Emer Purcell, Paul MacCotter, Julianne Nyhan and John Sheehan, eds. *Clerics, Kings and Vikings. Essays on Medieval Ireland in Honour of Donnchadh Ó Corráin.* Dublin: Four Courts, 2015.

# Risky, experimental, emergent: the timeliness and genius of CURIA and CELT

WILLARD McCARTY

CELT, Donnchadh Ó Corráin's Corpus of Electronic Texts project, arose from a long-standing collaboration between the Department of History and the computing centre, University College Cork. In the words of its original mission statement, its intention was to bring 'the wealth of Irish literary and historical culture (in Irish, Latin, Anglo-Norman French and English) to the Internet in a rigorously scholarly and user-friendly project for the widest possible range of readers and researchers'.[1] It is but one of many contributions Professor Ó Corráin has made to our knowledge of medieval Irish history, but it is the only one on which I am remotely qualified to comment. Here, from well outside his major fields, though from an historical perspective and with long commitment to the digital humanities, I intend to honour the man by saying what I think all of us can learn from what he did.

First a note about terminology and its social implications. For purposes of clarity I will distinguish 'digital humanities" (what digitally involved traditional scholars do) from 'humanities computing" (what humanists devoted to computing do). In other words, I here assume that (a) in order for the former to prosper and progress, others have to devote themselves to digital activities across the disciplines, and that (b) as Jaroslav Pelikan argued, both sets of 'colleagues in the research enterprise' must as a matter of justice and accuracy come together as colleagues on level ground.[2] The social and institutional changes necessary to make possible genuinely collegial research between humanities computing and the digital humanities are still in progress and instantiated in only a few places world-wide. In the period of concern here they were just beginning, quite painfully as I recall. But we now know from the success of academic appointments in humanities computing, from achievements of genuinely collaborative work and from PhD programmes in the digital humanities that Ó Corráin's efforts and those like his are bearing good fruit.

Second, a longer note about history. Here I will take as given that without a genuine history the proper study of a subject is not possible. Without a history, all one has are the facts and no idea of how they are related, or what choices were involved in making them, and so no reliable sense of where the subject is going. If the subject is technological, as mine is, then determinism is very difficult to avoid because the way we think and talk about technology is compromised by the discourse we inherit. The best guide I know here is Raymond Williams' in *Television: technology and cultural form*.[3] Once in play, technological determinism substitutes simple causality for the complex interrelations among the many strands of historical development that connect events together. Contingent events then seem inevitable, and we are deprived of responsibility and so of freedom. We are deprived of history.

**1** My thanks to Beatrix Färber and to Peter Flynn for much of the historical information included here. **2** Jaroslav Pelikan, *The idea of the university: a reexamination* (New Haven, 1992), p. 62. **3** First published by Routledge in 1974.

My subject is also difficult to talk about historically because it contains an unresolved paradox: the fusion of technology, for which progress is integral, and the humanities, which change stylistically, but orbit much the same enduring concerns. We in humanities computing have not done particularly well with this paradox to date, tending like all technologists to regard the technical past as inferior (which is in fact not always the case) and then to forget it. For reasons that historian Michael Mahoney has done more than anyone else to detail,[4] a genuine history of computing, and so a history of humanities computing, seem now to lie beyond our powers to write. The fact that computing is so very recent, poorly documented and unreflectively practised makes any attempt frustrating. We can at least do something now about the poor documentation and unreflective practice despite the unfavourable environment for scholarship in which we currently work.

The basic problem is not so much that given our heedless forgetfulness, wheels are reinvented – this is the common complaint of those who can remember the past – but that their invention is as a result of this forgetfulness deprived of its own intellectual context and so its substantial contributions lost. Hence everyone loses. This volume provides an occasion on which to recall a particular invention, CELT (together with the earlier CURIA project), and to place it in a provisional context that will allow us to recover its enduring value beyond the field it was designed to serve.

## *Basic facts*

CELT arose in 1997 out of Ó Corráin's CURIA (acronym for Cork University / Royal Irish Academy, denoting the collaboration). CURIA was founded in November 1991 with a grant from Professor Marianne McDonald, who supported CURIA's *Thesaurus Linguarum Hiberniae* as well as the better known *Thesaurus Linguae Graecae.* In response the UCC Computing Centre, in the person of Peter Flynn, established the first Web server in Ireland for it and the ninth in the world, three months after the World Wide Web itself went online. While (in the words of a student there at the time) most others had very little idea of computing, Ó Corráin was not only collaborating with Flynn to put UCC's potential for Irish medieval scholarship online, but was also teaching the basics of *Nota Bene*, a word processing program with particular capabilities for scholarly writing.

CURIA must have seemed an especially risky venture at the time. The work was not done in one of the great technological centres of the world nor at one of the world's top-rated institutions nor by the hand of a specialist in humanities computing but by a medieval historian and those with whom he found to collaborate at UCC – to list only those I know personally, Flynn (involved with CURIA from the outset), Beatrix Fäaerber, Mavis Cournane and Elva Johnson, the latter two of whom did their PhD research with Ó Corráin.[5]

**4** For Mahoney's writings on the subject, see www.princeton.edu/~hos/Mahoney/ accessed 9 May 2010. **5** According to a memo dated 15 July 1996, the administration of CURIA included the following individuals: Margaret Lantry (Managing Editor); Philip Irwin (Assistant to the Managing Editor); Professor Donnchadh Ó Corráin (Project Director); Dr Damian Bracken (Assistant to the Director). The scholars and technical assistants were: Jennifer Brannen, Mavis Cournane, Vibeke Diijkman, Dara Mac Domhnaill, Letitia MacSweeney, Audrey Murphy, Ivonne Tummers. Those formerly associated with the project were: Pamela Butler, Dr John Carey, Maurice Cronin, Teresina Flynn, Rachel Granville, Dr Kaarina Hollo, Dr Bart Jaski, Elva Johnston, Pádraig Kelliher, Leeann Lane, Robert Lotty, Seosamh Mac Muirí, Maeve McAllister, Jane McCarthy, Gillian McIntosh, Angela Malthouse, Dr Nicole Müller, Charles O'Hara, Aideen O'Leary Mlitt, Denis O'Sullivan, Regina Sexton, Donna Thornton, Professor H.L.G. Tristram. The Consultative Committee consisted of

Nearly 20 years after CURIA began, so much having happened meanwhile, we might be inclined to give these projects no more than a cursory though appreciative nod if we stumbled across them. We might judge them a commendable effort for the time, but regard them simply as part of the furniture, and not the most stylishly up-to-date furniture at that. Having discovered how unusual both were for their time, we might pause somewhat longer in admiration. But what could we learn on behalf of the digital humanities that would actually be of value now, in 2014?

## *Prior history*

Innovative CURIA and CELT certainly were, but they were established in the context of work that had been going on since 1949, when Fr Roberto Busa SJ began his great *Index Thomisticus* project, to provide a lemmatized concordance to all the writings of St Thomas Aquinas.[6] Let me briefly set the scene Ó Corráin entered in the early 1990s.

As Edward Vanhoutte has pointed out, Busa's foundational work was itself preceded by the great, though doomed, Machine Translation (MT) project, which arose out of work done for the war effort and which was powered both by curiosity and by the imperatives of the Cold War. Before MT began in an exchange of letters between Warren Weaver and Norbert Wiener two years earlier,[7] variously relevant work can be traced back as far as one wishes to go. Like Ó Corráin's later wise seizure of the day, Busa saw in the then very new developments in computing an opportunity to further medieval scholarship. Since then humanities computing has had about six decades to develop. But it did not begin to affect mainstream disciplines significantly until the introduction of the Web, in August 1991, which made possible the online resources and publication venues we now have. Although the Web lacked the technical sophistication and analytic power of some previous tools for humanists, such as the hypertext system FRESS[8], it offered an appealingly simple, easily accessible and visually appealing way of providing resources for scholarship. Common appeal triumphed over technical superiority as has happened in the past.

Before 1991, with very little capabilities for displaying and distributing images, the emphasis in humanities computing and among its disciplinary fellow travellers was confined mostly to textual data, which at the time seemed the low-lying fruit of the digital harvest. Since displaying text for reading is also a matter of imaging, computing was used chiefly in the early years for preparing concordances and *indices verborum* for print. After microcomputers (as they were then called) became commonplace and interactive concordancing thus practical for analytic purposes, text-analysis, including elementary stylometry, was also relatively easy to do. Text-analysis may be defined as the algorithmic generalizition of concordancing, extending the manipulation of 'running text', as the

Professor Marianne McDonald MRIA (UC San Diego), Chair; Professor Seamus Deane MRIA (Notre Dame); Professor L.G. Huxley MRIA (TCD); Dr Patricia Kelly (UCD); Professor J.J, Lee (UCC); Professor Pádraig Ó Riain MRIA (UCC); Professor V.J. Scattergood MRIA (TCD); Professor Barbara Wright MRIA (TCD). The Editorial Board consisted of Professor Máirtín Ó Murchú MRIA (DIAS); Mr Peter Flynn (UCC); Mr Martin Hayes (UCC); Dr Patricia Kelly (UCD); Professor Donnchadh Ó Corráin (UCC); Prof Máirín Ní Dhonnchadha UCG); Prof Breandán Ó Madagáin (UCG); Dr Michelle Ó Riordan (DIAS). **6** R. Busa, 'The annals of humanities computing: the Index Thomisticus', *Computers and the Humanities*, 14 (1980), 83–90. **7** Warren Weaver, 'Translation' in William N. Locke & A. Donald Booth (eds), *Machine translation of languages: fourteen essays* (Cambridge, MA, 1955). **8** Steven DeRose & Andries van Dam. 'Document structure and markup in the FRESS hypertext system', *Markup Languages: Theory and Practice*, 1:1 (1999), 7–32.

corpus linguists say, as far as we can devise algorithms to go. Text-analysis was a direct development of the centuries-old practice of making concordances, with its roots in biblical exegesis and philology, with obvious connections to two disciplines it influenced profoundly, lexicography and linguistics – hence modern dictionary-making practice and corpus linguistics, respectively. On the literary-critical side, text-analysis was inspired or imprinted by the then-ageing New Criticism, with its strongly positivistic conception of text.[9] But by that time literary criticism was moving rapidly away from New Critical preoccupations, stranding text-analysis among a small minority of computer-using scholars, who have persisted but continue to complain about their minority status. They have had little-to-no effect on mainstream criticism. Among this minority are the literary stylometricians, such as John Burrows, Hugh Craig and David Hoover, who have demonstrated beyond doubt the presence of statistically significant clues to readerly impressions, for example, of authorial voice. They are a small minority because the work they do requires an unusual assembly of talents and ascent of a very steep learning-curve.

My point is this. As many have subsequently realized, but Ó Corráin saw immediately once the Web had become available, the central computing task for the disciplines shifted from analytics to a great stocking of the virtual shelves. Doing that may at the time have been harvesting the very lowest lying fruit, but when food is scarce the imperative to pluck it is unarguable. For medieval Irish materials particularly – of strong interest but to a relatively small number of people scattered very widely around the world – the Web presented a great opportunity to strengthen the discipline. Ó Corráin's was a brilliant move, though, I'd suppose, quite risky at the time, when all that we know now was at most one set of possibilities among many, with its outcome unknown and potential for corruption by dreaded techno-science greatly feared. To most scholars, I recall, this was not an attractive way to spend one's time. Taking the risk leads me to my central point, to which I will return in a moment.

Before the Web, in the absence of any practical means of putting texts into wide circulation without significant technical overhead, energies of the few, mostly literary scholars, were focused on text-analysis. After the Web went online, the imperative to provide texts directly to the scholar's computer was both obvious and difficult to gainsay, though the obviousness of the benefits took most of us quite a while to see, and the ability of the ignorant to gainsay these benefits formidable. Attention to analysis, with its already poor showing, waned. Literary scholars could use texts more or less as they had before, but historians, linguists, lexicographers and others now could do much that had been largely out of reach previously. In the last few years attention to text-analysis has returned, in the form of experiments on large quantities of textual data (a.k.a. data-mining), more theoretically informed questioning of what text is (that it might be analysed) and probing of the capabilities of markup. From a literary-critical perspective we are still far from the point at which our best theories of text can be modelled in software, however. We are still struggling to understand how best to put resources most effectively online.

In 1991, when the Web began and Ó Corráin began using it, what he did with it was exactly right. Luck? Clearly not, if we mean, with the *OED*, 'Chance regarded as a cause or bestower of success and failure'. Insight? Clearly yes. But my main point goes beyond pure intellection, however admirable Ó Corráin's may be, to that which escapes the net of reason entirely.

**9** Susan Wittig, 'The computer and the concept of text', *Computers and the Humanities*, 11 (1978), 211–15; Jerome McGann, 'Texts in n-Dimensions and interpretation in a new key [discourse and interpretation in n-dimensions]', *Text Technology*, 2 (2003), 1–18.

## *Being bold*

Edmund Spenser, with his lands in Co. Cork and his argument for the destruction of Irish customs and language, is not the most popular of historical figures in Ó Corráin's homeland, but Spenser's quotation in *The Faerie Queene* of the old English proverb, 'Be bold, Be not too bold' (III.5.32) is apt for one of its greatest academics. This proverb points me to that which I wish here to celebrate: the derring-do of Ó Corráin's doing. Irish historians will, I suspect, have a different take on what is to be celebrated. And derring-do is not always easy to live with. But without it there is no life. My take is perhaps best illustrated by Norbert Wiener's normative declaration, in *The human use of human beings*, that 'Among the animals, man is a Peter Pan who never grows up'[10] – who never, unless out of sloth, social conditioning, fear or lack of imagination, shuns the genius of the humanities to enlarge 'the alternativeness of human possibility' and collapses into what is fixed, proper, approved, safe.

Of course it is an easy thing and product of simple hindsight to conclude from the career of the successful that the choices made can be fashioned into a moral lesson we must take to heart and teach to our students. Technological determinism, running underground if not also above it, makes such conclusions particularly appealing to the unwary. Does the historian know better? Perhaps not always, since the historian is also an ordinary person who lives in the world and so is subject to its powerful influences. The question of whether technology drives history, and if so, how and to what degree, is also an open one among those who think about it critically.[11] But Thomas Misa's observation, that technological determinism is especially favoured by two groups – philosophers of technology and military historians – qualifies their support: the former, he argues, because they are 'unconcerned with the messiness of the historical process',[12] the latter because their subject has in the past so often demonstrated the futility of resistance to better hardware. Those we call terrorists have shown, however, that low-tech cleverness can win over high-tech might. So, as Williams points out, determinism is clearly even under conditions which most favour it: a one-sided oversimplification of limits set and pressures exerted, by people and their inventions, 'within which variable social practices are profoundly affected but never necessarily controlled'.[13] Freedom of choice is possible – and can be realized if one has what it takes to take risks and know when to take them.

I cannot go any further with historiographical questions here. Rather the conclusion I draw is all about openness to opportunities for keeping alive, indeed nurturing those fundamental desires that pulled us into the academic life in the first place. Not just curiosity but also reprocity: *do ut des*, giving that others may give, as Professor Ó Corráin did in this and many other ways. Here's mine in return.

**10** Norbert Wiener, *The human use of human beings: cybernetics and society* (Boston, 1954), p. 58. **11** Cf. e.g. Merritt Roe Smith & Leo Marx (eds), *Does technology drive history? The dilemma of technological determinism* (Cambridge, MA, 1994). **12** Thomas J. Misa, 'How machines make history, and how historians (and others) help them to do so', *Science, Technology, and Human Values,* 13:3–4 (1988), 308–31 at 310. **13** Williams, *Television: technology and cultural form*, p. 133.