Thomas Haigh; Sebastian Gießmann

Defining Digitalities III: What's Digital About Digital Media?

2023

https://doi.org/10.25969/mediarep/20048

Veröffentlichungsversion / published version

Buch / book

MEDIEN DER KOOPERATION

Universität Siegen

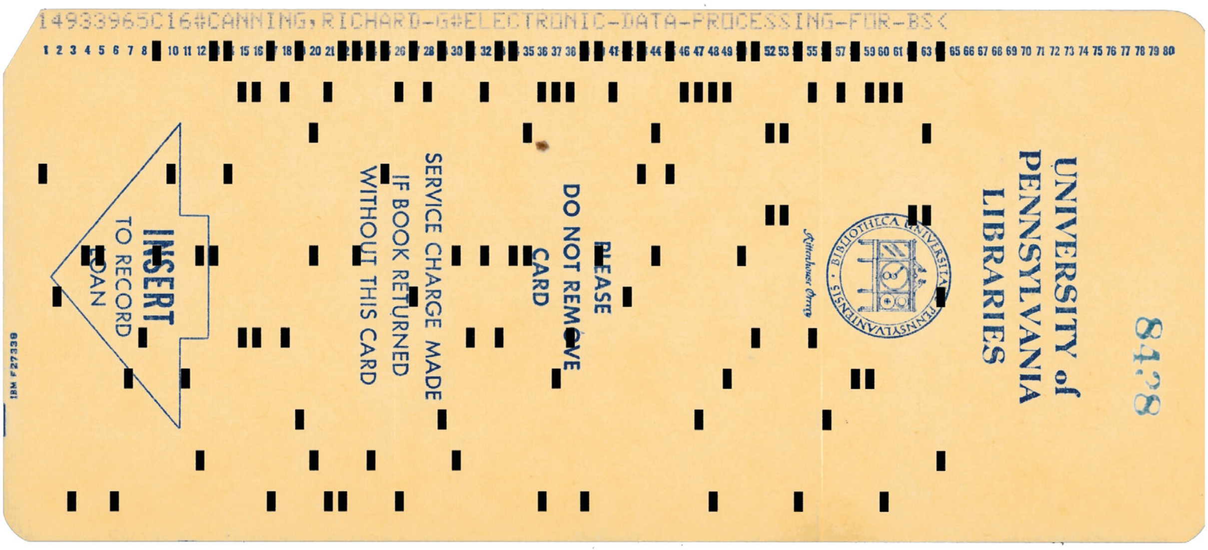
DFG

Defining Digitalities III:

What's Digital About Digital Media?

Thomas Haigh*+, Sebastian Gießmann*

**Siegen University, +University of Wisconsin, Milwaukee*



WORKING PAPER SERIES | NO. 32 | JULY 2023

Collaborative Research Center 1187 Media of Cooperation

Working Paper Series
Collaborative Research Center 1187 Media of Cooperation

Print-ISSN 2567–2509
Online-ISSN 2567–2517
DOI doi.org/10.25819/ubsi/10261
dspace.ub.uni-siegen.de/handle/ubsi/2454
URN urn:nbn:de:hbz:467-24543

This Working Paper Series is edited by the Collaborative Research Center Media of Cooperation and serves as a platform to circulate work in progress or preprints in order to encourage the exchange of ideas. Please contact the authors if you have any questions or comments.

The Working Papers are accessible online at:
https://www.mediacoop.uni-siegen.de/de/publikationen/working-papers-media-of-cooperation/

Print copies can be ordered by sending an email to:
workingpaperseries@sfb1187.uni-siegen.de

Publication is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 262513311 – SFB 1187.

Cover image: U. Penn punched card
Layout: Mattis Hunting

Universität Siegen
SFB 1187 Medien der Kooperation
Herrengarten 3
57072 Siegen, Germany
https://www.mediacoop.uni-siegen.de.sfb1187.uni-siegen.de
workingpaperseries@sfb1187.uni-siegen.de

Defining Digitalities III:
What's Digital About Digital Media?

Thomas Haigh*+, Sebastian Gießmann*
**Siegen University, +University of Wisconsin, Milwaukee*

Abstract: In this working paper we explore an alternative thread in the early development of media and medium as concepts: the origins of the idea of the storage medium in digital computing practices and communities of the 1940s and 1950s. While such practices were obscure at the time, they laid the technological foundation for today's range of digital media. We discuss digitality as a feature of the practices used to read and write symbols from a medium, not a physical property of the medium itself. We then move on to a discussion of the alphabet as itself digital, grounded in the work of Nelson Goodman. Engaging with the contributions of Matthew Kirschenbaum, we explore the limited interchangeability of representations between different encodings of the same symbols, connecting the purported immateriality of digitality to this actual fungibility of material representations.

Keywords: Claude Shannon; information theory; digitality; Matthew Kirschenbaum; Nelson Goodman

A modern reader is likely to approach the terms *communication* and *media* through the distinct scholarly traditions that have grown up around them. Scholars attempting to establish a science of communication trace their lineage back to the engineering work of Claude Shannon. In a series of seminal articles, Erhard Schüttpelz has shown how the notion of communication emerged out of the American military work to become central to information theory, cybernetics, social and anthropological research.[1] One major approach centered on cybernetics and information theory, the other on the study of mass communication and manipulation but both shared the *message* as an organizing concept.

In contrast, media theory takes as its point of departure Marshall McLuhan's 1958 dictum of the medium itself being the message. Yet, argues Schüttpelz, McLuhan and Edmund Carpenter's earlier *Explorations in Communication* had relied upon the paradigm of communication. *Media*, as an analytical category for media studies emerged out of the Toronto School's shared interest in orality and literacy, and its tendency to understand media as grammars of culture.

So what did *medium* mean to the historical actors of the 1940s and the early 1950s? The term comes, of course, from *communication medium* and hence prior to McLuhan's act of appropriation was part of the communications agenda. Yet the term was not as central as *information* or *communication* to Shannon's work, and Shannon preferred the term *channel* to *medium*. From the viewpoint of communications research the crucial technologies of the 1940s were telegraphy, telephony, and radio. All were synchronous and based on the transmission of messages. None involved the storage of information.

In this chapter we explore an alternative thread in the early development of *media* and *medium* as concepts: the origins of the idea of the *storage medium* in digital computing practices and communities of the 1940s and 1950s. While such practices were obscure at the time, they laid the technological foundation for today's range of digital media. We should say at the out-

1 Not all of them have been translated, alas. We already made reference to Erhard Schüttpelz, "'Get the message through': From the Channel of Communication to the Message of the Medium, 1945–1960", in Media, Culture, and Mediality: New Insights into the Current State of Research, ed. Ludwig Jäger, Erika Linz, and Irmela Schneider (Bielefeld: transcript, 2010):109-138.

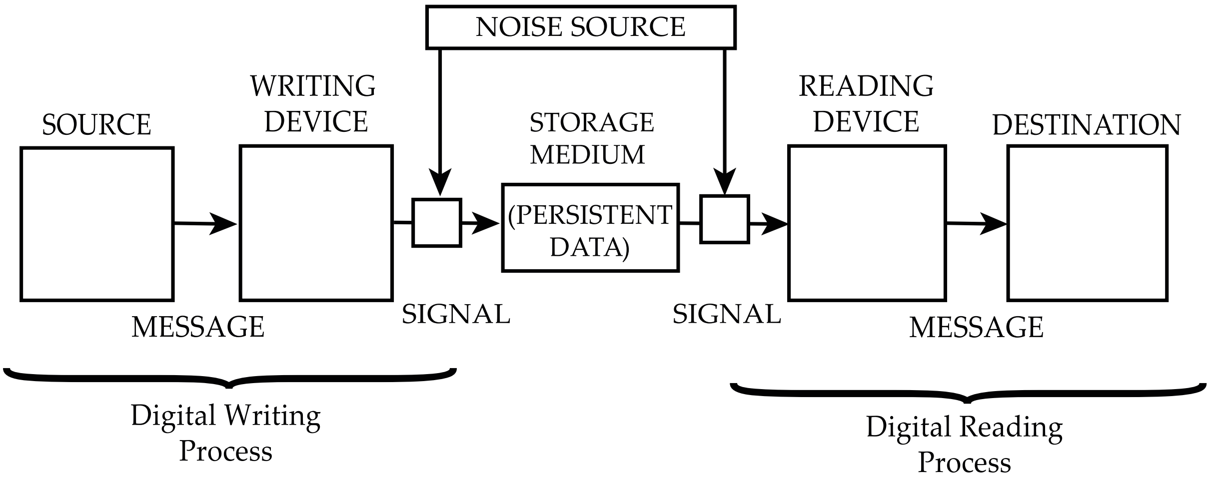


Figure 1: A reinterpretation of Shannon's model, in which the central box represents a digital storage medium such as a punched card deck or paper tape. Reading and writing processes occur in the same way as the standard model, but at different times.

set that digital media do not necessarily use numerical codes. They are digital in the broad non-numerical sense of Shannon's mathematical theory of communication, rather than the narrower and slightly earlier sense of numerical digitality.

Digital Storage vs. Digital Communication

The distinction between digital (storage) media and digital communication, if such a distinction can be meaningfully made, is one of temporality. The system described by Shannon, in which the signals read digitally by the receiver had been deliberately encoded and placed into a channel with the intention that they be immediately received and recoded, describes a special case of digital writing and reading. This model, based more than anything else on the practices of telegraphy, is most directly applicable to synchronous communication. A human operator taps out the message as a series of letters expressed in dots and dashes. The receiver transcribes the letters as they are received. In telegraphy information placed into the channel is not persistent. The receiver must match the pace of the sender and either note down or mentally buffer symbols as they are transmitted. The temporalities of sending and receiving are synchronized. This does not mean that communication is instantaneous. Even electrical impulses are not truly instantaneous, and transcription may lag slightly behind transmission.

The temporalities of digital storage media were more complex. Early computer systems applied digital reading and writing practices to media such as punched cards, magnetic tape, magnetic disks, and paper tape. All these media were stable: digital sequences written to them would persist indefinitely and could be read repeatedly as required. Other media, such as mercury delay lines and cathode ray storage, retained information for very short periods which necessitated a constant process of reading and rewriting the symbols stored in them to preserve information from one minute to the next.

Shannon had produced a theory of communication, not a theory of storage media. But with one minor adjustment his model fit perfectly with digital storage technologies. Imagine that instead of transmitting each symbol immediately to the receiver, the channel instead buffers the sequence of symbols for later retrieval. There is still a sender, a receiver, a process of encoding, and a process of decoding. But the message sits in the channel until the receiver is ready. Because the processes of reading and writing occur asynchronously, the channel of digital communication has become a medium for digital storage. Only then does the apparently confusing pairing of the static *storage*, a container in which something rests, make sense in conjunction with the active *medium*, the infrastructure through which messages move.

An example will make this less abstract. From the 1940s to the 1970s many computer systems used punched cards as a medium for the temporary storage of input data and program code. The source of the data was a person reading information from forms. The writing device was a key punch. The storage medium was a deck of punched cards in which the sequence of coded symbols keyed by the key punch operator accumulated. Eventually the deck of cards is mounted in a reading device and read digitally to reproduce the same series of coded symbols in a different medium. Early punched card readers used mechanical brushes to sense the presence or absence of holes in each position, later models were able to read more than a thousand cards a minute using photoelectric sensing. In either case, the reading process transcribed the bits from cards into electric impulses flowing down wires. The process would repeat inside the computer, where these impulses were read and transcribed into electronic storage.

The practice of storing encoded messages also has its roots in telegraphy, where paper tapes were used to record messages, and in some cases to transmit at high-speed messages punched onto tape in advance. For example, the German network of encrypted radio teleprinters targeted by the British Colossus devices could be used to transmit messages stored on paper tape. The same five track paper tape was a popular recording medium for the computers of the early-1950s, a cheap and readily available method of getting programs and data into computers and results out. We may distinguish further here between inherently persistent media like computer tape and those that must constantly be automatically read and rewritten to achieve temporarily stability such as a mercury delay line, cathode ray tube storage, or later a dynamic RAM chip. There are likewise differences in the affordances of purely sequential media such as tape and random-access media such as disks and core memory, and between media such as punched cards and paper tape that can be written only once, and those such as magnetic tape that can be overwritten.

In this conception (Figure 1) the storage medium replaces the unlabeled central box that in Shannon's standard model represents the communication channel. The medium is, essentially, a buffer between the transmitter and receiver. But it is also possible to view the same process as two distinct processes of communication using Shannon's standard model (Figure 3). The storage medium is the destination of the first act of communication, representing digital writing. It is the information source of the second act of digital communication, representing the process of digital reading.

Such processes occur constantly within digital computers, involving the transmission of coded messages between non-human actors. A processor communicates a string of symbols to the mechanism of a card punch; a magnetic tape drive reads codes and sends them along a wire as encoded impulses to the processor which reassembles them in a register; characters are transcribed from main memory into a register, which immediately switches from receiver to sender and copies them down a different wire leading to a printer. In each of these cases a message is passed from a sender to a receiver, it's just that both are mechanisms.

Again, an example makes our point clearer. Consider a process of data entry onto a computer as it might have been practiced in the late-1960s using a *key to tape* device. This replaced the key punch unit, logging keystrokes from an operator directly onto a reel of magnetic tape. The same reel was later mounted in a tape drive connected to a computer, from which chunks of data were read sequentially into the computer's processor.

Figure 2: A key to tape unit, introduced by Mohawk Data Systems in 1965.[2]

In the diagram below (Figure 3), the process is represented by two interlinked copies of Shannon's classic diagram. The first act of digital communication occurs inside the key to tape machine, as key presses are coded into electric impulses received by the tape mechanism and written onto the tape. The second act of digital communication occurs when the same tape, which is now the source rather than the destination, is read in a tape drive. Signals travel over a cable, are received by the input/output hardware of the computer and finish up transcribed into electronic memory of some kind it with.

The two examples given above are functionally equivalent: data moves from paper into a computer via an intermediate medium. The first diagram conceptualizes this as a single asynchronous act of message transmission in which a storage medium replaces Shannon's channel. In Shannon's model a communications channel has a storage capacity of just one symbol which must therefore be received as it is being transmitted. Digital reading and writing are therefore synchronous. Increasing that storage capacity by replacing the channel with a punched card deck or magnetic tape changes the temporality of communication. Reading and writing occur asynchronously.

2 https://georgecogar.com/2016/05/15/mds-1101-brochure/#jp-carousel-233

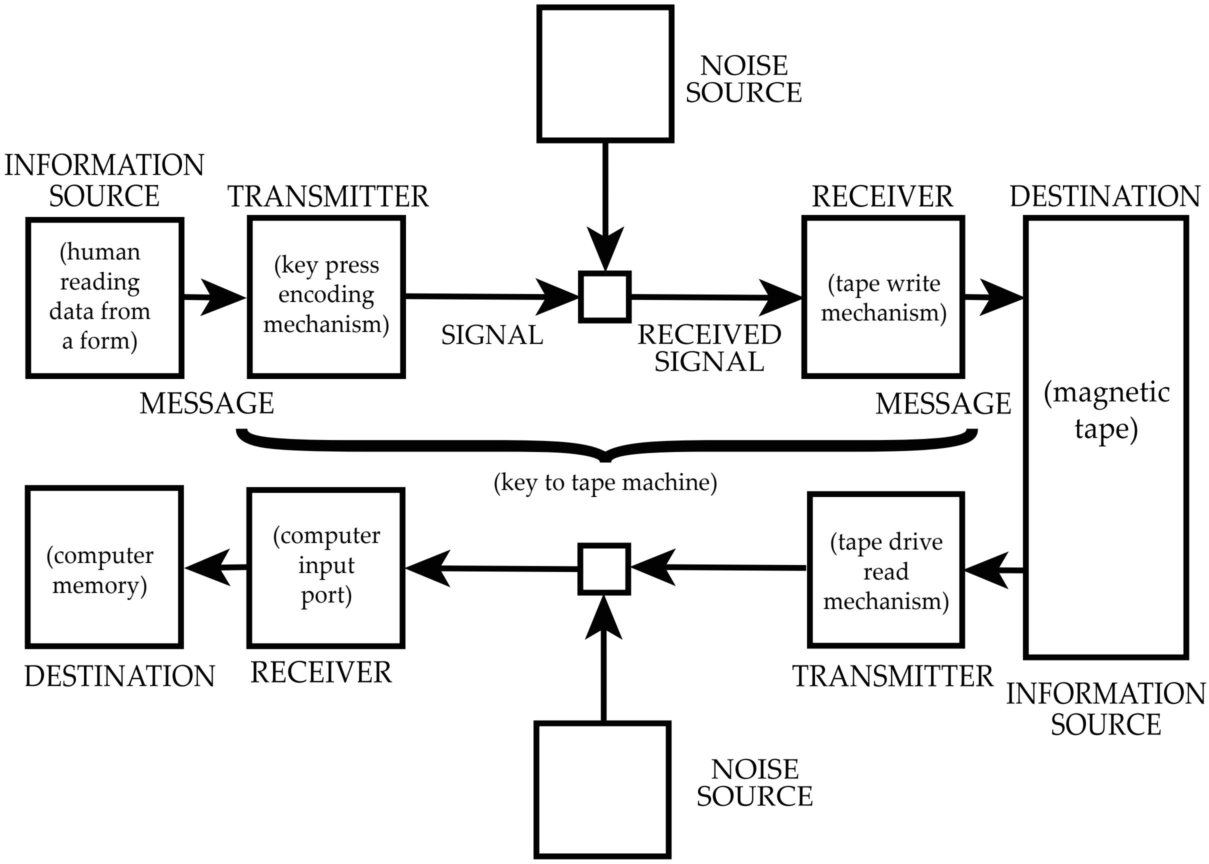


Figure 3: Each act of digital reading or digital writing can be conceived as a distinct communication process using Shannon's original model with no changes.

One might conclude that a communication channel is just a storage medium with a very small capacity. Yet, as the second diagram shows, it is just as informative to conceptualize the same example as describing two synchronous processes of communication each involving digital reading and writing. The acts of digital reading and writing in the first diagram each decompose in the second to become entire processes of communication including both reading and writing. The second diagram could be further decomposed into distinct acts of communication by following the progress of the bits though the computer's input circuits and via the main processor or an auxiliary processor, though one or more registers, and eventually into main memory. We conclude that the distinction between digital communication and digital media is a matter of perspective and temporality.

Digital Storage in Early Computing

This mapping of the technical processes of early electronic computing onto Shannon's model of communication might seem like an empty exercise. But in fact it illustrates the applications for which the ideas of Shannon, Hamming, and their fellow information theorists were most immediately and profoundly relevant. Having established the equivalence of storage and communication in this context, we will look more closely at the context of data storage in early digital computers. The language of media, mediation, and symbol processing was central to technical conceptions of modern computing from the very beginning for the pragmatic reason that these concepts were vital in the design of functioning computer systems.

Shannon was motivated by the potential of digital reading and writing for long distance communication. This drew directly on the wartime involvement of Bell Labs in encryption of digital voice communications. But Bell Labs was also a wartime center for computing research, and their concerns and techniques of telecommunications engineers overlapped greatly with those of computer engineers. A digital computer is a network of components in constant communication with each other. As they go about their work digital messages are constantly being encoded, decoded, and transcribed from one medium to another.

Determining the bandwidth of these channels was vital when maximizing the performance of computer designs. In later decades attention would shift to the connection of computers to remote terminals and the development of networks for the exchange of messages between computers. But from the very start, communications processes were vital to electronic digital computing. From this viewpoint, the convergence of computer engineering and digital communications took place in the 1940s, not the 1970s or 1980s.[3] Rather than being a radical discontinuity, one might see the development of computer networking as an extension of these connections over longer distances.[4] Conceptually it makes little difference whether information is being communicated within a processor chip or with a space probe on the fringes of the solar system. The same digital reading and writing processes that made computer networking possible had been taking place all along within digital computers and between digital computers and peripherals.

Most of the electronic digital computers built in the late-1940s and early-1950s employed several media to store and process both numbers and instruction codes. The majority of early US computers, including commercial models, were based on the design produced by John von Neumann's team at the Institute for Advanced Study at Princeton. That was a refinement of von Neumann's original description of the modern computer, the 1945 *First Draft of a Report on the EDVAC*, in which these media constituted two distinct "organs" of the machine: organ M for memory, the high-speed electronic storage used for the program instructions and data currently being processed which "requires perfectly distinct and independent registration and storage of digital or logical symbols" and organ R for recording, the slow but persistent medium used to load programs and data into memory and store the results of the computer's work. Von Neumann called organ R "the natural medium for long time storage of all the information obtained by the automatic device on various problems."[5] This organ was, in his terminology, "outside" the computer proper, so that operations of "input and output" would be "mediating the contact with outside." This distinction would appear to be the origin point for the now ubiquitous concepts of *computer storage* and of *input/output*.

Both storage and memory organs were essential to any useful digital computer. Every computer was coupled with at least one device able to read and write digitally to a permanent storage medium. Early systems used paper tape, punched cards, or magnetic tape for this purpose. Programs to be run were transcribed from this medium into a high-speed addressable memory: a delay line, a Williams tube, a Selectron tube, a rotating magnetic drum, or core memory. The instructions currently being executed and the numbers currently being manipulated were copied again, into register storage (a small amount of high speed memory inside the central processor, built using vacuum tubes).

We will not describe at length the many different objects that were manipulated to store bits and read to retrieve them; suffice it to say that the mechanisms used were varied and ingenious. Running even the simplest program required many operations to transcribe bits between these media, during which they were transiently embodied in yet other forms: electric pulses conducted by wires, electromagnetic waves picked up by read heads, or the motion of rods punching through card. Bits moved within and between these computers and their peripherals as digital messages of exactly the kind discussed by Shannon, each wire a communications channel. Determining the bandwidth of these channels was vital when maximizing the performance of computer designs.

The major challenges involved in building a usable digital computer centered on perfecting these mechanized practices of digital reading and writing. Both candidate memory technologies, cathode ray tubes and mercury delay lines, required lengthy periods of experimentation to become functional. Magnetic tape, the highest performance medium for long term storage, was also challenging. Between them these accounted for most of the challenges that delayed early computer projects, such as von Neumann's own computer at the Institute for Advanced Studies or the commercial Univac effort, years beyond their original schedules.

As Mara Mills has noted, "Although a growing scholarly corpus has now demonstrated the materiality of electronic/digital/computerized media, most authors continue to attribute a fantasy of disembodied communication to early cyberneticians and electrical engineers."[6] The historical record amply disproves this fantasy. Bell Labs studied information for the benefit of engineers, not cultural theorists. During the late-1940s and early 1950s the community of engineers involved in building computers and storage devices were the earliest and most enthusiastic adopters of information

3 Because mass market media and communication technologies remained analog for decades after Shannon's work, some media scholars have been slow to recognize the central importance of the work of Shannon, Hamming, and their colleagues to computer engineering. Bernard Geoghegan, for example, suggested that "When Shannon's theory of communication appeared, it was celebrated but also regarded as a theoretical study of little practical applicability." Geoghegan suggests that error detecting codes were of merely theoretical interest in the 1950s and were not widely implemented until the 1980s. Bernard Geoghegan, "Information", in Digital Keywords: A Vocabulary of Information Society & Culture, ed. Benjamin Peters (Princeton, NJ: Princeton University Press, 2016):173-183, quotation p. 179.

4 Sebastian Gießmann, The Connectivity of Things: Network Cultures Since 1832 (Cambridge, MA: MIT Press, forthcoming), chap. 9.

5 John von Neumann, "First Draft of a Report on the EDVAC," IEEE Annals of the History of Computing 15, no. 4 (October 1993):27-75. The comment on "mediation" is a section heading for 2.6. The quote on a "natural medium" is from section 2.9. Note that von Neumann's mention, in section 12.8, of the memory storing "digital or logical symbols" aligns with the initial, literal sense of digitality as applying only to representations of digits.

6 Mara Mills, "On Disability and Cybernetics: Hellen Keller, Norbert Wiener, and the Hearing Glove," differnces 22, no. 2&3 (2011):74-111, quote p. 78.

theory, because they were the community charged with figuring out rapid and reliable ways of making digital communication work. These acts of communication were not primarily between humans but between mechanisms, occurring constantly between a computer and its peripherals and within the different parts of the processor itself. Their creators adopted the vocabulary and practices pioneered at Bell Labs: bits, bandwidth, parity, coding schemes, error correction and detection, redundancy, and information. They relied on the techniques described by Shannon and his colleagues to make these enormously complex machines function reliably, for example by introducing redundancy when storing messages on tape so that errors could be detected (We will discuss below the deep connections of our perspective here to the work of Matthew Kirschenbaum)..

Information Becomes a Thing

The work of Shannon and his colleagues on information theory was central to the early development of what we now call computer science. In most European languages this discipline is named with some variant of the word information: Informatik, l'informatique, and so on. The global organization for computing researchers is the International Federation for Information Processing. Even in the US, information-centric names were proposed for the field that became computer science. Some early programs, such as the Department for Computer and Information Science at the University of Pennsylvania, or the Department of Communication Sciences at the University of Michigan took on names that reflected this heritage. Richard Hamming chaired the Association for Computing Machinery and was among the first winners of its flagship honor, the Turing Award.

The embrace of information theory by computer engineers was accompanied with a new meaning of the word information, which had traditionally been inseparable from the act of informing: information was only information if somebody was learning something from it. Information was a process, not a thing. Applying the term *information theory* to Shannon's work this fit with that usage: it explicitly described the transmission of a message from a sender to a receiver, a process during which the receiver was informed. Inside computers, information was constantly being sent and received without human involvement. The various components of a computer system were constantly informing each other as they sent bits back and forth. It took only a subtle linguistic and conceptual slippage to think of the data stored in computer files as *information* even when it was not being transmitted and received.

"Information" became, among other things, a synonym for facts or data – and in particular for digitally encoded, machine readable data.[7] Information became what linguist Geoffrey Nunberg memorably called an "inert substance" that could be stored, retrieved, or processed. We've come to think of anything processed by computers as information, creating a new sense of the word that information scientist Michael Buckland dubbed "information as a thing."[8]

Just as the concepts of information and bits moved from communications engineering to computing, the vocabulary of digital and analog moved the other way, quickly displacing Shannon's own preferred terms of discrete and continuous. Telephony became a hybrid of analog and digital systems, with analog transmission of voice data down landlines to handsets but increasing reliance on digital exchanges and long-distance transmission of information. Records, conventional television broadcasts, and audio tapes were all analog media, while compact disks, high-definition television, and DAT cassettes were digital. Thus the senses of analog and digital introduced to describe different approaches to the representation of quantities inside computers turned out to have much broader application. While they were still applied most often to different kinds of electronic systems, they apply conceptually to any methods used to encode information.

Alphabets are Digital

Shannon's reliance on textual examples raises the question: is text expressed in an alphabet always digital, or is it digital only when stored in computer readable form? Is a conventional printed book a digital storage medium? The answer depends on one's definition of digitality. According to the numerical sense of digitality, texts become digital only when transposed into numbers. In the Shannon-influenced, symbol-oriented sense of digitality, text has always been digital. Arabic numerals provide 10 symbols while the English alphabet provides 26 letters. Combining those 36 characters with upper- and lower-case variants, punctuation, and other symbols one might use six, seven, or eight bits to code a full set of symbols. [9] Unicode, intended to cover the needs of writing systems such as Braille, Old Persian, and kanji as well as Western European alphabets, uses up to 38 bits per character in its most common

7 Geoffrey Nunberg, "Farewell to the Information Age", in The Future of the Book (Berkeley: University of California Press, 1997): 103-138.

8 Michael Buckland, "Information As Thing," Journal of the American Society of Information Science 42, no. 5 (June 1991): 351-360.

9 Encoding and mechanically processing the character sets used by other languages was vital to the global spread of information technology. It poses a challenge not just for internal representation, but also for data entry via keyboards. Chinese text, in particular, was often mapped onto representations coded in Latin alphabets. Thomas S. Mullaney, "QWERTY in China: Chinese Computing and the Radical Alphabet," Technology and Culture 59, no. 4 (October 2018): S34-S65.

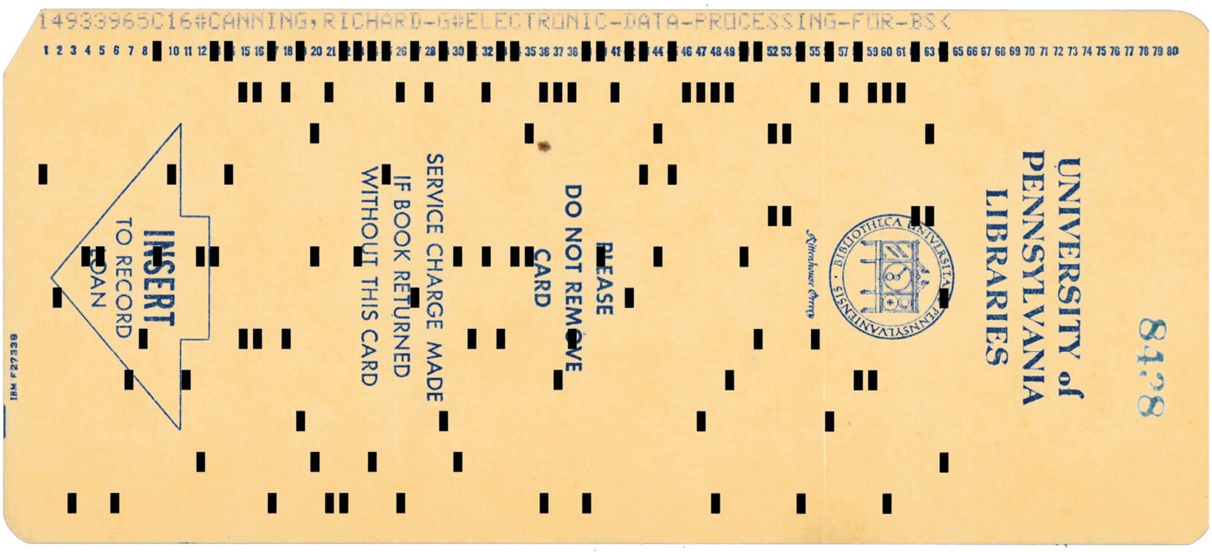


Figure 4: This IBM card was stored inside the cover of a library book to speed the checkout and return process. It encoded text as well as numbers on a card format originally designed to store 80 decimal digits, by using two extra rows at the top of the card and giving a unique meaning to every combination of a punch in one of the top three rows with a second punch in one of the nine lower rows. Printed text at the top duplicated this information in alphanumerical characters, a digital representation read more easily by humans.

format. But even that large set of symbols is most definitely finite.

This was recognized in practice long before Shannon formalized his theory of communication. We already discussed the original use of punched cards to represent digits. By the 1930s, however, IBM realized that it could extend its standard 80 column card format to represent text as well as numbers. There was, as in the card we saw earlier, room for 12 rows of punch positions. By using two extra rows and building equipment to recognize all possible combinations of a punch in one of the top three rows combined with a punch in one of the lower nine rows, IBM increased the number of character codes from 10 to 39. That was enough to represent all the upper case letters and digits, with three codes left over for &,-, and /. Later text card formats, such as the one shown in Figure 4, represented additional symbols using other combinations of punches. [10]

Text is most certainly not analog: an alphabet provides a fixed set of characters rather than a continuous range in which one letter blends into another. The philosopher Nelson Goodman recognized this in his 1968 book *Languages of Art: An Approach to A Theory of Symbols.*[11] Goodman defined languages as "symbol systems." Their characters could, as in typography, be written in many different ways as long as readers could reliably distinguish all the equivalent marks as representations of the same character. Characters must be disjoint, represented in ways clearly distinct from each other with "a wide neutral zone" between them. Goodman called this principle "finite differentiation." According to Goodman, "the syntactic requirements of disjointness and finite differentiation are met by our familiar alphabetical, numerical, binary, telegraphic, and basic musical notations."[12]

Goodman went on to connect his ideas to the concepts of analog and digital, reflecting the evolution of these categories away from their original roots when he wrote "Plainly, a digital system has nothing to do with digits, or an analog system with analogy." His analysis echoed Shannon's conception of coded symbols. It was, essentially, that analog systems provide an undifferentiated range of markings, which are "the very antithesis of a notational system.... A digital scheme, in contrast, is discontinuous throughout."[13]

We earlier suggested that digits written on a piece of paper must be considered digital, according to any reasonable definition of the term. But by the broader definitions of digital, which encompass all sequences of symbols rather than just sequences of digits, we must also recognize text as digital when printed on paper, or even when written by hand. To type a manuscript into a word processing program is not an act of digitization or analog to digital conversion; it is a transcription from one digital notation to another.

10 Once computers arrived, IBM added a "binary" card format that used all possible combinations of punches, effectively coding 4096 possible symbols in each row. This was far more efficient, but the volume of punching tended to cause both cards and card processing equipment to fall to pieces.

11 Nelson Goodman, Languages of Art: An Approach to a Theory of Symbols (Indiannapolis, IN: Bobbs-Merrill, 1968). Haigh is indebted to Matthew Kirschenbaum for suggesting Goodman's work to him in this context.

12 Ibid., 132-140. Quotation is from p. 140.

13 Ibid.,160-161.

Figure 5. To create texts using pieces of movable metal type, a compositor would select letters, punctuation, and spacing elements individually from bins in a type case and sequence them in a type stick. This is a digital writing practice, in as much as symbols are being selected and sequenced from a pre-defined set. Image by Wikimedia user Willi Heidelbach, used under CC BY 2.5 (https://creativecommons.org/licenses/by/2.5/).

Being digital does not necessarily mean that a medium is machine-readable, nor that it cannot be read by humans. What is, or is not, machine readable will in any event change with time. In the 1950s, for example, a project was launched to print account and routing numbers onto checks in a format readable both by humans and by computerized check clearing systems. That was realizable at the time only by writing the numbers in magnetic ink, using a highly stylized font that looked odd but recognizable to human eyes. By the 1990s, neural networks developed at Bell Labs could reliably read handwritten numbers to automatically process the amounts written in pen on checks as well as the codes printed in magnetic ink.

Given the prominence of Gutenberg, keyboards, and movable type in media history it seems reasonable to ask whether the reading and writing practices that evolved around printing in the early modern period were more digital, or differently digital, than those involved in reading and writing by hand.[14] With respect to reading we see no fundamental difference: typeset and handwritten text can both be hard to decipher and both can be read either in a simple way as a source of alphanumeric symbols or with richer and deeper practices. With respect to writing, though, we see a significant difference. Writing by hand, one can draw an essentially infinite number of possible shapes though it is generally the intention of the writer to make shapes than can each be clearly recognized as matching one of a small set of valid characters or punctuation marks. In contrast, when constructing a line of movable type a print worker is quite literally selecting (from designated buckets) and sequencing reusable symbols chosen from a finite set. The typesetting process was later automated with keyboard-driven machines, and on a more personal scale with the typewriter. Used as intended the manual typewriter delivers evenly spaced characters horizontally and vertically, though this can be subverted by using backspacing and manual adjustment of the paper position. We therefore view the typewriter as a mechanism designed to support digital writing practices rather than a mechanism that is in inherently digital.

It’s a matter of historical record that nobody called text *digital* before the invention of the modern computer. Thinking of characters as well as numbers as

14 For example, Friedrich A Kittler, Gramophone, Film, Typewriter (Stanford, CA: Stanford University Press, 1999).

digital reflects both the influence of Shannon and the reality of digital computers that used numbers to encode and process text. Seeing text as digital is thus a backward projection of categories. Yet once the categories are established there seems no plausible reason to deny that the alphabet was, in the modern sense, always digital.

The Interchangeability of Digital Representations

The idea that *the digital* is defined by its immateriality is both ridiculous and common, a combination that calls out for substantial historical and philosophical analysis. Yet there is something special about the relationship of bits to their material representations: different material representations are, from a certain viewpoint, interchangeable. The same sequence of symbols can be read from any of them. As Matthew Kirschenbaum has argued:

> two properties of digital computation—its allographic identity conditions and the implementation of mathematically reliable error detection and correction—are what ultimately account for the immaterial nature of digital expression. My point is not that this immateriality is chimerical or nonexistent, but rather that it exists as the end product of long traditions and trajectories of engineering that were deliberately undertaken to achieve and implement it.[15]

Analog representations of information could also be converted from one form to another. Mara Mills has explored analog technologies developed to transform audible speech into visual representations (sound spectrographs and audiograms) and tactile sensations (the so-called *hearing glove*).[16] But these conversions were approximate and ambiguous. In contrast, digital information can be copied from one medium to another without any loss of data, and the same sequence of symbols can be recovered from each medium. Transcribe the text of the book into a text file, save that file, compress it, email it, download it, decompress it, and print it out. The different representations have different affordances and must be read in different ways. Yet each can be read in a way that produces the same sequence of symbols.[17] Discussion of *digital formats*, vague as it often is, gestures towards the truth of this experience: digital content can be downloaded over a network and, if sufficient bandwidth is available, experienced just as if it had been accessed from a local disk. "The real virtues of digital instruments," Goodman insisted, "are those of notational systems: definiteness and repeatability of readings."[18] As the information is constantly and automatically transcribed and transcribed from one digital representation to another it is easy to lose track of its materiality entirely and assume that it really does live in an immaterial cloud of data floating somewhere in the heavens.

What we call the *interchangeability of representations* has been called *cultural transcoding* by Lev Manovich. In his *Language of New Media*, transcoding figures prominently as the fifth, and most important principle of new media. It builds upon the four preceding principles of numerical representation, modularity, automation, and variability. "In new media lingo," Manovich writes, "to 'transcode' something is to translate it to another format." While making representations interchangeable, digital computation affords for the astonishing variability in media practices that Manovich maps. To Manovich transcoding is what is turning the computer into a media machine that turns slightly older media into computer data. [19]

Different material representations of the same text are only interchangeable from certain viewpoints, and for certain purposes. A book is not the same thing as a text file stored in memory, which is not the same thing as a compressed file on a hard disk. The printed text would differ materially depending on whether one used a dot matrix printer or a modern laser printer. Neither would replicate the original book. For one thing, any illustrations or annotations would be lost, as would the possibility of studying the book to learn details of the process by which it had been printed or about the history of typography. Likewise, no two performances of the same musical score are identical, however talented and careful the musicians. If blessed with superhuman abilities one might listen to the performance of a musician and write out the corresponding score. The two scores and the performance can all thus be read digitally as interchangeable representations of the same sequence of coded symbols. Yet the classical music industry spent many decades creating and selling new recordings of old works. To interpret a score the performer adds a great deal that is not written on the page.

Goodman suggested that this interchangeability across physical representations was an inherent fea-

15 Matthew Kirschenbaum, Mechanisms: New Media and the Forensic Imagination (Cambridge, MA: MIT Press, 2007), 137.

16 Mills, "On Disability and Cybernetics: Hellen Keller, Norbert Wiener, and the Hearing Glove", Mara Mills, "Deaf Jam: From Inscription to Reproduction to Information," Social Text, no. 102 (Spring 2010):35-58, Mara Mills, "Deafening: Noise and the Engineering of Communication in the Telephone System," Grey Room, no. 43 (Spring 2011): 118-143.

17 Kirschenbaum offers a longer and more worked through discussion of the possible sequence of representations and translations a text might go through in the process of composition in Matthew G Kirschenbaum, Bitstreams: The Future of Digital Literary Heritage (Philadelphia, PA: University of Pennsylvania Press, 2021), 1-5.

18 Goodman, Languages of Art: An Approach to a Theory of Symbols, 161.

19 Lev Manovich, The Language of New Media (Cambridge, MA: MIT Press, 2001), 45-47.

ture of notational systems. Drawing on his background in visual art, he pointed out that paintings can be faked but poems can't, because the same poem might be written out by hand or printed in different styles on different kinds of paper. "All that matters is what may be called sameness of spelling: exact correspondence as sequences of letters, spaces, and punctuation marks."[20] Musical scores use a different system of notation which, from Goodman's viewpoint had the same fundamental characteristics. He called systems of this kind "allographic," the term invoked by Kirschenbaum above as a key characteristic of digital representations, unlike "autographic" works such as paintings where any copy differs from the original in fundamental ways.

Many readers might indeed view a handwritten manuscript, a typewritten page, a plain text computer file, a magazine publication, and a volume printed using one of a variety of unremarkable typefaces as authentic representations of the same poem, even if they preferred the aesthetics of one or another format. However, Kirschenbaum has recently given two examples that implicitly challenge Goodman's assertion that poetry relies only on the affordances of text and punctuation symbols. Both poets made use of the affordances of early Apple Macintosh personal computers. William H. Dickey used the HyperCard authoring tool to produced fourteen "HyperPoems" that mix text with graphics and react to each reader's mouse clicks. Edward Kamau Braithwaite published his poems on paper, but their meaning is conveyed in part by the aesthetics of custom screen fonts that pixelated heavily when resized on his early Macintosh. Some editions sacrificed much of this typography, but Kirschenbaum insists that even minor changes inherent in the publishing workflow "would invariably compromise the original integrity of the work."[21] This is not, however, to say that the text used in more conventional poems is not digital. Rather, some poets have produced work that requires far more complex (and unstable) platforms of digital representation than plain text to faithfully distribute. Braithwaite used a digital computer to produce work that could, at the time, be redistributed only by the analog process of photographing his original printed manuscripts onto film.

Kirchenbaum's examples highlight the extent to which media from which the same sequence of symbols can be read are interchangeable for some purposes, or for some readers, but not for others. For applications where typography matters one might, as some manuscript preservation programs do, deal with this by making an extremely high-resolution scan of the book rather than transcribing the text into a computer file. That would produce an entirely different digital representation of the same object, and one that might serve the needs of more kinds of readers. But even that would not substitute for the all the uses of the original. In this context, Anthony Grafton invoked the story of a researcher who traced the spread of cholera by sniffing letters, to see which had been perfumed with vinegar before sending (vinegar having, allegedly, been believed to prevent the spread of the disease).[22] More prosaically, a printout would be more suitable than a file on a USB stick suitable for lighting a fire with.

Digitality is in the Ear of the Beholder

The fact that we can write on, print on, draw on, or punch an IBM card provides a nice illustration of a more general truth: digitality lies not in an object but the way the object is read. We might, for example, admire the position of the holes on a punched card as a piece of abstract art, reading them in an analog fashion. Plenty of people enjoy the aesthetics of tattoos written in characters they cannot read. They appreciate them as brush strokes and shapes. We sometimes admire characters we can read in a similarly abstract manner, coveting the shape of a letter in a particularly handsome typeface.

The same is true of other media. A cassette recorder is itself neither digital not analog. I (Haigh) have in my lab a beautifully engineered Sony audio cassette recorder that I have used to record hundreds of hours of oral history interviews. The only digital thing in it is the mechanical tape counter. My microphone turns fluctuations the volume of sound into fluctuations in electric current; the tape write head turns those into magnetic fields which in turn realign particles the tape as it passes underneath it. The amplitude of the sound, the resistance, the current, and the magnetic field are all analogous to each other. When I play a tape back the process reverses: magnetic field to electricity to sound. The stored sounds of the interview fill the air.

Yesterday I plugged the headphone jack of the same tape recorder into an Acorn BBC Microcomputer manufactured in 1982. I placed a standard audio cassette labelled *Rocket Raid* into the player, entered *TAPE followed by CHAIN "" on the computer keyboard, and pushed play. Played on a speaker, the tape yields a series of atonal warbles and howls separated by short intervals of silence. That's the analog interpretation, something a music critic might approach as an experimental work akin to Lou Reed's inscrutable *Metal Machine Music*. Read digitally by the appropriate control circuits and algorithms, rather than turned into vibrations in the air and my inner ear, the same oscillations were instead read as a series of digitally coded symbols that loaded about 12 kilobytes of data into the computer's RAM.

20 Goodman, Languages of Art: An Approach to a Theory of Symbols, 115.
21 Bitstreams, p 64, ch. 2.

22 Anthony Grafton, "Further Reading: Digitization and its Discontents," The New Yorker, November 5 2007. He took the story from John Seely Brown and Paul Duguid, The Social Life of Information (Boston: Harvard Business School Press, 2000).

Having counted its progress block by block in hexadecimal digits, after about five minutes the computer sprang to life and offered me a passable imitation of the arcade classic *Scramble*. Digitality here is not in the tape cassette, nor anywhere in the tape player, but in the reading of the output of the tape player. (Fax machines and modems used similar techniques to read and write digital data as electrical oscillations that could make their way reliably over telephone lines optimized for analog voice signals).

I could have loaded the same game in a few seconds from a floppy disk, another process of magnetic reading, or from the Gotek floppy emulator next to it which would have fetched the same sequence of bytes from a flash memory chip. Each medium can be read to produce the same sequence of symbols and hence put the same bytes into the same memory locations, but they differ greatly in affordances such as speed, cost, random access, reliability and so on. Just like my tape recorder, disk drive controllers convert analog signals produced by magnetic read heads into patterns of encoded bits.

That's true of hard as well as floppy disks. In early personal computers, such as the IBM PC XT, the hard drive output analog signals. [23] The job of turning them into bits and bytes was the responsibility of the controller card, plugged into the motherboard at the end of a long ribbon cable. Upgrading the stock IBM controller card with a replacement that used a more efficient system of encoding (RLL rather than MFM) could upgrade the capacity of the XT's standard 10MB drive to 15MB, though every sector of the disk would have to be demarcated anew on its surface, in a "low level format" that took hours to complete.

Conclusion

To reiterate the main themes of this paper, the concept of a digital storage medium is as old as the digital computer itself and is central to the possibility of its existence. Digital storage is digital in the same sense that digital communication is digital, i.e. because it involves the reading and writing of coded symbols, no the literal reading and writing of digits. Yet digitality is a feature of the practices used to read and write symbols from a medium, not a physical property of the medium itself. Because different digital reading practices can be used to reliably extract the same sequences of coded symbols from entirely different media, we say that those different representations are interchangeable. For example, reading practices exist that can derive the same symbol of coded symbols from a printed book and a Project Gutenberg text file encoded magnetically on a hard disk drive. Of course, any object can be read in many different ways, so for other purposes the two objects will not be interchangeable.

As Kirschenbaum observed, because many kinds of digital reading processes take place on scales invisible to the human eye, and because processes of automatic digital reading and writing are often coupled to constantly and invisibly transcribed symbol sequences between media, these affordances of digitality can create the illusion of immateriality. Redundant digital encodings that allow the recognition and correction of errors improve the reliability of digital reading and the creation of perfect copies. Modern digital media technologies read and write information in forms invisible to human senses: radio waves, light pulses, flash memory cells packed together on a microscopic scale, electrical impulses, or magnetic bands. Information is always material, in that it consists of symbols read digitally from one or another corner of the physical universe. But when it is being transcribed so rapidly and automatically from one representation to another we lose track of that materiality.

23 The technologies of hard disk drive recording were given a thorough examination in Kirschenbaum, Mechanisms: New Media and the Forensic Imagination, 86-96.

References

Brown, John Seely, and Paul Duguid. *The Social Life of Information*. Boston: Harvard Business School Press, 2000.

Goodman, Nelson. *Languages of Art: An Approach to a Theory of Symbols*. Indiannapolis, IN: Bobbs-Merrill, 1968.

Grafton, Anthony. "Further Reading: Digitization and its Discontents."*The New Yorker*, November 5 2007.

Kirschenbaum, Matthew G. *Mechanisms: New Media and the Forensic Imagination*. Cambridge, MA: MIT Press, 2007.

———. *Bitstreams: The Future of Digital Literary Heritage*. Philadelphia, PA: University of Pennsylvania Press, 2021.

Kittler, Friedrich A. *Gramophone, Film, Typewriter*. Stanford, CA: Stanford University Press, 1999.

Manovich, Lev. *The Language of New Media*. Cambridge, MA: MIT Press, 2001.

Mills, Mara. "Deaf Jam: From Inscription to Reproduction to Information." *Social Text*, no. 102 (Spring 2010): 35-58.

———. "Deafening: Noise and the Engineering of Communication in the Telephone System." *Grey Room*, no. 43 (Spring 2011): 118-143.

———. "On Disability and Cybernetics: Hellen Keller, Norbert Wiener, and the Hearing Glove." *differnces* 22, no. 2&3 2011): 74-111.

Mullaney, Thomas S. "QWERTY in China: Chinese Computing and the Radical Alphabet." *Technology and Culture* 59, no. 4 (October 2018): S34-S65.

Schüttpelz, Erhard. "'Get the message through': From the Channel of Communication to the Message of the Medium, 1945–1960." In *Media, Culture, and Mediality: New Insights into the Current State of Research*, edited by Ludwig Jäger, Erika Linz and Irmela Schneider, 109-138. Bielefeld: transcript, 2010.

von Neumann, John. "First Draft of a Report on the EDVAC." *IEEE Annals of the History of Computing* 15, no. 4 (October 1993): 27-75.

Authors

Thomas Haigh is a Professor of History and Computer Science at the University of Wisconsin—Milwaukee and visiting Comenius Professor at Siegen University. Haigh has published extensively on many aspects of the history of computing and won several prizes for his articles. He is the primary author of *A New History of Modern Computing* (MIT, 2021) and *ENIAC in Action* (MIT, 2016) and the editor of *Histories of Computing* (Harvard 2011) and *Exploring the Early Digital* (Springer, 2019). Learn more at www.tomandmaria.com/tom.

Sebastian Gießmann is Reader in Media Theory at the University of Siegen. In 2023, he serves as visiting professor for cultural techniques and history of knowledge at Berlin's Humboldt University. His book *Connectivity of Things: Network Cultures Since 1832* is forthcoming in MIT Press's Infrastructures series. Gießmann's work intertwines practice theory (which he helped to establish within media studies), cultural techniques, Science and Technology Studies, and grounded histories of (digital) media. He is principal investigator of a major research project on the history of network infrastructures within *Media of Cooperation*.