7. Digital Methods

From Challenges to *Bildung*

Bernhard Rieder & Theo Röhle

In a previous text (Rieder & Röhle 2012) we argued that the existing traditions of the humanities and social sciences, including their particularities, interests and methods, are currently encountering an object – the computer – that is characterized by its own logics, logistics, styles of reasoning (Hacking 1992), habits, (best) practices, modes of valorisation, actor-networks and institutions. The computer may well be a contained technical object, but its accumulated history and therefore its substance is full of heterogeneous elements that constitute a type of *a priori* that cannot be easily ignored. Now that various attempts are being made to build 'digital' versions or extensions of long-established disciplines, this encounter marks a moment of destabilization and deterritorialization, a moment that implies significant contingency and different possible outcomes. Although it remains doubtful that even Kuhn's 'normal science' (1962) was ever truly settled, this is a moment that provokes and requires far-reaching debate and inquiry into the practice, meaning and purpose of our academic disciplines.

The encounter between the humanities and computing plays out in different ways in different arenas, but needs to be addressed in principle as well as in relation to particular settings. The fact that after 50 years of experimentation many of the fundamental questions remain deeply controversial can be seen as an indicator for how close these questions come to core debates about the means and purposes of scholarly work. While terms like 'digital humanities', 'Cultural Analytics', 'digital methods' or 'web science' can play the role of buzzwords, their proliferation can be seen as indicator for a 'computational turn' (Berry 2011a) that runs deeper than a simple rise of quantitative or 'scientific' modes of analysis. Large and often unusual data sets, advanced visualization techniques and fuzzy processing have led some of those who have held numbers, calculations and computers at a safe distance for a long time to warm up to new computational possibilities. Our core question was therefore: If these new methods are more than just another set of tools in our arsenal, how do we deal with the fundamental transformations that challenge established epistemological practices and paradigms?

The starting point for our previous investigation was the concept of 'method'. Defined by the Oxford English Dictionary as 'pursuit of knowledge, mode of investigation', we are also reminded that this pursuit is both systematic and orderly. Additionally, method is directed and purposeful: specific decisions are tied to specific goals. Like a blueprint or recipe, research methods guide behaviour and even if some of our approaches are only moderately explicit, their commonality allows us to share experience and establish reference points that provide orientation – even when there is little agreement on utility and validity.

Although we are wary of Tom Scheinfeldt's assessment of ours as a 'post-theoretical age' (Cohen 2010), his diagnosis of a 'methodological moment' is certainly appropriate. Coming from German academic tradition, we developed our perspective against a backdrop of decades of *Methodenstreit* ('methods dispute'), beginning with Weber's endorsement of sociology as an 'understanding' (*verstehend*) rather than an 'explaining' (*erklärend*) discipline, which later morphed into the famous *Positivismusstreit* ('positivism dispute') between Adorno and Popper. Part of this was the sometimes profoundly paralysing and sterile opposition between quantitative and qualitative research methods in empirical social science. While not truly analogue to Snow's 'two cultures' problem (1959), there are certainly parallels here that point towards different ways of knowing and thinking – styles of reasoning – caught up in larger normative horizons, as seen in the altercations between 'critical' and 'administrative' types of research, epitomized by the clash between Adorno and Lazarsfeld.

Our refusal to cede to simple oppositions is built on an anti-essentialist approach to many of the concepts that appear in these debates. Computation, quantification, algorithm, visualization, graph, data analysis, statistics, software, and so forth, are terms that point to concepts – but also to objects, practices and skill sets – that we consider to have considerable internal heterogeneity and variation. That does not mean that they are not caught up in particular configurations and constellations that are productive in very specific ways in terms of knowledge and power; but it means that the spaces of design and 'appropriation' (Akrich 1998) of computational methods afford considerable leeway and do not translate into or perform singular logics. Even if 'the digital' has become a dominant passage point, it works like a meat grinder: the shredded material does not come out as a single thread, but as many.[1] To connect back to the *Methodenstreit*: compu-

1 For a detailed investigation into different types of digital processing, see Winkler (2015) (where the meat grinder is actually used metaphorically on the cover).

tational methods can be both deductive and inductive (see e.g. Tukey's (1962) concept of exploratory data analysis), both quantitative and qualitative in outlook, both critical and administrative. But these spaces of movement, of epistemic freedom have to be constructed and defended, sometimes by forging alliances, sometimes by demarcation; certainly through a better understanding of what computers can actually contribute to knowledge production and of the ways they produce this epistemic 'surplus'.

If digital technology is set to change the way scholars work with their material, how they 'see' it and interact with it, a pressing question is how these methods affect the way we generate, present and legitimize knowledge in the humanities and social sciences. In what way are the technical properties of these tools constitutive of the knowledge generated? What are the technical and intellectual skills we need to master? What does it mean to be a scholar in a digital age? To a large extent, the answers to these questions depend on how well we are able to critically assess the methodological transformations we are currently witnessing.

As a growing range of investigations into the status of (big) data (e.g. Gitelman 2013; Elmer, Langlois & Redden 2015; Amoore & Piotukh 2015), as well as ongoing discussions in the digital humanities (Gold 2012; Arthur & Bode 2014; Svensson & Goldberg 2015) suggest, there is something deeply complicated about this methodological moment. We argue that, if some of the criticism being phrased towards the wider field of digital humanities and social sciences is indeed justified, this should not be seen as discouragement, but as a challenge, in the most engaging sense of the term.

In this chapter, we want to shortly summarize what we consider to be five central challenges before interrogating Berry's concept of 'digital *Bildung*' (Berry 2011a) as a means of facing these challenges. Our goal in this discussion is, maybe paradoxically, to move the spotlight from 'the digital' and the associated über-skill, programming, to the plethora of concepts and *knowledges* mobilized in digital tools. To this end, we discuss three examples that allow us to both concretise and complicate the debate.

Five Challenges

In our previous paper (Rieder & Röhle 2012), we presented a non-exhaustive list of broad issue clusters that we believe have to be addressed if we want to productively integrate the new methods without surrendering control over the conceptual infrastructure of our work. Our question was not how to conduct 'good' digital research in the narrow sense: we were not concerned with specific methodological pitfalls or 'errors' in data collection, or with the choices and applications of methodological tools, but with the larger ramifications of digital research inside the field of the humanities and social sciences. In that sense, we wanted to tackle the challenges faced by even the 'best' work in the field.

A first challenge, which we called 'The Lure of Objectivity', raised the question why computational tools have sparked such a tremendous amount of interest when it comes to studying social or cultural matters. One explanation might be the notion that the computer is able to reach beyond human particularities and into the realm of objectivity. We discussed the fascination that the ideal of detached, mechanical reasoning was able to induce historically and asked whether this fascination might keep us from laying bare the many explicit and implicit decisions that went into our tools and instruments. Questions of bias and subjectivity, which the computer was thought to do away with, enter anew on a less tangible plane – through the choices concerning modes of formalization and algorithmic procedures, as well as through the various ways data processing can mask partiality (see Barocas & Selbst 2015). This becomes an especially pressing problem when studying commercial social media platforms. Considering the 'politics of circulation' (Beer 2013) that these platforms are embedded in and the resulting elaborate ecosystems of API regulations (Bucher 2013; Puschmann & Burgess 2014; Rieder et al. 2015), issues of preselection constitute a major methodological dilemma. The challenge is thus to accept the fact that, on an epistemological level, computational methods often create complications rather than resolve them.

Under the heading 'The Power of Visual Evidence', we discussed the role of visual output, such as depictions of network topologies, timelines or enriched cartographies. Since these visualizations possess spectacular aesthetic – and thus rhetorical – qualities, we asked how the argumentative power of images could (or should) be criticized. We stressed the tradition of critical inquiry into the use of images that the humanities have fostered over the years, but remarked that the situation now has indeed changed, since digital humanists themselves produce and rely on images as evidence and heuristic devices. The challenge is thus to maintain a productive self-reflexive inquiry into our own visual practices, i.e. to acknowledge how analysis and cognition are both partial and interwoven with power relations – both currently and historically (Halpern 2015) – without abandoning the promise of gaining insights via visual forms (Drucker 2014: 130-137).

'Black-boxing' referred to our ability to understand the method, to see how it works, which assumptions it is built on, to reproduce and criticize it. Despite the fact that writing software forces us to make procedures explicit by laying them out in computer code, 'readability' is by no means guaranteed. However, an open process of scrutiny is one of the pillars of scholarship and, in the end, of scholarship's claim to social legitimacy. We argued that this problem presents itself on at least three different levels: a) concerning the practical possibility to access the most obvious layer of functional specification, i.e. a tool's source code; b) concerning the ability to understand the code and, even more importantly, the ability to grasp its epistemological ramifications, and c) concerning methods that become opaque despite being fully explicit, such as techniques issued from the field of machine learning, where the connections made between inputs and outputs can no longer be easily retraced by human observers. This point really concerns the question how the epistemological surplus that is provided by computation can be specified, controlled and relayed to others without falling victim to the sometimes deceptive simplicity of graphical user interfaces and shiny visualizations.

We identified 'Institutional Perturbations' as a fourth set of challenges. We saw a chance that, given the growing need for computational expertise, the humanities may increasingly hire researchers from computer-adept disciplines. Also, computational methods may have advantages in settings where even humanistic research is increasingly financed on a project basis – which implies very particular pragmatics based on structured time frames, planned expectations and identifiable 'deliverables'. The challenge, we argued, is to develop a sensibility for such wider repercussions of methodological innovation. In many areas there is an argument to be made for the confident defense of methods that are based on principles other than mechanized 'persistent plodding' (Wang 1963: 93).

The fifth issue we highlighted was 'The Quest for Universalism'. Here, we argued that the establishing of pervasive concepts and principles becomes increasingly common whenever computers come into play. When reality is perceived to adhere to a specifiable system of rules, the computer appears to be the quintessential tool to represent this system and to calculate its dynamics. The epistemological commitments and reductive nature of the underlying models are often 'forgotten' when it comes to the explanations derived from them. Instead, the scope of the explanations is extended indefinitely, reminiscent of the universalist aspirations running through historical discourses on computation. Concepts from network science are a case in point. The challenge is, thus, to arrive at a more adequate demarcation of the explanatory reach of formal models, e.g. by combining different methodological configurations, both digital and non-digital.

In terms of a conclusion, we continue to advocate involvement with the new methods. By involvement, we mean both the actual application of these methods and a critical reflection of such uses. We thus argue for a transfer of the concept of 'critical technical practice', proposed by Agre (1997a), to the scholarly domain: a practice that oscillates between concrete technical work and methodological reflexivity. Current approaches that draw on Agre's concept hold a lot of promise in this regard. As Matt Ratto, Sara Ann Wylie and Krik Jalbert (2014) argue, actual engagement with materiality – what they call 'critical making' – can be a productive complement to the traditional linguistic forms of knowledge production, also in fields such as STS and media studies. Rather than developing methods with a clear goal in mind, the design process can be a means to advance a more inquisitive attitude towards our digital environments – 'bringing unconscious aspects of experience to conscious awareness, thereby making them available for conscious choice', as Sengers et al. state in their outline of 'reflective design' (2005: 50).

In what follows, we want to focus specifically on the challenge of black boxing and, more generally, on the role of digital tools in emergent research constellations. All of these challenges, however, connect more or less directly to the question what we need to know in order to make this critical, reflective, inquisitive and nuanced practice a reality. We thus turn to the matter of knowledge and skill, which has been discussed with particular vigor in the digital humanities, often with a focus on programming as the watershed expertise that separates 'who's in and who's out' (Ramsay 2011). We consider this emphasis to encode a somewhat reductive understanding of computing and suggest a deeper appreciation of both conceptual and technical knowledge and practice in the face of an ever increasing arsenal of digital methods.

From Challenges to *Bildung*

In this section, we approach the question of the challenges for and to (digital) humanities and social sciences through the lens of Berry's notion of 'digital *Bildung*', 'a liberal arts that is "for all humans"' (2011b: 20), although we will focus on digital humanists and social scientists rather than a general public.[2] Our question is what we need to know to become digital

2 Berry's description of digital Bildung as 'a rolling process of reflexive thinking and collaborative rethinking' (2011b: 22) seems to share many characteristics with design traditions that invoke Donald Schön's notion of 'reflection-in-action' (1983), as Agre (1997b: 10) also does.

scholars able to 'examine, theorise, criticise and imagine' (*ibid.*: 169) research methodology – the systematic and reasoned pursuit of knowledge – that is caught up in computation. Ultimately, we believe that this debate remains vague and superficial without a concrete set of references. We will therefore discuss three examples, which we hope will contribute to a more in-depth discussion of how the challenges we identified can be related to a broader notion of digital *Bildung*.

A key question in this discussion is whether it is possible (or desirable) to train 'computationally enlightened' humanists who will themselves actually write the computational methods they will apply in their analyses. We hold that this notion is tempting, but ultimately unrealistic and even potentially problematic: while anybody can learn to write a bit of code in a couple of days, the practice of programming or software development requires far-reaching acculturation and many, many hours of practice. If we consider disposable time as a limited resource, the priority given to programming may actually come to the detriment of other technical and conceptual skills that facilitate the critical understanding of computational procedures. The singular focus on code may detract from what is actually coded.

Because for any experienced programmer, code may well be the medium of expression but, just like a writer attempts to say something through language, the meaning expressed through programming is functionality; and while the two cannot be fully separated, programmers and computer scientists generally reason on a conceptual level that is certainly circumscribed by the requirements of mechanical computation – what one of us has called the 'shadow of computation' (Rieder 2012) – but expressible in various forms, from systematized vocabulary and conversation to flowcharts and, more often than not, mathematical notation. While implementation is certainly not irrelevant, the methodological core, the very definition of what computation adds resides in what the program does. This functional level can be of daunting complexity, even if many sophisticated techniques can be boiled down to a small number of central ideas. Subsuming these ideas under the broad notion of 'the digital' locks the analysis to a surface view that risks hiding the methodological substance or rationale of the work performed by methods rendered in software. Facing the challenges outlined above depends, at least in part, on whether we are able to get to the conceptual core of the computational techniques we are using. Only then can we assess the potentials, limitations and styles of reasoning held by the tools we integrate into our research configurations.

To flesh out this argument in more depth, we turn to three examples that allow for a nuanced approach and highlight the difficulty of setting overarching principles. In all of these examples, we ask what 'understanding' a computational technique would mean.

Statistics

Since the empirical social sciences have been using digital tools as integral part of their work for decades, applied statistics is a good place to start. One of the most widely used software packages in the Social Sciences is SPSS (formerly Statistical Package for the Social Sciences) and the significant reliance by researchers on this program begs the questions to what extent these scholars are capable of 'understanding' – or even seek to understand – the considerable methodological and epistemological choices and commitments made by the various analytical techniques provided. If we consider, for example, regression analysis, a technique that is extremely productive (literally, no endorsement implied) in academic research as well as in business and government, as a means to produce an epistemic surplus, how would we go about understanding more precisely what the technique and its intellectual contribution consists of?

The source code of SPSS is not available, but the way the software calculates its analytical measures is well documented in mathematical notation and relies on established and much discussed constructs such as the Pearson coefficient for correlation (r) or established regression techniques. Looking at an open-source alternative such as PSPP (no acronymic expansion), what would we actually gain from reading the source code instead of simply consulting the documentation and checking the research papers it refers to?

While a critique of the standardization and streamlining of research through widely available software packages is important and raises many concerns,[3] it does not tell us how epistemological agency can be wrestled back from tools that make exceedingly complex methodological procedures available through simple graphical interfaces. A critique of digital tools is incomplete without a critique of their *users* and the wider settings they are embedded in. As banal as it may sound, what is required to understand and use SPSS reflectively – or any statistics package for that matter – is a robust understanding of statistics and probability theory, not a crash

3 See Uprichard et al. 2008 for an in-depth discussion of the significance of SPSS for sociology.

course in Java. What is black boxed in such a tool is not merely a set of calculative procedures, which are, in the end, sufficiently well documented, but statistics as a field that has not only its own epistemological substance, but many internal debates, contradictions and divergences. The 'thirteen ways to look at the correlation coefficient' identified by Rodgers and Nicewander (1988) and the debates around null hypothesis testing, which Gigerenzer, Krauss and Vitouch (2004) refer to as the 'null ritual', are just two of many examples for the quite fundamental disagreements in the practice of applied statistics. While software can be designed in a way that highlights these divergences, it is too much to ask of a program to carry the weight of providing an education in the field it is mechanizing. This raises and complicates the question of the educational embedding of digital tools. If students and researchers are trained in using these tools without considerable attention being paid to the conceptual spaces they mobilize, the outcomes can be highly problematic. Digital *Bildung* thus requires attentiveness not just to the software form, but to the actual concepts and methods expressed and made operational through computational procedures.

Network Analysis

A very similar argument can be made for the popular field of network visualization. It is again important to notice that the point and line form comes with its own epistemic commitments and implications, and graph analysis and visualization tools like Gephi (Bastian et al. 2009) further structure the research process. But where do we go from there? If we consider that graph theory still provides powerful and interesting means to analyse a data set, what would critical analytical practice look like? For example, how can we consider the layout algorithms that transform n-dimensional adjacency matrices[4] into two-dimensional network diagrams? These artefacts interpose themselves as mediators because each algorithm reveals the graph differently, highlighting specific aspects of its structure, thus producing a specific *interpretation*.

There are different families of algorithms – most approaches are based on force simulations, but other strategies such as simulated annealing exist as well – but even the same algorithm, fed with different parameters, can

4 An adjacency matrix is a way of representing a graph as a special kind of table (a square matrix) that specifies which nodes are connected to each other.

produce quite different outcomes. If we apply the ForceAtlas2 algorithm (Jacomy et al. 2014) to a graph file, should we go to Gephi's source repository on Github and search for the ForceAtlas2.java file and try to make sense of it? What would we find there? A few hundred lines of Java code that implement a highly iterative simulation of attracting and repulsing forces that makes ample use of the notion of 'swinging' (in a very literal sense) to find an 'optimal' position for nodes on the canvas without getting stuck in local optima.[5] It is very naive to believe that anybody who has not had considerable training in both programming and simulation modelling can say anything meaningful about how ForceAtlas2 is implementing the force-direction concept differently from its historical and conceptual ancestor, the work of Fruchterman and Reingold; and much less how these differences affect spatialisation in concrete circumstances. How will properties of nodes and topological structure affect positions on the map? Which aspects of the latent structures in the data does the diagram reveal?

Even with the required training, testing and running the algorithm on different data sets with different parameters is a necessity to begin to understand how outcomes relate to instances of computation because no human brain can anticipate the result space of even simple functions iterated thousands of times. The problem, again, comes from the fact that tools such as Gephi have made network analysis accessible to broad audiences that happily produce network diagrams without having acquired robust understanding of the concepts and techniques the software mobilizes. This more often than not leads to a lack of awareness of the layers of mediation network analysis implies and thus to limited or essentialist readings of the produced outputs that miss its artificial, *analytical* character. A network visualization is closer to a correlation coefficient than to a geographical map and needs to be treated accordingly.

We would again argue that the critical mastery of the methodological substance introduced by the software would be best served by studying material on graph theory, graph spatialisation and, in particular, literature on concrete analytical applications. Looking into the history and state of the art of sociometrics and network science would be helpful to acquire 'graph literacy'. To be even more concrete, an in-depth study of Linton Freeman's The *Development of Social Network Analysis* (2004) would be a good start. Inevitably, spending considerable amounts of time trying

5 Consider an analogue problem: a simple algorithm for hill climbing consisting of 'always go up' will end up on top of a hill (a local optimum), but not necessarily on the highest one (the global optimum). Swinging counteracts a similar problem of getting 'stuck' in a local optimum.

out different algorithms on different data sets to build understanding of the specific ways they interpret a graph is crucial. Reflective practice requires much more than a critical attitude, it requires *deeper* involvement with the associated knowledge spaces to make sense of possibilities and limitations.

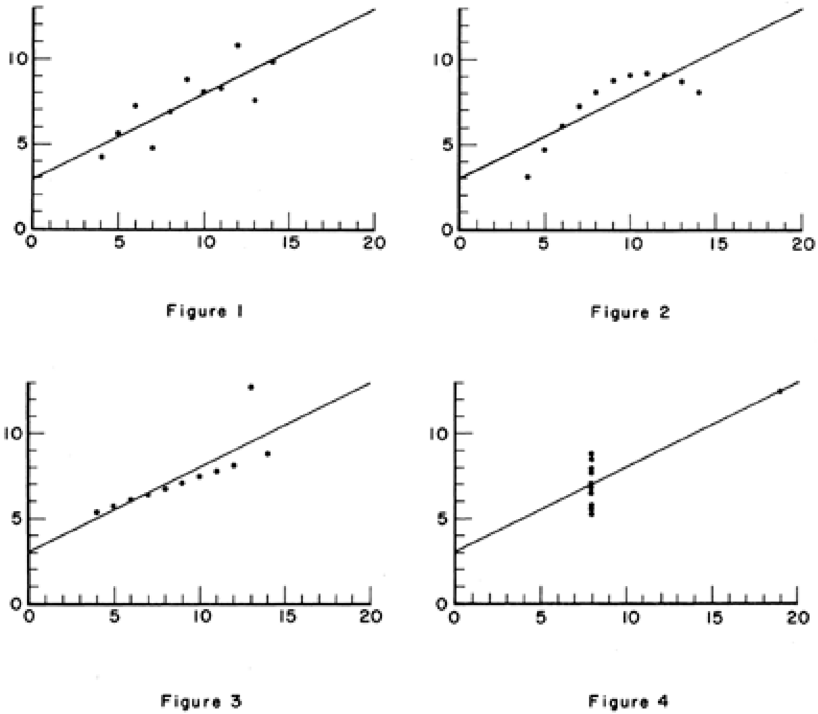
Thousands of Images

These two examples are certainly not fully representative of the tools used in the field, but the argument can be extended beyond the more complex software packages just discussed. The work that Lev Manovich (2012) has done under the label 'Cultural Analytics' can serve as an example: to order black and white manga images on a scatter plot, Manovich uses 'entropy calculated over greyscale values of all pixels in a page' for the y-axis, and after pointing to the history of the entropy concept explains what this measure expresses in terms of the images in question: 'If an image consists of a few monochrome areas, its entropy will be low. In contrast, if an image has lots of texture and details and its colours [...] vary significantly from place to place, its entropy will be high.' (Manovich 2012: 266).

Independently of what we think about what Manovich is doing with these images in intellectual terms (Art history? Image science?), it is his considerable training and experience in working with digital images that allows him to confidently relate a mathematical measure to actual visual properties of the images in question. We are not qualified to say whether the results Manovich gets from this operation are truly useful for his analytical goals, but this is not the question here. What matters is that the skill applied in this example is the capacity to reason on images in formal or mathematical terms, to connect these terms to visual properties of the image as it is perceived by humans, and to derive an epistemic surplus from the whole operation. What would we gain from looking at the source code of Manovich's script? Perhaps we would find an error. Perhaps we could come up with a more efficient implementation. But although Manovich does not provide the used measure in mathematical notation (why not?), his reference to Claude Shannon is a good reason to believe that the entropy measure in question is something like -sum(p * log2(p)), where p contains the image's histogram in 256 bins if the image is encoded in 8-bit.

Now, just like Anscombe's famous four data sets (1973) that are quite different in structure but have the same statistical properties, a very synthetic measure like entropy, which expresses something about a complex object

Fig. 7.1: The four scatter plots from Anscombe (1973). They have identical values for number of observations, mean of the x's, mean of the y's, regression coefficient of y on x, equation of regression line, sum of squares of x, regression sum of squares, residual sum of squares of y, estimated standard error of bi, and multiple r2. Anscombe uses them in an argument for the usefulness of visualization in statistics.



such as an image in a single number, can label a very large number of very different images with the same value. Thus, Manovich not only had to commit to the entropy measure as such, but also to the entropy measure as it reacts to the data set in question. From what we understand, a greyscale gradient would have a very high entropy value since the histogram does not contain any information on how the colours are spatially distributed; it's a simple occurrence count for every colour. Would a certain colouring style in a manga thus 'break' the measure? For certain data sets – Barnett Newman or Piet Mondrian maybe? – the measure could be completely useless because the salient element would be the arrangement of surfaces rather than the probability distribution of colours.

There is no doubt that programming skills are useful in this context. But entropy is not a 'programming' concept; it is, like most statistical measures, a means to summarize data, a means to speak about data from a very particular vantage point. It is reductive, certainly, but reductive in a *specific way* and therein lies its epistemic character. As a concept, entropy ties into the complex histories of information theory and statistics[6], and reflective use will have to attend to these connections.

This is the work digital humanists and social scientists have to do and they cannot easily delegate it to computer science collaborators or hired programmers. Notice that this is a complex technical discussion that does not contain a single question about programming. Any somewhat capable programmer could produce a script from the specification 'calculate entropy from greyscale histogram' and in environments like MATLAB there are even predefined functions that do all the work for us. The actual methodological 'content' and commitment is simply not a question of 'software', first and foremost. Certainly, we can only do this because there is software in the first place, and interfaces hide and cement our commitments, but the knowledge required to judge the method in question is only in very small part related to the question of code; rather, it spans a space from information theory to art history and visual studies in a way that certainly involves abstraction, but of a different kind than programming implies.

Conclusions

While our three examples might be considered very specific, we think that similar arguments could be made for a wide variety of cases where software performs a *method*. While methodological concepts and techniques enter in negotiation with implementation, the 'content' of software is a procedure *expressed* in code, not simply code. We can certainly find cases where the mathematical dimension of a tool is completely trivial, but we would argue that in most of the tools that are used by digital scholars, significant methodological work is performed by techniques that have their origins in the conceptual substance of disciplines such as statistics, information science, sociometrics, computer science and – quite often – mathematics.

6 For an account of these histories that is accessible to and interesting for humanists, see for example Christian Kassung's (2001) contextualization of Robert Musil's 'Man Without Qualities' within modern physics, esp. pp. 132-260.

And this is the crux, here. Although we fully agree with Berry (2012) that digital *Bildung* – in particular for the digital humanist, but also beyond – would benefit from 'iteracy [...] defined broadly as communicative competence in reading, writing and executing computer code', the focus on programming as 'writing code' rather than 'implementing a technique' runs the risk of missing this more conceptual level that is, in our view, both epistemologically more relevant to scholarship relying on digital tools and, in many cases, more accessible in terms of skills to acquire.

While our evidence is only anecdotal, we notice in much of the humanities a desire to explain technology as quickly as possible through something else. Instrumental rationality, cybernetic utopias, neoliberalism, phantasies of perfect control, positivism, revenue maximization, and so forth. These assessments may ultimately be enlightening and meaningful at a very broad level of analysis. But if we want to meet the challenges of computational methods, we have to encounter technology as technology for at least a little while. Paradoxically, the one-sided focus on the 'digital' aspect of computational methods and, in conjunction, on programming as the *Via Regia* to digital enlightenment implies a reductionism that, again, serves to keep technology 'small'. There is no doubt that programming skills and 'iteracy' are extremely valuable and a way to ease into some of the harsher complexities involved in computational methods. But we hope to have demonstrated through our examples that the tools we have come to use mobilize wide arrays of knowledge that we should only grudgingly compress into the supposedly coherent category of 'the digital'. The problem of black boxing does not begin with the opacity of computer code, but with the desire to banish technology from the 'world of signification' (Simondon 1958: 10).[7] Behind the laudable efforts to increase levels of technical capacity lies the dangerous phantasm that technology's epistemologies are ultimately 'thin' and that once programming skill has been acquired, mastery and control return.

We believe, on the contrary, that any nontrivial software tool implies thick layers of mediation that connect to computation as such, certainly, but in most cases also imply concepts, methods and styles of reasoning adapted from various other domains. We can critique the standardization of research through software all we want, but, to put it bluntly, there is no critical practice of statistics without considerable knowledge of statistics,

7 'Culture is out of balance because it recognizes certain objects, such as the aesthetic object, and grants them the right of residence in the world of meaning, while it relegates other objects, and in particular technical objects, to the world without structure of those things that do not have a meaning, only a use.' (Simondon 1958: 10, authors' translation).

independently of the question which tools are used. The problem of *Bildung* cannot be reduced to the acquisition of a set of skills. What Simondon (1958) calls 'culture technique' (technical acculturation) should not be limited to technical training, but needs to start with the recognition that technology constitutes a fundamental way of relating to the world and human diversity goes hand in hand with technological pluralism (cf. p. 218).

We have to be able to think *with* and *in* technology as a medium of expressing a will and a means to know. This is not only necessary to decide when to apply what techniques and to interpret the results they produce; it is also necessary to decide where the computational is superfluous, deceptive or simply sucking up to some funding agency's idea of 'innovative' research. Digital methods are here to stay and to go beyond the simplistic reflexes of enthusiasm and rejection we need to engage in critical practice that is aware of the shocking amounts of knowledge we have stuffed into our tools.

References

Agre, Philip. 1997a. "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI." *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide,* ed. G.C. Bowker et al., 131-157. Mahwah, NJ: Lawrence Erlbaum Associates.

—. 1997b. *Computation and Human Experience.* Cambridge; New York: Cambridge University Press.

Akrich, Madeleine. 1998. "Les utilisateurs, acteurs de l'innovation." *Education Permanente* 134: 79-90.

Amoore, Louise & Volha Piotukh. 2015. *Algorithmic Life. Calculative Devices in the Age of Big Data.* London: Routledge.

Anscombe, Francis John. 1973. "Graphs in Statistical Analysis." *The American Statistician* 27 (1): 17-23.

Arthur, Paul L. and Katherine Bode, eds. 2014. *Advancing Digital Humanities. Research, Methods, Theories.* Houndsmills, New York: Palgrave Macmillan.

Barocas, Solon & Andrew D. Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899.

Bastian, Mathieu, Sebastien Heymann & Mathieu Jacomy. 2009. "Gephi: an open source software for exploring and manipulating networks." *ICWSM* 8: 361-362.

Beer, David. 2013. *Popular Culture and New Media. The Politics of Circulation.* New York: Palgrave Macmillan.

Berry, David M. 2011a. "The computational turn: Thinking about the digital humanities." *Culture Machine* 12 (2).

—. 2011b. *Philosophy of Software. Code and Mediation in the Digital Age.* New York: Palgrave Macmillan.

—. 2012. "Iteracy: Reading, Writing and Running Code." http://stunlaw.blogspot.nl/2011/09/iteracy-reading-writing-and-running.html.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioural Sciences.* 2nd edition. Hillsdale: Lawrence Erlbaum Associates.

Cohen, Patricia. 2010. "Digital keys for unlocking the humanities' riches." *New York Times,* November 17, http://www.nytimes.com/2010/11/17/arts/17digital.html.

Drucker, Johanna. 2014. *Graphesis. Visual Forms of Knowledge Production.* Cambridge, MA: Harvard University Press.

Freeman, Linton. 2004. *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver: Empirical Press.

Gigerenzer, Gerd, Stefan Krauss & Oliver Vitouch. 2004. "The Null Ritual. What You Always Wanted to Know About Significance Testing but Were Afraid to Ask." In *The Sage Handbook of Quantitative Methodology for the Social Sciences*, ed. David Kaplan, 391–408. Thousand Oaks, CA: Sage.

Hacking, Ian. 1992. "Statistical language, statistical truth and statistical reason: The self-authentication of a style of scientific reasoning." In *The social dimensions of science*, ed. E. McMullin, 130-157. Notre Dame: University of Notre Dame Press.

Halpern, Orit. 2014. *Beautiful Data. A History of Vision and Reason since 1945.* Durham; London: Duke University Press.

Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann & Mathieu Bastian. 2014. "ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software." *PLoS one* 9 (6): e98679.

Kassung, Christian. 2001. *EntropieGeschichten. Robert Musils "Der Mann ohne Eigenschaften" im Diskurs der modernen Physik*. Munich/Paderborn: Fink.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Manovich, Lev. 2012. "How to Compare One Million Images." In *Understanding Digital Humanities*, ed. David M. Berry, 248-278. New York: Palgrave Macmillan.

Ramsay, Stephen. 2011. "Who's In and Who's Out." Stephen Ramsay (blog), 8 January. http://stephenramsay.us/text/2011/01/08/whos-in-and-whos-out/.

Ratto, Matt, Sara Ann Wylie & Kirk Jalbert. 2014. "Introduction to the Special Forum on Critical Making as Research Program." *The Information Society* 30 (2): 85-95.

Rieder, Bernhard. 2012. "What is in PageRank? A Historical and Conceptual Investigation of a Recursive Status Index." *Computational Culture* 2, http://computationalculture.net/article/what_is_in_pagerank.

Rieder, Bernhard & Theo Röhle. 2012. "Digital Methods: Five Challenges." In *Understanding Digital Humanities,* ed. David M. Berry, 67-85. New York: Palgrave Macmillan.

Rieder, Bernhard, Rasha Abdulla, Thomas Poell, Robbert Woltering & Liesbeth Zack. 2015. "Data critique and analytical opportunities for very large Facebook Pages: Lessons learned from exploring 'We are all Khaled Said'." *Big Data & Society* 2 (2): 2053951715614980.

Rodgers, Joseph Lee & W. Alan Nicewander. 1988. "Thirteen Ways to Look at the Correlation Coefficient." *The American Statistician* 42 (1): 59–66.

Schön, Donald A. 1983. *The Reflective Practitioner. How Professionals Think in Action.* New York: Basic Books.

Simondon, Gilbert. 1958. *Du mode d'existence des objets techniques*. Paris: Aubier.

Snow, C.P. 1959. *The Two Cultures*. Cambridge, MA: Cambridge University Press.

Svensson, Patrik & David T. Goldberg, eds. 2015. *Between Humanities and the Digital*. Cambridge, Mass.: MIT Press.

Tukey, John W. 1962. "The Future of Data Analysis." *The Annals of Mathematical Statistics* 33: 11–67.

Uprichard, Emma, Roger Burrows & David Byrne. 2008. "SPSS as an 'inscription device': from causality to description?" *The Sociological Review* 56 (4): 606-622.

Wang, Hao. 1963. "Toward Mechanical Mathematics." In *The modelling of Mind*, ed. K.M. Sayre & F.J. Crosson. South Bend, IN: Notre Dame University Press.

Winkler, Hartmut. 2015. *Prozessieren. Die dritte, vernachlässigte Medienfunktion*. Munich/Paderborn: Fink.