Bridging the Skills Gap: Innovation in Journalism Education

Martin J. Chorley
Cardiff University, UK

Glyn Mottershead
University of London, UK

Abstract

In this disrupted media landscape as publishers look to increase their online output and deliver news in more innovative and engaging ways, a clear skills gap has become evident between traditional journalism education and training and the emerging requirements of the news industry. The move beyond big data, massive data leaks, and towards the future of large scale societal algorithmic influence has led to the need to develop specialist skills to allow journalism to work with this information deluge. The MSc Computational and Data Journalism at Cardiff University is a one-year postgraduate course which was designed to meet some of those needs. The joint-honours programme, now in its fifth year, takes students with a wide range of backgrounds: working journalists looking to re-train, journalism graduates wanting to specialise, software engineers or computer scientists looking to broaden their skill set, or graduates with no background in either journalism or computer science looking for a change in direction. This paper presents details of the course ethos and structure, including the teaching methods used, and personal reflections and learnings from the first five years of the programme.

What Do We Mean by Computational Journalism?

Given that the work of the newsroom is an ever-evolving process, it is not accurate to say that working with data within news is a new thing – in fact many working in the field in the UK will reference the (then) Manchester Guardian as a forerunner in the field back in 1821 (‘The first Guardian data journalism: May 5, 1821’, 2011).

The development of this work in to what would later become known as data journalism and its sibling discipline computational journalism has a wide range of definitions. Some see computational journalism as “the combination of algorithms, data, and knowledge from the social sciences to supplement the accountability function of journalism. In some ways computational journalism builds on two familiar approaches, computer-assisted reporting (CAR) and the use of social science tools in journalism championed by Phil Meyer…. Like these models, computational journalism aims to enable reporters to explore increasingly large amounts of structured and unstructured information as they search for stories.” (Turner & Hamilton, 2009). Others focus on the discipline as “the application of computing and computational thinking to the activities of journalism including information gathering, organization and sensemaking, communication and presentation, and dissemination and public response to news information.” (Diakopoulos, 2011). Flew and colleagues offer one of the simpler definitions: “At its simplest, computational
journalism involves the application of computing to journalism. This means not simply the application of computing technologies to journalism, journalism has of course long been engaged with information and communication technologies (ICTs) throughout the modern era but active engagement with techniques for the large-scale manipulation of data using computing software to enable new ways to access, organise and present information. “ (Flew, Spurgeon, Daniel, & Swift, 2012)

To us as we consider the training and educational needs of postgraduate journalism students entering the newsroom, it is the application of problem-solving skills and understanding of data drawn from computer science as applied to the journalistic traditions of sense making, storytelling, computer assisted reporting and dissemination, combined with the ability to communicate clearly with a community. However, rather than arguing about how the field is defined the focus of this paper is to look at how a university team has successfully introduced a programme that allows students to develop a cutting-edge skillset within this field that has allowed them to work within data and special project teams within major news rooms.

Introduction

The School of Journalism, Media and Culture at Cardiff University has an almost half-century long tradition of postgraduate education aimed at delivering work-ready journalists into newsrooms, with programmes focused on (written) news, broadcast, magazine and international journalism. These programmes have long embraced the ‘new’, bringing in digital and social-media skills at a point at which these skills were only beginning to be embraced by newsrooms. Meanwhile, the School of Computer Science and Informatics at Cardiff University also has a long tradition of post-graduate education in which graduates with little or no technical background are upskilled in the coding, software development and general computational skills required by the IT industry.

An educational establishment has many forces acting upon its curriculum: In a ‘research-led’ institution there is a desire that academic research pushing at the boundaries of the field should make its way into the classroom. Many subject areas have professional accrediting bodies who require coverage of certain topics to certain depths. Finally, of course, the needs of employers must be taken into account, as students exiting programmes expect to be able to find employment in their chosen field and expect that their University education will provide them with the necessary skills for this.

Discussions with employers in the news industry over many years have presented a compelling case for a new skills-gap in the newsroom. Some of these conversations became the backbone of chapters in the 21st Century Journalism Handbook (Holmes, Hadwin, & Mottershead, 2012). This research, and subsequent discussions highlighted the increasing complexity of large-scale data-driven investigations, the desire to innovate to develop products utilising new technologies, and the need to deliver the news in more dynamic and flexible ways which has led to a need for individuals that sit somewhere on the wide spectrum between computationally savvy journalists and software developers with an understanding of and appreciation for writing stories.
But it isn’t just in industry that this skills gap had been identified. In the US there has long been a tradition of technologically-augmented journalism, which means that some J-schools have been teaching data journalism for some time. However, this was not the case in Europe at the time of programme development. This skills gap within the academy, as Mensing says, was an issue that needed to be challenged as the “continuation of the status quo within journalism schools is increasingly untenable. Developing new models more appropriate to the needs of this age, as opposed to that of the media industries of the past, is a compelling obligation for journalism educators and scholars.” (2010). It is into this gap that the MSc in Computational Journalism (as it was then titled) was launched in 2014.

Designed in consultation with figures from industry and academia, the programme retains a practical focus and aims to fill the skills gap identified by many employers and deliver technically-capable graduates able to work across a wide spectrum of roles. These graduates will self-identify along a newsroom skills spectrum with some clearly aligning themselves with the more story-based aspects of the role and going on to work as a ‘data journalist’, others will follow a more technical route and go on to work as a ‘newsroom developer’ or in other related roles.

**Designing the Course**

No QAA subject benchmark exists for Journalism at the postgraduate level\(^1\). Meanwhile, the QAA subject benchmark for Computing (QAA, 2011), while a useful guide, is not entirely relevant. There are aspects of being a computer scientist or qualities of a computer science course that are not relevant to a being Computational and Data Journalist. There are some previous areas we were able to look at for guidance, (for instance (Yarnall, Johnson, Rinne, & Ranney, 2008) examines how journalism students are taught CAR\(^2\) data analysis skills), but these were again limited in relevance, being prior versions of what is actually being used in the news industry.

In order to ensure academic rigour and standards, we had only the general level 7 Masters descriptors from the Framework for Higher Education Qualifications in England, Wales and Northern Ireland (QAA, 2008), and the Credit and Qualifications Framework for Wales (Higher Education Funding Council For Wales, 2015) to use as a guideline, and it was these that were leaned on heavily in the validation documents for the Programme Approval.

Given the practice focused nature of the course, it was deemed more relevant and natural to consult practitioners on the required skills, knowledge and qualities of potential graduates, rather than relying upon what could be perhaps a more academic assessment of the discipline. This external industry input helped guide our thinking about how an ‘educated person’ (Boyer, 1995) from this course should look, and helped us devise the overall learning outcomes for the course. It also allowed the course team to develop a network of peers that could not only offer ongoing

---

1. Other postgraduate courses rely on benchmarks from professional organisations such as the NCTJ, however, these bodies are not relevant for a course in Computational Journalism.

2. Computer Assisted Reporting, the forerunner of today’s Data Journalism.
guidance to the support team, but become part of the network of the students themselves (Lave & Wenger, 1991) in order to bring in the wider knowledge of those in currently working in the industry or in hiring positions.

The core foundational data journalism and coding knowledge taught in the first semester is absolutely critical, a view held by Berret & Phillips (2016) in their recent report on teaching Computational & Data Journalism in the US. As they say: “The bottom line is that to do more advanced data journalism, its practitioners need, at a minimum, to understand how programming works” and “Above all, we recommend that all programs have a required foundational course in data journalism, teaching basic principles of data analysis…” (Berret & Phillips, 2016). This is most interesting, as the report highlights the same key skills and knowledge requirements that we identified in creating our course 5 years ago³. Comparing the structure of our MSc to their suggested programme structures shows a commonality of modules and topics which validates our approach further. The coding aspect of the course is a particular concern for those students entering the course from a non-technical background, as ‘programming’ itself can be considered a threshold concept (Meyer & Land, 2003) that can be difficult for students to grasp, although in our experience so far this has not proven to be an insurmountable barrier to most of the students on our course.

Course Ethos

There has been a lot of discussion of the differing ethos around the training and education of journalists (Cutter & Marken, 2005; De Burgh, 2005; Reese, 1999) – on one level we require the classic ideas of public service, objectivity, autonomy, immediacy and ethics as the core of journalism “pragmatically speaking, journalism within the context of professional education and industry training means the preparation of students for a career working in news media organizations and studying the work of those people editorially responsible for different types of storytelling in a wide range of news media.” (Deuze, 2005), but this must be in a way that recognises the changing nature of the industry (Lewis, 2015) and helps drive technological development in the newsroom (Pavlík, 2013) and allows a continuous interrogation of emerging technologies (Berret & Phillips, 2016)

The course design aimed not only to leverage modern journalistic techniques and interdisciplinarity but to look towards more innovative teaching methods and philosophies. The core thinking of the development team focussed around ideas such as constructive alignment, connectivism (Boers, Ercan, Rinsdorf, & Vaagan, 2012; Siemens, 2005) and computational thinking (Wing, 2006). There is also a strong focus around problem-based teaching, and self-directed learning which is facilitated via the programme leaders. A number of classes are also flipped (Bergmann & Sams, 2014; McGregor, 2015; Talbert, 2017) and there is a heavy reliance on practical activities following outside contact theory delivery.

³. The initial 2 years of our course included the module ‘Digital Journalism’ in the first semester, this has now been replaced with ‘Data Journalism’.
As outlined previously the diverse nature of the group of students accepted on to the programme means they have to engage with a varied syllabus designed to impart the skills, knowledge and professional reflective capability required to take up jobs across a wide range of areas in Computational and Data Journalism.

Introductory courses in programming, web development, data analysis, data journalism and reporting issues are combined with elective modules in data visualisation, human computer interaction, and various academic and practice focused studies in areas of journalism which are taught between the School of Journalism, Media and Culture and the School of Computer Science and Informatics. The taught part of the course is finished with a team-based investigative project delivered between both schools, before the whole course is capped with the student’s individual masters project where students are mentored by colleagues from the partner schools as issues arise in the workflow.

Another key element, although not credit-bearing, was the creation of a group tutorial space known as “The Lab”. This is something recommended during the programme approval process; the approval panel were concerned that the course would feel very split as the respective classes were taught at the different schools. To get around this idea that they do journalism and computer science, this weekly tutorial slot allows students to work with the course leaders to harmonise the skills learned each week to ensure students see themselves as computational journalists. In the early part of each semester this is initially a problem-solving space around technical elements or news writing but over time develops into an informal project space where students and staff work together to look at issues and develop their skills in research, Freedom of Information Applications and coding – as well as team work and project management.

The formal aspects of the course are supplemented with additional activities outside of the classroom, ranging from visits to academic/practice conferences, to hackdays and innovation workshops involving organisations such as the BBC, Dow Jones, and the NHS. Students are also encouraged to work on side projects, pieces of work outside of the syllabus that can be used as part of their developing professional portfolio.

Reflections

A particular challenge has been delivering a disruptive framework for this innovative educational offer within the traditional University milieu, across two academic schools with differing aims, expectations and constraints.

A unique feature of this course is that despite these operational issues, it is a truly joint-honours programme, with significant commitment and contribution from both the journalistic and computer science teams which focusses on real-world issues and employability. This process has required the course leaders to engage with networks and keep abreast of new skills as they emerge in order to keep the syllabus up-to-date. This can be quite challenging, but real-world examples of data stories or news applications are outlined each week. The use of tools such as Slack and Github also allow easy sharing and conversations among staff and students – allowing them to
become part of a networked learning community. The team has also been scraping journalism job adverts as part of a longitudinal study of role development, which feeds back into curriculum development (Mottershead & Chorley, 2017).

Another major issue is the perception of students themselves, particularly from those coming from an Arts and Humanities backgrounds – they perceive themselves to be creative rather than technical - something not uncommon among journalists and journalism students (Jackson, Thorsen, & Reardon, 2019). One of the key educational methodologies implemented to support the transition of students away from this perception is the idea of applied learning – not focussing on theory for the sake of theory, but to show students what could be achieved and how the underlying theory supports the skills they were developing. For instance, when introducing R programming in the Data Journalism class it was decided not to approach computer science theory and programming logic but instead to show how journalists who can code could solve a wide range of problems around data acquisition and simple visualisation. The reasons for this were straightforward: the students were already learning more theory and programming skills in their Python and Web Applications classes, the applied learning in journalism classes could then be used to draw together the elements across the schools and a discussion of how the Python skills they were learning clearly aligned with the much more applied R coding they were testing in journalism classes.

The Lab was also used in a similar way, to shows the similarities between the schools and the way the module elements were designed to reinforce the idea that computational thinking (Wing, 2006) - a field of problem-solving within computer science – is actually closely aligned with journalistic investigative thinking (De Burgh, 2008; Hunter, 2011; Spark, 1999) and works well within a news room workflow.

Being a new course, potentially of niche (or at least moderately limited) appeal, resource pressures meant that we had to fit as much as possible within existing courses and modules during the programme design phase. We did not have a blank slate upon which we could design a course that was tightly integrated, pedagogically focused, or rigorously slavish to any particular model of learning or curriculum design. Instead we needed to select modules from the existing portfolio of courses offered by both schools, proposing new modules only where there was no alternative. This reduced the risk of the new course, primarily by reducing the cost required to be invested. However, it also meant that we were forced to use existing modules from other courses over which we had limited control.

Following the assessment of the existing modules available, gaps were revealed that could only be filled by new modules. Reflection on the Intended Learning Outcomes for the modules during the first year (particularly when attempting to create assessments for the first time to evaluate students’ abilities to meet these learning outcomes) revealed that initially some of the ILOs were actually difficult to assess. They had been written during the early stages of design for the MSc CompJ, and had not been adjusted since. These ILOs were therefore re-written and made more directly assessable (while not specifying the nature or details of assessments), a clear use of the Constructive Alignment model (Biggs, 1996) in order to reflect upon module outcomes and assessment strategies and adjust accordingly to ensure they are both in line. The nature of the
practical coursework in these modules have also been shown in our group project to be preferred by the students themselves.

The most significant addition of a new module as part of the programme is the module focused on ‘Digital Investigation’. In this module, students work in self-directed teams (as close to a real-world ‘investigation’ team as possible). They are responsible for managing their own projects using an Agile methodology (Fowler & Highsmith, 2001; Maher, 2009), thus giving them exposure to how real news teams operate. They are also responsible for self-managing the learning and skills acquisition required in order to successfully complete their project. Teaching staff are on hand throughout the module to deliver support and learning where required, and to help ensure projects maintain some direction, but otherwise the students are very much on their own. Using a Connectivist learning model (Siemens, 2005), it is assumed that the student teams will work as networks, sharing and gaining knowledge together from a diverse set of sources in a form of peer learning (Havnes, 2008). This model has a particular focus on the use of technology, and the temporary nature of 'knowledge' and the need for continual learning. In this module learning and social aspects of the projects are managed through a set of technological tools such as Slack⁴ for communication, Trello⁵ for project management, and Github⁶ for code and resource sharing.

Obviously, the course was designed with a particular set of assumptions. One of these was around the makeup of the student intake. An initial assumption held that the course would skew mostly towards home students, with international students as a minority. In actuality, the situation has been reversed across the five years of the programme. This has led us to make changes to the course. The harmonisation seminars have been skewed more towards learning about local knowledge, and English language requirements. Language has also become more of an issue during assessments, which have had their assessment criteria adjusted to take into account the fact that one of the learning activities included within assessments might be a drive to improve English language skills. The fact that there are more international students than expected is particularly something to be aware of with regard to the group investigative project. As a group based project, teams are necessarily formed from students of different cultural backgrounds and competencies (Stahl, Maznevski, Voigt, & Jonsen, 2010). This can lead to concern among students that working in multi-cultural groups can have a negative effect on marks, although this is not actually the case (De Vita, 2002), something we have seen clearly in assessments so far. A further assumption was that as a technical course intake would be skewed towards male students. In actuality, the first year was 75% female, and the second year was an equal split male/female. Every year since has been skewed slightly female. The issue of gender equality in STEM subjects is well considered (Baker, 2013; Cheryan, Plaut, Davies, & Steele, 2009), but so far we have found no issues with gender participation in the new course, in fact quite the contrary.

---

Outcomes

The course has seen a growth in the number of students applying to study on the programme over the last five years as it has become easier for people to understand the field due to the major news stories that have broken and an increasing interest in areas such as algorithmic accountability.

As far as the course team is concerned, the major success has been how the programme’s graduates have fared. Students on the programme have taken up paid internships either as part of their dissertation project or before, with employers such as The Financial Times, The Guardian, Full Fact, BBC News Labs, Reach (Trinity Mirror), Dow Jones and CNN.

Graduates have gone on to work for organisations including the UK’s Office of National Statistics (ONS), The Guardian, The Times (of London), India Spend, Switzerland’s NZZ (Neue Zürcher Zeitung), and The Bureau of Investigative Journalism. Others take up positions outside media, working as code-literate graduates in other fields.

Recommendations

From the experience of the programme leaders over the first 5 years of teaching this programme, there are a number of key recommendations:

1. Ensure strong participation, delivery and contact with students from both the Journalistic and Computer Science sides of the programme. This has many benefits. Firstly, for those students going on to work on the more technical or development sides of the newsroom, the stronger formal training in Computer Science competencies provides them with a solid foundation which may not be gained purely from a Journalistic practitioner training in computational methods. Secondly, the exposure to other (non-journalist) students at the same stage in their careers, who are also facing similar challenges and experiences as they begin to take on coding and Computer Science, allows students to see that their experience with taking on a threshold concept such as computer programming is not unique, and there is not something particular about journalists learning to code that somehow makes their situation harder than for others.

2. Ensure the gap is closed between the two disciplines. While fully immersing students in both the computational and journalism has benefits as described above, it can lead to an inability to see how the two relate and work together, particularly in early stages of the course. It is therefore essential either through the data journalism instruction or some form of combined seminars to provide space for students to explore both sides of their education within the context of the other.
3. Provide ‘as close to real-world experience’ as possible. The simulation of a real teamwork environments through group assessment and large-scale projects provides essential experience for students preparing to enter the workplace.

There is little doubt that there is a need for further technical and computational training of journalists as the newsroom diversifies to cope with the informational, algorithmic and data deluge which must be presented to the public in ever-adapting forms. It is hoped that the lessons learnt from the first five years of providing this training will be of use to others looking to do the same.

**About the Authors**

**Martin Chorley** is a Senior Lecturer in the School of Computer Science and Informatics, Cardiff University. He is the co-director of the MSc Computational and Data Journalism, and is also Director of Learning and Teaching for the School. He is co-founder of the European Data and Computational Journalism Conference. His research interests cover aspects of Computer Science education, particularly cross-disciplinary education, as well as the areas of Computer Science where people and computers meet.

**Glyn Mottershead** is a Senior Lecturer in Data Journalism at City, University of London. Prior to that he was founding co-director of the MSc in Computational and Data Journalism, a joint honours course in journalism and computer science at Cardiff University. He is also a co-founder of the European Data and Computational Journalism Conference, and his interests include data journalism, journalism and software development, social media and production journalism.

**References**


QAA. (2011). *Subject benchmark statement Master’s degrees in computing.*


