computational thinking, *Communications of the ACM*, Vol. 49, No. 3 pp 33-35

Winnett, Robert and Rayner, Gordon (2011) *No Expenses Spared*, London: Random House

Yang, Fan, and Du, Ying Roselyn (2016) Storytelling in the age of Big Data: Hong Kong students' readiness and attitude towards data journalism, *Asia Pacific Media Educator*, Vol. 26, No. 2 pp 148-162