

Between the Chairs: An Interdisciplinary Career

Thaller, Manfred

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Thaller, M. (2017). Between the Chairs: An Interdisciplinary Career. *Historical Social Research, Supplement*, 29, 7-109.
<https://doi.org/10.12759/hsr.suppl.29.2017.7-109>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

Historical Social Research Historische Sozialforschung

Manfred Thaller:

Between the Chairs.
An Interdisciplinary Career

doi: 10.12759/hsr.suppl.29.2017.7-109

Published in:

Historical Social Research Supplement 29 (2017)

Cite as:

Manfred Thaller. 2017. Between the Chairs. An Interdisciplinary Career.
Historical Social Research Supplement 29: 7-109
doi: 10.12759/hsr.suppl.29.2017.7-109.

Historical Social Research

Historische Sozialforschung

Other articles published in this Supplement:

Manfred Thaller

Between the Chairs. An Interdisciplinary Career.

doi: [10.12759/hsr.suppl.29.2017.7-109](https://doi.org/10.12759/hsr.suppl.29.2017.7-109)

Manfred Thaller

Automation on Parnassus. CLIO – A Databank Oriented System for Historians [1980].

doi: [10.12759/hsr.suppl.29.2017.113-137](https://doi.org/10.12759/hsr.suppl.29.2017.113-137)

Manfred Thaller

Ungefähre Exaktheit. Theoretische Grundlagen und praktische Möglichkeiten einer Formulierung historischer Quellen als Produkte ‚unscharfer‘ Systeme [1984].

doi: [10.12759/hsr.suppl.29.2017.138-159](https://doi.org/10.12759/hsr.suppl.29.2017.138-159)

Manfred Thaller

Vorüberlegungen für einen internationalen Workshop über die Schaffung, Verbindung und Nutzung großer interdisziplinärer Quellenbanken in den historischen Wissenschaften [1986].

doi: [10.12759/hsr.suppl.29.2017.160-177](https://doi.org/10.12759/hsr.suppl.29.2017.160-177)

Manfred Thaller

Entzauberungen: Die Entwicklung einer fachspezifischen historischen Datenverarbeitung in der Bundesrepublik [1990].

doi: [10.12759/hsr.suppl.29.2017.178-192](https://doi.org/10.12759/hsr.suppl.29.2017.178-192)

Manfred Thaller

The Need for a Theory of Historical Computing [1991].

doi: [10.12759/hsr.suppl.29.2017.193-202](https://doi.org/10.12759/hsr.suppl.29.2017.193-202)

Manfred Thaller

The Need for Standards: Data Modelling and Exchange [1991].

doi: [10.12759/hsr.suppl.29.2017.203-220](https://doi.org/10.12759/hsr.suppl.29.2017.203-220)

Manfred Thaller

Von der Mißverständlichkeit des Selbstverständlichen. Beobachtungen zur Diskussion über die Nützlichkeit formaler Verfahren in der Geschichtswissenschaft [1992].

doi: [10.12759/hsr.suppl.29.2017.221-242](https://doi.org/10.12759/hsr.suppl.29.2017.221-242)

Manfred Thaller

The Archive on Top of your Desk. An Introduction to Self-Documenting Image Files [1993].

doi: [10.12759/hsr.suppl.29.2017.243-259](https://doi.org/10.12759/hsr.suppl.29.2017.243-259)

Manfred Thaller

Historical Information Science: Is there such a Thing? New Comments on an old Idea [1993].

doi: [10.12759/hsr.suppl.29.2017.260-286](https://doi.org/10.12759/hsr.suppl.29.2017.260-286)

Manfred Thaller

Source Oriented Data Processing and Quantification: Distrustful Brothers [1995]

doi: [10.12759/hsr.suppl.29.2017.287-306](https://doi.org/10.12759/hsr.suppl.29.2017.287-306)

Manfred Thaller

From the Digitized to the Digital Library [2001].

doi: [10.12759/hsr.suppl.29.2017.307-319](https://doi.org/10.12759/hsr.suppl.29.2017.307-319)

Manfred Thaller

Reproduktion, Erschließung, Edition, Interpretation: Ihre Beziehungen in einer digitalen Welt [2005].

doi: [10.12759/hsr.suppl.29.2017.320-343](https://doi.org/10.12759/hsr.suppl.29.2017.320-343)

Manfred Thaller

The Cologne Information Model: Representing Information Persistently [2009].

doi: [10.12759/hsr.suppl.29.2017.344-356](https://doi.org/10.12759/hsr.suppl.29.2017.344-356)

Between the Chairs. An Interdisciplinary Career

*Manfred Thaller**

Abstract: »Zwischen den Stühlen. Eine interdisziplinäre Karriere«. The author was one of the earliest representatives of computer applications within historical research in Germany, later being appointed to the first professorship for computer applications in the Humanities in Germany outside of linguistics. The following text describes his experiences as part of that development, which lead from the beginnings in the seventies to the current state of "Digital Humanities". His view on this development of an interdisciplinary area left him with rather mixed memories: behind a sparkling front story of an enfolding field, he frequently had the feeling that there was a tendency to ignore the huge epistemic potential of a serious attempt to apply computer science to the field of history in favor of glamorous but shallow short term goals.

Keywords: History of Digital Humanities, Historical Studies and Computer Science, Interdisciplinarity.

For the usual introductory round at newly constituted committees, where one is supposed to present oneself within two or three sentences, I used in recent years quite frequently one or the other variation of "Forty years ago, I received a PhD in modern history. A year ago, I retired from a professorship of Computer Science applied to the Humanities. A lot happened in between." The reactions to that minimalistic CV have convinced me that this kind of career between traditional lines is still sufficiently rare to make it worthwhile to describe how it can happen – the more so, as in hindsight it appears, at least to me, quite straightforward and consistent.

Being invited to write an academic autobiography in a series like this, with quite a few very impressive scholars before me, is undoubtedly a great honor. And vanity would tend to let one assume that such an honor indicates personal success. This is somewhat irritating, as my own view is considerably more subdued: While I may have influenced a few things, most of the time the developments I started did not really go towards the goals I had. So while the interdisciplinary career looked extremely logical to me, the feeling of an uncomfortable emergency seat right between the plush chairs of established disciplines never quite left me. Describing what I wanted to achieve, rather than what I did, is therefore also a subject for the following pages.

* Manfred Thaller, Rosenhof 16b, D 50321 Brühl, Germany; manfred.thaller@uni-koeln.de.

These two themes – an interdisciplinarity, which appeared completely organic to me and simmering frustration about achieving slightly different things from those that I wanted to achieve – are twin red lines going through all of my life in academia. This quiet biographical flow was broken a number of times, however, by relocations which have been quite drastic, as in all cases I did not only change my place of residence, but moved into a new job profile, which changed academic as well as social orientation completely.

1. First Life: Getting Away

I was born in 1950 in Feldbach in Austria, very close to the triangle of her borders with Hungary and Yugoslavia. Administratively and even more so in the self-perception of its inhabitants, it was a town – albeit with a bit less than 4.000 inhabitants in the fifties, so a town only when contrasted with the truly rural surroundings, where even in the early fifties quite a few of the numerous extremely small scale peasants used oxen, or occasionally cows, to pull the plow or transport part of their produce to the store run by my family since the early years of the 19th century, buying for the return journey from the grocery store to which the agricultural trading was associated. Things have changed: The surviving farmers have converted themselves into producers of high-end regional food, which today can be found across luxury food stores all over Europe, and the grocery was transformed by my brother into a shopping mall of regional visibility.

But when I grew up there, it was very much a border province remote and detached, indeed. Quite generally one can observe that people coming from such very small towns and communities orient themselves in their later lives to one of two extremes: either they find it impossible to leave what they perceive as the support and warmth of the nest; or they cannot get away quickly enough from what they see as suffocating. I'm a hard case of the second syndrome.

Part of the reason for that may have been that growing up, I already lived through two quite radical relocations. Calling my mother overly protective would be quite an understatement. Therefore, somewhat discouraged from social contacts, I was a quite voracious reader already before the age of ten. In 1960 going to a gymnasium in my smallish town still implied to go to boarding school. Being overly protective never prevented my mother's conviction that both her sons were unblemished geniuses and my father's family had certainly always assumed that secondary education was something the children should consider. So at age ten, I was transferred to a public school a few hours away by train, where over protection shattered. I learned within four years to reliably hit the ankle of one of my buddies with a stone across a distance of 20 meters, which certainly made me fitter for life; but while I learned valuable lessons about defending myself, I remained sufficiently much of an outsider to be one of the more intensive customers of the school's library – stocked heavily with antiquity and history, as befits a Humanistic Gymnasium. Unfortunately this was very useful to inspire dreams about the past, or increase the quality of contributions to the school newspaper; not so for creating insights into the irregular verbs of classical Greek, so I had to leave at fourteen.

Minor points, like failure at school, never could diminish the belief of my mother in her son. So I was transferred to another gymnasium which thanks to the explosive expansion of public education in the sixties had in the meantime been established in a neighboring town. For unknown reasons that transfer helped me and starting as repeater of a year I graduated from gymnasium five years later with distinction. Still, being transferred at the age of fourteen, again, into a completely new peer group, did not really help to turn me into a social lion. So reading remained as primary hobby. This was facilitated by the existence of what at the time I considered a quite respectable family library with a couple of hundred volumes. But, while my parents most certainly understood themselves as members of the educated classes, and considered traditional high culture and literary tastes as a given, they were not truly actively involved: so modern literature was rare, while all the main classics of the canon of European literature of the 19th and early 20th century were present, augmented by a few not quite so canonical writers of the same time.

That intensively read canon created a lot of benevolence from my teachers in German and History classes, but much more important were two other texts, which in a sense have been responsible for all that came after. For some reasons forgotten, since early childhood I had always expected to become an engineer, made more concrete later by a few popular books on nuclear physics into becoming a physicist. So until about seventeen, whenever I was asked what I would do after school, I answered “study physics.” (Not going to university was never considered for mother’s acknowledged genius.) But sometime at the age of fifteen or sixteen I discovered in the library of the family the *Propyläen Weltgeschichte* in ten lexicon-format volumes, a world’s history according to the modernized German model of a *Universalgeschichte*, going from the stone age to yesterday – for all areas of the globe. (Even if Europe was decidedly more equal than the rest.) I managed to read these ten huge volumes from line one of the first page of the first volume right through to the last line of the “outlook” at the end of the last one. And the fascination for this epic of the apes crossing the borders of the *conditio humana* and finally building the world of today never again left me. This might have been just one more novel, when read along with the collected works of nineteenth century novelists. But at roughly the same time, I also encountered the memoirs of Winston Churchill. Again, I managed to read them completely, from the opening page of the first volume to the closing page of the last. While the epic might have been considered as literature, this record of an event which in the sixties was still omnipresent in the discourse, arising out of individual letters, telegrams and other sources commented upon by a writer who – at least for his fifteen or sixteen year old reader – seemed to do his best to be objective, impressed very strongly upon me that telling a story was not enough, you should feel obliged to reality.

So at some afternoon, I wondered myself, why I would want to go for physics, when I was so much fascinated by history. And that course of study was chosen.

I’ve already described that at the time I grew up there, Eastern Styria, where my home town is located, was very much on the fringes of Europe. Not only economically, but deeply rooted in a very ancient agricultural conservatism. The family’s library being highly respectable, though slightly out of date, was barely compatible with that: but reading *anything* written in the sixties somewhere else would had to

have striven hard, not to look dangerously progressive to the eyes of a convinced Eastern Styrian. So the traditional skepticism of a late teen about the wisdom of the state of things could not help to be intensified for any voracious reader. Details might be amusing; let me just mention that when I had to go to the Austrian army for mandatory military service after gymnasium, I managed to have a reasonably good military track record, being among the first of my cohort to be promoted to the exalted rank of private first class – and at the same time being known to my company officer as an activist of the referendum for the abolition of the Austrian army.

When I finished this mandatory military career, I was completely sure that studying history was what I wanted to do; that my future would be outside of my home town; and that prevalent opinions about how things should be done were probably wrong.

And so I went to university.

2. Second Life: Becoming a Professional

Going to the University of Graz, approx. 50 km away from my hometown, probably seems to be not much of an escape for someone professing a strong drive to get away from home. From the perspective of a town with slightly less than 4.000 inhabitants, a city with 200.000 seemed to be a good step, however, so my desire to leave the limits of my home town was quite satisfied, for a beginning.

Studying history brought about a reality shock: my *universalhistorische* dreams had let me think that I would learn about the developments that lead to the battle at the Talas (751, where Arab expansion met Chinese and other outcomes *might* have changed the world completely) or the pre-Columbian civilizations. Reality presented me with a historical institute which was constituted by professorships for Medieval history (mostly German, Austrian of particular importance), Modern history (mostly diplomatic, Austrian of particular importance) and Austrian history. A professorship for Economic and Social history *did* exist – but I was not yet able to understand its potential and the holder of this professorship being an arch conservative, who focused on making sure that the candidates for a degree qualifying for teaching history at secondary schools did not stray too far from the right way of thinking, was no help in enlightening me. An professorship for Ancient History in a separate institute did exist, but that was so much connected to the canon of antiquity which I had already assimilated during secondary education that it was a bit too close to the mainstream for me. The most exotic subject available was a chair for Ancient Oriental studies, which had no incumbent, but where two readers offered lectures. So I settled on a degree in “History,” with obligatory courses at the professorships for Ancient, Medieval, Modern and Austrian History and augmented that with as much Ancient Orient as I could fit into the schedule (some philosophy courses being obligatory at the time).

This situation, where all of the exotic questions I had dreamed about had to remain dreams for the time being, was certainly a disappointment. But that history was a good choice and a proud discipline was never in doubt. Given a significant glut of students, the opening lectures rather tended to emphasize that this was a very serious field, which required a lot from people, so you should think twice to embark

upon it. A historian had to be able to read French and Italian, in modern history even English (sic), no need to say anything about Latin; an Austrian historian actually should be able to read at least one Slavic language; a good knowledge of Geography was supposed to be mandatory, as was art history and legal history, and all sorts of other disciplines would be extremely useful. It *very* soon turned out that this was the theory and after the opening lectures one would quickly discover that the lecturers themselves did not have quite the broad knowledge they had declared mandatory. But if one wanted, it was certainly easy to understand that history was a central discipline and a good historian was part of the academic elite by definition.

At least for me, this was an encouraging climate – even if with my personal conviction that “prevalent opinions about how things should be done were probably wrong” I was quite skeptical and critical of my academic teachers, almost all of which were quite conservative. On paper the formal requirements were rather high: in my folder from the university I find 29 certificates of examinations taken after term-long courses. But most of these examinations were oral ones, at the most a quarter of an hour in length; professors who had to pass two hundred students through these exams at the end of a term being overjoyed if one of them somehow indicated that he had actually read something on the topic which was not mentioned during the lecture. And while most of the lectures were deeply rooted in the mainstream literature of a conservative understanding of the discipline, the library of the university provided up to date access to the much more modern literature of the day. And the many extremely lightweight exams did definitely not prevent you from reading.

This is not to say that I did not learn a lot from my academic teachers, even if in some cases I understood only much later what.

Friedrich Hausmann, editor of the charters of Konrad III and his son Henry in the *Monumenta Germaniae Historica*, introduced me to traditional medieval studies. More important in the long term was probably that he impressed me immensely by his understanding that research needs researchers who get their hands dirty in politicking, if it shall thrive. In 1975 he founded at Graz the institute of “*Historische Grundwissenschaften*” (something like “pure” or “fundamental” historical research), which basically implemented the canon of “*Historische Hilfswissenschaften*” (auxiliary historical studies) of the medievalist tradition. With some amusement I have encountered in recent years some reflections by others of the epistemological position towards history this change of name is allegedly related to. As he told the story to the students in his seminars, the point was simply that as dean he encountered an initiative of the Austrian government to advance “fundamental” or “pure” research – so in 48 hours the required new subdiscipline was invented. And, while he certainly was highly traditionalist, nevertheless in his later years he became a patron of computer supported approaches to medieval studies in Graz – from the very simple conviction that the best available is just marginally good enough for the study of history. This pragmatism was always a model for me.

Much more important for my academic orientation was Herwig Ebner, medievalist as well, who forced his students to read medieval sources in the archival original, as well as in edited form. His seminar on medieval Styrian charters in the local state archive became one of the more thorough impressions of those years. More abstractly, it impressed me extremely, to be shown how one could extract

structures, both economic and social, by a systematic analysis from medieval charters and other “mass” sources of medieval times, like registers of property holdings. While I followed only a few seminars taught by him, two of my closer friends wrote doctoral theses along these lines supervised by him.

Hermann Wiesflecker, author of a highly traditionalist biography of Maximilian I., functioned (unknowingly) more as an intellectual punching ball for me. Seeing himself as the last historian who truly understood Ranke, and infuriating me again and again by his resulting opinions, he was a constant motivation for me to read up what people after Ranke had thought and nowadays thought about proper historical methodology. Even if I almost never risked to confront him with such sacrilegious heresies.

Alexander Novotny, chair of modern history, was an extremely charming person. Specialist for the late nineteenth century, he was impressive by his knowledge about his era, but rather close to his retirement, there was not very much I learned from him concretely. Though he became the supervisor of my PhD, a contradiction to be explained a bit later.

One of the more important things I learned studying history, however, is not connected to any of the holders of chairs at the institute at the time. The lecturers of the historical institute (and in a sense also of the institute for ancient history) offered at that time a rather thorough introduction into not so much the abstract methodology of history, but the way in which a historian had to work, where to find information and how to organize it. That probably was the best possible pragmatic augmentation of the high flying enthusiasm for universal history which had brought me to the study of the field.

My engagement with Oriental Studies was somewhat brief and only lasted for four terms: the enthusiasm for universal history just mentioned had led me to the study of Aramaic as the *Lingua Franca* between classical antiquity and the adjacent culture to the East. But the requirement of Hebrew on top of Akkadian and Aramaic seemed to be a bit stiff for a minor (though responsible for study programs myself nowadays, I understand in the meantime why this was considered necessary). So in a sense that was a dead end – but nevertheless, these four terms influenced me more than may be apparent. Asked in my sixth week at university to present a paper orally about “The cosmological significance of board games in ancient Ur”, and wrangling in exercises with Sumerian and Acadian I learned how different cultures can be – and my deep distrust against the naivety of certain approaches towards sources of the last two centuries is strongly rooted in this fundamental experience that concepts and meanings are significantly different between cultures and periods, even if on surface they may look self-evident.

And to finish the list of things that influenced me during the time at the university in Graz: In 1970, just a few months before I entered the university, there appeared a German translation of Somervell’s abridgment of Toynbee’s *Study of History* – which soon after fascinated me, building directly on my love affair with *Universalgeschichte*.

As I wrote, formally the requirements at Graz were rather high. But in reality the individual exams were simple. So I managed to fulfill all the formal requirements within four terms. In the seventies, Austrian Universities still functioned under the traditional rules for Academic degrees: So at the philosophical faculty, the only

degree was a doctorate (besides a state exam qualifying to teach at a gymnasium). No Master or Bachelor being available; but also no Bachelor or Master required to start on your thesis. This being so, I went to Alexander Novotny, professor of modern history at the very end of my third term and enquired whether he might be willing to supervise a doctoral thesis of mine, starting after the end of my fourth term, when all mandatory requirements would be fulfilled.

Why to the professor with whom I actually had the loosest contact during the preceding terms? Well, if you (a) are fascinated by the great synthesis *à la Universalgeschichte* or Toynbee, (b) have a deep distrust against naive interpretations of surface texts, derived from ancient oriental shocks and (c) are fascinated by the way a painstaking examination of the details of medieval sources, can derive outlines of structures and processes from sources, which never have been written to describe them, how do you find a theme for your doctoral thesis? My solution was to look at newspapers in a new way: at my time at Graz University, a “newspaper thesis” was something slightly sneered upon, as a variety existed, which simply reported what the leaders of newspapers with a clear political profile reported about some historical developments or events. I had the feeling that this was missing the point of newspapers as historical sources, as besides these obvious surface level stories, which were usually not surprising, they contained a great wealth of details on all sorts of the development of what at that time I had not yet learned to call mentality. Unearthing a hidden structure from a series of such, trying to confront that mass of information truly at the same level of details as my co-doctoral students applied to their medieval sources, might be a preliminary exercise to base high level generalizations directly upon the study of mass sources, overcoming the irritating defect of *Universalgeschichte* or Toynbee to be unable to connect directly the high order interpretation with what the surviving sources say. As avid reader of a family library with a slightly obsolete focus, I had, additionally, the feeling that the late 19th - early 20th century was the period with which I was most familiar by far.

So I proposed to my supervisor-to-be to try to systematically compare the way in which the information offered by contemporary newspapers created, maintained and changed the majority opinion about the United States in Austria, Germany and England in comparison, between 1840 and 1941 (“United States” simply as they were the most exotic topic still covered by the scope of the study of history available in Graz). The supervision by Mr. Novotny was quite light, one might say superficial. But that he agreed to further and encourage what from his point of view might easily have been an Utopian project by a confused student, I appreciate extraordinarily.

After three years in libraries in Bremen, London, Munich and Vienna, creating ca. 35.000 index cards out of seven hundred years of newspapers and augmenting them by excerpts from a bibliography of approx. 650 items, I delivered 1230 pages, including some 7000 footnotes (Thaller 1975).

On the formal level, this gave me a PhD at the age of 25; at the intellectual one both the conviction that trying to derive structures – specifically of the development of the mentality – from mass sources was a fascinating challenge, as well as the suspicion that there must be a more convenient and rewarding way to do so than I had done, in what eventually was an enormous apparatus supporting some rather short theses on the development of mentality out of the information available.

Having a PhD at 25 has one big advantage: you are not really very much concerned about your pension, so your drive for a secure income is not particularly developed. My parents not exerting any pressure to get into a proper job, I was not searching very intensively (besides taking an unforgettable three months overland trip all the way to India and back). And out of my only partial satisfaction with my thesis, despite it having been rated as outstanding, arose the feeling that I would have to learn much more about sociology, the social sciences being at that time the disciplines with which historians seemed to be most compatible and complementary.

Austria has in Vienna an Institute of Advanced Studies, which offered at that time a formal postdoctoral education, among other fields of study in (empirical) Sociology, where you could apply for a two year scholarship, independent of the field you had studied before, provided you passed an entrance examination in sociology (based on a rather long paper on a topic of your own choosing) and written exams in mathematics and statistics. I had not had any contact with mathematics since gymnasium, nor any at all with statistics. Nevertheless, teaching myself from various text books, I managed to pass both written exams as second best in that year, very much to my surprise.

So I went to Vienna.

These two years provided a quite intensive contact with sociology, in my case sociology of medicine and particularly social mobility, my major topic on which I delivered both of the annual theses expected. The postdoctoral training gave me the possibility to listen to a group of excellent international sociologists and related scholars, Antony Giddens from Cambridge, Aaron Cicourel from San Diego, Anatol Rapoport from Toronto (albeit he is stretching the definition of a sociologist a bit) and quite a few more.

But while these proved useful in later years, when trying to act as intermediary between historians and sociologists, my Viennese time was much more important for me in another aspect. While not all my fellow postdocs felt equally happy about it, the institute at that time was unequivocally dedicated to an empirical approach to sociology. As a result, we had a mandatory course in SPSS. This fascinated me enormously, simply thinking back to my 35.000 index cards and their endless resorting. But more than this practical consideration, I was fascinated by the abstract concept of SPSS as an embodiment of a specific methodology, which turned abstract statistical methods directly into operative tools. At the same time, my heavy exposure to sources between the ancient orient and 1941 convinced me in a very short time that the way between a historical source and a statistical data set was much longer and more convoluted, than that between a questionnaire and such a data set. Therefore, for a historian to enjoy the same kind of tools embodying a methodological approach would need a different kind of software. With a 26-year-old's combination of naivety and optimism, I decided to write such an SPSS for historians. As transferring my 35.000 index cards to punched cards seemed to be a bit excessive as a first exercise, I fell back to the medieval charters which had impressed me in my first terms and decided to define a way to enter "Standardized Computer Readable *Regesta*" into a computer and provide tools for handling them. For this I consulted the staff of the computing center of the Institute for Advanced Studies – who pointed me on the one hand to a manual of the General Inquirer,

unfortunately not able to run on the institute's hardware, and to Ralph Griswold's tutorial for SNOBOL 4, a programming language rumor had it, was good for people who wanted to handle texts.

I seriously started with my computer readable *regesta*, or rather with learning what I needed to program the software for their handling. But this hobby was quite soon overshadowed by a very surprising development. One of my friends from Graz, who had written one of those theses deriving medieval structures from charters, Gerhard Jaritz, had in the meantime found employment at the Institute for Medieval Material Culture (*Institut für Mittelalterliche Realienskunde*) of the Austrian Academy of Science. This institute supported as its flagship project a huge photo archive, which was built by systematic campaigns to photograph all surviving pictorial sources from the Middle Ages, to study the material culture – clothes, furniture, buildings – of medieval times. Geographic coverage was “Austria”; the borders of the modern republic not being particularly relevant in the 15th, 16th, and preceding centuries, the definition being tacitly extended eventually to cover more or less the area considered Austrian at the eve of World War I. To administer this archive, the institute had contacted IBM which resulted on the one hand in a first version of a flexible scheme of hierarchical descriptions of images, and on the other in a complex thesaurus of descriptive terms for the description of the content of the images. While an impressive achievement, even today, this thesaurus unfortunately turned out to be basically useless: it pre-assumed a precision of knowledge about what was what in a medieval image, which technical term, e.g., to apply to which only partially visible item of clothing that it was never applied. Beyond that, while the thesaurus was created by some sort of IBM's support for research program, any data base solution to be implemented by IBM was far beyond the limits of available funding.

This story I heard from Gerhard Jaritz one evening, when I had told him about my plans for medieval charters and we wondered, if I could not be useful for the institute where he worked. Being still 26, I found it quite obvious that I could outdo IBM and approx. 10 weeks after I had submitted my first SPSS job, I signed a contract to develop a software solution for the Institute for Material Culture. The SPSS for medieval charters was shelved and an SPSS for medieval pictorial sources replaced it. As using computers was bad enough in the eyes of the more conservative – i.e. the vast majority of – Austrian medievalists, their Franco-Italian language bias should at least not be offended on top of that by English, so the command language chosen was Latin (Thaller 1980-1). Looking back at it forty years later, I am surprised myself – but the software actually worked and after a few iterations, some of which we will encounter on the following pages, the image descriptions entered on punched cards from early 1977 onward, are still part of an active database on the internet, augmented by the images in the meantime.

That the medieval charters had been shelved was also connected to another development. At the Institute for Advanced Studies, postdocs were expected to write a quite extensive annual thesis. I had originally intended to work on strategies of members of the upper class in the first Austrian Republic (1918- 1938) to achieve material advantages by using alleged cultural interests to re-direct certain types of funding, but after spending six weeks in the *Allgemeine Verwaltungsarchiv* before finally getting access to the first set of documents, I gave up in frustration and

decided to work instead on the social mobility in the same period in Vienna. For this, like almost all similar research, I relied on the marriage registers of selected parishes. Being very much aware of reliability problems of encoding, I decided to transcribe the occupational and topographic information of the registers verbatim and write a set of programs with which I could later encode these terms by a successively developed code book.

While many things are different in 2017 from 1977, one seems to be a constant: the number of people who have a substantial background in History, as well as in the tools of the information technology of the time, is much smaller than the demand for them. So word of mouth about my activities lead to further part time jobs. As a result, I wrote in parallel, software to:

- 1) Encode my marriage registers.
- 2) Administer the Krems image data base.
- 3) Administer the quite extensive collection of census lists for (demographically oriented) family history by the projects directed by Michael Mitterauer at the University of Vienna.
- 4) Extract information on persons mentioned in a huge set of records on the management of big companies, which existed as an early example of an electronic record from which books would be printed. This to analyze interlocking directorates, i.e., the way in which people having management or board positions in one company also occurred in boards of other companies. All in the context of a German-Austrian project by Rolf Ziegler, at that time holding a chair in sociology at the University of Vienna.

That, together with my postdoc fellowship, these contracts all together, amounted to a contract time of approx. 250 % of the official working hours, can be ascribed as part of the self-ascribed working capacity of somebody 26-28 years of age. What really was important about it, intellectually, was that at a rather early stage of my career I was prevented from overspecialization, being forced to work on completely different subject areas at the same time. And more prosaically, it had another effect. My original specialization in the history of the mentality had raised little, if any, interest. Attempts to talk about historical methodology at Vienna with Heinrich Lutz, resulted in his response to my praise of Fernand Braudel's *La Méditerranée et le Monde Méditerranéen à l'Epoque de Philippe II*, by the statement that from Tolstoi's War and Peace more about history was to be learned than from all that the *Annales* ever published. And an interview with Fritz Fellner, where I mentioned the then young Bielefeld school of *Historische Sozialwissenschaft*, resulted in a veritable explosion about "those young people who make Wehler their god."

On the other hand, I applied for and was offered a contract by the Institute for Advanced Studies in Vienna; participation in a workshop of the family history project of Michael Mitterauer lead to an offer of a position at the Max-Planck-Institute for History in Göttingen; Rolf Ziegler offered a possibility to move with him to Munich, where he had just accepted a new professorship; a group of people at my old university at Graz, notably Walter Höflechner, indicated a possible position in a project there. What all these offers had in common was that they required a profile with different mixtures of historical, sociological and information technology knowledge. For my original research interests in the kind of history of the men-

tality I described, exactly no jobs were offered. This made the basic decision to orient myself towards an interdisciplinary profile extraordinarily simple.

Of the possibilities offered, Göttingen had the highest proportion of history; at the same time it looked like a logical extension of my flight from home; and the knowledge to join a young team there, which was definitely as convinced as I was that prevalent opinions about how things should be done were probably wrong, also helped.

And so I went to the Max-Planck-Institute at Göttingen.

3. Third Life: Defining Myself

Göttingen was a new world, indeed. While in Graz and Vienna, as at the vast majority of German universities at the time, history was still dominated by quite traditional views and strongly rooted within the German historiographical tradition, Göttingen, very much due to the influence of Rudolf Vierhaus, the then director of the Modern History department of the Institute, was intentionally open to a historiographical concept space which included Anglo-American and French traditions at least as strongly as the German ones.

When I arrived in 1978, I was supposed to support a bundle of projects represented by Peter Kriedte, Hans Medick, and Jürgen Schlumbohm, and, a bit apart, David Sabean. Kriedte, Medick, and Schlumbohm had just published jointly *Industrialisierung vor der Industrialisierung* (Industrialization before Industrialization: Rural Industry in the Genesis of Capitalism) which certainly put “proto-industry” on the hot list of historical research topics of the time. After publishing this book, they now intended to engage into in-depth studies of three presumably proto-industrial communities to verify – or in any case assess – the validity of their concepts, by micro historical studies. This was supposed to be based on what was later called an “extended family reconstitution.” Roughly speaking: take the marriage, birth and death registers of a historical community over a period of preferably not less than 200 years and reconstruct out of them all the possible genealogical relationships. Connect these reconstituted families with all census records available, to check the correctness of the basic reconstitution and augment it by a list of the non-related persons living in the household; connect them further to lists of taxation, to get an idea of the economical position; augment this by links to protocols of proceedings from local courts, to show their involvement in infringements upon law and expected behavior; add further itemized lists of testaments and inventories of property to fine-tune the indicators of economic standing; and generally connect everything you can get hold of which mentions people identifiable within the system. And analyze it.

For the protoindustrialization group at the time of my arrival, this was a plan. David Sabean already had boxes and boxes of forms, containing painstaking excerpts from the archives of “his” village, Neckarhausen. And he had certainly the most elaborate visions of what analysis should accomplish: I remember particularly the idea that it should be able to check, whether the appearance of root crops in testaments might be related to a higher incidence of illegitimate children, at least among the servants of the household, as the higher need for light manual labor

implied by raising root crops might be a reason to accept extramarital sexual relationships more relaxedly, as they would increase the labor force available within the household.

Incidentally my varied projects in Vienna had prepared me quite well for that. The image archive had familiarized me with complex data bases; history of the family research provided an understanding of the steps involved in handling census-derived data sets; marriage registers were one of the most central sources in the mix; and identifying people in business registers provided a background in the problems of comparing names with potentially different spellings. What probably was more important was that these different experiences had made it very clear to me that I could not possibly write individual programs for the large number of possible computational steps involved, but that I would have to find an abstract data model which could support a software system able to perform these tasks, controlled by a higher level command language, finally implementing the “SPSS for history” thought about earlier. Relatively soon it became clear that this model would have to deal with two issues: on the one hand an extremely flexible structure for the segments of data extracted from the sources and their relationships. On the other hand support for the strange notational forms occurring in historical sources, where currencies appear in all forms except decimal, measures of volume may change, depending on the produce they are applied to, and calendar dates range from such related to the nearest feast of a Saint through to the French revolutionary calendar appearing in documents from areas being part of France at the time, and just about everything can be given as an approximation only, like January 5th-February 2nd.

The data model behind that was made explicit only a few years later; in its first form it was simply derived from the requirements to develop something flexible enough to host all the sources implied by the projects I was supposed to support; and after my Vienna experiences I had at the back of my mind also the notion that whatsoever had been needed for those projects would have to be supported by the software-to-be as well. How to solve practical problems – input formats for structured and semi-structured data, approaches to record linkage, orthographic creativity notwithstanding – could partially be derived from the emerging literature on computing for historians and the Humanities in general, though most of it simply had to be invented. But the ample library resources of a Max-Planck-Institute made it easy to collect more or less systematically such literature as was available, Historical Methods Newsletter, renamed Historical Methods later, probably being the most useful item.

But while this scope existed, the job I held was still the classical case of a technical support job for a bundle of research projects justified by virtue of *their* historical goals, even if my own quite intensive historical background made the communication very much easier than it would have been, had I had a background in contemporary computer science.

That I could go beyond the role of technical project slave was the result of three closely related decisions and developments.

(1) For a technical support job of that type provided by somebody with a historical background usually one of two outcomes is possible. Many people find after some time that their technical qualifications have risen to a level where the job

opportunities provided in business or industry are so much better than those on the project slave market, they more or less melancholically quit academia after some time. The other possibility is to collect so much sympathy from the people in the project supported that after a few years you are allowed to show by your own content-defined historical research project that you can also be trusted with work valued more highly. In my case that was made a bit harder, as the only kind of project I could have fit organically into the historical research environment in which I was embedded, would have to be one in the field of socioeconomic history described, where the relevant literature was mushrooming at the time. While at the very least after Vienna, the scope of my familiarity with a quite wide variety of fields of historical enquiry was probably uncharacteristically wide, nevertheless the only field where I really could claim extensive familiarity with was my own peculiar background of history of the mentality. To get widely read in socioeconomic history, or microhistory, at the same time as building a really challenging software system, getting myself an on the job training in applied computer science and software technology, was completely unrealistic. So the only future I could look forward to if I wanted to stay in academia was to be a bad historian, tolerated because also being a reasonable amateur programmer.

This was not really attractive. So I intentionally and knowingly gambled that there was room for an academic specialization in computer technology for historical research. And if such a field did not exist, I'd have to invent it.

(2) I am really convinced that it is not possible to imagine a more inspiring atmosphere to develop the abstractions needed for the support of the type of research described, then at the Max-Planck-Institute at the end of the seventies. Feeling to be part of a young group that was finally doing history as it should be done was stimulating in the extreme. Nevertheless, my colleagues looked at the software to be developed as a solution to their concrete problems. Being the only person with a technical background around made it extremely difficult to discuss any kind of technical decision or problem. This became apparent very soon; at the same time I discovered that Max-Planck's budget was happy to pay for extensive postage and the high speed printers of the computing center made it cheap and easy to print texts. So, starting my job in Göttingen on October 1st of 1978, already on January 1st of 1979 I started distributing a "system bulletin" for the software system which in the meantime had been baptized CLIO, after the muse of history, describing the basic decisions about capabilities needed and data formats to be handled (Thaller 2017a [1980], reprinted in this volume, Thaller 1980-3, 1981-1, 1984-2, 1985-1, 1987-1, 1987-5, 1988-2). This was mailed – unasked for – to 120 or 200 persons, basically everybody who appeared in a publication that made it probable that they would be interested in applying computers to historical research and presumably read German. As a mechanism to get in touch with a community, within which I could discuss my technical concepts in the sense of software development, it was a total flop. But the notion of a program package for historians, which would make discipline specific solutions available at the same level of complexity as the statistical packages of the day did for statistics, raised much more interest also from people I had never heard of, than I ever expected. So after the fourth and final of these system bulletins had been mailed in the summer of 1980, I was very widely known in the German speaking scene, and slightly beyond, for all that concerned compu-

ting in the historical disciplines; so widely, actually that the ensuing correspondences and other contacts took up so much time that no fifth edition of the bulletin ever has been distributed.

(3) By far the most important reason for my emancipation from the status of technical slave, however, I owe to the director of the Max-Planck-Institute, Rudolf Vierhaus. As mentioned above, I arrived at a time when three of the micro historical case studies I was hired to support were on the drawing boards and the fourth consisted of a huge pile of boxes with paper forms. For the first three, the only real requirement at the beginning was, therefore, to design sensible input formats for all sorts of quirks which might occur in the data and the fourth case study might proceed a bit faster, but here, too, a lengthy stage of data input must precede anything else. On top of that, the tradition of structured historical research projects, distributing spade work to student assistants, did not exist yet at the time. So the notion that data entry should not be done by the highly qualified authors of micro historical studies themselves, but by data entry assistants was new and needed quite some time to be developed. And in most cases, there were urgent other manuscripts to be finished by my colleagues, before they really could engage upon the case studies to be supported by me. This meant that within a very few months I found myself with early software versions, which desperately needed data to test them, without much of these data arriving from the projects I was supposed to support. I first looked to my Vienna heritage and converted the data from some of the projects pursued there, notably the medieval image descriptions, into the input formats of the new software taking shape.

The bulletin described under the previous item also lead to cautious enquiries of other people, whether what I described could possibly be made available to them. This made a lot of sense to me, as I could test what I had developed, well before the projects in the institute had their data ready. But this was clearly an age before UNIX. My software had been written in PL/1, which was much more efficient than the SNOBOL used in Vienna – and, besides the problems I will mention a little bit further, a marvelous tool, offering all sorts of possibilities. But portability between operating systems and hardware platforms, in the pre-UNIX days incomparably more variant than today, was not a strong issue. More directly: porting a PL/1 program at the best of times was a major headache, at times not so good a veritable nightmare. On top of that historians at many universities found it quite difficult at that time, to convince their computing centers to allow them access to anything considered serious resources. This was never a problem in Göttingen; for reasons which would require a digression into the innards of research funding, the computing center at Göttingen, known as GDWG, one of the two high performance computing centers run by the Max-Planck-Gesellschaft, was actively happy if one of the local Max-Planck-Institutes claimed resources. So the obvious, but less than orthodox solution was simply to invite external projects to use my software on the hardware in Göttingen giving results to them and test data to me.

This is where I have to be immensely grateful to Rudolf Vierhaus: one could of course have argued that the technical slave should restrict himself to support the projects for which he was hired, rather than quite a few external ones. But, fortunately, though never involved in computer supported research himself, he followed the implicit logic that any of the academic research staff he was responsible for

could do more or less whatsoever he or she intended to do. Provided, there was some kind of proof that what they did was really at the cutting edge. This proof seems to have been provided by the feedback he got from others. So in a sense, he was the first person I convinced that a distinct field of study did exist which merited to be explored not to serve others, but in its own right, described provisionally as “computer technology for historical research”. – Only much later I learned that he not only suffered my invitations to other projects to work with my software in Göttingen, but actively encouraged some external research groups to approach me for that purpose.

In any case this meant that sometime during the early eighties, I had shifted from a support function to the representative of an independent line of research labeled *Historische Fachinformatik* – computer science applied to historical research – at the Max-Planck-Institute, embodied in the software system known as CLIO. While this still was a bit uncertain as long as I held temporary contracts, from 1983 onward I had tenure and could plan for the longer view.

Those plans were built around a major shift in the “vision” behind my work. The original inspiration had been “SPSS for historians”: tools which implemented specific methods. The discovery that before applying any kind of analytic tool whatsoever, technical support was already needed for the preparation of data material much more intensively when dealing with historical data than with those of contemporary sociological studies, led to a complete re-orientation of that concept. My assumption was now that an encompassing software system for historical research would first of all have to provide for a representation of the informational content of a historical source in electronic form as complete as possible, well before any decision was made which analytic tools could be most appropriately applied to that source for a given study. I used the term “source oriented data processing” (see later Thaller 1994-1) for this concept, as an intentional juxtaposition to what I perceived as the “method oriented” approach, building a software system out of modules each implementing a specific statistical method.

While the concept of a “software environment” became current only much later, it is useful to describe what has been built as CLIO already in those early years. The “representation of the information contained within in a source” was handled by two main components.

On the one hand the text of a historical source was supposed to be transcribed in a more or less structured way, to be administered by a system offering the typical retrieval capabilities of a data base system. The basic building block for that purpose was an “element” – unlike the scalar terminal field of a database, however, this was always perceived to be a vector of “items.” So, if a conceptual field had more than one manifestation (or a few hundred) this could be modeled. The classical example from socioeconomic history always being a person with more than one occupation – like a peasant also being a weaver, where no historian examining the source would be able to determine, which of these two occupations was the more important indicator for the social position of that person. “Elements” however were not simply arrays, but bundles of arrays, as each element (or alternatively, each item) was supposed to exist potentially in three different “aspects”, most frequently used for alternate readings and annotations of the transcribing historian. The length of items was always variable; databases where 99% of the occurrences of a specific

element contained less than ten characters, but a few went beyond thousands, being quite frequent.

These elements could be grouped together, such groups vaguely resembling records of data bases, variation, however, being again optimized – data bases, where all instances of a group contained one specific element, but a few hundred of elements each appeared in only one of the fifty thousand instances of this group were extreme cases but existed. The groups in turn could be organized in hierarchies – depth unrestricted, number of instances of a particular group subordinated to another one unrestricted and variable. One group was distinguished as “document,” defined as being the major subdivision of the content of a database; databases consisting of exactly one document existing as well as such holding tens of thousands.

Together with these data structures, a CLIO data base also contained various “logical objects” which implemented a specific type of procedural knowledge: How to handle a specific type of calendar dates, how to apply a specific variety of a Soundex algorithm for name comparison, or how to convert a historical currency into numbers one could work with in a statistical program. Connecting such procedural knowledge to a specific element designated this field as having a specific data type – influencing sort order, evaluation of queries etc.

Once data had been prepared in such a database, it offered the obvious, a retrieval system and a system for the generation of reports for the creation of indexes and counts. It also offered an export facility, which supported the export of subsets of the data contained in such structures into records suitable for the analysis by SPSS or other statistical packages, services rendered consisting of features mapping the highly variant structures (or rather: parts thereof) into regular statistical cases and translating verbal expressions into appropriate codes via a system of computer supported code books and thesauruses. Which also allowed me to show occasionally that between a statistical and a hermeneutic approach there was much less of a contradiction as frequently claimed (Thaller 1982-2).

And, most distinctive for the original community of researchers interested in micro history, a system for record linkage, which allowed the identification of persons occurring in one data base within another by user defined rules, supported by localized algorithms for name comparison, like Soundex. To allow an iterative approach for that, comparisons were based on “catalogs” of persons occurring in a data base. Once a person was identified in another data base, its entry in the catalog could be removed, excluding him from further attempts at connecting the remainder, without changing the underlying data base representing the source. On the other hand, the link between the occurrence of a person in one database and another could be activated as a “bridge” between the two databases, allowing retrieval operations – or the production of data sets for statistical analysis – which drew on both data bases (On the vision of κλειω from the micro historical / prosopographic point of view see Thaller 1985-3, 1986-1).

This approach on the most practical level simply asked for highly flexible data structures. On a slightly higher level for practical, algorithmic solutions for specific types of content: how do you do arithmetic with non-decimal currencies? On the abstract level, there was always the tantalizing vision that as far as a traditional software system can be expressed as a set of statements in the theory of crisp sets, a software system appropriate for historical research would have to be expressed as

a set of statements in the theory of fuzzy sets (Thaller 1984-3, Thaller 2017b [1984], reprinted in this volume). Tantalizing, because there was always the hope, *those* would be the problems I could dig into, as soon as I had demonstrated sufficient usefulness of the lower levels. “Sufficient” turning out to be one of those moving targets ...

The implementation of this system was my main business between 1978 and 1986; and “testing it”, by allowing other people to use it with my help in Göttingen created a growing net of external contacts. Most intensively at that time with the University of Bielefeld: here the projects of Wolfgang Mager and Dietrich Ebeling, based also on a family reconstitution used Clio at the same time, as the proto industrialization projects at the Max Planck Institute. That it is not self-explanatory that what could be seen as a competing research project could use the same software and even hardware resources as the projects for which it had been developed, became only apparent to me much later – a good illustration of the climate existing at the Max-Planck-Institute under the directorship of Rudolf Vierhaus. The reports on activities delivered to the advisory board of the institute speak of 30 external users, besides a few more projects to support at the Max Planck Institute. My contacts from the time in Vienna still quite active – medieval images (Thaller 1984-1) and quantitative family history – made sure that I could work with material of many periods, medieval pottery and some attempts at automatic classification of its shapes, with Brigitte Cech as a partner, probably farthest away from the original social economic micro history bias.

While my emerging data model for historical sources benefited much from that, those years may have been even more important, by substantiating my profile as a specialist for “computer technology for historical research” as an emerging subject. I managed to raise substantial interest: At the climax of this part of my career, I have been teaching term long courses as a guest lecturer concurrently at the universities of Graz, Hamburg, and Munich in the summer term of 1985; at the universities of Göttingen, Munich, Siegen, and Vienna in the winter term of 1985 / 1986; and at the universities of Göttingen and Salzburg in the summer term of 1986. Rereading the report I delivered to the scientific board of the Max-Planck-Institute I discover that I promised at that time that after summer 1986 I would accept invitations to deliver regular courses only at the University of Göttingen anymore. A promise which was not kept completely. In any case, while details are not interesting, during the eighties and early nineties, I had the possibility to teach regular courses at about a dozen universities in Germany and Austria.

This provided, on the one hand, a reality check for the research fellow at a leading research institute. On the other hand it forced me, to read up also on areas of the application of computers to the Humanities, which went clearly beyond history and very much expanded my horizon.

Having ample resources, both bibliographical as well as related to computing power, collaborating with more than a dozen research projects in Germany and increasingly beyond, teaching at many universities: This might still have made me an academic computer freak who invented a small world of his own. That I developed a wider perspective than that of an individual scholar, I have to be grateful for to two communities.

The one was the possibility to teach at the “Quantkurs,” started by Gerhard Botz at the University of Linz in 1978 from 1979 onward, also when it moved with him to Salzburg in 1981. As a matter of fact, I am the only person who participated in this summer school more often than Gerhard Botz himself, as I was part of it also in its twilight years in Bergen, Norway, where it petered out in 1999. This summer school was particularly during the eighties an extremely important hotbed for “people who disagreed with the way in which things had always been done” (Thaller 1988-3, 1988-4, 1988-5). Teachers – mainly early career researchers like myself – as well as students – at the beginning either other early career people or doctoral students – constituted a vibrant community, which continued the discussions regularly into the small hours in Salzburg’s beer cellars and bars. People like Joseph Ehmer, Sylvia Hahn, Albert Müller, and Gerald Sprengnagel and at the beginning Gerhard Jaritz, later his students, created a wonderful climate for trying new things, cooperating also beyond the two weeks in autumn (Thaller 1984E, 1989E). (I restrict myself here to the “Quant” group, focusing on quantitative methods; the “Qual” branch (oral history) and the image branch certainly provided a wonderful background for the notion that methodological innovation was not restricted to quantification, but the discussions between the branches were never quite as intensive as those within the branches, the quantitative one not only being the oldest, but the one with the most tightly knit community within it).

The other context preventing me from becoming a local freak was provided by Quantum or the *Zentrum für Historische Sozialforschung* – and there predominantly Wilhelm Heinz Schröder. This embedded me into the “quantitative” empirical core of the German paradigm of *Historische Sozialforschung*, only loosely connected, but connected, to the much longer visible, primarily theoretical concept of an interdisciplinary connection between History and the Social Sciences represented by the Bielefeld school. In the “Cologne group,” as it was usually known, I participated in a number of summer schools as well as acting for some time as a regular contributor to the journal *“Historical Social Research / Historische Sozialforschung (HSR)”* in the eighties (Thaller 1981-2, 1981-3, 1982-4, 1982-5, 1982-6, 1982-7, 1983-3, 1983-4, 1983-5, 1983-6, 1984-5, 1984-6).

In both these reference communities I originally entered as a “quantifier”; more bluntly: as somebody who was able to teach how to handle SPSS (Thaller 1982-1, 1985-2). That was exciting, for a time. But it was actually slightly out of focus. In my own work I was following the notion that historical sources provided information which was not immediately ready for statistical analysis, but could with the help of computers be prepared for such analysis and the tools to be used for that purpose could be used to produce rather important services long before statistical tools themselves could be applied. In contrast to that, all my practical teaching was restricted to software available at the various universities and summer schools where I taught. Which simply meant: statistics packages. And in the eighties even cluster analysis, with which I experimented quite intensively (Thaller 1982-2), was challenging the computer resources available at these sites so much that teaching their use was a pain. And on the theoretical level, I was decidedly unhappy that the equation “computer usage in the Humanities = statistics” was taken for granted by many of my colleagues.

Together with two other developments, this brought me into a kind of crisis in the middle of the eighties:

(1) I formulated my own interest as an interest in “formal methods,” operationalized by me as “all methods which are so precisely defined that they allow the derivation of algorithmic solutions”, explicitly aiming to connect the community of quantitative social history, with the community of literary and linguistic computing and that of archaeological computing, only very vaguely aware of each other. (Therefore also an attempt to get closer to computer linguistics, soon aborted: Thaller 1987E.)

(2) If I wanted to promote any of this on a practical level outside of Göttingen, I had to convert my software into a form in which it also could be applied by others (and, hopefully, also developed further by others).

(3) Having developed the conceptual model behind the software I created between 1978 and 1985 alongside of the software, it needed a more systematic implementation in any case. Which would have to be done quite radically: To port the hundred thousand lines of PL/1 I had written in those years unto other brands of mainframes had never really worked; and while visionary in many respects, PL/1 was a nightmare in others. There have been all too many times when the sunrise coincided with my recognition that a bug was not my fault, but simply a bug in the compiler – completely and hopelessly inaccessible to me.

To react to these problems, which I considered to be a collection of dead ends, I started to work in 1985 towards two aims: On the one hand, I wanted to bring together an international community, which would combine representatives of the groups so far related mainly to individual disciplines. (My own background in a very self-assured, possibly imperial, tradition of historical research, considered it quite evident that historical research was a very suitable umbrella over a much larger part of the Humanities, than the narrower understanding of history implemented by the typical historical department at a university.) On the other, I decided to re-implement my CLIO system from scratch – re-baptizing it κλειω, to signal continuity as well as change.

Both projects of course happened in the context of 1985:

In the seventies and early eighties most historical research communities, and indeed many sub communities of Computing in the Humanities as well, considered applying computers as almost identical to the application of statistical methods. (Typesetting explicitly excluded, as it very rarely was based on methodological reflections, but mainly intended to keep traditional printed media affordable.) During the eighties, even before the appearance of the PC, however, “data bases” started to replace “statistics” as index fossil of Humanities Computing.

That in many cases statistics were not applied for any methodological reason, but because computers were employed with the hope of allowing a faster orientation in an otherwise unwieldy mass of sources and the only software that was easily applicable was statistical, is part of the history of the field (Thaller 1990-3, Thaller 2017d [1990], reprinted in this volume), not of my biography. That the satisfaction not to have to create statistical codes but simply to type in what was there, frequently turned into a rude disappointment later, when it turned out that the lists of “simply what was there” could hardly be interpreted, belongs into the same category.

Be that as it may: My attempts to create a community across the traditional dividing lines lead to an international workshop at Göttingen, supported by the Volkswagen foundation, in 1985 under the name of “*International workshop on standardization and exchange of machine-readable data in the historical disciplines*” (Thaller 1986E, Thaller 2017c [1986], in this volume; Thaller 1986-2). Lou Burnard in private communications has mentioned a number of times that this was the event which for the first time brought him in touch with the necessity of standardization to become able to exchange data. This initiative was much more successful than I ever imagined: 1986 a second instance at the University of Graz was of a much larger size, a conference, rather than a workshop. And when a third instance of this mini-series was organized in Paris, it definitely was a full blown international conference. Both these follow up conferences were chances to learn that the community was much wider than I had expected. The Paris conference brought me in touch with Jean Philippe Genet; Graz laid the groundwork for the relationship with Ingo Kropač – one of the most extended cooperative relationships I ever encountered and the only one, which resulted in concrete software contributions to κλειώ beyond my own.

The second project to break out of various dead ends was also influenced by the time, technically in this case. Switching from PL/1 to C promised that future sunrises would not find me brought up against a compiler’s impenetrable wall. (Rather against mistakes introduced by myself. A big improvement, psychologically, and the difference between impossible and difficult to solve.) Much more important for the future was that with the starting wildfire spread of UNIX, it became feasible to transfer the system from mainframes to PCs – the first C / UNIX version actually started its life on a 286 PC, running the very first UNIX distributed by IBM (XENIX).

The decision to reimplement the software unfortunately did not improve the support I could lend to the projects at the Max-Planck-Institute. During my early years, those projects were so deep into data entry that I was rather desperate for data to test. Now, unfortunately, their most intensive phases of analysis met a system designer who was totally distracted by reimplementation. Even if that support ceased to be my most direct responsibility: as my research activities acquired their distinct profile, I had already earlier been able to draw upon the support of an assistant – Michael Goerke – and in 1988 my support responsibilities for the projects of the institute were transferred to Peter Becker, for whom a new position was created.

The contacts to the outside world, were intensified considerably: Getting more computing power on an IBM 6150 under AIX (feeling like your very own mainframe at the time) provided the base for the implementation of a first publicly available version of κλειώ – documented by a systems manual printed not only in German (Thaller 1987-6, 1988-7), but also in, well, Germlish, I am afraid, the English translation being my very own (Thaller 1987-7); and also French, the translation by Josef Smets (Thaller 1988-8) being funded by Michel Parisse, then director of the *Mission Historique Française en Allemagne*, since 1977 located at the Max-Planck-Institute. In the tradition of the “system bulletins,” which made my original ideas visible, the manuals were mailed to a few hundred interested people. And this first version *was* portable; to a whole range of UNIX systems – and even to MS-DOS, starting with the 286 processor.

These results, together with my track record so far, convinced the Volkswagen foundation in 1988 to fund – outside of all regular funding lines – for three years a project to develop software for the historical disciplines. This project was small, compared to some of the ones I succeeded to get funded later in my life, but in some ways it was the one I still can identify with most completely. The assumption that the historical disciplines require software solutions that reflect the specific properties of their sources, *was* the drive behind my whole academic life: This is the only project that clearly and directly was dedicated to that purpose. All the other ones were more or less successful attempts to fit what I wanted to do into the requirements of various funding schemes.

Additionally there were a few developments in 1987, the year before the project started, which allowed for an intensity of preparation for the project which was highly unusual. On the one hand I could dedicate my regular seminars I had started teaching in the meantime at Göttingen University to introductions into my ideas about “source oriented data processing”, including practical exercises with the first versions of κλειω – and the staff for the project was almost completely recruited from students in this seminar. On the other hand, I had the honor and pleasure to have a visiting professorship at the Hebrew University in Jerusalem in 1987. I am afraid my tracks in Jerusalem were not exactly deep: The first 50% of the term were covered by a very strictly organized strike of the students, so I was scarcely visible for the university. For myself, however, those were probably the most productive three months of my life. Besides writing large parts of the manual, which was so significant in preparing the project, and working on the software itself, I wrote altogether almost 200 pages of various articles for edited volumes, journals and *Festschriften* – which I printed together as a collection of “Jerusalem papers” (Thaller 2017e [1991], in this volume, 1987-3, 1987-4, 1988-1, 1988-3, 1988-4, 1988-5, 1989-4, 1990-1). For me as a blueprint for how to progress and a very good way to define for the community of interested partners, what my emerging methodological concept consisted of.

The κλειω project proposed to do five things:

- 1) To develop the software further and document it persistently, including appropriate tutorials.
- 2) Offer the software to the community within a support environment, including regular summer schools advising on its usage and a help desk.
- 3) Develop a completely new, additional software package, called Standard Format Exchange Program (STANFEP) to convert between various input formats (HGR B6, HGR A6).
- 4) Establish a series of printed support materials, introducing the most appropriate computational techniques to apply to specific groups of sources.
- 5) Establish a series of machine-readable data for secondary analysis, available on floppy disks.

Some of these need a bit of explanation.

In my early years at Göttingen, as mentioned, I had cooperated with a relatively large number of projects, which had made historical data machine-readable according to many different concepts, and received data for further experiments from many more. To make any experiments with κλειω using such data, required obviously to convert those input formats into something κλειω would be able to process.

(Therefore already: Thaller 1983-1.) Not really trivial, between data sets that had been prepared on punched cards, where capitalization had to be encoded in upper case only data, and the today largely forgotten Earls Colne project, which used a convention of embedded markup inserted into transcribed full texts, which in my opinion had a much more consistent textual model than the TEI. This was one of the motivations that lead to the *standardization and exchange of machine-readable data* series of workshops (Thaller 1986-2, 1987-2, 1988-3). My reaction to that variance had been that different input formats were appropriate for different sources and purposes, so the solution should be to have a software system in the background that supported a consistent conceptual model of data as arising out of historical sources (κλειω) and a frontend able to map all sorts of encoding into the formats supporting that conceptual structure. – Which on a more theoretical level was supposed to help bridging the gap between approaches in Humanities Computing analyzing full texts and such which focused on data manually extracted and abstracted out of them (cf. Thaller 2017f [1991], in this volume).

This was supposed to help preparing machine-readable texts / machine-readable data which could be purchased routinely just like printed editions (HGRC 1-HGRC 18). Cheaply; but in any case without any previous organizational efforts. Which reflected my own experience that until very recently, and possibly still in some cases, what all the data archives / text archives / repositories for secondary analysis had in common was a dark secret: that they contained data, which were in principle available for secondary analysis, but where the depositor reserved the right to review (and possibly deny) the license to do so. Which is the reason that the other thing this type of institutions had in common was the veil of secrecy over the number of data sets actually re-used. Rather few people being courageous enough to submit a detailed enquiry to a researcher for the license to reuse such data, particularly when the contact information available was a ten year old mailing address.

The mechanism for the distribution of such data sets to be embedded into a group of series of publications: series B containing the technical documentation (HGR B1, HGR B5, HGR B11, but also HGR A1, HGR A7, HGR A23), series A “*Quellenkunden*” closing a gap in the literature available. Each of the volumes in this series should give a detailed introduction, what types of analysis would be useful for a specific type of sources – church registers, letters, testaments, inventories at the time of marriage etc. – and how to implement them with κλειω (HGR A2, HGR A3, HGR A4, HGR A5, HGR A8, HGR A9, HGR A29).

These series were published as “*Halbgraue Reihe*.” The name was a kind of practical joke: these were supposed to be cheaply available booklets which could actually be afforded by graduate students, the SAGE university papers being the model. This clearly was not possible with any of the large publishing houses in Germany so it had to be published as “gray literature.” You nevertheless should be able to get it via regular bookshops: So it was *halbgrau* “semi gray”, as it was distributed as commission by a specialized publisher, the SCRIPTA MERCANTURAE Verlag. The results were quite substantial: series A being repurposed later for reasons described below, series A and B together produced 38 volumes, roughly 7600 pages, published between 1988 and 1997.

The project started in 1988, with an extremely young team, Martin Gierl, Thomas Grotum, Kathrin Homann, Wolfgang Levermann, Thomas Werner; at the begin-

ning also supported by Peter Becker and, seconded from the City Archive Regensburg, but very closely connected to the team, Jürgen Nemitz.

The next three years saw intensive work. The software development progressed, the connected publication series grew. To support a more general view of the teaching available in the application of computer related methods and tools in the German speaking countries, a joint announcement of the summer schools at Cologne, Göttingen, and Salzburg was published. The Göttingen summer school, a two week introduction into the usage of κλειω, peaked at 120 participants in 1990.

Parallel to that, the conference series on standardization and exchange of machine-readable data brought me in touch with a much wider international community. Peter Denley and Deian Hopkin had issued in 1986 a call for papers for a small conference on “History and Computing,” which had such an unexpected response that it was repeated in 1987. These two conferences resulted in the foundation of the “Association for History and Computing” after the second of these two conferences (subsuming the community behind the “standardization and exchange” conferences). Personally, I was on the one hand elated at seeing how wide the field was, wider than I ever had imagined; on the other hand, an irritation grew which I had already felt earlier at ALLC/ACH conferences; at the conference series “data bases in the humanities and the social sciences” with which Joe Raben tried to counter what he considered an inadmissible narrowing of the definition of Computing and the Humanities by the ACH/ALLC of the eighties; and indeed even at the “standardization and exchange” series. What was the intellectual communality behind all of this?

There were, of course, answers available within sub communities: The hard core quantifiers considered statistical reasoning a self-explanatory methodological goal. Quite significantly, their main proponents were simply offended by the idea that the use of computers could be targeted at anything else than statistical analysis. Quite typical is the sentence *„La véritable hystérie de la ‚banque de données‘ est la simple traduction de cette naïve idéologie.“* Alain Guerreau’s damnation of the assumption that computational methods could have an analytic value, even if *not* targeted at statistical analysis, expressed in his review of the conference volume of the History and Computing conference of 1986.

For me the intellectual goal seemed to be quite clear as expressed in my concepts of source oriented data processing and the inherent fuzziness of historical data introduced above. So a criticism like Guerreau’s did not convince me, as in my view it ignored the usefulness of computer technology for all too many of the problems to be solved, before statistical analysis had the quality of data it required (Thaller 2017j [1995], in this volume). But to solve all problems for the application of procedures with truly analytic power, a lot of progress in the application of well understood forms of computing would be necessary. *Progress*, not repetition. That is, the contributions of 1997 should be able to build upon solutions derived from what we learned at the conferences of 1987. And this is what increasingly irritated me: I heard and read many interesting communications. But I learned little from them that would have made it easier for future projects to apply more sophisticated computational methods than the ones presented at the current conference or workshop. This, at the same time, also triggered my strong involvement in teaching, both at universities as well as at summer schools: Only if the fundamentals already well

understood were covered, before people engaged in their new projects, we would be able to see more sophistication in the future.

Starting to work with computers when computer science was very much in its infancy, and software technology in particular was rather a craft than a science, self-taught computer people were rather the rule than the exception. At the same time I still had remnants of the classical perspective of a *Historische Hilfswissenschaft* as a concept of the necessity, to solve systematically general problems of handling the peculiarities of historical sources, which were more fundamental than the clarification of questions arising out of individual content driven studies. (That the *Hilfswissenschaften*, or, rather: the way in which they presented themselves to the community, may at that stage already have been fossilized beyond recovery is not the point here.) So my vision of the relationships between history and computer science was simple: As historical sources were more complex than the data structures administered by available software allowed, there was a clear need for a specialist who developed software which was up to these complexities. And if no appropriate data structures existed, they needed to be invented. An approach for which I coined the term *Historische Fachinformatik*, “Historical Computer Science” in my English publications.

The irritations I felt at the content of the conference series I started from, a few paragraphs above, arose from the fact that I saw very little of that vision – “Historical Computing,” with a very few exceptions, seemingly consisting of the application of whatever computer science had developed to some sort of interesting topic in history. And progress consisted in the hope that whatsoever computer science would have developed in 1997 would be more advanced than what existed in 1987. Whether it was more appropriate, one could only hope for.

This contradiction was quite central for me: until my retirement, my main seminar at the University at Cologne had the general title “Humanities Computer Science v. Humanities Computing”.

During the late eighties, besides the κλειω project, I had the feeling that this vision of a tiny subdiscipline, somewhere between a software-technology-focused computer science and a method focused understanding of historical research, needed to be represented more strongly internationally if that dependency on what computer people might come up with, should be ended. For me this was an intellectual agenda, but as I had to use κλειω as an example, I am afraid that my attempts to argue for a more robust and self-assured position of the historical disciplines v. mainstream computer technology (cf. Thaller 2017e [1991], in this volume) came across for many people just as a concern for a pet project – all the more so, as the relational data bases, which I argued against as unable to support the data structures needed for historical research, undoubtedly *were* part of the main stream of computer science. So my argument for a stronger involvement with computer science, may have been seen by some people as a lack of understanding of what the state of the art of computer science actually was. I admit that during the last few years, I have been gnashing my teeth quite frequently, when projects in the Digital Humanities discover *now* from the NoSQL movement that graph based data bases have quite a few advantages – but I am getting ahead of myself again.

Back to my attempts to explain my point of view at the end of the eighties. 1987 I had been an enthusiastic supporter of the foundation of the Association for History

and Computing, so I found it natural to draw the next of its conferences to Germany. Still thinking that it would be a good idea to create a broad view of disciplinary relationships, I allied myself with Heinrich Best and Ekkehard Mochmann, to organize – under the name of “Cologne Computer Conference” – an umbrella for events from three conference series (Thaller 1991E): the third conference of the AHC, the next conference in Joe Raben’s “data bases in the humanities and the social sciences” series and the 1988 instance of the series of “International Federation of Data Organizations for the Social Sciences” annual conferences. With more than 500 participants and more than 200 presentations this was one of the largest conferences in this specific interdisciplinary domain until quite recently, but my motive in creating a platform where hard core quantifiers and other types of computer using historians and adepts of historical social science inspired research could recognize what they had in common more or less failed. In any case, it strengthened my links with the association, as did my presence at many of the events organized by members and annual conferences, frequently accompanied by members of the κλειω project organizing project sessions.

As a result of this engagement, I was elected as president of the international association for the term 1992-1994, cooperating closely with the secretaries-general, Peter Denley first, Leen Breure later. The association and I have to be grateful for their unceasing efforts, to create an international association, which had a unified international face, but completely different forms of organization in individual countries. I certainly very much like to remember the enthusiasm of those years, at which the conferences of the Association of History and Computing drew larger audiences than those of the ACH/ALLC. And I am extremely grateful, to the many people beyond Peter Denley and Leen Breure I met and had intensive contacts with, before and during my term as a president.

As president – besides being invited to give more or less inspiring talks at many regional and national conferences – I was mainly involved with the question, how an association could become more substantial than a framework for the organization of annual international conferences. Which I perceived as a necessity, if the more aggressive kind of interdisciplinarity I envisaged as Historical (later: Humanities) Computer Science should ever become realistic. The vehicle for this became the “workshops” of the AHC. Many people complained – and many still do – that conferences are so big that it becomes rather difficult to discuss themes in detail; on the other hand small workshops appear frequently productive for the participants, but find it hard to make their results known to the wider public. The proposed solution was to form subgroups of members interested in specific topics, which would meet between the annual conferences at various locations in workshops organized by different conveners, work in depth on a specific topic and report their results at the annual conferences in dedicated sessions, which were guaranteed, independent of any reviews. Furthermore, there would be a guarantee for these workshops that their results would be published as books of about 200 pages, available easily – usually contained within the conference fee – at the next annual conference. The enthusiasm arising from the response the group received at the conference hopefully carrying it forward to continue to work on the subject also during the next year. The books of about 200 pages could be guaranteed, as I decided to reorient the series A of the *Halbgraue Reihe* I described above towards that purpose.

Did the solution work? Well, yes, sort of. Seventeen such workshop volumes were published between 1991 and 1996, dealing with widely varied topics – modeling of historical data, curricular issues, OCR, statistical analysis of occupations, historical cartography and many others (HGR A11, HGR A12, HGR A13, HGR A14, HGR A15, HGR A17, HGR A18, HGR A19, HGR A20, HGR A21, HGR A24, HGR A25, HGR A26, HGR A27, HGR A28, HGR A30). What failed completely was the hope that this solution would become self-supporting: as long as it was actively “marketed” by me it worked. But it never developed into a stage where the workshops would continue without somebody acting as a series editor in the background – politely nagging would be conveners, quite besides the usual editorial joys of getting contributions in time for the books to be ready at the time of the conference.

Still, if my only motivation would have been to strengthen the Association for History and Computing, I would have been most happy with the results and certainly have continued to support the workshop concept when my term as president ended. The real disappointment, however, was that my motivation for the engagement in the association – to create a platform where a more ambitious approach towards the interdisciplinary field between computer science and the historical disciplines could evolve (Thaller 1992-4, 1995-1 in Thaller 1995E) – basically failed. I am aware that the following sentences are extremely harsh, but need to be written, if I try to remain honest. A problem which has plagued almost all attempts to organize interdisciplinary activities within Humanities (or Historical) Computing earlier and Digital Humanities today, is that only a very small number of the participants in the conferences are really interested in interdisciplinarity as such. Many more seek a platform to present results, which get their meaning from within the disciplines that provide the content. Which for me raises a question. If you have gained results within a discipline, which you are interested in because of their significance for that discipline and not because they have a methodological value beyond these disciplinary confines: why do you not simply present them at a conference of that discipline? Over the forty years of my involvement, there have been extremely few Humanities Computing or Digital Humanities events with more than twenty participants, where a general discussion did not include the statement: “The most important thing we have to do is to convince the (insert your favorite Humanities discipline) establishment that our results are important for them.” Could it be that what the speakers really wanted to say is “My discipline does not take me serious. Please help.”? For some strange reason, I have been invited to give many papers to audiences, where I very clearly said that I did NOT consider myself as a member of the discipline present, but as a specialist in interdisciplinary work, steadily more pronouncedly as a specialist of Humanities Computer Science the older I got, and most of the time I felt being taken serious by most of the audience.

Be that as it may: personally, after investing quite a bit of work into the organization of an interdisciplinary community, I found it easier to convince funding bodies directly that an interdisciplinary project I planned merited to be funded, even if it had a heavy proportion of technical components, than I found it to convince allegedly interdisciplinary communities that a joint effort at systematic methodological discussion was worthwhile. So my involvement in the self-organization of such communities almost stopped after the middle of the nineties.

Parallel to this involvement with the international interdisciplinary community, my involvement with computer science in a number of ways became much more intensive. Not so much institutionally: while I most of the time found polite interest when talking to computer scientists, software technology at the time, as mentioned before, was at least as much of a craft than a science. And with the need to implement a large software system to realize the concepts I described as source oriented computing, my reading of computer science texts was highly selective and there were only a few cherished excursions into the core literature of the field. And very few of my early publications have ever been visible to computer – or even information – science (Thaller 1982-3, 1989-3).

This changed a bit with two projects, both in some way related to IBM. Alarmed by the problems created by the rapidly reduced importance of the classical mainframes in the early years of the move to PCs on the one hand, UNIX based workstations on the other, IBM started a big marketing drive into the academic market at the end of the eighties. Being visible as IT aware within the Humanities, I was approached in 1988 by IBM for a “project”. This was quite important for me, though I actually misunderstood the situation completely. What IBM wanted were pretty pictures in a glossy brochure dedicated to the wonderful things a highly respected institution like a Max-Planck-Institute did with the help of IBM machines.

Today I understand that, had I proposed a “project” where the institute received a couple of their most recent PCs, pledged to employ them in worldwide correspondence and were willing to produce pretty photographs for the eight or sixteen page color brochure, my primary contact at IBM would have been perfectly happy. I on the other hand lived in the blessed believe that “the industry” was actually taking us serious and asked for the than brand new 200 MB WORM drive from IBM, for which I proposed to create a software library allowing to produce a full text indexing system which would allow rapid access to large amounts of text on the WORM, providing for “virtual updates”, which meant that some parts of the index files could be kept changeable on the hard disk, to allow for “personalized” versions. While this specific medium never became used widely enough to be significant, the implementation – supported by IBM by a grant in the low five digit range – lead me to a re-invention of a suffix-tree-like data structure which stayed useful for a long time (Thaller 1989-2). Personally, while I somehow managed to overlook the existing literature on suffix trees (or possibly did not understand its significance for my problem), this forced me to look much deeper into the lower levels of file handling and system software.

Both, the contact to IBM, as well as the closer understanding of lower programming levels, proved very useful when another situation arose at the very end of the eighties. I have mentioned that the very first IT work for which I ever have been paid, related to the creation of an image data base for the Institute for Medieval Material Culture (*Institut für Mittelalterliche Realienkunde*) of the Austrian Academy of Science, which found a thesaurus developed for them by IBM too unwieldy. Nevertheless their relationship with IBM continued in a loose way. My cooperation with the institute in general and Gerhard Jaritz in particular has continued ever since, so they were very well aware that in the meantime I used a workstation from IBM under AIX as my primary platform. One day I was informed that IBM Austria

had proposed a project to them to use one of their work stations to support their image archive – this time not as a database of descriptive terms, but as a database containing the images themselves and using the existing descriptions to access them. For that purpose, they would make available a system called “Image Assistant,” developed at the then IBM Scientific Center at Winchester. This system offered a library of C routines for the actual image handling, offering a wide variety of operations for image enhancement. Would I be able to extend κλειω to support a data type “image”, based on that? I would and I did. Of which, admittedly, I am still quite vain. The system within 12 months or so, offered the first version of a fully graphic user interface for an image enhancement system, which not only supported batch operations, but which also intimately connected it with the database administering the descriptions of the images together with the images, which could be stored on a variety of devices, including a CD jukebox. The images could be selected by a complex query operation and then submitted to batch processing of the selected images, by a macro built out of filtering primitives which could be built and tested interactively.

The reason I am very vain of that is that at that time that meant that the first version of the whole interface had to be implemented in the naked X system, only a second one embracing the luxuries of Xt and only a third one the ultimate in user friendliness, the Motif widget set. In hindsight I understand, remembering all the difficulties encountered, why the people at Winchester never went beyond a command line tool that allowed applying the filters of Image Assistant to images addressed by their file names. Which is probably a pity for IBM: within the infighting between various branches of IBM, Image Assistant never achieved “product” status, which means the world never really became aware of what at the time may have been one of the best libraries of image analytic functions prepared for general usage. While the implementation of that graphics interface gave a big boost to a number of important lines I worked on, as described below, it had one serious disadvantage: Image Assistant was decidedly proprietary. Κλειω’s traditional components at that time ran on six or seven brands of Unix plus MS DOS Windows (plus partial ports to a number of more exotic operating systems). The image components worked on AIX. After a couple of highly frustrating attempts to get access to other proprietary image processing libraries for other brands of Unix (Linux only slowly appearing at the time), I started studying technical image processing in earnest and simply created a call compatible re-implementation of Image Assistant, where success was defined as the ability to produce the bit-identical version of the result of a filter operation done beforehand with Image Assistant. This worked on all brands of Unix – and after some time even under MS Windows, though the necessity to install the Motif Widgets on the Windows PC made that a somewhat tricky operation.

This experience and endeavor had three important effects for the further development of my work.

(a) My philosophy of source oriented data processing had so far assumed that there was an initial stage of purely human work, where the historian transcribed sources (or created an intellectual description of an image or object). This should result in a representation of the original source, which represented that source as closely as possible and minimized the interpretation of the source taking place. The

following stages of historical work should be supported by various tools: The data base engine responsible for assembling the chunks of information needed; the various “logical objects” providing tools to relieve the historian from routine work, like decoding some fanciful way of expressing a calendar date. The methodological importance of image processing for me was the vision that the always somewhat doubtful concept of “minimizing the interpretation during transcription” could be made much more robust.

The first step should now be a purely mechanical reproduction, provided by scanning, which would put the source directly onto the screen, where all transcriptions could then be checked against interpretation errors directly against the virtual original. And a whole new set of visual tools could enhance that first level of interpretation of the signs during transcription: quite early a set of tools to administer repertoires of character forms typical for a specific scribe was available.

Already before the possibilities thus opened by image processing, I had used the concept of a functional replacement of printed editions by digital data sets (Thaller 1988-6). This was the idea behind the series “C” of the *Halbgraue Reihe*, the sets of floppy disks containing machine readable data prepared for secondary analysis. But with reproductions of a manuscript right on the screen available together with a transcriptions ... or possibly reproductions of various manuscript witnesses of one text ... the idea of a fully digital edition, as a complete functional replacement of printed editions in the core domain of editorial scholarship, the edition of full texts, rather than administrative sources for type of analysis mainly based within historical social research, was quite natural, in my opinion. Which is why in the fall of 1989 I arranged for an AIX work station to be carried to the *Quantkurs*, the Salzburg summer school and offered a seminar on “dynamic editions”, possibly the first attempt for looking systematically at the notion of a digital edition which was not destined to prepare, but to replace a printed one.

(b) “Arranging for an AIX work station to be brought to the summer school” pointed at the second important side effect of this move to the handling of images. As mentioned, access to Winchester’s “Image Assistant” had been arranged by IBM Austria for the benefit of my Austrian friends at the Academy of Science, or, rather, its institute in Krems. This made it easy for me to get access to equipment for public presentations which was extremely rare at Humanities facilities at the time, but which IBM tried to push into the Academic market quite determinedly. Not only in Austria, but worldwide; particularly in Germany, where my name was still known in academic marketing because of my involvement with the early WORMs.

As a result, I received perfect support with advanced IBM equipment for public demonstrations, 23” screens together with the required workstation, often transported for me to the site of the public presentation. This allowed me to present the image facilities to large audiences at many places over large parts of Europe. A portfolio crystallized out of the various experiments: enhancement of deleted portions of a medieval charter; extractions of shapes of objects from a medieval painting; detailed connections of zones of an image with a structured description (known as “annotations” nowadays) and a few more. In hindsight I wonder, how much of that I actually got across: even on a 23” screen not all that much can be recognized, when you put it on a podium in front of 200 or 300 people. (We are talking about the very early nineties: data projectors for screen resolutions well beyond a megapixel

were not invented yet.) In any case, it gave me and my projects considerable visibility (Thaller 1992-3, 1994-3, 1994-4, 1997-1, 1998-2).

(c) The third major effect of this incursion into graphic user interfaces on the one hand and bit level programming on the other was that my understanding of the architecture of complex programming systems became much deeper, particularly by studying the X Windows libraries and the emerging layers of toolkits based upon them. This looped back into my previous experiences with engineering a data base upon the assumption that historical data might indeed be something significantly different from the kind of data from the hard sciences – and particularly from engineering – which had very much influenced the development of computer science. That had led to a data model for κλειω which was built upon a graph; even if hierarchies were so prevalent in the kind of sources most central to social history that they were much favored by some aspects of that model. But that was a data model, which was still responsible for ordering the relationships between atomic chunks of information. Some of them deviated from the traditional data types – calendar dates being fuzzified intervals, numbers as well – but texts still were completely traditional strings. If one looked at the way, a text where different substrings carried different visual attributes was handled in X Windows; was the classical string not naive in the first place? The text of an edition, where the text was actually transmitted over time as a set of variants was a good example for the limitations of the naive concept that a conceptual string was simply an array. Might it be that well below the notion of fuzzifiable data models, we also needed a more complex understanding of what a text as such should be represented as? A graph instead of a character string, e.g., this graph being on another level of abstraction different from the graph model of the data base itself? (Thaller 2017h [1993], in this volume.)

These three perspectives together looked very bright. The more so, as in the meantime the interest in κλειω clearly went beyond the German speaking countries. While in the late eighties, I had – besides the summer schools in Göttingen and Salzburg – classes at a number of German universities where I taught hands on introductions in the use of the system, in the early nineties I also had the honor to teach the use of the system abroad, most importantly at the than Westfield College, supported by Peter Denley, at Utrecht, supported by Leen Breure and at Odense, supported by Hans Jørgen Marker.

Peter Denley's interest was so substantial that he became the most important driving force behind the "English Version of Kleio" project, for which a consortium of institutions, including the University of Southampton, Queen Mary and Westfield College, London, and the Institute of Historical Research, have agreed to support the development of an English version of Kleio, a package for source oriented data processing developed at the Max-Planck-Institute für Geschichte, Göttingen. (Support including first and foremost the organization of funding for a professional translation of the technical manual (HGR B11) as well as a genuinely new English language tutorial.) This new English tutorial (HGR A23) was written by Matthew Woollard and Peter Denley, a bit influenced by an MA course I taught in the spring term 1993 as a visiting professorial fellow at than Queen Mary and Westfield, Queen Mary having integrated Westfield a bit earlier (and later renamed itself into simple "Queen Mary" again).

This also created an intensive contact to Frank and Jean Colson, then at the University of Southampton, which continued well beyond this specific project. But my gratitude in that period belongs to Peter Denley, first and foremost. Not only for organizing a very memorably term at Queen Mary and Westfield, but being a wonderful connection into the not-social history parts of the historical disciplines, where his engagement was perfect proof that even at that time formal approaches could become useful far beyond the social history province of the discipline.

In 1978 the decision had been made to implement the command language of CLIO's forerunner in Latin, not to vulnerate linguistic sensibilities of Austria's more conservative medievalists. Now 15 years later, as part of the "English Version of Kleio" project, the Latin command language was replaced by an English one, as this seemed to be a precondition for the more international setting. As relatively early all language related parts of κλειω were exported into text files from which they were dynamically loaded, providing a new surface language (and English as well as German feedback by error and other messages) was quite simple. A small tool, which translated existing command scripts from Latin into English, was easily written: that it was baptized "Boadicea" may provide amusement to some readers.

So, in 1993 the outlook for κλειω as the practical implementation of my theoretical model of "source oriented data processing" looked extremely bright and opened a great future for it? Well, not really. The project funded by Volkswagen foundation in 1988 had ended after three years and, as it was a "special case", outside of all of their funding programs, a continuation of the project was impossible – that they found possibilities to add a small amount of money at the end, which allowed the continued employment of one of the original project group, Thomas Grotum, and a new student assistant, Gabriele Gross, for some time was a very big concession already. But this did not change the basic situation that what had been a multi-person project group, which was supposed to be responsible for nothing but the κλειω project, collapsed again to myself and such funding as I could acquire from various sources. Or, more precisely: as much funding, granted for content driven projects, which I could redirect at support activities for the software system which was supposed to support those content oriented goals.

As far as the development of the system went, not very much seemed to change. Κλειω had always been programmed by me; that with the end of the dedicated funding STANFEP could not really be developed further, to allow the processing of a wide variety of input formats, was unfortunate, but not central. But what broke away was on the one hand the man power for systematic testing of new features and much more important yet than that, the possibility to document such new features and to provide help-desk-type support to the user community.

So the next years saw a shift from a dedicated development project to a whole bundle of projects, which focused on the co-operation with other institutions. This worked up to a degree: I described above, how significant the image processing facilities have been for the development of my conceptions. That the manual describing them was authored by Gerhard Jaritz in Krems (HGR A22), shows clearly that the notion of "cooperative development" worked; but it can also be seen as an indication, how difficult support for the system had become, when even such a central feature could not be documented in Göttingen itself anymore.

The strongest support for κλειω on a practical level always had come from Ingo Kropač and his Department for *Historische Fachinformatik und Dokumentation* at the University of Graz. At the start of our co-operation Ursula Leiter-Köhrer had been employed there, who was the only person who contributed software functionality, which still is part of the core of the κλειω system. But in the end Ingo Kropač faced a very similar problem to my own: software development for historical research as such found it very difficult – increasingly difficult – to get funded, unless connected with projects demonstrating the immediate usefulness of the developments. And this demonstration of the immediate usefulness of the developments ate up so many resources that the developments which had taken place, could scarcely be brought into a form, documented and supported, which made them really useful for others.

As I know today, this is also the generic reason, why the World Wide Web is a veritable graveyard of projects funded by the European Union, which had submitted wonderful sustainability plans which nobody bothered about anymore the day after the final evaluation of the project had been completed. Which is also the way almost all of the projects went, where joined funding has allowed a computer science department to develop solutions for a content partner. But both of these traditions were not yet known to me in 1991.

Encouraged by the results of the co-operation with Ingo Kropač and Gerhard Jaritz I developed the notion of the “Historical Workstation Project” (Thaller 1991-2, 1991-3, 1992-1) as an organizational framework to keep my implementations alive. The concept here has been that a generalized system – κλειω – could provide access to data bases, which were close enough to the original that they could be analyzed without recourse to specific documentation for the individual data bases. And various contributions from other developers could fit into the overall software environment, to enhance the overall capabilities. The most successful example had been a re-implementation of the Latin lemmatization, which had been developed under the direction of Andrea Bozzi at the *Istituto Linguistica Computazionale* in Pisa in the early eighties. That had been ported by me from PL/1 into C and connected to κλειω, where it was an optional enhancement for the full text retrieval subsystem. (κλειω in principle would today be classified as a system for semi-structured text; as the length of the text in a “field” was virtually unlimited, however, could hold the complete text of a novel, leave alone a shorter text like a letter or medieval charter, it had also acquired a menu driven full text facility early on.) But the hope that out of such co-operations a truly distributed development could be arrived at, failed.

In some cases that clearly came from misunderstandings by me. I had been a driving force behind the acquisition of one of the first Kurzweil Data Entry Machines (KDEM) in Germany in 1984 at the computing center in Göttingen. (I had written a hardware review of the KDEM in 1983 (Thaller 1983-2, 1983-4).) Padre Busa honored me by his interest in my work from a very early stage onwards. In 1986 he invited me to a two day lecture series to the *Università Cattolica* in Milano, where I was flattered by him acting as translator of my lectures for two full days. Sometime during these two days we discussed the possibilities of the KDEM at some detail. A discussion he remembered, when a few years later he planned to process the texts of Ignatius of Loyola according to a similar model as the ground-

breaking Thomas of Aquinas project. We rather soon agreed that I would supervise a few students paid by the *Societas Jesu* to make the complete works of Ignatius machine-readable; Padre Busa then took care of having them proof read by the nuns at some monastery in Milano. Within the *Societas Jesu* he recruited support for a project to publish the works of Ignatius as an electronic full text edition, most visibly at a big conference at the *Universita Gregoriana* in Roma in 1992, big enough to attract the attendance of the then Italian Prime Minister Giulio Andreotti. The only Prime Minister I had the honor to address directly with one of my presentations. Personally, I would not have chosen Giulio Andreotti.

After having worked with Padre Busa for the digitization, I was also invited to prepare a full text system for the distribution of the texts. And here we managed to misunderstand us completely. I had dreams of cooperation with some bright young Jesuits studying computer science at one of the Jesuit universities, in broadening the support base for the historical workstation concept; he – and particularly his supporters in the *Societas Jesu* – basically wanted a software provider, to be paid for a final product. As a result, while I spent two weeks at the Institute of Jesuit Sources at St. Louis, Missouri, to implement a prototype for a textual retrieval system, hoping to get into touch with my dreamed of partner implementer, my prototype was never really used. The data were published afterwards on electronic media of the time by some mainline delivery system, which simply ignored most of the linguistic encoding Padre Busa's other co-operators had added during pre-processing, which had been the focus of my efforts.

I would like to stress that I am not complaining about my partners in that case, least of all Padre Busa himself. Whatsoever you read about his mixture of charm and intellectual brilliance is an understatement. I took this as an example of a number of cases, where I tried to get development partnerships off ground, while my potential partners simply wanted a software developer selling things to them, because: if this model would not even work with Padre Busa involved, with whom could it have worked?

So this time in the early nineties was a strange experience. On the one hand, my international reputation grew – while I spent the spring term of 1993 as visiting professorial fellow in London, winter term of the same year saw me as visiting professor at the *Istituto Universitario Europeo* in the hills above Firenze, and my exposure as president of the international Association for History and Computing brought me invited lectures all over the place. All of this created a steady flow of visitors to Göttingen, where I had many stimulating discussions about what the purpose of using computers in history and the Humanities should be. The most prominent one being Willard McCarty, with whom I had a truly inspiring talk, when he visited Göttingen as part of a fact finding mission on the state of the art. These were also the years, where the *Halbgraue Reihe*, which I mentioned earlier, flourished.

But at the same time, my base at the Max-Planck-Institut für Geschichte in Göttingen at the very least did not get strengthened, but rather eroded, when Rudolf Vierhaus retired from the directorship. While he supported my work, his successor respected it. The quote I remember most vividly from my conversations with the director of the medieval department of the institute, which at that time was responsible for administrative decisions, was: “This is a wonderful idea; but you know, we

have no money.” Which in a sense put some seeds into the ground: learning how to attract third party funding became very useful ten years later.

Κλειώ during those years underwent a development, which proved irrecoverable later: it acquired an ever increasing number of features and became optimized more and more, the underlying engineering model being implemented clearer and cleaner. But there was simply no time to document it; so the many new features were introduced mainly for the benefit of individual partners, or individual projects, and useless for the rest of the world. That was augmented by a heavy strategic error. When project funding for κλειώ petered out in 1991, there was still a strong user community, going back to the summer schools from 1988-1991 and the projects that sprang from their participants. From the user group existing at the time came the strong recommendation, to augment the scripting language with which the system was driven, by a menu oriented user interface. With some doubts, I invested a lot of time – and almost a year of a student assistant – to implement and document such an interface (HGR B10). Which almost never was used afterwards – except by one project described below. I should either have stuck to perfecting the data base engine and the scripting language or invest into the existing graphical user interface, which sprang from the image analysis modules, accepting that for a few years it would not be accessible to the majority of the users. Using the menus, without understanding the underlying structures was simply impossible, unless you used them for an extremely well defined purpose – so they did not really make the software more accessible. But the last bits of funding dedicated unconditionally to system development were spent on them.

So many of my attempts to broaden the concept of “data bases” to include services usually connected with other areas of research were not really applied by larger numbers of users (Thaller 1990-2, 1994-2, 1996-1, 2000-2). And, even worse: there actually have been developed two software libraries by Wolfgang Levermann, which in some way already point to the contextual programming considerations, which I will point to in the very last chapter of this text, which I never managed to integrate into κλειώ (HGR B4, HGR B8).

I mentioned that during that time I started to learn how to attract third party funding. Three projects have to be mentioned, which very much shaped the last years of my time at Göttingen.

One evening during the *Quantkurs* at Salzburg in 1990 I was approached by a young Polish historian, Jan Parcer, who was working at the archive of the memorial museum at the former concentration camp in Auschwitz, Oświęcim now. The changes of the political situation in 1990 had made contacts across the borders infinitely easier – and lead to the discovery that during the economic crises during the last decades of the regime, the memorial had fallen far behind Western museum standards, and even the substance of the buildings had deteriorated alarmingly. Gerhard Botz, who already in the difficult years at the end of the *Ancien Régime* had regular contacts to the memorial, had invited some of the younger staff of it to all three sections of the summer school. Jan Parcer enquired whether I could imagine using a database to administer the content of the archive of the memorial, describing the situation of the memorial and its holdings. If historians from Germany and Austria are approached about this specific part of their historical heritage and do not feel an implicit urge to provide assistance, as far as possible within their

historical specialty, something must be wrong with them. Beyond this automatic predisposition I was intrigued by the situation Jan Parcer described: at a time when heritage institutions in Germany permanently had to worry about staff, Oświęcim had enormous funding problems for absolutely everything *but* staff; and Jan considered it quite feasible that a working group of six people or so could be dedicated to the handling of a systematic conversion of the holdings of the archive into digital form provided hardware would materialize somehow. In the tradition of the Salzburg summer school, the initial conversation took place at some beer cellar, so it was a napkin on which what I considered a useful hardware configuration for such a project was drawn. Software not being a cost factor, as that was supposed to be κλειω.

Many projects have been discussed in the beer cellars of Salzburg, but Auschwitz had to be a special case, for which I felt immediately a personal responsibility. So I started to write letters and received positive replies specifically from *Volkswagen*, where the director of the board always had a strong feeling of responsibility towards this part of the past: the company has a long tradition of sending every year some of the apprentices to a seminar in Oświęcim, where they visit the memorial and are taught about the history of the concentration camp. Together with some other donations, mainly channeled via the *Stifterverband*, sufficiently much money was raised to bring a state of the art AIX server, together with six or eight terminals and periphery as required to the memorial. (As I do not plan on visiting the USA in the foreseeable future anyway, I may disclose now that at the time this was strictly speaking illegal – the ban on hardware of that class to be exported to the Warsaw pact countries, as being of strategic military value, being raised only a few months later.)

While I personally spent a few weeks or months at Oświęcim over the next few years, setting up the software and optimizing it for the usage intended, the major part of day-to-day practical co-operation was handled by Thomas Grotum, a member of the staff of the original κλειω project, who in the meantime had graduated in contemporary history.

What the group in Oświęcim accomplished over the next few years, with the support of Thomas Grotum, was one of the most impressive performances I have ever been privileged to watch. Starting virtually at zero, Jan Parcer geared the working group up to a speed, where within a few years he produced a whole set of “memorial books” starting with a publication on the Sinti and Roma in 1993 and creating a whole rush of other documentations in 1995. My respect for all the Polish colleagues concerned is great.

Unfortunately he, and they, have been far too successful. Jan Parcer’s group had been created from scratch as part of the archive and never included under the umbrella of the “research division” of the memorial. This created jealousies. Even worse: he himself presented the results of the working group to the national and international community. With such success that sometime in 1996 he was explicitly prohibited to accept an invitation to present one of his later publications to the (Polish) pope at Rome, who already had received one of the earlier memorial books, the Polish press praising Jan Parcer for this, rather than the directors of the memorial. Another invitation was simply irresistible to a Polish researcher, so he presented the memorial book in Rome nevertheless; the directors of the memorial

responded by firing him and disbanding his working group for disobedience. The fallout of these events was sufficiently severe that unfortunately even a much wider project was killed by it, which in the meantime had been funded by the *Volkswagenstiftung*, despite the name completely independent from the car company which was so important at the beginning. That project would have tried to connect the archives of all of the major memorials in Poland; but after the explosion at Oświęcim this project more or less dissolved in distrust – and had lost with Jan Parcer the Polish coordinator needed desperately.

That German voices would be raised critically to the way in which the Polish memorial at Oświęcim acted, was impossible then and is not really possible today. Let it suffice to say that while my feeling of responsibility towards this specific part for the German/Austrian history is undiminished, my appreciation of what all too appropriately has been called the “memorial industry” remains severely limited after experiencing its operation in that case.

This project arose out of the development and application of the κλειω software. Out of my other main activity at the beginning of the nineties, the involvement with the Association for History and Computing, another major “project” arose. While the Oświęcim project in hindsight always leaves a bad taste, this, fortunately is one I am really unreservedly happy about.

In 1990 Wilhelm Heinz Schröder asked me, whether I would be willing to accompany him and Heinrich Best to Russia on a journey to make contacts with Russian colleagues. This was an initiative of Ivan Dmitrievich Kovalchenko whom Wilhelm Schröder knew from various encounters in international groups connected with quantitative methods in history. This two weeks trip through Russia in the most intensive phase of the change of the political and economic system produced many interesting and some strange experiences; but lead, most of all, to a personal encounter with Leonid I. Borodkin, director of a laboratory for quantitative methods at the Lomonosov Moscow State University. While at that time in the west high end PCs were becoming ubiquitous at history departments, resulting in a flood of pie charts of doubtful statistical and even more doubtful intellectual value, I saw in this laboratory a micro-computer that was clearly stone age even then, employed to calculate analytic values for economic history after a fuzzy logic model. I was truly impressed and thought that it would be of great advantage to both communities in academia, the Eastern as well as the Western one, if the possibilities of modern equipment would become available for the former and the seriousness of methodological reflection to the later.

The first step in that direction consisted in a workshop for representatives of research groups of universities in the just recently dissolved Soviet Union in January 1992 in Salzburg, under the auspices of the Association for History and Computing, which tried to create a framework for future co-operation (HGR A15, HGR A 21). This is not a story of the AHC, but a retelling of my personal experiences, so I have to leave out the details of the very strong reaction from the membership, providing equipment for Russian universities and support grants for the participation in conferences in the West. I myself convinced the *Stifterverband* to fund the beginning of a series of autumn schools at the Lomonosov University in Moscow from 1992-1996, the much larger later instances funded by *Volkswagenstiftung*. These events made use of the informal networks in the former Soviet Union, where most of the

important researchers in the field of quantitative history had been trained at Lomonosov and knew each other quite well. In 1996, therefore, this autumn school drew 80 participants from 30 institutions spread over four of the countries of the former Soviet Union. The content taught at the summer school consisted of various aspects of computer technology made accessible by then modern PCs, taught by a changing team of honorary Western teachers, Peter Doorn, Jan Oldervoll and Kevin Schürer being most frequently appearing besides myself, while quite a few others appeared at individual years.

These autumn schools were embedded into other activities, which at Göttingen were administered by Wolfgang Levermann, who, holding a doctorate in the History of Russia, had originally started at the institute as part of the κλειω team and then became responsible for the – IT heavy – final stages of the Dahlmann Waitz bibliography, at that time already running for decades. With the Russian projects, he was particularly responsible for a publishing program I had managed to convince *Volkswagenstiftung* to fund: our Russian colleagues in the early nineties quite definitely wanted to change the way in which history would be taught at Russian universities, but this was difficult if only the old text books would be available. So we reached an agreement, by which *Volkswagenstiftung* funded State of the Art printing equipment for a Russian archive, which on the other side provided all other costs for using this equipment to produce a series of ca. ten text books, under the scientific advice of Leonid Borodkin and his group.

Activities which were strengthened in 1995/96 by a joined project funded within the TEMPUS framework of the European Commission which connected the quantitative methods departments at the universities of Baku, Azerbaijan (Irada Rafizade), Barnaul, Russia (Vladimir Vladimirov), Dnepropetrovsk, Ukraine (Vitaly V. Podgaetsky), Minsk, Belarus (Vladimir N. Sidortzov), Moscow, Russia (Leonid Borodkin) and St. Petersburg, Russia (Sergei G. Kaschenko) with the *Istituto Universitario Europeo* at Firenze, Italy (Michael Goerke), the Max Planck Institute at Göttingen, the *Institut für Mittelalterliche Realienkunde* at Krems, Austria (Gerhard Jaritz), *Rijksuniversiteit* Leiden, the Netherlands (Peter Doorn), and *Universiteit* Utrecht (Leen Breure). Within pairs of Eastern and Western partners the application of specific information technologies were experimentally introduced in the laboratories for quantitative methods participating.

In parallel to these activities, I remember fondly a summer school in Tartu, Estonia, taught together with Gerhard Jaritz, which brought me in first contact with Raivo Ruusalepp, now a prominent representative of Estonia in many archival projects, which was also later continued by an (in that case: extremely small) project.

I am happy to say that the network which was behind the Moscow summer schools supported the creation of national branches of the Association for History and Computing in the countries of the former Soviet Union which are still going strong today, when the international Association has ceased functioning (HGR A15, HGR A21). It is astonishing, how much effect ridiculously small amounts of money can have, if they are spent within the context of networks like this, where all participants have a personal interest in success. Of all the activities I had to give up when I moved first to Bergen in Norway and then to Cologne in Germany, as described in the next chapters, this is the one I am most sorry that I could not continue.

My final activities in Göttingen were mainly dedicated to the “Duderstadt project” (Thaller 1999E, 1999-2), funded, again, by *Volkswagenstiftung* and running between spring 1996 and spring 1999. Originally, it was simply intended to digitize systematically an interconnected group of archival materials, covering the period 1397-1650 in the city archive of the small city of Duderstadt, near to Göttingen. Most central were the *Amtsbücher*, the central administrative series, within which the *Rechnungsbücher* – account books – form an unbroken series. Altogether these sources constituted roughly 77.000 pages. Originally, these were supposed to be digitized and made locally available – I had at that time some experience with CD jukeboxes (already Thaller 2017h [1995], in this volume). I will always be very grateful to Stefan Aumann, who convinced me to make the material available in the developing world wide web, which let the visibility of the project explode and laid the groundwork for many of my later projects. He worked for the project with me at the Max-Planck-Institute; Hans-Heinrich Ebeling, then director of the city archive of Duderstadt and Hans-Reinhard Fricke forming the archival part of the team. At some time in 1997 this was, for a short period, the largest collection of archival material available in the internet, even larger than the Library of Congress’ “Washington Papers” have been at that stage. And most people, who remember that project, simply understand it as that, a pioneering digitization project for archival material.

It was a pioneering digitization project for archival material, showing that since the project of the *Archivo General de Indias* in 1992 was attempted as a visionary trial out of the range of normal institutions, systematic digitization of whole archival series could now, five years later, be implemented within a budget in the lower six digit range at a humble city archive. And opening up completely new levels of accessibility, thanks to the WWW. This was the effect the project had; but this result was not the original vision of the project.

My original idea had been to show that between the publication of raw archival material and the creation of scholarly editions, a continuum existed (see already Thaller 1992-2). Digitized archival material was not only more easily accessible than the original: with the help of image enhancement it could actually show potentially more than the originals (unless you wanted to apply destructive techniques to them, as had been done at the end of the 19th century in some cases). And: you could potentially connect transcriptions to the images, extract the repertoires of individual scribes from them (tools for which existed in κλειῶ already) and even connect them to critical editions. To prove the later, a very small subproject existed, in which Malte Rehbein created a prototype for a system which should administer a critical edition on the screen in such a way that it was the choice of the reader – not the privilege of the editor – to decide, which witnesses were to be given the greatest weight.

It was clear that this way to handle archival material was new, and would therefore be looked at doubtfully by parts of the medievalist community of historians. To reduce that skepticism, the project plan proposed a series of three workshops during the duration, where the results should be discussed with the communities of archivists and historians. At the first of these workshops these three goals – archival digitization, tools to enhance the accessibility of the material by image processing technologies, possibilities for editorial paradigms – were presented as roughly

equally important. But it turned out that the interest of the participants, the number of which grew rapidly between the years, was more than anything else raised by the implications of the simple possibility to have the archival stuff accessible, immediately, right at your own desk. Taking up the response of the possible user communities, therefore, the other original goals were emphasized less and less during the second and third year of the project and the accompanying workshops.

That the immediate accessibility of archival material in digital form at your own desk was a fascinating prospect, which at the very least merited dedicated discussions, was almost unanimous – and so the Duderstadt project was remembered simply as a digitization project. The consensus was *almost* unanimous: I remember one archivist participant at workshop one or two who got up and said in no uncertain terms that he did not know, whether this was a good idea or not, but one thing he was certain about was that *if* this was a good idea, it should be left for the *Archivschule* Marburg (the national institution for the training of archivists in Germany) and plain historians should definitely not dare to involve themselves in the archivists domain. Some echoes of that will appear in the chapter describing the time after my return to Germany.

The Duderstadt project was actually finished only after I had moved to Bergen in Norway. The reason for that move was at the end of the day an act of capitulation. As mentioned above, my experiences in the nineties had been extremely contradictory: I enjoyed, both within the historical community proper, leave alone the community of computer using historians and the Humanities' computing world, great visibility. But while I was able to attract funding for many worthwhile things, the core of my academic interest, research on computational methods appropriate to the nature of historical sources as a field in and by itself, was almost impossible.

I have mentioned above that while Rudolf Vierhaus as director of the Max-Planck-Institute supported my work, his successor, Hartmut Lehmann, respected it. The problem may be illustrated by a story which breaks the chronology, as it happened in 2000, after my return to Germany unto the chair at the University at Cologne. During the last ten years or so, my research at the Max-Planck-Institute was described in the annual reports under a separate main heading, *Historische Fachinformatik*, the term I had coined to describe my field. In 2000 an emissary of Hartmut Lehmann approached me and told me that since I went to Norway a member of the scientific advisory board had constantly complained that it was very unfortunate that the activities in *Historische Fachinformatik* had ceased after I left. That emissary himself was nominally responsible for the continuation of these activities. To quiet this nagging voice of the scientific advisory board, I was, via that emissary, offered to get a special position on the advisory board, “but I should not be afraid of the work load, if I never would turn up, it would not be a problem”. My talent as a fig leave being restricted, I did not accept this exciting proposal. Unfortunately this had a negative effect. That emissary afterwards refused all offers by a number of people to take over the administration of the heritage of the Duderstadt server, still running quietly on the server hardware purchased in early 1996. When that hardware broke after fifteen years of faithful service, the project disappeared from the internet. (Actually a few years *after* the Max-Planck-Institute for History had ceased to exist. Said emissary still working at the successor institute.) This hap-

pened much later, but it probably exemplifies an atmosphere in which I found myself less than enthusiastically supported in the middle of the nineties.

So I was willing to draw the line and embark on something completely new.

4. Fourth Life: Research Administration

My move to Bergen was clearly understood by me to redraw radically my own profile. While I had some hope to continue my own research agenda, it was clear that that would be a secondary concern and from now on I would mainly have to coordinate work done by others, hoping of course to be able to influence the priorities this work would have to take – and possibly motivate projects which worked towards the overall goal of having a closer connection between fundamentally new software designs and history; or rather: the Humanities, as I also accepted that my responsibility in Bergen would not be one to history, but to all disciplines of the Humanities.

This radical break had been prepared by a long relationship which arose originally from my agenda in the late eighties. Since 1987 I had met and cooperated with Jan Oldervoll, originally from Bergen, then working for some time at the university of Tromsø, before he moved back again to Bergen. Jan Oldervoll was one of the very few people, who also had started his own career by interdisciplinary work in which he himself produced quite complex software, though he always remained much closer to content oriented work than I did. We met with increasing frequency in the context of the Association for History and Computing, as described above, including the organization of workshops and the activities around the Moscow summer schools. After participating in the Salzburg summer school once or twice, Jan Oldervoll even was willing to organize a continuation of the quantitative / computationally oriented part of it when Gerhard Botz moved from Salzburg to Vienna. Jan had spent some time in Göttingen with a fellowship from the *Ruhrigas Stiftung*, where we programmed along side by side on our different software systems with many exchanges (HGR B9).

After a conference in Luxemburg in 1995 on the usage of computers in the history classroom, where we had the barely hidden agenda of trying to connect that community closer to the Association of History and Computing, he and his wife, who also worked at the Historical Institute of the University in Bergen, gave me a lift partway home. During that journey they asked me whether I would be willing to teach in a part time position at the University in Bergen. More precisely as *førsteamanuensis II*. Translating “*førsteamanuensis*” is simple – it is usually translated as associate professor. The “II” is more complicated, as I know no other university system with that feature. Formally it is a 20% position. Within the Norwegian university system it is officially used – or was, at least, at the time – to connect lecturers to a university, who either have primary positions in the industry or come from abroad and should be attracted for some lectures to Norway. There is considerable discretion, how these positions are interpreted: some are truly working appointments, some are almost completely honorary, to attract a big name. In my case, I was quite willing to put serious effort into it, as I did not have at that time the possibility in Göttingen to supervise graduate work, leave alone doctoral stu-

dents. So I accepted a position as *førsteamanuensis* II in “Historical Computer Science”.

I am happy to remember the *hovedfag* (roughly: the Norwegian pre-Bologna master degree) thesis of Anette Skogseth Clausen at the end of the nineties, dealing with possibilities of digital technologies for the access to / edition of archival resources on Bergen from the City Archive in Lübeck. At some stage I hoped to extend that model and build, by successive theses, an interconnected system between archival holdings related to the history of Bergen and various harbor cities around the historical trade networks around the Northern and Baltic seas. But this involvement in graduate studies gives a distorted picture; in some ways teaching took me back to my own roots in history: a seminar (*mellomfag*) on the year 1848 in Europe being the one drawing the highest number of students during my teaching time at Bergen, where the computational components were rather slim, indeed.

My duties in Bergen were rather unexpectedly light – I seem to have been the only one who did not understand that these 20% positions are usually mainly honorary in nature. So while spending there quite some time, often staying as guest at the home of the Oldervolls, my impact on the students as well as on my teaching profile remained rather low. In any case, I was present there and Jan dropped my name, when the faculty found itself with an organizational problem in 1997.

Bergen had since 1972 been the home of the “Norwegian Computing Center for the Humanities”, funded by the national funding agency for research in Norway, the Norwegian Research Council, *Norges Forskningsråd*. This was an institution which in the seventies reflected the technical need to rely on centralized computing centers for any serious work, so having a national center was probably just seen as a logical step. In any case that center was supposed to develop its own research profile, but at the same time also to act as a service unit, supporting research at other Norwegian universities. While the center was visible very much, particularly internationally, with the spread of the notion that computer usage in the Humanities could take place at individual desks – and more prosaically, because other universities thought, they could spend the money just as well as Bergen – support for the national center dropped and national funding ceased in 1992. So the University in Bergen inherited a national center, which had this double role of acting as a service unit, as well as a research institution in its own right. Inherited that is, the unit, but not its funding.

Bergen also, independently of the Computing Center for the Humanities, was the home of one of the most significant early projects for digital editions “*Wittgenstein's Nachlass* – The Bergen Electronic Edition”, organized by Claus Huitfeldt who founded in 1990 the “Wittgenstein Archives” at the University in Bergen as an organizational framework for the edition, which he kept alive by ingenious efforts at funding. It is truly admirable how he managed to do so, alongside the intellectual effort required to manage the edition, inventing a very advanced encoding system for the texts and implementing software for its support.

In 1980 Norway encountered a disaster in her off shore oil industry, where at one of the offshore installations, the Alexander L. Kielland platform, an accident killed 123 people. Among causes for the high number of victims communication problems were suspected, including the problem that the technical documentation was almost completely in English, because a Norwegian terminology for the oil industry simply did not exist. As a result in 1981 the Norwegian term bank, *Norsk*

termbank, was created: a research unit with 20 staff at the peak dedicated to a thorough attempt to create a consistent Norwegian terminology for the oil industry. This institution was attached to the University in Bergen.

In 1997 all three units had arrived at a crisis: the Norwegian Computing Center for the Humanities found it difficult, to redefine its role as a service unit for the university; the Wittgenstein archive faced the problems of moving from an ongoing project, working on an electronic edition, towards a permanent institution maintaining the results; the term bank simply had accomplished what it had been created for and looked for a solution which would allow the continued existence of a center of lexicographic and linguistic competence, with a strong reliance on computing technologies. The university – or rather: its Humanities' faculty – noticed that all three units had a much closer relationship to information technology as usual in the Humanities, so the idea sprung up, to combine these three units into a joined program, the Humanities Information Technology Research Center, HIT center for short. As tradition was heavy in all three cases, the faculty looked for an external specialist in Humanities Computing to analyze the situation from the outside and create a blueprint for a solution.

Jan Oldervoll asked me in 1997 whether I would be willing to take that task on and act as preliminary director to oversee the beginning of the merger; I accepted and started in September 1997 with a six month contract to solve the task described above. My proposals were accepted as a strategy for creating such a center; as its formation took longer than expected, I was invited to serve as preliminary director somewhat longer; I applied for the position of founding director, connected with the permanent position as a professor at the faculty; and was hired, effective March 1st 1999.

The details of that merger would be of greater interest for a history of the field at the University in Bergen. For my biography the main effects were more connected to the overall scope of that experience, rather than to the result of individual decisions.

- 1) While in Göttingen I had broad interest in *all* fields of Humanities Computing, it was still an interest focused on history. There I looked at examples of the applications of information technologies in other disciplines as inspirations for historians. Bergen presented me with a view point, where the communalities of the problems of applying IT to the disciplines of the Humanities clearly outweighed the specifics raised by the subset of the historical fields of enquiry.
- 2) My involvement with strategic planning, both my own on behalf of the HIT center, as well as the national one I had to relate the Bergen plans to, very much impressed upon me the problems of large scale and long range planning within research funding.
- 3) Norway's academic community is extremely oriented towards international contacts. That can occasionally quite literally lead to situations, where colleagues from the same university, which would see no reason to co-operate at their home institutions, are brought together in workshops and conferences a few thousand miles from home. A side effect of this extraordinarily strong focus on international contacts is a superb infrastructure for the acquisition of project funding from the European commission. I am not quite sure, whether

this is profitable from a funding point of view for Norway. But it provided an excellent introduction for me into the chances and pitfalls of European funding.

Looking back at the never ending discussions on how to solve the underlying problem of the HIT center – how to negotiate permanent funding out of project oriented funding concepts – and at the frequent travels required to make the center as visible internationally, as it was expected by the outlook described, it is almost surprising that somewhere between these administrative tasks there remains also a set of activities which related directly to my own intellectual agenda and which in hindsight create a connecting line between the time before and after Bergen. Most surprising that even κλειώ continued to play a role.

One could roughly say that these activities connected my relationship to XML and my interests in the digitization of manuscripts.

My relationship to the TEI had always been strenuous at best. While not being at the Poughkeepsie meeting, I was present at the Chicago meeting of 1989, representing the Association for History and Computing. I probably was a somewhat queer guest, as I may have been the only one present, who was familiar with information technology and *not* quietly assuming that SGML was all we needed, pointing instead to the need of a more precise understanding of what problems Humanities' texts really posed, before a technical solution was chosen. But in any case, I still was invited to join the Text Representation Committee of the TEI, where I participated in the meeting in Oxford in April 1990. I was still not talking about SGML, but rather about textual properties and at one stage protested against the discussions about specific tags for specific parts of books, not texts. Should we not rather address specific properties of historical texts? Many of them contained parts, where a deterministic interpretation was not possible, as, e.g., ambiguous or plainly unknown abbreviations. To which a very distinguished TEI specialist present replied, oh well that couple of abbreviations in sixteenth century texts we all know pretty well that is no problem. I resigned from the committee after that meeting, quoting the heavy work load of the Auschwitz project as an excuse.

While that was a reason for serious mistrust against the textual model behind the endeavor, for my own plans in the nineties the TEI was simply straightforwardly detrimental: The promise of this early days was that if you only encoded your texts according to a TEI based upon SGML, you never needed to bother about writing software any more. SGML being an industry standard, industry would do the rest. So, world, please stop developing software for the Humanities!

Nevertheless, the homepage of the TEI acknowledges that in 1999 the TEI consortium was founded by the universities of Virginia and Bergen. Bergen doing so, because the acting director of the HIT center had convinced the steering board of the center that this was a meaningful strategic involvement, which could strengthen Bergen's international reputation and visibility. I mentioned that one of the things I learned in Bergen was that you had occasionally to accept and actively support solutions on the strategic level, even if you did definitely not believe in them otherwise.

Which had a strange echo, quite some time after Bergen. Despite - or possibly: because - it was known that I doubted (and doubt) the wisdom of providing an encoding scheme which is not underpinned by a conceptual model, I was elected as

a member of the technical council of the TEI for 2008 and 2009. I had thought I was invited to discuss the conceptual foundations; as it turned out, the technical council's role consisted of discussing the usage of individual elements and tags. I probably should have tried more energetically to bring the discussion to what I considered more central. As it is, I basically listened quietly to two years' worth of phone conferences on tag usage, without ever really raising my voice. This paragraph as a belated apology to the people who had voted for me during the elections.

But back to Bergen. The one area where I hoped to bring my own research interests into play was the field of manuscript digitization. One of my earliest efforts at drawing funding to the center consisted in an application for European funding of a "Workshop for the Creation of a Server for Illuminated Humanist Manuscripts", which was granted and took place in Ravenna in 1998. The workshop, with ca. 40 participants, brought together representatives of most institutions I had been related with by my work on the processing of images and manuscripts with a group of Italian libraries including, besides the major local libraries organized by the *Rete Bibliotecaria di Romagna* (i.e. the *Istituzione Biblioteca Classense* of Ravenna and the *Biblioteca Malatestiana* of Cesena), the *Biblioteca Nazionale Centrale* of Florence (BNCF). These groups were augmented by representatives of some publishers and library and heritage institutions, 20 institutions spread over most of Europe. The plan was to create a model for the systematic digital publication of illuminated manuscripts, with an emphasis on the 15th century and Renaissance, assuming that the transfer of manuscripts from Constantinople after 1453 to Italy was a particular fitting subject for a European project, as these manuscripts are frequently cited as a root for the spread of the Renaissance over all of the continent.

My plan as director of the HIT center was simple: Such a project would be perfectly suitable to put Bergen strongly onto the map as a major European center for the digital handling of manuscripts. On the map of Europe as well as on the Norwegian map of institutions fulfilling the national need to project an image of excellence abroad. Personally I could there have a project, which was much closer to my own interests than most of what I had to supervise.

That these specific libraries were selected, was the responsibility of a very good friend, with whom I had the privilege of long discussions, virtually over decades: I cannot recall at which conference in the eighties I first met Dino Buzzetti of the *Università di Bologna*, but he has been one of the most inspiring partners in the discussion of what constituted appropriate information structure for Humanities' texts I had the pleasure to meet. He visited me for prolonged stays in Göttingen, Bergen and Cologne and we met at an enormous number of conferences all over Europe. This workshop was just one of extremely many occasions our paths crossed. It was probably the most promising avenue to create the possibility of a real co-operation between producers of Humanities' related software. Dino Buzzetti for years always attempted to create contacts, where my efforts could be combined with that of other persons or working groups towards an implementation of broader systems.

The workshop was most promising in that respect, as out of it a major funding proposal for a European project was developed. In hindsight I am almost relieved that it did *not* get funded. At that time I had no experience yet, which extraordinary amount of resources is needed, not only to direct such a project, but do it according

to the concepts of the European Commission, so the resources assigned to coordination would have been completely inadequate. A smaller project evolved, however: A project not targeted at academia, but at the labor market. The ECHT (European Cultural Heritage Technician) project assumed that the application of information technology to the cultural heritage would create an opportunity for new job profiles, also for people without full academic training. To develop and test training schedules for such technicians, the three libraries mentioned would join forces with a number of institutions in Europe providing training for the labor market and HIT center would provide the technical expertise necessary for that. It was an ill-fated project. Sometime during the decision stage, the service provider on which the European commission relied for the administration of this funding line was closed down by the police for suspicions of fraud. After learning that we would be funded, we were later asked, whether we still had copies of the complete communications with the commission, as their copies were inaccessible, being kept closed by the police. But even so, we should start with the work immediately, simply using the budget the partners had pledged as their own contribution; which all the small units providing job training could definitely not afford – and risk – to do. And when things finally went underway, I already had left Bergen.

While not many of our partners delivered anything tangible under these conditions, some digitization work actually *was* carried out at the participating libraries. What was more important for me was that I became seriously interested in the new XML technologies, as whatsoever I might think about the TEI, it was completely clear that this was a technology, which would dominate the future of information technology in the cultural heritage domain. So at HIT center an “XML Tutorial for European Cultural Heritage Technicians” was produced by Vemund Olstad; and privately I started to think seriously about the possibility to process XML by κλειω.

As it turned out, this private interest was extremely important for the next stages in my career. My wife, Sigrid Amedick, had accepted a position as director of the library of the Max-Planck-Institute for legal history in Frankfurt in 1999. There she “inherited” a huge digitization project, which was in the middle of a digitization campaign eventually reaching 4.316 volumes with 1.351.223 pages – and only an extremely sketchy concept on how that stuff would eventually be made available in the internet. The structural metadata, the tables of content that is, were encoded according to E-bind, a very early XML standard, later replaced by METS in the library community. Bibliographic metadata existed in the format used internally by Sisis, a system used for OPACs, today marketed by OCLC. While some vague agreements existed that the library of the University of Göttingen would provide expertise, and it was clear that the institute, like all Max-Planck-Institutes could use the services of the Max-Planck-co-owned computing center also in Göttingen, it was neither clear, how the raw images of the digitization process could be converted into the JPEGs required for the internet, nor was it clear, how the two types of metadata would be combined and turned into a library server.

Already in the context of the Auschwitz and Duderstadt projects I had accepted that to use κλειω it would become necessary to import data from the native formats of other database systems. In the case of these two projects, this meant that I created import facilities for the than fairly ubiquitous dBase database files; much more important in the long run was that during that process I had created a general

framework into which import facilities for other formats could be inserted quite easily. And I even considered this as a conceptual challenge. If the *internal* data structures supported by κλειω were duly “general” in the sense this *epitheton* is coveted by computer scientists, it should be possible to import data existing in almost any sort of *external* data structure. E.g. data encoded according to a specific XML format or data available from a software system administering OPACs.

So, during the final months of my time at Bergen, while I worked as research manager during the day, my evenings were spent as a software engineer again, who implemented in κλειω import facilities for XML – with a possibility to describe the format in the data definition language of the system – and for Sisis. And created some features allowing κλειω to function as a CGI server. All three tasks turned out to be accomplished astonishingly easily: the internal data structure *was* general indeed and the command language was structured in a way, in which it could easily be used to formulate scripts stored on a server. Taken together with the capabilities of κλειω for batch processing of images that meant that I could set up from the evenings at my desk in Norway, what in 2000 was the first (or possibly: second) collection of digitized prints in Germany breaking the 1 million page ceiling.

Some remarks above already indicated that my time in Bergen was short indeed. At about the same time in the fall of 1997 when I accepted the invitation to provide a plan for the HIT center and act as temporary director, I had also submitted an application for a new professorship in *Historisch-Kulturwissenschaftliche Informationsverarbeitung* in the University at Cologne. Some explanations on this term which is almost as incomprehensible to a native speaker of German, than it is to somebody without any German, follow below. Let’s just say for the time being that it is some way of referring to Humanities Computing. I did not hear anything about this, until the end of April 1998, when I just prepared to take up my second contract as temporary director. Then a note arrived that I had been shortlisted and was expected in Cologne ten days later for the usual presentation to and interview with the committee of the faculty. I presented. More silence. Almost exactly one year later, at the end of April 1999, a few months after I had started at my permanent position in Bergen, I got a phone call, inviting me very urgently to submit some personal details, which the ministry of research needed, to process the offer of the faculty to accept me on the chair just mentioned.

I accepted. Why? The first and foremost reason was simply that I had spent at least ten years of claiming that the interdisciplinary field between the Humanities and Computer Science *did* constitute a distinct academic field and needed institutional support. Turning down the first professorial position in at least the German speaking world which had been created for it was psychologically almost impossible.

The second reason was simply that I had realized in the meantime that any illusion I had about continuing my own research profile alongside with the job of keeping an institution alive based on precarious funding, was exactly that: an illusion. I actively wanted to change my orientation towards research management when I went from Göttingen to Bergen, but I had to realize that that was strictly a one way street: two years more of what I did and I was an academic administrator for the rest of my life, no chance getting back into the trenches of research, ever.

And so I went to the Rhine.

5. Fifth Life: Professor

5.1. Teaching

When describing the process which drew me to Norway, I mentioned my great interest in becoming able to supervise students' theses. While I lost track of the number of universities at which I gave courses in various subfields of Humanities Computing, almost all of them were voluntary; which at the end of the day means that there was always an element of negotiation how much effort beyond the "I just want to get a rough idea what all this is about" could really be expected. Which severely restricted the level at which things could be taught, and never allowed me to take any pre-existing knowledge for granted. Only if I had a say in what students would be taught in term one, I could expect to progress in term two to such things as I considered truly interesting – which would turn the students into truly competent partners to discuss the broader issues at stake between the Humanities and Computer Science.

People with that competence are needed in my opinion for two reasons. The loftier one: interdisciplinarity by most people is considered to be a social relationship between *two* people, while my approach requires it to be a personality trait of *one* researcher. Interdisciplinarity happens within the head.

The more emotional one: over the years, discussions about Computing and the Humanities, Digital Humanities or whatsoever other label people come up with, increasingly gave and give me the feeling to be part of a remake of *Groundhog Day*. How many statements which complained that non-traditional publications were not appreciated by the community I have listened to, I cannot remember. Too many, by far. And to list all the topics which came and come up with this regularity, year by year, decade by decade, perceivably day by day, without any noticeable progress over time, would exhaust the reader's patience.

But this might just be accepted as one more nuisance which happens in academia (or in life as such, actually). Much more severe is that many of the advances in the interdisciplinary literature also give you this Groundhog Day feeling. I mentioned above that in the very first issue of the "system bulletin" with which I started to establish my reputation, there was a discussion of how calendar dates as they appeared in historical sources could and should be supported by software. I am not sure, how frequently I have heard solutions of the very same problem described in the meantime – be it by excited newbies at conferences, describing that solution in the software system they started using a year ago, be it by grave professors of computer science, who explain what they did for their co-operation partners. And almost none of them ever has heard that somebody else worked on those issues before.

One of the Groundhog Day remake topics at Digital Humanities conferences is, of course, exactly the topic I am touching – how and what should be taught in "Digital Humanities courses"? As the ebb and flow of academic interest in computer applications in the Humanities went, the number of courses taught as part of standard degrees is almost limitless. The number of degree programs, particular MA programs, promising explicit interdisciplinary training in the area, is smaller, but I

still would not be surprised, if there have been a thousand since it all began, world-wide. So, why do we still have this Groundhog Day remake discussions, then?

The problem in my opinion is that almost all of these programs and courses are invented with the intention of convincing the students addressed, and even more the grantees of the local university that by supporting this, they are catapulting their training or their institution to the very cutting edge of academic development. If we had thousand degree programs so far, at least hundred have been announced as being the very first one world-wide. The problem with that is that they slightest hint that any element of this course or program might be less than the ultimate possible, damages the picture presented to the local students or grantees. Therefore, it was seemingly impossible, to agree upon the slightest forms of a binding consensus on quality criteria for History and Computing courses in the early nineties and equally impossible, to create such a consensus for Digital Humanities degrees in the German speaking countries between 2010 and 2015. If there is no consensus on the minimal requirements of such courses and degrees, however, there is no downward limit to these requirements. I'm still quoting the class offered in the eighties in Germany on "Computer Science for German Studies: Wordstar 2000" as an example how ridiculous things can get. This is the reason why Computer Scientists find it so often extremely hard to take Digital Humanities scholars serious. And this is why such "cutting edge" courses and programs usually collapse within 24 months, when the students notice that what they have been taught by those enthusiastic young teachers at the university is actually standard knowledge of many of their untaught drinking buddies.

Finally the possibility to teach students in a way, where they could reach a truly professional level!

And so I went to the Rhine.

It might look like, as if such enthusiasm could only lead to disappointment. Surprisingly enough, the situation I found at Cologne was rather more promising in *some* ways, than I ever had expected. The convoluted denomination of the professorship *Historisch-Kulturwissenschaftliche Informationsverarbeitung* did hide an astonishingly strong local involvement. Cutting the prehistory short, in the nineties the University at Cologne had decided to create one professorship in, essentially, computational linguistics, specifically for a local pioneer, Jürgen Rolshoven. Computational linguistics in the nineties was still sorting itself out as an academic discipline in Germany, so all sorts of names were used for the existing professorships, particularly at universities with only one of them. Why Jürgen Rolshoven decided upon *Sprachliche Informationsverarbeitung* (linguistic information processing) I do not know; in any case his concept is decidedly linguistic, Chomskyan essentially, and bears absolutely no relationship to library / information science which the term *Informationsverarbeitung* evokes in the ear of many German listeners, and most attempts at its translation in the ear of their English speaking counterparts.

One professorship in a subject is a nice thing to have for a faculty, but it is not enough, if you really want to implement a degree program. Therefore sometime in the late nineties the philosophical faculty decided to establish an academic subject (and full degree course, leading up to the doctorate) *Informationsverarbeitung* which was to be implemented by two professorships. Jürgen Rolshoven's for the linguistic part of the Humanities, another one for the rest. "Professor for the rest" is

a denomination which would be even more incomprehensible than *Historisch-Kulturwissenschaftliche Informationsverarbeitung* so the later term (literally: “information processing in historical cultural studies”) was chosen.

As described above, the process which led to my recruitment was very protracted, so when I arrived, a course plan already existed (Thaller 2005-2). Germany at that time still had the pre-Bologna *Magister* degree, with no true equivalence of a Bachelor existing. A *Magister* was required to study three subjects – say history, English studies and philosophy – where the third one was more or less an appendix, but the first two carried almost equal weight, except at the very end the *Magisterarbeit* (thesis) deciding what was the main discipline. In that system *Informationsverarbeitung* was introduced as an academic subject on equal rights with the big traditional disciplines like history, English studies, archaeology etc. This indeed fulfilled exactly my belief and intention, to study the problems of applying computer science to the whole breadth of the Humanities, the peculiarities of the sources they analyse having so much in common that from a computational point of view they outweigh the differences of the research questions arising in and constituting the agenda of the various disciplines.

The study plan started with a rather thorough introduction into the basics of computer technology both with regards to hardware and software, insisted on strictly enforced training in practical programming, which may go beyond what some computer science degrees demand, included a general introduction into the main stream technologies of the day and asked for a general introduction into computational linguistics. This introductory level would prepare students for the main seminars, where they could combine the knowledge they had acquired in their second subject – history, English studies, archaeology – with their computational skills into truly interdisciplinary exercises, leading up to a thesis which explored either the application of information technology to a specific problem arising from one of the Humanities’ disciplines or dig deeper into the question which modification of underlying computational technologies might be required for their application to the Humanities in general.

This framework I found at Cologne when I came there, and I fully endorse and support it, right until today.

Things have changed with the Bologna process in many details, but the general framework is still valid. The details of that transformation are not relevant for this description of my years in Cologne; let it suffice to say that of course also Cologne suffers from the way in which the BA / MA system was implemented in Germany in the spirit of “let’s do what we always did, and change as little as feasible” which runs particularly strong in Cologne, which always had, and in many ways still has, one of the most conservative philosophical faculties of the country. So the wisdom of “provide bachelors with a broad education, which qualifies them for a broad range of specializing master courses” has never been understood and led to the same plethora of overly specialized bachelor degrees preparing for exactly one master degree, which plagues the post-Bologna universities of many European countries.

Before or after Bologna: the two professorships in Cologne were supposed to cooperate in the provision of the courses for the concept. This was possible; but was hampered by the fact that Jürgen Rolshoven and I have very different methodologi-

cal outlooks. He assumes that linguistics is the backbone of any application of computational methods in the Humanities; I assume that linguistics is one of many fields, which is useful for some Humanities' disciplines and completely pointless to others. Training in formal languages is extremely useful, if you analyse a grammar, presumably useful for many types of philological phenomena, but not really helpful, if you want to support art historical explorations of images algorithmically. (Particularly if it has taken up so much time in your training that you never heard anything about other media than texts.)

This meant that, fully endorsing the general model described, I had to find ways and means to extend it. One of the chosen ways was simply the installation of classes, which provided some parallels of the basic training modules, replacing one emphasis by another. But this was not always sufficient. So I became instrumental in the installation of a degree course in media studies. There the students were offered a specialization in *Medieninformatik* (computer science for media studies) which covered roughly 40% of their academic effort, alternatives within media studies being specializations like psychology of the media, media law and so on. This allowed me to teach in these courses visual programming at a level, which simply could not be forced into the curriculum of *Informationsverarbeitung*. Structurally it basically meant: try to reuse as many modules from the framework of *Informationsverarbeitung* as possible and augment it with specialized classes as needed.

On the master level, this concept of "try to create modules in such a way that they can be used in different contexts" was extended. For some years we participated in a European master degree taught jointly by the universities of Coimbra, Portugal; Cologne; Graz, Austria; Lecce, Italy; Turku, Finland. The concept behind this European degree has been that each of these universities would provide a master degree, which was providing a specific specialization, but taught in such a way that the second term would have to be spent at one of the other universities, participating in courses reflecting the specialization of that one. This term abroad was prepared by classes taught via the internet. The supervision of a thesis springing from the content of the term abroad by teachers of that university was possible.

These extensions of the original concepts proved quite successful. Of the hundred new first year students which sat in the introductory lecture each winter term towards the end of my active time, more than sixty percent came from *Medieninformatik*, only the smaller rest from *Informationsverarbeitung*. That worked simply by the time hallowed academic principle of self-exploitation. While a professor in Cologne is expected to teach nine hours each week, I taught twenty and occasionally a bit more in the terms immediately before retirement. I have to admit that that implies the traditional concept of poodle didactics still running strong in German universities: when you leave a lecture hall with hundred students you shake things off, like a poodle leaving a pond, and turn to other matters. The notion that teaching a class means that you have a personal one-to-one interview with each of the students every week, as I observed it at Queen Mary and Westfield, never has taken root in the country.

Was it worthwhile? I still believe very strongly in the framework I described for Cologne. And if I want to sell it, I can always say that to the best of my knowledge none of the graduates is unemployed. The majority actually drifting into employ-

ment already during their final terms. And the record for the speed with which a master graduate topped my professorial salary by his is at something like three months after graduation. This clearly underlines that people are able to find professional employment on the base of a training in information technology acquired in a Humanities context. On the other hand, I could describe all of this as a failure: the vast majority of the graduates left the Humanities for employment in the industry, the major part of the minority went into libraries, computing centers at universities and similar provinces of the job market, only the rest of the rest getting into the IT support of the Humanities.

Which I have to qualify: “rest of the rest” is a quantitative, not a qualitative statement. That Reinhard Förtsch’s *Arachne* data base, certainly one of the most significant projects in its class in archaeology world-wide, was staffed predominantly by former students of mine is enough to do any teacher proud. And equally proud I am of some of the doctoral theses that have been completed during my time in Cologne: let me mention Sven Schlarb’s thesis on the usage of fuzzy methods for the classification of pre-historical objects (KBGF 1), the thesis of Patrick Gunia, designing and implementing a city builder, which is not build upon geometric but semantic descriptions of the buildings to be created¹. And never to forget about Jan Wieners, whose thesis implemented a gaming rule engine which is able to support completely different types of games, which otherwise have always been handled by separate, type specific, engines².

Nevertheless, I very much hope that my successor finds it possible to change this ratio between fields of employment and I am encouraged by promising stories I hear about the development. Much could probably have been done by me, if I had managed to invest more heavily in cooperative projects with colleagues from the other disciplines or in direct cooperation with cultural heritage institutions. I attempted such cooperation at the beginning of my time in Cologne, but the results were not very encouraging, as usually the time span between a seminar taught together with some institution and the time somebody was starting to write a *Magisterarbeit* following things up, was simply too long to keep the connection alive. The other reason, why graduates usually drifted away from the university or at least its Humanistic side, is a simple one. After congratulating a student of mine on his thesis, I asked whether he now would get more permanently connected to the academic department, where he had worked for years as a student assistant. I got the reply: “No, I’ve been treated as an idiot long enough.” The problem is related to the complex relationship between humanists who understand that they need computer support, but are afraid that they might be overwhelmed by the people providing that support, where insecurity leads to condescending behavior. Condescending behavior to a student from a traditional Humanities discipline will be accepted by that

¹ Gunia, Patrick: *Der Semantic Building Modeler – Ein System zur prozeduralen Erzeugung von 3D-Gebäudemodellen*. PhD thesis 2013, Universität zu Köln. <<http://nbn-resolving.de/urn:nbn:de:hbz:38-54410>> (accessibility checked: July 19th, 2017).

² Wieners, Jan Gerrit: *SpoookyJS. Ein multiagentenbasiertes JavaScript-Framework zur flexiblen Implementation digitaler browserbasierter Brettspiele und spielübergreifender künstlicher Intelligenz*. PhD thesis 2014, Universität zu Köln. <<http://nbn-resolving.de/urn:nbn:de:hbz:38-59711>> (accessibility checked: July 19th, 2017).

student as part of the price to be paid if you want to get into the profession. A student being aware that leaving the university will immediately raise the salary received finds it much harder to stomach.

A word of thanks is needed here. Many people have taken over *Lehraufträge* during my time in Cologne, teaching these degrees would not have been possible without them. (A *Lehrauftrag* basically means that you are scandalously underpaid, or not paid at all, to teach one or two classes a week. The model works, as teaching experience is an important part of an academic vita in Germany.) I can only express my gratitude to two of them, fortunately usually also employed within a project or on a related position, so they did not have to rely completely on what is more an offense, than a salary, the *Lehrauftragsvergütung*. Without the enthusiasm in teaching provided by Susanne Kurz and Jan Wieners for more than ten years each, it would not have been possible. Thank you.

5.2 Research

Teaching twenty hours, when you are required to teach nine, can be interpreted as a romantic example of complete loyalty to one's academic ideals. I've never claimed sainthood and, while there was a strong element of self-commitment in that engagement, there was also a more mundane reason for it: money. My funding situation in Cologne has been quite peculiar from the start and only during the very last years before my retirement I was able to draw significant university funding for teaching, as the assignment of certain funds was for the very first time connected to the number of students taught. Being interested in attracting students from the start, suddenly paid off and fed back into my willingness to accept as many students as possible, probably too many at the end.

The peculiarity of my funding situation goes back to the plans of the faculty in the nineties. As mentioned, at that time there was the idea of a subject / discipline of *Informationsverarbeitung* represented by two professors, for whom two assistants, rooms and a noticeable budget had been assigned. The one professorship already existed. The second was filled by the extremely protracted process which I described.

When I arrived in Cologne, at the very first time I talked to my future colleague, Jürgen Rolshoven, under four eyes, he told me in a grave voice that Cologne university was still run on a completely personal level, particularly all sorts of resources would be privately owned by individual professors and when I wanted to have access to resources of what kind whatsoever, I would have to get them all by myself. His resources were his. So there were two professors on equal terms, who were supposed to implement a study program on equal terms between them. One assisted by quite some office space, two assistants, a flock of student assistants and a sizable annual budget. The other by a one person office located at some other institute, as in the *Informationsverarbeitung* space nothing was available and something like 1700 German marks a term, for absolutely everything starting from pencils to useful things like a PC, even if there was some allowance for starting in the position. And absolutely no secretarial or library infrastructure of an institute on which I could draw.

Why do I speak about “the peculiarity of my funding situation” and do not simply say that it was disastrous? I could probably have started my time in Cologne by a protracted campaign for “fair funding”, which had all the potential of one of these faculty battles, in which the champions see themselves as epic heroes and everybody else oscillates between finding them hilarious and obnoxious.

I mentioned that part of my private research agenda in Bergen had culminated in a funding proposal to the European Commission for a project to develop a joined server for illuminated manuscripts. This failed; but it left me with a complete blueprint for such a project. When I had accepted the position in Cologne in the summer of 1999, I had activated some contact to a small group of students in Cologne, which I had met during one of the workshops of the Duderstadt project and asked them, whether they knew of potential objects for digitization projects in Cologne. They created a contact to the director of the *Erzbischöfliche Diözesan- und Dombibliothek*, Juan Antonio Cervelló-Margaleff, who was at that time responsible for a manuscript collection which still contains the complete core of the medieval library of the cathedral, roughly 400 codices, representing about 150.000 pages, including some of the most valuable illuminated manuscripts of Germany.

I had at that time some experience with the curators of museums and libraries, who were plagued by the double suspicion that on the one hand digitization would destroy their objects, while at the other somebody would earn millions from prints derived from the digital images in the internet, of which the institution holding the originals would not see a single cent. Therefore I came prepared for a long and difficult discussion. I almost did not believe my ears, when I described what a digitization of the manuscripts could do and Cervelló-Margaleff did simply answer “Yes, we have to do that.” instead of starting on the usual long list of worries.

So, still in Bergen, I wrote an application to the German National Research Council, known as *Deutsche Forschungsgemeinschaft* or DFG, for a project, which would digitize the complete collection at resolutions which are still considered extremely high today and present the material in the internet. There were relatively good chances for such an application, as the DFG had a funding line asking explicitly for pilots of digitization projects. That the notion to digitize a complete *collection* at very high resolutions was slightly unusual at the time, when the mainstream digitizers wrote quite decidedly about high end digitization as something which always would be restricted to a very small group of manuscript “treasures”, is best illustrated by a story told to me later by a member of the group reviewing that application. Another very prominent member had swung the opinion in favor of the project with the statement that I would never be able to carry it through, but my failure would at least clarify what the limits for that kind of digitization were.

While the administratively relevant letter arrived only a few weeks later, I was told about the positive decision of the funding committee on the phone in the week immediately before I took up my position in Cologne. So, to come back to the funding situation, I might be without a budget – but I was part of the small group at the faculty who had what at that time was considered a substantial project from third party funding; and I had the experience from both, Göttingen and Bergen, that acquiring external funding was a routine part of research. So getting the money I would need from external sources looked much more simple and attractive than getting into a mudslinging match at the faculty. I actually found the extremely poor

material support at the faculty so natural at that time that I did not even try to get additional office space for the staff to be employed in the illuminated manuscript project.

I have to be extremely grateful to Wilhelm Heinz Schröder. When I moved to Cologne he welcomed me more warmly than many of my future colleagues and immediately offered me office space at the *Zentrum für Historische Sozialforschung* for the staff of that project. They profited from it unreservedly, as those rooms were much better than almost all office space I could possibly have gotten within the university. Structurally I wonder today, whether it may have been a mistake to accept that invitation: it did hide for some time from the faculty how much was actually happening around its newest member.

It is overdue, however, to change an impression, which may have arisen by the above: I most certainly was made felt welcome by many colleagues at the faculty and I never could have achieved anything without their support. While a polite non-relationship with my institutionally closest colleague existed right through to my retirement, there is a long list of colleagues with whom I enjoyed very close connections. I have to be grateful for the hospitality of the historical institute during my early years. Henner v. Hesberg went out of his way to arrange for rooms to be renovated at the archaeological institute of the university to provide me there with the nucleus of office space I needed to finally have people working in my projects not only at walking, but at chatting distance. And out of this at the beginning purely topographical relationship grew the one with my closest colleague of the years in Cologne: Reinhard Förtsch, who helped me immensely to get integrated into an environment, where IT applications in the archaeological branch of the Humanities became a common concern. And became a very good friend besides, whom I'd not like to have missed.

Indeed, I cannot even start to list all of the colleagues, with whom at one stage or other common seminars or small scale projects have been realized. I *have* to mention two relationships however: Wolfgang Schmitz, as director of the library of the university went out of his way to participate in and support some of my early projects in the area of retrospective digitization of library holdings. And, while the Computing Center of the University always proved most helpful from the very beginning under its director Wolfgang Trier, his successor Ulrich Lang, who besides his directorship holds a chair for visualization and computer graphics at the institute of Computer Science, arranged for access of my students to his lectures in the late years, when the number of students exploded and the necessary teaching capacity could not have been provided otherwise.

Still, despite all of these valuable and enjoyable personal contacts: that a lot could happen materially, I owed to a rather successful history of attracting further external funding, beyond the very first cathedral project. As a footnote which is amusing in hindsight, I have to add that my funding base was not only dismal within the university, but also, initially with the research ministry of North-Rhine Westphalia, the state of Germany within which Cologne is located. When I negotiated about my professorship, I had interviews there, where I also discussed the technical equipment I would need. I mentioned that I hoped to train students in algorithmic image analysis and would need a hundred or two hundred thousand German marks to acquire a small park of work stations, if I should be able to do so. Great enthusi-

asm in the ministry, two or three officers getting excited and promising that after I got to Cologne I just should send in a two page summary of what I wanted to do and the money would be there. I got to Cologne and I sent a summary, even five pages instead of two. Nothing. When I called a few weeks later, the most enthusiastic participant of that ministerial group summed up the situation with the memorable sentence: “But we did not understand that you are in the wrong faculty!”. Funny in itself, but funnier because at the very same time the prime minister of North-Rhine Westphalia ran a campaign to improve with unorthodox means the qualifications in information technology of young people, rather than encouraging the immigration of foreign experts. (*Kinder statt Inder*, the pun getting lost in the translation “children instead of Indians”). Ever since, when I listen to another speech in which politicians emphasize how absolutely crucial interdisciplinary and unorthodox work is in academia, I fervently hope that they manage to convince their own ministerial staff to read the manuscript of their speech.

In any case: my endeavors to recruit external funding allowed me between 2001 and 2015 to spend roughly 300.000-350.000 € of third party funding annually. Before ca. 2012 that was balanced by a regular budget of just a few thousand € annually from the university, which in some years I actually forgot to spend. Depending on what level of funding you still consider substantial enough to count as a “project” there were twenty or thirty of them for which I received funding from external funding organizations. This was comfortable, but it meant that my research biography in Cologne was very much determined by what I could get funded, not by what I would have always considered to be the most meaningful next step. What in the end is more frustrating with this long line of projects is something else, however. Like the Duderstadt project, many of these projects had quite some visibility and impact on others. Unfortunately, another experience of the Duderstadt project repeated itself as well: the projects usually achieved visibility for other reasons than I had organizing them. The digitization projects were *intended* to explore the possibilities to reduce the distance between completely unedited archival documents and edited texts; their *actual* impact was mainly in proving the feasibility of large scale digitization.

Short note on research: having fond memories of the *Halbgraue Reihe*, I did start a publication series entitled “*Kölner Beiträge zu einer geisteswissenschaftlichen Fachinformatik*”. But while it has been useful to document some of our projects, it never gained the importance or visibility of the earlier series (KBGF 1, 2, 3, 4, 5).

5.3 Digital Libraries

The deplored discrepancy between what I intended to achieve also exists for the project which was destined to digitize the complete holdings of the medieval library of the cathedral in Cologne, the *Codices Electronici Ecclesiae Coloniensis* (CEEC) project (Thaller 2017k [2001], in this volume; Thaller 2001E, 2002-1, 2003E). As mentioned, it was based on contacts created by some students of history in Cologne, who graduated as masters a short time before I came to the university: Torsten Schaßan and Patrick Sahle, both quite logically turning into the staff for the project.

The library, in the meantime under the directorship of Heinz Finger, supported the project by providing us liberal access to all the material.

Torsten Schaßan and Patrick Sahle had very good knowledge about the library, particularly about the plethora of catalogs, both internal to the library as well as externally published by codicological researchers. This codicological literature overlaps to some degree; what in my opinion made it much more interesting was that this tradition also shows, how the learned opinion about the codices changed and developed over time. We therefore agreed that we would try to implement what could be described as a synoptic view of the layers of codicological knowledge contained within those catalogs and treatises. Torsten Schaßan and Patrick Sahle convinced me that the TEI should be used for the encoding; a significantly enhanced κλειω provided the data base engine for the web server. From a TEI point of view the encoding of alternative views of different codicologists required some changes to the encoding scheme, but in general the metadata are surprisingly close to the standard.

I had two goals with the project:

- 1) To prove that against the ruling opinion of the time that “high resolution digitization will forever be restricted to a few carefully selected high end manuscripts”, high resolution digitization could be applied to mass digitization campaigns.
- 2) To prove that such digitized codices could be connected to a standard of codicological / editorial annotations, which were potentially contradictory and could develop over time. A development which need *not* prevent the immediate accessibility of the manuscript material based upon preliminary and dynamically changing metadata.

The first of these worked perfectly; the next years were filled with presentations, where astonished audiences saw that you really could get large amounts of manuscripts accessible rather quickly. So far the project enjoyed great support from the research community: if you taught paleography based on third-generation Xerox copies, the usefulness of 6000 x 8000 pixel 24 bit images did not need much explanation.

The second failed spectacularly. It should have supported two concepts.

If manuscript descriptions – metadata that is – could develop over time, you could start with a bare minimum. More or less, with a reproduction of the manuscripts as such, allowing you to browse through the document and read it. This depended on the possibility to get the cost of digitization down to the lowest possible number, with the vision in mind that at the beginning you could work your way through manuscript material without any detailed metadata, just as you did in an unknown analog archive. Browse through it, hopefully get lucky – and dig deeper, if you got so. This met the stiffest possible resistance by the curatorial community possible. I should not have been surprised about that. I mentioned that my very first seminar on “dynamic editions” was at Salzburg in 1989. In that or the next year, we had a meeting there, where we discussed with a curator of the *Albertina*, the museum in Vienna which houses one of the world’s largest collections of old master prints, whether it would be possible to make some of the less easily accessible ones on a CD edition. It essentially broke down, when she exclaimed at one stage “I really do not understand, why people should have access to *our* prints, before *we*

have published them.” This hidden assumption that the true purpose of cultural heritage collections is to provide intellectual raw material – and a reason for funding – for curatorial staff transcends all divisions between archives, libraries and museums. So it is not really surprising that the greatest resistance against the systematic digitization of manuscripts came from the German community of manuscript catalogers, emphatically upholding the mantra that all manuscripts could only be digitized after they had been cataloged according to appropriate standards. Appropriate standards implying roughly one person month of academic staff per manuscript. As this is guaranteed in the meantime, the resistance against digitization as a *second* step has softened considerably. Manuscripts from which no cataloger has received a salary yet, are of course still inaccessible digitally.

The other important concrete goal behind the abstract approach was the assumption that editions – and even more so, codicological descriptions – are *never* finished and there are many questions about them, where the differences between the interpretations are not a nuisance, but possibly more enlightening about the meaning of a document, than any of the two interpretations would be in their own right. Therefore, the representations of manuscript related material in the internet should not try to re-implement the concept of a final, correct, catalog or edition, as it exists in print, but open up the search for a more fluid, dynamic concept, which truly makes use of the digital medium (Thaller 2017 [2005], in this volume; Thaller 2004-2, 2006-1, 2007-1). I clearly failed to get that opinion across. That also holds true for what may have been the most important outcome of the Cologne cathedral library project. I am truly proud to be able to claim that the fascinating work which Christoph Flüeler at Fribourg has done around the e-codices project has drawn some inspiration, and actually the software for the very first version, from the project at Cologne. Unfortunately, pressure from the community moved him early on to try to provide catalogs which represented the traditional view of catalogs. As a result, while I admire this project very much and have tried to support it within my possibilities, I became disengaged with the technical solutions rather early on.

The high visibility of the CEEC project lead to a quite large number of smaller projects, where I could provide a server for an existing digital collection, but as the contradiction between what I wanted to achieve and what resulted remained as it has been describe above, some hopes I had originally to engage in systematic digitization activities was relatively soon given up. An exception was the attempt at a “*verteilte digitale Inkunabelbibliothek*” (vdIb) which was assumed to be a proof of concept solution for a broker surface, which connected a project digitizing ca. 600 *incunabula* at Cologne with another collection of roughly the same size of digitized *incunabula* of the *Herzog-August-Bibliothek* in Wolfenbüttel, uniting two high end digitization projects which were based on different technical as well as conceptual solutions under one user interface. I enjoyed very much the co-operation with Thomas Stäcker that made that possible together with the support received from the library of the university in Cologne, but as it became clear that in this field the structural problems I described for manuscripts were just the same, I later withdrew from digitization projects completely. Though in my very first years at Cologne I had used them to argue once more for a more aggressive stance of the Humanities towards information technology: (Thaller 2003-1, 2003-2, 2003-3, 2008-6). And they emphasized my opinion that structured, semi-structured and unstructured data

should be understood to be much more closely related than usual (Thaller 2000-1, 2004-1).

There has been, in my project portfolio in Cologne another high-visibility project, which does not really fit into the other categories used, but fits best under the heading of digital libraries. Since 2008 (until 2013) I have been involved with the VD18 (*Verzeichnis der im deutschen Sprachraum erschienenen Drucke des 18. Jahrhunderts*) (Thaller 2008-5) a very high quality national bibliography of all prints in areas where German was the primary language (even if the prints were in other languages than German) for the eighteenth century, part of a family of similar bibliographies relating to earlier centuries – VD16, VD17. When this project was contemplated, as the precursor, the VD17, drew towards completion, there were plans for two types of innovation against these precursors.

On the one hand, besides the creation of high end bibliographic records, the prints – books, leaflets, journals – would be digitized as images. On the other hand, a weak point of the precursor projects should be rectified: the IT support systems for these earlier bibliographies had not been integrated really well into the overall library system. German libraries, like those in other countries, rely since a few decades to a high degree on central catalogs of their content, which are accessed at the individual libraries by local systems. This is an obvious advantage: non librarians may be surprised by it, but creating a professional library record for a catalog costs quite frequently more than the actual book. So if, say, fifty university libraries buy a copy of the same book, cataloging it only once creates huge savings. Germany has a federal structure, where the individual states of Germany jealously guard their prerogatives. As a result Germany has not one central cataloging system, but six *Verbundsysteme* or *Verbünde*, as they are called. VD18 – unlike its predecessors – should be integrated as closely as possible into these, so users should ideally be able to find the catalog entry for a 18th century item in their local library system and connect directly to the digitized print.

Our task consisted of three steps: (1) integrate the catalog entries of the *Verbundsysteme* relevant to the VD18 into one consistent data base, (2) identify all duplicates of catalog entries, so it was clear, how many different prints there were, for most of them more than one copy existing in Germany's libraries and (3) distributing the identified prints more or less evenly between the cooperating libraries, so each of them would be responsible for the handling of a roughly equal number of prints. "Handling" meaning: create a high end quality bibliographic entry and digitize the print.

This was not particularly simple: the library systems of the *Verbünde* contained catalog information which was inherited from the conversion of the original catalog entries of the libraries they supported, so the differences between two different catalog entries for the same print could be astonishingly large. Technically, I think, we performed quite nicely: an error rate of about 5% for data of the quality received still seems to me to be quite an achievement. General knowledge of κλειω was in the meantime sufficiently small that in my last practical work with it I administered the batch jobs for the actual selection of the titles assigned to the participating libraries myself. This did not particularly enhance my intellectual insights, but it certainly served my vanity that the core of the data base engine optimized for a 286 in the early nineties allowed me 25 years later comfortably to administer a database

containing 2.063.946 catalog entries on a quite moderate PC under Linux at a comfortable speed, despite the checks for duplicate entries being quite complex.

The work on the actual algorithm for this identification of duplicates was done by Jan Wieners, implementing essentially a two level Levenshtein solution (one level to identify similarities between the strings of characters making up the words in a title, the other identifying differences between the strings of words in the title). The quality of his programming clearly reflected his earlier work (he had written a *Magisterarbeit* which implemented an OCR system from scratch, which could read incunabula, even if it did so with a high failure rate).

Politically it was frustrating: many of the *Verbände*, which had always claimed an identification of duplicates would be impossible, were deeply distrustful that when somebody worked with these data, they would lose control over them, so the overall technical structure was clearly less effective, than the one we proposed originally. This experience was not made more attractive by a few of the participating libraries being terribly afraid that their importance would not be duly reflected in the project. One of them spent 18 months more or less directly accusing me that in bad faith I had assigned too few titles to them for processing, to belittle their importance within the German world of libraries. I leave aside details which read abstruse to the outsider, but there were sufficiently many conflicts under the hood that I decided not to publish any detailed technical report at the end at all.

As in other cases, retelling these political dimensions overshadows the many nice personal experiences I had with many of the partners in the project. So I should not close this lament without expressing my gratitude to Heiner Schnellung who, as director of the library of the university at Halle was the driving initial force behind the whole VD18 concept and his vice-director Dorothea Sommer, who coordinated the project in the time when I was connected to it. Both have impressed upon me, how effective and pragmatic projects in the library segment of cultural heritage *can* be run.

5.4 Heterogeneous Systems

The *verteilte digitale Inkunabelbibliothek* was not only the last of my projects in the digitization area proper; it was also related to another research interest. One of my guiding interests had always been a software environment, which would allow administering material which was just as variable and irregular as historical sources. One data base should be able to combine documents where the majority consisted of just of a small number of short fields with a term or two, but a few of them might contain multi-page full texts.

This lead rather early on to a concept which I have propagated over the years as a sideline in many projects, the “autonomous (digital cultural heritage) object” (already Thaller 1993-2 and Thaller 2017h. [1993], in this volume). The idea behind them was that in the ideal world it should be possible to represent any kind of content in a digital representation, where individual objects could be created in arbitrarily diverse complexity, completely independent of each other, but any subset of them could still be loaded effortlessly into an integrated information system. The library of *incunabula* was for me a study how parts of the problems connected with that vision could be implemented across the internet, with the documents to be

integrated hosted by different servers. As such it was in some way an implementation of concepts developed by me earlier in the context of a project funded by the European commission – E-Culture Net – organized by Kim Veltmann, then at the University of Maastricht.

While both of these were relatively straightforward exercises, my vision was pushed forward most widely by the *Prometheus* project, funded by the German federal ministry of federal research, directed by Holger Simon, then at the University at Cologne. This was the last project, for which I myself contributed substantial programming support within κλειω. (Chronologically it started already in 2001, earlier than the two quoted above; they were more “advanced”, in integrating different servers dynamically. But much more simple with regard to the heterogeneity of the data bases to be integrated.)

The rationale for this project was quite straightforward. All art history institutes of Germany traditionally house large collections of slides with works of art, obviously overlapping to a considerable degree. Should it not be possible, to digitize these slide archives and interconnect them in the digital world, so they could be used as a unified system for art historians all over Germany? Preferably connected to some graphical user interface, which would allow the online creation of multi-projector slide shows, the primary tool for presentations in art history teaching? In a sense the major technical obstacle for such a solution was that the vision arrived a bit late – the majority of German art history institutes had already started on digitization, when the project started in 2001. And most of them, of course, had started with their very own solution, using not only software which was perfectly incompatible with any other, but also coding schemes and conceptual models which had nothing in common.

On the surface, again, the project was a considerable success: Within 4 years, we integrated something like 20 image data bases. Though some dynamics were provided, this was basically a static integration: i.e., each of the image databases was supposed to dump the descriptions of the images into some format, which was sufficiently non-proprietary not to make it illegal to program input modules being able to read them and map the underlying structure into one that could be handled by κλειω. That system would then support a web interface, which allowed the selection of images and the administration of virtual slide shows, giving the users the feeling that they would be using a monolithic database. This went sufficiently well that a successor to the original implementation on another platform, for which I have not been responsible, is still running today and generating sufficient income that *Prometheus* is more or less a self-supporting institution.

Unfortunately here again, what I wanted to achieve and what became a big public success were quite different. What I had *intended* to do, was to engage deeply into the possibilities of integrating crossovers between the data bases, so you could within *Prometheus* start with the metadata scheme you considered most useful and get access to all images, which could be mapped from their native categories unto that metadata scheme, with controllable degrees of fuzziness. This unfortunately, would have required some determined effort, while almost all of the consortium members were willing to accept an exceedingly small common conceptual denominator, allowing access with the most easily comparable categories, concentrating

technical effort at details of the interface like the width of separating areas between individual slides on the light desk allowing assembling the slide shows.

I completely agree that the later concern may have been an important sales motivation; but it was not exactly what I considered an interesting intellectual challenge in the semantics of image descriptions for art history. So I left the consortium, when the initial funding for the project ended.

5.5 Collaborative Systems

Sometime during the early presentations of the *verteilte digitale Inkunabelbibliothek* I heard about a project called “*Monasterium*”, which intended to digitize and present in the web medieval charters located in some of the archives in monasteries in Austria, which were basically inaccessible physically, as the locations were remote and the dwindling number of monks residing there reduced the opening times extraordinarily. While there was considerable interest in the metadata available, the primary interest was clearly in the digitized documents themselves. And if no metadata could be made available, at least the scans of the charters could. This, of course, came very close to some of my visions about archival material to be made available “raw” and the descriptions and annotations than being added dynamically later.

At the same time, I looked for a subject for a master thesis which should be connected to historical source material and its handling. This led to a contact with Thomas Aigner, the driving force behind the original *Monasterium* project, to whom I proposed to assign a master thesis to Benjamin Burkard to assign the creation of a system by which descriptions / transcriptions of the charters could be entered online by volunteers, with some precautions for editorial supervision of the contributed material. Burkard’s solution was rather convincing and so there started a co-operation which went on right unto my retirement. After some initial work, it was decided that some people working at Cologne should also become responsible for the main server behind the data base with the images of the manuscripts and some tools to enhance those images, quite beyond the editorial system.

In my academic biography this project was quite important – basically because it signaled a significant break. In the projects mentioned previously (and quite a few not mentioned here) I still had the illusion that as professor I could do actual software development work. Sometime between 2004 and 2006 I capitulated – or finally became wise enough to understand that this *was* an illusion.

Monasterium was a success: at the moment it offers some 500.000 medieval charters in the editorial environment, provided by about 60 archival institutions from 10 European countries. It long ceased to concentrate only on remote monasterial archives and contains content from some of Europe’s largest and most respected archival institutions. This is mainly the success of Thomas Aigner and his incredible skill at archival diplomacy and the propagation of enthusiasm, who widened the scope more and more.

Technically it is based upon eXist and the bulk of the current software version has been developed by Jochen Graf, working at Cologne, with a number of significant contributions by others. The overall technical co-ordination has been provided by Georg Vogeler now at the university in Graz.

My personal connection to it changed over the years more and more to an overall supervision and administration of the various sub projects which have been required to provide continuing funding for the development work (Thaller 2011-2). Intellectually it is a partial fulfillment of some of the visions of a new approach to the handling of historical sources, but it is not really part of my research persona.

5.6 Long Term Preservation

As mentioned, sometime between 2004 and 2006 I understood that I would not have the possibility to realize my goals by creating software solutions myself. Indeed, I contemplated at the time to stop applying for research money and fall back upon the more traditional academic role of writing papers, rather than being responsible for software specific for the discipline. *Monasterium* has been part of that change of perspective. But in 2003 I was approached by Seamus Ross, then at the University of Glasgow, whether I would be interested in aspects of the long term preservation of digital material, specifically, whether I would be interested to join the large scale European Delos project to contribute to its long term preservation aspects (Thaller 2007-2). Indeed preservation had also come up on the horizon as a side effect of arguing for systematic digitization already (Thaller 2005-3, 2011-1).

Until 2006 that seemed to go well together with my ideas to move back to a more theoretical approach. I had the idea that my work on “autonomous objects” might fit well here: if digital content can be smoothly integrated into many different environments, it has to be organized in a way, where it is fairly robust – so it might be a good start to survive for longer periods. This theoretical angle was well supported by the focus of the preservation discussions of ca. 2004 and onwards, when at the first summer schools on the subject my angle – how can we make survive digital content for hundred years? For thousand years? – was taken quite well, even if it may have appeared overly academic to the practitioners present.

This somewhat detached approach towards long term preservation was partially lost in 2006, when I accepted the invitation to become partner, a rather substantial partner, of the consortium of the Planets project, which promised to create distributed *service oriented architecture* for digital preservation. Avoiding details which are scarcely meaningful for a reader who has not encountered long term preservation discussions so far, I will just focus on the role of Cologne in the project.

When you preserve digital files, say image files, over long periods, it is quite clear that from time to time you will have to adapt them to the changing technical environment. The most obvious case: the format of an image file may become obsolete, or may for other reasons be not qualified for long term storage. So at some time you have to convert the format of that file into another format. If this is done correctly, the user watching the image stored in the file format of 2067 should see no difference from what the creator of that file has seen in 2017. It is surprisingly difficult, however, to guarantee that a file converted from one format – say: a PNG – to another – say: a TIFF – actually does contain all the information which is needed to render (less precise, but more easily understood: display) the converted image exactly as the original one. Most of the possible calamities which can befall a file during conversion can be discovered when a human observer compares the displays generated from the two files on two screens side by side. But, besides that

“most” is not really satisfactory that model becomes completely unrealistic, if you talk about millions of files, rather than just a few.

So part one of the tasks for Cologne was, to define an abstract model (plus a format implementing that model) that would allow to hold all the information – more correctly: all the data – contained in an image file. If that was possible, you could create a tool, which was able to compare two of these representations and report on the degree to which the two agreed.

What was still missing then, was a tool, which could convert a file in PNG to that generalized representation and another tool, which could convert a TIFF into the same format. In the long term preservation discussion the number of 17.000 more important file formats is frequently specified. I probably do not have to explain, why the notion of creating 17.000 tools to extract the information from these files into the abstract models applicable (“image”, “text”, “sound” etc.) is not very realistic. We therefore decided that we would use just *one* extraction program, which, however, would be able to read scripts in a formal language, which described how to extract data from a specific file format into the generalized model representing the data contained in that file. So to support PNG you had to translate the human readable file specification into this formal language, to support PDF you had to do the same for this much longer specification – or a subset of it.

This was achieved: three XML-based representational languages for the specification of (1) file formats, (2) content of broad basic types and (3) the parameters to be used and weighted during comparison have been defined and the corresponding software tools for their usage implemented. The whole thing was tested on TIFF and PNG for which the format specification was completely translated, DOCX and PDF, where only small subsets of the formats specifications were translated into the formal specification and a couple of other file formats for other media types, of which only subsets were translated into the formal specification. I am permanently grateful to the project team which realized the project, for which I just provided the general concepts and some abstract underlying theory. There have been too many people involved to name them all: my special thanks are due to Volker Heydegger, Jan Schnasse and Johanna Puhl (Thaller 2009E, 2008-1, 2008-2, 2008-3, 2008-4, 2010-1).

Why have you probably never heard of the project? One problem was connected to the approach of the Planets project: to use a *service oriented architecture* in 2006 was not a guarantee, but a long way towards getting a project through the peer review process and receive funding. Whether it was a good approach to *implement* the project, was a completely different matter. That with some tools this architecture meant that you would transfer thousands of files, each of which contained hundred megabytes, across the network to have a tool residing on a server performing a service which needed less than a millisecond processor time to complete what it had to do to each of these files is a good indication. But this points to a more generic problem. “Long term preservation” was a buzzword in the technical and infrastructural discussions of the first decade of this century, but few, if any, of the libraries or archives of the time actually had the data and the prerequisites needed to employ a functioning preservation toolbox. So practical problems would become visible only at the very end of the project.

On the theoretical level the result, an abstract model for the representation of the information hidden in the data within a digital file, the project was quite useful for my own research – you will find significant traces of it in a later chapter. But on the practical level, I found myself with rather complex software tools, for which we had no environment, particularly no convincing test data, to apply them to. And about the probability that tools of that complexity work without extensive field tests, I had no illusions whatsoever.

At the end, the inevitable happened: The Cologne project team dissolved at the end of the project and while the software tools waited patiently on a server which was switched off in 2010, but remained ready to be restarted, it was clear after some years that those modules and tools would never be used for real work.

In 2009, when I still hoped to find an environment where they could be applied productively, I was approached by Wolf-Rüdiger Schleidgen, at that time assigned to the *Staatskanzlei* of North-Rhine Westphalia (roughly: North-Rhine Westphalia's prime-minister's office; apologies to German readers for the only partially valid translation). We knew each other since 2001, when he invited me to give a lecture on the CEEC project for archivists. After that, particularly while he was later in charge of the *Hauptstaatsarchiv* (state archive) of Düsseldorf, we co-operated in a few very small projects, including one to develop a proof-of-concept solution to connect digitized archival material to an internet-based version of the finding aids of the archive he was responsible for.

In 2009, in any case, he invited me to participate in a meeting, where the *Staatssekretär* of the *Staatskanzlei* asked for a solution, which would at the same time make the cultural heritage of North-Rhine Westphalia more visible as well as safer. The political ideas at the time were somewhat vague and I am jumping ahead with the following description. But what essentially evolved was the idea that North-Rhine Westphalia should create an infrastructure, where any kind of institution which kept digital material relevant for the cultural heritage – in the broadest possible sense – could send that digital material to this infrastructure and be sure that it would be saved according to the state of the art of digital preservation. And, as ingesting such digital material into a proper digital archive would require the data to be processed and validated in any case, at the same time digital derivatives would be produced, which could be included into web portals. Web portals of the institutions which owned the material, a state-wide portal showing the cultural heritage of North-Rhine Westphalia; but the system should also be able to act as a pre-aggregator for the Europeana and the *Deutsche Digitale Bibliothek*, converting the original metadata into something these larger systems could digest easily.

As mentioned, I've been jumping ahead: This was, in any case, the project which appeared after a preparatory process in 2009 and 2010 which involved a feasibility study by a consulting agency, a change of government in North-Rhine Westphalia and endless discussions about the responsibilities of individual project partners. Though prerogatives may be a better term than responsibilities: when the concept was for the first time presented as above, the first question came from a partner who wanted to make sure that it would be possible to exclude data from the presentation components, using the preservation part of the system only; the second from a partner, the *Landesarchiv* (a successor institution to the *Hauptstaatsarchiv* mentioned above) which enquired whether it would be possible to keep all archival

data on hardware on the premises of the *Landesarchiv*, as this was allegedly legally required.

Be that as it may – towards the end of 2010 it was decided that the ministry which had become responsible for the whole thing in the meantime would fund a *Vorprojekt* (preparatory project) which would essentially implement a prototype of the solution which, a central wish of mine, would be extensively tested. For that purpose, three computer centers, one from the research library sphere of the state of North-Rhine Westphalia, one supporting her cultural heritage institutions and the computing center of University at Cologne, to represent academia, were supposed to run three instances of the software and administer three redundant copies of the material on heterogeneous long term storage media, the three instances being permanently synchronized. All of this to operate together under the name of the *Digitale Archiv NRW* (DANRW) (Thaller 2013E, 2013-1).

This started with the beginning of 2011 and went on in different phases until spring 2014, leaving me with two parallel, but completely contradictory experiences.

On the one hand, after I had finally given up my hope to get involved in practical software development myself, the project was one of the most satisfactory I had the privilege to manage. The team – most centrally: Sebastian Cuy, Daniel de Oliveira, Martin Fischer, Thomas Kleinke, Jens Peters, Johanna Puhl and Lisa Rau – worked together extremely well. We went a few times to the iRods team at Chapel Hill in Northern Carolina, activating contacts I had made during earlier projects with Reagan Moore, and may have been the first German site who tried to use iRods quite intensively, before its meantime prominence, in a project expected to reach production level at the latest in 2012. I definitely cherish that experience.

On the other hand, these years have been an unmitigated political nightmare, which may to a large degree be responsible for the acerbic tone which parts of this text exude. I have never before or after worked in such a climate of profound distrust of central project partners. The main responsibility for this belonged certainly to the *Landesarchiv* of North-Rhine-Westphalia. They had at the time of the start of the project no practical experience with digital preservation whatsoever and their main point was to somehow keep all matters of digital archiving under the closest possible control nevertheless, lest the might lose some of their institutional responsibilities.

This was quite frequently a reason for great amusement for the technical partners. I fondly remember the occasion when the technical representative of the *Landesarchiv* left a meeting of the technical committee of the project with the statement, he would doubt, whether we really had any idea what extraordinary amounts of computing power we would need. *He* had recently employed *two* PCs with *Quad* processors – *Quad* processors! – to reformat some images, and things went on for days. Everybody remained silent, until he had left the room – and nobody had the heart to point out to him that the computing center at University at Cologne employed in its high performance cluster CHEOPS at the time roughly 10.000 cores, giving it a respectable place on the list of the world's 500 fastest computers when that cluster started production. At some other date we were informed he had recently been to a workshop on long term preservation and had learned there how complex things were, as e.g. “significant properties” of digital

files existed. Leaving early again, nobody told him that the lady keeping the minutes of the meeting had actually been the responsible editor of an ontology for the handling of significant properties in digital preservation.

But, funny and entertaining as many of these archival utterances have been, the political effect was disastrous. Particularly, as it managed to poison the whole climate in the consortium, up to the point where one of the libraries in it, which had by far the largest amount of digital material, suddenly refused to provide these data for ingest testing, as it was not sufficiently clear, whether they would remain in legal control of them, if such data would be processed at the computing center of another university.

Wolf-Rüdiger Schleidgen left in disgust into early retirement towards the end of 2011 after fighting a running battle against this climate for two years since 2009. This was no problem for our “pre-project”, which despite all inventive reasons why we could not get the test data promised and promised again, went along to a technical solution, which was fully endorsed when reviewed by an external consultant in 2014. His successor certainly backed as; but while for Wolf-Rüdiger Schleidgen the *Digitale Archiv NRW* defined his job at the ministry, for his successor it was just one more responsibility added to a portfolio full already.

Which meant that that successor did not retire, but capitulated in 2013 in another way, by which much of the responsibilities were delegated to an external consulting unit, with a long record of co-operating with the various institutions of the state as well as the cities of North-Rhine Westphalia, which should be brought into the boat, as their municipal archives were to be supported by the project as well. No technical problem, but a development which led to the breakdown of *my* political agenda in long term preservation.

I always had the feeling that the way in which the complexity of digital long term preservation was hyped in recent years, created an atmosphere where everybody was willing implicitly to assume that the solution of such a complex problem *must* be extremely expensive. (An assumption quite useful, if you wanted an excuse to justify large project budgets.) And *my* political agenda behind DANRW has been to show that that is not necessarily true, if you have a solution were it is clearly defined, what resources are needed, and then check which institutions can provide these resources as cheaply as possible. When the cities entered the game that rationale was changed: it was clear that the current suppliers of computing services for them should become the technical backbone. Or, more concretely: it was not the public anymore, which looked for the most effective suppliers of computing resources, but the suppliers were nominated and then asked, how much the public was expected to pay them to achieve what they as suppliers thought necessary. Which I do consider a bad idea, not only as an academic, but also as a tax payer. So I left the project.

In any case, there is a happy ending of sorts to this highly frustrating story. The core of the technical development team has been hired at the end of the project by one of the selected providers of technical services to maintain and apply the software developed. A few more political games have delayed the application of the system, but it seems that in the near future what has been basically ready in 2013 will actually be employed.

5.7 Infrastructures

Digital long term preservation, digital libraries, heterogeneous networks and collaborative systems for the annotation and editing of sources have all at one time or the other been in the focus of the discussions about the necessity for research infrastructures for the Humanities.

I was involved in a number of projects dealing with infrastructures: the German National Research council (DFG) honored me twice by funding a project for the evaluation of its digitization activities, the first in 2003/05 leading to a report where a detailed summary of the 190 page report was published on the website of the research council (Thaller 2005E). It was a pleasure to work on this with a highly motivated group, including, among others, Alexander Czmiel, Pia Janczak and Susanne Kurz. Another highly motivated group, notably Simone Görl (Kronenwett, nowadays) and Katharina Mahler produced another such report between 2011 and 2012. This later one, unfortunately, was never made public, as some of the committees of the DFG considered parts of it too sensitive and we could not agree, how far an abbreviated version would still have public value. In both cases these reports were based on extensive interviews, mainly within Germany, but also with selective representatives of other European funding agencies and libraries, involved in the funding of digitization activities within Europe.

But my main contribution to this species of advisory reports on information infrastructure was certainly a set of recommendations for the future development of the systems responsible to supply academia within North-Rhine Westphalia with information. This quite massive volume (Thaller 2011E), resting on the work of Simone Görl (Kronenwett, nowadays) and Johanna Puhl, has a history which may illustrate some of my positions better than the actual volume.

As a member of the advisory board of one of North-Rhine Westphalia's library systems, I had a discussion with Friedrich Bode, then responsible in the research ministry of the country for the libraries and computing centers of the state. He emphasized that a succinct strategy for the further development of these institutions, particularly the libraries, would be needed. I did agree, pointed out, however that even Clausewitz would have found it hard to develop a strategy based solely on the command "Wage war!". At least the name of the enemy would have been appreciated. Less metaphorical: strategy concepts are pointless, if you do not know what that strategy should achieve. So I proposed *first* to define, what the information landscape in the state *should* look like in 2025 and only *then* to discuss a strategy how to get there. (2025, as I claim that the trends for the further development in digital technologies can be fairly reliably projected for 15 years, given the time between the first appearance of a technology in papers and conferences and its routine application in the field.) To which he replied: Well, fine, write a funding application for a project to create a definition of such strategic goals.

Not considering this requirement that a strategy needs an explicit goal, is in my opinion the main reason why so spectacularly little comes out of the "lets develop a strategy" fevers, which afflict academia occasionally, politicians suffering from that affliction seemingly permanently.

My connection to "infrastructures" therefore rests primarily on my active involvement with fields, which are usually considered part of the infrastructure dis-

discussion, on the one hand; on the other on my responsibility for a couple of studies, which try to advise on the further development of infrastructures. Besides that I was relatively little involved in the more recent type of projects, which are explicitly dedicated to the development of research infrastructure for the Humanities. With two exceptions.

When describing my work at the University at Cologne, I already mentioned that over the years, as the number of students multiplied, the number of concrete collaborative efforts with my colleagues at the philosophical faculty actually decreased. At the same time, it was clear that not only at archaeology, where I was working closely together with Reinhard Förtsch, a lot of projects were going on, which were of obvious interest for anybody having a broad concept of what constituted the application of information technology in the Humanities. To provide a platform for a more extensive knowledge of what was going on at Cologne, I managed to fund Patrick Sahle for a few months starting in 2009 to prepare a concept for a more active collaboration within Cologne as well as to prepare a broader platform for the discussion of the teaching of information technology in the Humanities. This resulted within Cologne in the creation of a first version of the Cologne Center of eHumanities, mainly as a discussion platform for common interests and possible joint funding applications later. It is a lasting achievement of Andreas Speer, who succeeded me as *Sprecher* (chair) of the center at the turn of 2012/2013 to have attracted considerable funding from the faculty, so visibility and substance have increased not by an order, but by orders of magnitude.

More important for me, personally, was that out of the tentative work in 2009 there also resulted a working group of most of the institutions and persons interested in Germany in the teaching of information technology related degrees – and such skills also below degree level. This was a revival of attempts in the nineties at defining common requirements for such interdisciplinary degrees and courses in the domain of History and Computing, leaving traces as volumes in the *Halbgraue Reihe* mentioned above. This time the culminating effort was a catalog of courses in Digital Humanities taught at, roughly, universities of German speaking countries. This was presented as a result of a working group of the *Digital Humanities im deutschsprachigen Raum* (DHd) association at its annual conference in 2016, Zoe Schubert having co-ordinated the later stages of that work. I very much enjoyed meeting and discussing curricular matters again, after more than twenty years. Nevertheless, as with almost anything that brings me in contact with organized Digital Humanities nowadays, the Groundhog Day remake feeling arose again. It was impossible at the early nineties to convince my colleagues that the profits to be gained of a clear definition of what a quality curriculum in History and Computing should contain (HGR A12, HGR A17) at the very minimum outweighed the risks that a local dean, university president or similar might be disappointed that the degree or course taught at their university was not incomparably superior to all others. It was impossible in 2015, again, to agree on any kind of minimal requirements for a quality Digital Humanities degree.

I have not strayed from infrastructural involvement by the abstract above, as the early phase of these revived curricular efforts, possibly together with a certain notoriety in the field, got me invited into DARIAH-DE, the German part of the biggest European research infrastructure program for the Digital Humanities in

2011. As with any very large project, the contributions of individual members are rather small, so I will not report here in which sub projects, besides curricular development, Cologne was participating.

My own feelings about these infrastructure projects are still ambiguous, despite having been part of one. It is certainly true that these projects have created a greater visibility for what is called Digital Humanities today. But for me two questions remain without answer:

A traffic infrastructure is built by road builders. It is built to accommodate cars built by a completely different class of engineers. That all cars built by engineers fit onto the existing roads are guaranteed by a rather small set of agreed upon restrictions. Virtually *all* infrastructure projects I am aware of today, are trying to develop digital tools according to their own specifications; they are *not* appropriate to guarantee the usefulness of tools developed by others. At the very least that makes the infrastructure metaphor inappropriate in my opinion.

Secondly: I am deeply distrustful about the notion of “infrastructural research”, when connected with projects which are supposed to create infrastructures. If you commission a group to build a road from A to B, you do not expect them to spend a substantial amount of the funds available on the development of a new type of tarmac. You expect them to use trusted and proven types of building materials, to finish the project in time and within budget, with a precisely predictable capacity. Again, the metaphor fails.

As long as these defects are not removed, I find it hard to take infrastructure projects *quite* serious. It is very hard, not to suspect that they are just borrowing the infrastructure metaphor to fund technical and methodological developments, which otherwise, as I have experienced all too well, are very hard to fund.

5.8 Κλειω

What became of the guiding theme of my third life in the fifth? When I went to Cologne I was willing to change my life radically, from being primarily defined by my own research interest towards the connection of such research interests to broad academic teaching, to help establish work between Computer Science and the Humanities as its own well defined field, some brand of Computer Science applied to the Humanities. But I *did* hope to be able at the same time to use my preliminary work as a platform for this.

In some sense that was realized: if κλειω would not have been there in the background, I would never have been able to realize such highly visible projects like the CEEC, the technical backbone of *Prometheus* or *the verteilte digitale Inkunabel Bibliothek* and quite a few others basically without university funding within less than five years. And the capabilities of the system have been extended considerably for the sake of these projects.

In another sense, it was a total failure: I was neither able, to attract students to the system in seminars, nor was I able to train people to create the low level software libraries which would have been needed to go ahead with the technical solutions. (There have been exceptions: Johanna Neumann wrote an unpublished MA thesis implementing a very simplified version of some of the ideas behind the ex-

tended string concept; Jan Schnasse and Elona Weiper (Chudobkaite, at the time) wrote unpublished MA theses exploring the idea of self-documenting images).

In itself this is unremarkable; nevertheless, I would like to point to the reasons for it, as they represent a conflict, which may be important for others to realize.

That *Informationsverarbeitung*, despite all short comings in its implementation, was acknowledged as a self-contained academic discipline in Cologne is important, if we take the interdisciplinary work between Computer Science and the Humanities serious. And in its Cologne implementation it was certainly successful, attracting 100 new students each year at the end of my time. But there is a dilemma there, which remains unsolved. The pressure to understand this as a field, which derives much of its justification from the immediate applicability of its results to other academic disciplines gets somehow transformed into the notion that the students educated must excel in their immediately applicable technical skills. I certainly subscribed to that notion, which is presumably one of the reasons, why so many of the graduates found immediate employment somewhere in the IT industries.

But it turned out to be almost impossible, when students started their studies for a significant time getting the ability to realize immediately useful applications, to redirect them towards advanced work later, which requires to dig deeper and work on projects, which are based on technologies beyond the bulk of the literature. In the case of κλειω this was my personal problem. But I have the strong feeling that it will be a hindrance also for most other advanced development projects in this interdisciplinary field. The mindset required to optimize the usability of a website right now, and the mindset required to care about what technology *might* provide for the Humanities in ten years' time are quite different.

5.9 Funding Politics

As I mentioned in the section on my research activities at Cologne, the importance of third party funding for the research topics chosen in this my fifth life can scarcely be overestimated. Actually that influence went even deeper, than it may have been apparent so far. Reviewers of funding proposals for interdisciplinary work are hard to find and I personally consider the willingness to participate in peer reviews a straightforward moral obligation of the citizens of the academic republic. I am completely serious when I say that in my opinion everybody refusing to review a funding application, when asked to do so by a funding agency, should be barred from receiving any funding from at least that agency for at least five years, unless *very* unusual circumstances exist.

That mental framework may have given me more than my share of reviewing activities. I was a member of the *Bibliotheksausschuss* (later: *Ausschuss für wissenschaftliche Bibliotheken und Informationssysteme*), the committee responsible for libraries in the context of the German national research council (DFG) and some of its sub-committees between 2000 and 2008. My interdisciplinary qualification being sufficiently unusual that between 2002 and 2004 I found myself actually in three such sub-committees – which translated in this two years into approximately 20 to 25% of my networking time being spent for the DFG. Even withdrawing myself from the DFG later, I was a member of regular reviewing bodies in Austria, Belgium, Germany and Switzerland during the years between 2010 and 2016, and

did take part in more than occasional reviewing during that period also in France, Italy and The Netherlands. In 2014 this culminated in my being involved intimately in the reviewing of approximately 250 funding proposals. Besides doing some reviewing for the European Commission, culminating with my membership in one of the Advisory Groups of Horizon 2020.

Obviously, there is not much I can report about this part of my activities without conflicts with confidentiality requirements. I do mention them, however, as the experience I gained reading all these funding proposals and discussing their relative merit in committees of very variable composition has very much influenced my outlook at and interpretation of interdisciplinarity in research.

That I was involved in so many of these committees reviewing funding proposals started, as mentioned, with my interdisciplinary qualification being somewhat unusual; at the same time it is owed to the fact that the necessity for the application of information technology in the cultural heritage area tended to become self-evident during recent years and its application to problems of the Humanities has definitely become fashionable since about 2010, far beyond the groups and people interested in it at the turn of the century.

I certainly appreciate that this increasing recognition of the necessity for interdisciplinary work has been mirrored by a willingness to fund such work. And when I came to Cologne in 2000 I was definitely eager to get involved in the decision processes about such funding, to be able to steer interdisciplinary work into the direction I considered right. But, when I describe below, why I will not follow up my academic career with the traditional pattern of an academic retiree by trying to prolong his influence upon his university and the self-organization of academia, the experiences from these funding committees and the insights I gained there about the Humanities' mindset have been one of the most important reasons.

6. So What?

My five lives have been good ones. If anybody would have told me in 1976, when I first came to be actively involved to the question of whether there might be a possibility to systematically apply electronic data processing, as information technologies were known then, to the Humanities and what public visibility this field would have at the time of my retirement, I probably would politely have canceled further social contacts, as such a person was obviously suffering from a mild psychic disorder. It has been good to be part of what after all is an astonishing success story.

Nevertheless, my fifth life, as a professor within organized academia, is ended. I do not plan to appear at my university beyond what is required by the implicit obligations arising from supervising the last stages of the academic work of students who started when I've been active. While I will fulfill all obligations which are the result of earlier or expiring memberships in funding and peer review committees and procedures, I will under no circumstance get involved in any further review processes. I will never again submit a funding proposal myself. I may occasionally present some aspects of what I consider the core of my methodological opinions, but I will avoid all further contacts to organized "Digital Humanities". The later not, because I believe that there is no valuable and clear concept for such

a field (Thaller 2017-1, 2016E, 2017E) though many questions are open (Thaller 2012-3, 2014-1, 2015-1, 2016-3), but because I think, they all too frequently avoid the issues I consider crucial (Thaller 2016-1, 2016-2).

Some of my plans for the future involve interdisciplinary work, still. A bit more on that further below.

But such a further life being whatsoever it becomes, I definitely intend to draw a very clear dividing line between the previous lives – and specifically the last one – and the future. Why?

It is a bad old tradition of elderly people to grumble along that the world is becoming a worse place and I'd like to take pains to make very clear that I do not share this opinion. Most of the changes I've been living through since 1950 have been changes to the better, nostalgic memories being exactly that: nostalgia.

But there are some tendencies, in the Humanities and the handling of our cultural heritage which I have wrestled with for a very long time now, and I feel it is time to retire from that wrestling match. So I am *not* trying to say that the Humanities are going to the dogs – it is just that there are traits in them, which I have accepted with gnashing teeth when I was twenty, hoping things would change. Advancing towards seventy my patience is exhausted.

As should have become clear, when I described my motivation for studying history, I am definitely an offender by conviction: I studied history, because I considered it important. And while my original motive may have been simple curiosity, this feeling of importance soon ran much deeper. My problem with many historians and indeed with much of the Humanities is that many colleagues seem to assume that the aesthetic of a result is sufficient to make it important. During the years of the project at archives of the former concentration camp at Auschwitz, this only deepened. I never quite understand how historians, even specialists for contemporary history or the Shoah itself, can be satisfied by leaving students in dismay, shock or simply being humanly concerned. Auschwitz, in my opinion has once and for ever put upon an historian the solemn obligation to try to *explain*; and not simply to cry and commiserate. The later, in one form or other, is a requirement for anyone with a sound system of morals, but clearly not enough for people who claim the past as their profession. And the past is indivisible. As long as we cannot explain how people can become willing to commit atrocities, we cannot exclude that even the most unlikely cause can be connected with it – and the crucial explanation may at the end be found in a working model for the mechanism for the distribution of motives between literatures of different languages in times long past, just as well as in phenomena seemingly more directly connected to what we want to explain. And that populist renderings of history, be they related to perceived human constants from antiquity, medieval glories or fairy tale versions of more recent epochs, are at the very heart of post factual populism should in my opinion make the need for a gold standard narrative of history self-explanatory.

In any case, the type of historical research and the Humanities I identify myself with, are serious business and have an obligation towards society. My problem is that a tendency I was amused by when I was twenty, being sure it would be overcome in time, is still running strong: there are all too many people who seem to understand the Humanities as an endeavor that allows smartly dressed people at Vernissages to converse politely about some aesthetically pleasing topics.

The two preceding paragraphs may seem to be very far from the question, what is the most appropriate approach to apply information technology to the Humanities. For me it is a core issue: if historical questions are significant, they need to be studied in such a way that derivations from the available sources must exploit them to the limit and try to exploit them in a way which provides a protocol how to reach the most precise intersubjective agreement – or an equally precise disagreement – about their meaning. (Cf. already Thaller 2017g [1992], in this volume) This is the domain, where my vision for the usage of information technology by Humanists lives. Understanding the information contained in Humanities' sources is much too serious to leave it to the engineers and hard scientists. If all we need is a source for nice topics for a relaxed but intellectually vibrant conversation that analytic commitment is redundant, of course. And writing history with a quill may provide haptic pleasures that make an excellent topic alongside a glass of wine.

While this has been a problem for me with the Humanities, it is worse when I encounter many cultural heritage institutions. There is a romanticism around archives, libraries and museums that I find it extremely hard to cope with, both when discussing things with curators as well as when talking to the users of their collections. I heard the chief curator of one of Germany's largest collection of *incunabula* argue against their digitization, as "we know very well that all what our users are interested in are their bindings". I have spent a few years on wondering, how a sane mind can reconcile this opinion with the responsibility to supervise quite a few academic level salaries every year going into the painstaking cataloging of this – allegedly – totally uninteresting content. And I am truly relieved that retirement means I do not have to try to understand this mentality anymore.

My suspicion is that this aesthetic streak, which is so prevalent among many Humanists, is also at the core of the frustrating relationship between the Humanities and information technology. It is an interesting experience what happens when – as in recent years – funding suddenly becomes available relatively easily for interdisciplinary projects between the Humanities and information technology. The usual reaction for the academic organizations of Humanities' disciplines when such a situation arrives is to claim that it is important to make sure that it is "not technology which drives the selection of projects for such funding, but Humanities content or: Humanities' real problems". This sounds wonderful; in reality it means that the traditional academic organizations try to make sure that projects can get funded, which get away with as little technical innovation as humanly possible. And I am absolutely tired to hear in the selection committee of an interdisciplinary funding line, again and once more, the sentence "yes it may be that the information technology employed is totally trivial – but than it is such a wonderful topic". Saying all those critical things let me confirm that my deep believe in the necessity for the Humanities in general and the study of historical problems in particular has never wavered. But I am deeply disappointed that so many members of the humanistic parts of academia see information technology as a threat to established and cozy habits, rather than as a chance to be embraced to advance the field. There is such a thing as intellectual cowardice.

And the underlying romanticism has a tendency to cling stubbornly to paraphernalia. The app.crit., as we know it, was not an intellectual achievement. It was a (typographic) technical solution for a conceptual problem. That another medium

than print should not ape print, but create another technical solution for the underlying problem, should be self-evident. That it is not makes it difficult for me to avoid the impression that it is exceedingly hard for many Humanities' researchers clearly to separate between *substantialiter* and *accidentaliter*. Or concentrate on *why* they do things, rather than *how* the always did them.

I became tired of that, during the reviewing of a couple of hundred research proposals, but besides being tired of it, I am afraid that here lurks a problem, which may go well beyond the Humanities and actually hint at a societal problem with new technologies.

Many Humanists seem to be mortally afraid that if they allow computer scientists to have an influence on their projects, they would lose control. As a result, they usually are very insistent on applying technologies, which are immediately transparent (or seem to be) to the Humanist. The down side of that is that interdisciplinary projects tend to apply procedures, which are actually very well understood in engineering already. Which means that the actual innovation in software and information technology always happens far away from the humanistic disciplines and the triumphs of Digital Humanities usually consist of the application of relatively old technologies which have been prepared with completely different problem areas in mind, but are superficially adapted to the needs of the humanistic disciplines after they have already hardened into software systems ... to which at the end the questions of the Humanities have to be adapted, as they were never asked, when the development was still enough in flux to influence the way in which solutions are arrived at. It seems to be a strange contradiction. On the one hand, Humanists insist very much that they, not "the engineers" have control of a project; on the other, they believe that what those engineers have developed at an earlier stage – relational data bases, XML –, is divinely inspired, so it must be applied as it exists. There may be better examples for the over compensation of an inferiority complex, but you have to search for them.

I mentioned that I am afraid that this actually hints at a societal problem. Let me illustrate this with a final provocation, derived from my experiences in the field of digital preservation. The provocation is simply that I assume about 95% of all the funding spent in the last fifteen years on digital preservation has been a waste of money. The reason: the institutions, which have neglected the problem of the long term viability for digital media for as long, as they possibly could, were suddenly forced to try to solve it. According to the paradigm that "the engineers" must not take over, all preservation projects needed a significant percentage of the project, which could be safely assigned to people without any specific engineering knowledge. As a result, we have now the most wonderful organizational schemes how to organize a heritage institution so that the digital media are documented, supported by metadata and directed through a complex work flow. That the digital data, which are administered in that way, are still based on engineering solutions, all of which were designed without any concept of persistency that goes beyond temporarily switching a computer system off, was safely ignored in the process. So we have completely persistent archives of completely volatile digital objects now. Phrasing it a bit cynical: when one looks at the development of digital preservation since ca. 2000, it is hard to avoid the impression that the primary purpose was the preservation of the organizational institutional framework of heritage institutions,

rather than that of their content. A position towards technology which – appropriately generalizing – has indeed the potential to become a problem for society far beyond those little institutional pondlets.

But, grumblings of a retired academic and provocations aside: overall I did enjoy my fifth life. But I am definitely not sorry to leave it behind and try a new one, far away from organized academia and funding considerations.

7. Next Life: My Very Own Ivory Tower

Most of my lives past have suffered, the more recent the more so, from the fact that I have been quite successful in my undertakings, but the successes I achieved, were not those I intended to achieve. Talking about the future one can concentrate on what *should* be – and hope that an intentional step away from the strange world of institutionalized academia allows one to concentrate on clearly defined and purely intellectual interests, without all sorts of influences changing the intended course into different directions.

So, assuming that senility will wait a bit, or at least start in mild forms only, and hoping that the ability to ignore all external deadlines will not mean losing the drive to get things done, the following describes what I hope to do next.

I described how in my early years in academia the big challenge that intrigued me was the quest to underpin high end generalizations, required by *universalhistorische* visions, by meticulous work with sources which were many levels of generality away from these high end generalizations. Even if wisdom or resignation has lead me away from this extremely broad view, two elements of this youthful excitement remain. I am still fascinated by my understanding of *history* as a *regina humaniora* and I still perceive the necessity to base all historical research squarely in meticulously analyzed sources.

When I am trying to derive a personal research interest from that, however, a very important change of view has taken place since I submitted my first SPSS job. Then I saw the formalisms which were necessary to apply computational techniques to historical sources as a requirement of computer science. A few decades later, I consider the precision required by algorithmic work as so natural that I am willing to define my historical interest by the specific properties of the information contained in and provided by historical sources.

7.1 What is Peculiar about *Historical* Information?

So from the point of view of information theory “history” – or: the part of the Humanities I am interested in – is defined as follows.

1) *History is the study of all phenomena which require the analysis of information created by human endeavor where the creators of that information cannot – or shall not – be asked about the meaning of the information left behind.*

The “cannot” in the sentence above may seem to be trivial; the “shall not” is a methodological postulate, deeply rooted in at least the German tradition of histori-

cal scholarship. In the typology of historical sources of Johann Gustav Droysen there exists the central differentiation between an *Überrest* (remainder) and a *Tradition* (tradition). The later describes sources which have been created intentionally for posterity: the monument of a great general, the text of a chronicle. The former type of source – *Überrest* is left behind unintentionally, without considerations of posterity: the manure of a cavalry horse, the text of a register for taxation.

Überreste always take precedent over traditions. How great a general was is wide open to interpretation; whether the manure of his cavalry contains oats or hay indicates military capabilities which are much less so. The application of this principle to sorts of text is left as an exercise to the reader. With other words, history is not interested in the way our predecessors wanted to be remembered, but in the way our understanding of the world leads us to see them. (Which of course does not exclude the study of the way they wanted to be remembered; not to believe them, but to understand their motivations.)

In the days of Johann Gustav Droysen historians had learned that the way in which people present themselves should always be distrusted. They seem not to have understood yet that this also means that historians always should mistrust themselves, as well. This lack of distrust against one's own capabilities to distance oneself from "evident" contemporary assumptions of course lead Leopold von Ranke to his claim to tell *wie es eigentlich gewesen ist* ("how things actually were" or "show what essentially happened", depending on the type of Ranke *exegeisis* you believe in). That claim makes sense only, if we assume that we ourselves are completely independent of all unchallenged opinions of our own time. That all persons are so deeply embedded into the – not eternal – intellectual and societal norms of their contemporary society that any "objectivity" of human interpretation has all the charm and intellectual resilience of a fairy tale, was simply beyond the 19th century.

And I emphatically and explicitly want to repudiate and exorcise the spectres of objectivity still occasionally wafting through the methodological discussions of history. However, while we have permanently to distrust our own unconscious interpretations of the content of sources, we nevertheless have to be able to communicate and discuss our conscious interpretations – not the least, as only that gives us a chance to discover what unconsciously may lurk behind them. Any statement in academia must be open to discussion and disagreement. All statements which are not, are part of either artistic or theological creativity. But a community, which simply makes statements and disagrees about them, is somewhere between sterile and ridiculous. The community of historical research needs rules to decide about disagreements: not to decide which of two interpretations is the correct one, but to drill down into the chains and levels of interpretations of sources until we progress from "your theory is wrong" to "your interpretation of the manor court decisions in the sixteen fifties is based on the assumption that the sentences were actually executed – this is doubtful in the light of xyz". Without any mechanism which allows as a last recourse this level of progress in intersubjective agreement, historical research is meaningless in my eyes.

Or:

2) *Research is defined by a system of rules for the communication of disagreements about interpretations of information; out of their application an in-*

crease in the body of intersubjectively agreed upon interpretations arises. Research is defined by a consensus about acceptable models of argumentation, not about the knowledge domain to which these models are applied.

There is a corollary to this, which may not be immediately apparent: I consider the differences between the hard sciences and the Humanities as rather trivial. Both are part of a rationalist way to confront the world I subscribe to. Their and my common enemies are creationism, esotericism and post-factual thinking.

There is a consequence of the sentences (1) and (2), which may be even less apparent, but which leads to a way to approach information theory, which departs quite radically from the one which is behind some of the most fundamental concepts behind current information technology.

We can modernize the *Überrest* discussion by saying that the topic of history is the unintentional information contained within a source, rather than the intentional one. When a medieval city prohibits for the fifteenth time that citizens spend more than ten guilders on a specific item of clothing, the historically *most* important information is that citizens *did* spend more than that, as otherwise there would be no need to repeat that prohibition.

While I write these pages concentrating on *historical* studies, I would like to point out that looking at the unintentional part of the information conveyed by Humanities material, is spread throughout the disciplines: the discovery that the usage of function words is a better indicator of personal style, than that of words carrying meaning is a point in case; Giovanni Morelli's 19th century discovery that the way in which small recurring details, like the proportions of a hand, are an excellent way to identify an artist, another.

Most computer science text books, and much text on information theory, start with the "father of information theory", Claude Shannon – whose significant paper was incidentally entitled "A Mathematical Theory of Communication", *not* "A Mathematical Theory of Information". As a result, information is usually described as the content of the data transmitted by a sender to a recipient. And communication is deemed successful, if the recipient receives exactly the data transmitted. Which in information theory is then extended to the requirement that the receiving cognitive system is receiving exactly the information which the sending receiving system did send. We can leave the precise relationship between data and information aside for the moment and simplify the situation by saying that in communication a sender transmits information intentionally to a recipient.

This is, of course, exactly *not* what a historian wants to achieve, even though the sender – recipient model may at first look to be quite appropriately describing what happens when a 21st century historian deciphers a cuneiform text. Historians do *not* want to know, what a person in earlier times wanted to tell them; they want to know, what at the times of creating a source was so obvious that it did not need saying. (Yes, *CUM grano salis*.)

When we talk about the information to be drawn from textual data, there is another concept which is quoted almost as regularly in introductions than Shannon's. According to Ferdinand de Saussure the differentiation between the signifier and the signified is a key step on the way to understand the structure of language. Applying our "history is the interpretation of the non-intentional information in a source" we could say that we have here a problem very similar to that we discov-

ered with Shannon. Speakers will use a signifier to signal the signified concept in their world view; for historians later quite frequently that signifier will signify some completely different signified concept in *their* world view.

Finally: while it is tempting, to consider the interpretation of textual data as the primary domain of the historian, at the very least we should not forget that non-textual sources are so important that any theory of the information to be processed for the sake of historical research must be general enough to allow for *all* types of information. And even for texts we do not understand, as the language, dialect, jargon or writing system is only imperfectly understood. (Read nineteenth century texts written in shorthand, if you believe that the assumption that modern texts are straightforwardly understood is generally applicable.)

Therefore:

3) Historical sources should not be interpreted as signs for an unequivocal message. They are tokens which have to be interpreted by a historian. Such interpretations have to be documented as precisely as possible, as a concrete disagreement about the interpretation of a specific token indicates a focus for the necessary clarification of intersubjective understanding.

7.2 Consequences for the Representation of Information

If a token from the past can be interpreted in many different ways, no such interpretation may change the token. Therefore we need a way to represent information, which clearly distinguishes between a token and its various and potentially conflicting interpretations. If a historical source is transferred into a digital information system, the token itself must be the closest possible representation of the physical original, with the least possible influence of any cognitive system in the process creating that representation. The interpretation of such a token is the domain of the cognitive system. The scanned image of a text can be an example for the representation for a token; the transcription added to the image as an annotation, as example for an interpretation. This is true *in principle*. A willful distortion of the scanning process can turn the representation of the token into an interpretation. This is an explicit attempt to cheat, does not vulnerate the principle, however.

The separation of a token and its interpretation is actually very deeply rooted in information systems already. Whether the bit string “01000001” is interpreted as the character “A” or the number “65” is stored as an independent chunk of data advising the program whether to handle this specific byte as a character or a number. (That ultimately this advice is stored in the form of another set of bit strings is a consequence of the Von Neumann architecture.)

To discuss and implement this understanding of information, I propose:

4) Information taken from historical sources can be represented as a specific configuration of tokens with a specific geometry.

5) Any subset of the tokens within such a configuration can be connected to arbitrarily many interpretations.

6) Interpretations themselves are represented in information systems according to the same model. The overall model is recursive.

7) *Any two subsets of tokens within the same type of configuration can be compared according to one or more inherent metrics.*

8) *Any two interpretations on the same interpretative dimension can be compared according to one or more inherent metrics.*

This concept may look quite abstract, possibly esoteric, at first look. I would like to point out, therefore that it has been underlying the XML based representational languages implemented at Cologne for the Planets project (XCEL and specifically XCDL). A very short summary has also been presented at DH2012 in Hamburg (Thaller 2012-1). Though the description of the model underlying the XCEL/XCDL languages is contained in this volume, a few short comments on the statements above seem to be appropriate.

re (4): A typical data object for historical research will usually consist of a high number of atomic tokens. Such data objects appear in different categories: texts, images, image series (*aka* videos), 3-D models etc. These categories imply a specific type of token, which is above the bit level, but still at a level close to the hardware (character codes, pixels, drawing instructions etc.) which are connected in a specific token configuration implying a clear geometry (one dimensional series of character codes, two or three dimensional arrays of pixels etc.).

re (5): Such an interpretation can be that a certain substring in a character string is expected to be a proper name; that the pixels within a polygon defined on the plane in which the pixels are contained, show a specific person; that a subset of the drawing instructions in a 3D model represent a column.

re (6): The relationships described are fully recursive. A substring may be interpreted by an image; an area within that image may be interpreted by a string.

re (7): Any two sets of tokens which share the same token configuration must be comparable, giving similarities between them. (There are of course binary similarities: equality v. non equality.) For tokens these comparisons operate on the binary values representing the token in the technical system. While for one dimensional traditional character strings these comparisons are trivial, already for a two dimensional pixel configuration they are not. So far there is no consensus about the relative similarity between two image pairs where (a) in one case all pixels are identical on the red and blue color band, with a given numeric difference in the green color band and (b) in another case all pixels are identical on the red and green color band, with a given numeric difference in the blue color band.

re (8): *Any* interpretation assigns a position on an interpretative dimension to the set of tokens it applies. Interpretations which map onto the same interpretative dimension can be compared; interpretations which do not cannot. This does *not* vulnerate recursiveness. A character string cannot be compared on the token level with a bitmap; though both may be related to exactly the same interpretation as two occurrences of the same proper name, which therefore can be compared on the interpretative level. An interpretation of a token set as a point in time cannot be compared to the interpretation of a token set as a metaphor; though both interpretations may be related to the same token set which can be compared to other token sets on the representational level.

This concept of information is related to and possibly extends two approaches to an understanding of information which have been developed over the years.

7.3 Information and Context

Keith Devlin's *infor*, which he introduced as basic unit for a mathematical treatment of information, introduces at a very low level the notion of time and space as integral constituents of the very concept of information:

*The majority of real life 'facts' pertain only to a certain region of space and a certain interval of time. In order to deal with these kind of facts, I need to introduce both spatial and temporal locations and to allow locations to figure in the argument places of relations*³.

There is a straightforward way to connect these two dimensions onto the model I try to describe in these pages. That a specific token relates to a given time and space, is an obvious application of the interpretation concept. I therefore propose to extend the "*facts' pertain only to a certain region of space and a certain interval of time*" statement to "*facts' pertain only to a certain interpretative context*". From the point of view of Devlin's theory that means that the time and space arguments in the definition of a situation / *infor* would have to be replaced by the set of applicable contextual ... probably *infos*.

A simple example of that can be shown, if we translate his situation⁴.

$s_0 \models \langle\langle \text{running, Jon, Palo Alto, 1.1.88, 1} \rangle\rangle$

"It is true – 1- that Jon has been running on January 1st 1988 in Palo Alto."

into German:

$s_0 \models \langle\langle \text{laufen, Jon, Palo Alto, 1.1.88, 1} \rangle\rangle$

While in the English case it is quite clear what s_0 describes, in the German one it is dependent on whether the translation presumably done by an Austrian native speaker of German is read by an Austrian or a German native speaker of German. In the first case, it would re-translate, correctly, to:

$s_0 \models \langle\langle \text{running, Jon, Palo Alto, 1.1.88, 1} \rangle\rangle$

while in the second a re-translation might give

$s_0 \models \langle\langle \text{walking, Jon, Palo Alto, 1.1.88, 1} \rangle\rangle$

as *laufen* translates into running in Austrian German, but (usually) into walking in German German.

For Devlin, who avoids the historical dimension, this is a relatively trivial problem. For a historical information system, where we incur such ambiguities permanently, we need a concept to take care of them. (Introducing the interpretative context mentioned above.) If we take the historical dimension serious, we have to change Devlin's concept also in another detail. If the situation above is a token taken from a historical source, where a lost English original has been translated by a writer about whose background we are ignorant, all we could say would actually be

"There is a probability p_1 that Jon has been running on January 1st 1988 in Palo Alto; there is another probability p_2 that he has been walking there."

This can be translated into the requirement to turn Devlin's binary "polarities" (truth values) into continuous ones, which may be governed by *some* restriction like

³ Devlin, Keith: *Logic and Information*, Cambridge: Cambridge University Press, p. 23.

⁴ Ibid. p. 49.

p_2 always being $1 - p_1$, but could also be governed by other restrictions or by an restriction we are not aware of.

For the model of information in historical sources developed here, the result of this is:

9) *All existing interpretative dimensions of two sets of tokens form a context, which has to be considered, when comparing the two sets of tokens.*

Let us take a step back at this moment, as this statement is more dangerous than it looks. In the texts

$t_1 ::=$ “*Lorem ipsum dolor sit amet. **Henricus episcopus** de aliquo fecit aliquis, ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.*”

and

$t_2 ::=$ “*Lorem ipsum dolor sit amet. **Henricus comes** de aliquo fecit aliquis, ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat*”

any comparison of the token sets “**Henricus episcopus**” and “**Henricus comes**” on the interpretative level would not only have to take into account that we have a sequence of two token sets here, where in both cases a subset interpreted as proper name is followed by a subset interpreted as an status indicator, but also all interpretative information which applies to the larger token set representing the complete embedding text(s). Which can be quite tricky: If t_1 predates t_2 the probability that both *Henrici* are identical is much lower than if t_2 predates t_1 . It is entirely feasible that a *comes* may be referred to as *episcopus* later in life, while it is almost beyond imagination that an *episcopus* would be called a *comes*. Except, of course, the two token sets – historical texts – have been written by partisans of two opposing political camps, the second one of which never accepted *Henricus* being raised to the bishopric in question.

The last example, of course could also be considered as an argument for a traditional Humanist being sceptical about the possibility to formalize the reasoning of an historian as “you have always to consider everything”. Indeed: ultimately the interpretative context relating to two such contexts is essentially the sum total of our knowledge of the past. The polemic of the traditional Humanist that a computer based comparison will never be able to consider “everything”, can at the polemic level be answered quite easily: no human being will be able either. We have to admit, however that human researchers do a much better job at considering *many* things, though not *everything*, and may claim that their comparisons can therefore be more trustworthy as any done by a technical system. Which of course does not convince me, after what I wrote about intersubjectivity, as it is almost always impossible to document appropriately which of the many versions of many things a specific human researcher has *actually* taken into account.

Nevertheless, what to evaluate of a potentially unlimited context *is* a problem, when we take the notion serious that an information system for the handling of historical or humanistic questions should not simply use an underlying information model which comes from other backgrounds, but shall fit the specific properties of historical information.

7.4 Consequences for the Processing of Information

The argument of the Humanities scholar skeptical about the ability of an information system to consider all of the information related to a given question, was given at such length here, as it leads to what I consider a basic misunderstanding about the nature of information, as usual, at least information as encountered by historians.

A fundamental difference between the handling of the problem that a source can only be interpreted within the totality of a conceptually unlimited context by humans and technical information systems is that the understanding which is fixed in a printed book is the result of a cognitive process, which has been going on for a long time. (And while the proud author may think that the result fixed in print is final, most authors will usually admit that after a few more years, they would have described the result slightly differently, as that cognitive process has continued after printing was done, frequently unconsciously.) In contrast the analytic step for which a technical information system is harnessed is delimited sharply: if a digital system is asked to evaluate whether two persons mentioned in two documents may be identical, it is expected to give an answer within a short – and in any case: finite – period of time. And it is most certainly not expected to come back a week later and announce that after due consideration it has changed its opinion.

At the intuitive level, this is simply so. At this intuitive level this is presumably the result of psychological expectations about the behavior of information systems, created by the experience how they work today. I would like to point out however that this behavior is actually grounded in information theory on the one hand, and in some of the basic concepts of computer science at the other.

I am aware that quite some criticism has been raised against the ladder of knowledge concept; nevertheless, most people will still accept its premises, modified more or less. To recapitulate:

Most treatments of information will assume today that at the sign level there exist *data* which turn into *information* when they are processed by a cognitive system. If this information is then within the cognitive system embedded into a context gained by experience it can trigger actions, at which stage knowledge has arrived. The examples usually given for this relationship are quite convincing at first look: the number “25” is pure data; being by a cognitive system translated into a “room temperature of 25° centigrade” information arrives; previous experience provides the *knowledge* that now is probably a good time to get rid of your jacket, social situation not prohibiting doing so, as you do not just have a fleeting impression of warmth, but it actually *is* warm enough that that action is appropriate from a health point of view.

Looking a bit closer, however, the steps between “data – information – knowledge” become much less clear. A few pages ago we have recapitulated that whether a bit string represents a character or a number is already a question of interpretation. So the decision that we are talking about the number 25 and not the ASCII character “end of medium” is *already* a cognitive achievement which produces something we could call information if we build our example slightly differently. The problem here is that the sequence “data – information – knowledge is

usually understood as a discrete – genuinely digital – relationship, while in all probability it is a continuous – analog? – one.

Fortunately a concept of information exists, which can easily be extended to model this continuous concept of information quite nicely. Already in the seventies of the last century Langefors⁵ proposed his infological equation:

$$I = i(D, S, t)$$

which is to be read as saying that information I is the result of a function $i()$, which interprets some data D in the context of a suitable semantic background S at a time t .

I have already earlier shown (Thaller 2017m [2009], in this volume) that we can derive from this model through a number of intermediate steps the form

$$I_x = i(I_{x-\alpha}, s(I_{x-\beta}, t_1), t_2)$$

to be read as follows:

The information I at time x is the result of a function $i()$ which has interpreted the information I available at the time $x-\alpha$ in the light of the result of the semantic function $s()$ applied to the information I available at the time $x-\beta$. This semantic function $s()$ has been applied for a period of time t_1 ; the central information generating function $i()$ has been applied for a period of time t_2 .

This concept has one big advantage: it converts the static model of information as existing in discrete units in a timeless universe, into a dynamic entity permanently changing over time as long as the two functions are running. This concept seems to me attractive, as I think it describes exactly what happens to the historian and humanist who contemplates a problem, much of the time subconsciously, over a long period; it is quite dangerous, however, as it vulnerates one of the most basic assumptions of computer science.

That assumption is that data, representing information, exist as discrete units somewhere, usually called data structures in computer science. Such data structures are occasionally processed and converted into other data structures, the processes usually called algorithms. A set of data structures plus a set of algorithms operating upon them resulting in an information system.

Summarizing my foregoing reflections upon Devlin and Langefors, I propose:

10) An information system fit for the handling of historical sources should exist as a set of permanently running processes, which try to remove contradictions between tokens. Such tokens are used to represent data. They do not directly map into information. Information is represented by a snapshot of the state of a specific subset of the concurrently running processes.

Or, more briefly: data structures do not represent information; the state of algorithms does.

A full implementation of that model means of course that the addition of a another set of tokens to an information system will wake up some of the algorithms ... and may therefore mean that the system changes its opinion about a question it has been asked beforehand, hopefully being able to notify the cognitive system which has asked that question originally. (The “cognitive system”, frequently, but not

⁵ Langefors, Börje: Theoretical Analysis of Information Systems, Göteborg 1973.

necessarily, being a human, as presumably many such information systems work together to serve the human end user.)

7.5 Some Consequences of the Model Described so Far

Trying to summarize sentences (4) to (10) into one statement, I arrive at:

11) The data in the totality of historical sources, or any subset thereof, forms a mutual context for the interpretation of any set of specific items contained therein. It can be envisaged as a set of n-dimensional configurations of tokens representing physically existing sources, each of which exists in an m-dimensional universe of interpretative assumptions. Information arises out of these data by permanently running processes, which try to minimize contradictions and inconsistencies between subsets of the data.

12) This model is both, a conceptual one for the hermeneutic “understanding” of historical interpretation, as well as a technical one for future information systems supporting historical analysis.

If we take these sentences serious, a few things follow on highly different levels of generality.

(a) Very concrete:

If any of the above makes sense, embedded markup is a dead end street, as it completely mixes up the representation of evidence of the past and its interpretation.

(b) More abstract:

In principle I think that the concept of sets of tokens representing sources which exist in an interpretative universe of interpretations of subsets of these, goes a long way to describe historical work. There remains one phenomenon, however, which can be described only with some difficulty.

The concept of token configurations, as discussed above, implies character strings as a basic building block. They offer a clear one dimensional geometry, which makes it easy to connect various interpretations to various substrings via appropriate co-ordinates (offsets from the start of the string that is). This is also very convenient, as it provides a way of connecting representation and interpretation, which can easily be extended to two dimensions – images – and forms of representation at higher dimensions.

What is missing here, however, is an intermediate level between the level of tokens represented in the digital space and tentative interpretations of the meaning of these tokens: a level of tentative theses about the relationship between tokens. Such theses are usually expressed in the Humanities in the form of critical editions. I would not like to extend arguments here which are described in detail elsewhere, but I have argued since almost a quarter century now that to process the texts of critical editions a new concept for “strings” would be needed, which solves a number of problems, including the problem of textual variation.

The “representation in an interpretative universe” model derived in the pages above can on the conceptual *level* probably be extended reasonably easily, to include a class of interpretations about the relationship between representations. However: to solve the problem of textual variation, I have always argued that texts

should not be represented as one-dimensional arrays, but as graphs of atomic tokens. The arguments given for this in the paper introducing the notion of “extended strings” (Thaller 2017i [1995], in this volume) appear still valid to me. While it is extremely easy to link an interpretative item to a specific substring, however, I have so far no proposal to offer, how reliably to connect an interpretative item to a sub-graph. This is an unresolved problem in the model I have derived here.

(c) Very fundamental and abstract:

While I point only very occasionally at it, there are very few steps towards the implementation of any of the above, where binary reasoning does not become unsuitable extremely soon. I see few chances for the implementation of any information system based on the preceding considerations, if the application of indeterminate and fuzzy reasoning to information systems does not become much more straightforward than it is now.

7.6 Pipe Dreams and Reality

This “representation in an interpretative universe” model can probably be relatively easily accepted by many historians. The major challenge is contained in sentence (12) above, where I claim that it is not only fit as a conceptual model (analogy, metaphor) for the reasoning of a human historian, but also a blueprint for the implementation of technical systems.

A technical system, which would keep a few trillion tokens representing sources of the past connected, recalculating all conditional connections between them, as soon as an interpretative item changes or additional tokens representing data are added, is slightly beyond the immediate horizon.

There are some building blocks, however, for that kind of information system, which in my opinion are crucial and could be realized within a reasonably short time. I propose to focus here on extensions on the level of programming languages or operating systems. This makes the argument probably very hard to follow for most Humanities-oriented readers which have read up to this point. However:

- If you solve a problem at the level of an application, available to the end users of that application – e.g. a data base “Spurious people of the 13th century”, <http://sp10c.someuniversity.terra> – *every other* application to be developed, ever, has to reinvent the solution.
- If you solve a problem at the level of an application system – e.g. a specific data base system like Neo4J – it is easily available for all applications realized with the help of that system; *every* application realized in *another* application system to be developed ever, has still to reinvent the solution.
- If you solve a problem at the level of a programming language – e.g. Java – it is easily available for all applications realized with the help of any application system realized with the help of that programming language; though *every* application realized in *another programming language* to be developed ever, has still to reinvent the solution.

A result of this relationship, or rather the tradition at solving technical problems for the Humanities at the level of single applications, is the reason for one of the most irritating and most frequent *Ground Hog Day Remake* experiences throughout my previous lives. I said so above, but it really is a burning nuisance, so let me repeat: I

cannot count the number of times, I have listened to or read descriptions how people have solved the problem that calendar dates in historical documents cannot be handled as current dates can. And the *really* frustrating thing is that despite these problems have been solved innumerable times, the next thousand projects trying to handle calendar dates in historical documents will have to do it all over again.

Let me conclude with a few examples of such building blocks, which might be fit achievements of a next life; unless those achievements again turn out to be others than planned.

Context oriented programming

According to current programming paradigms, variables are logically independent of each other, unless they are explicitly part of a higher order data structure, like an object in the object oriented programming paradigm.

A programming paradigm, which assigns a “context” – similar to a scope – to a variable, would considerably ease the implementation of interactions between seemingly independent variables. Such a mechanism would, e.g., allow to declare a variable “time sensitive”. As soon as such a variable is used in an operation, the executable should check: (a) whether in the currently active context a time dimension has been activated, (b) if so consult the variable representing that time dimension and (c) select the appropriate way to execute the operation in question.

Extended Data Types

In current programming languages there is an extremely small number of data types, which are built into a specific programming language. Some data types, like *time* are not directly supported by many programming languages, are indirectly supported, however, as they are sufficiently common to be supported by interfaces to general services provided on the level of operating systems.

What we need is a generic interface for Humanities’ specific data types – historical calendar dates, e.g. – which can be easily called upon within a programming language. And can be called upon *outside* of any inheritance system of an object hierarchy explicit or implied by the programming language, so there are no restrictions imposed by such hierarchies.

Fuzzy and Vague Reasoning

There exist quite a few solutions to support vague reasoning or fuzzy predicates. However, these are contained in very highly specialized libraries and add-ons, each of which has sufficiently many idiosyncrasies that they are almost never employed in software written for Humanistic purposes. We need ways to include such constructions into main stream programming languages, with the absolute minimum of changes to their basic programming model possible.

And Then ...

I hope that the three preceding examples are sufficiently concrete to illustrate, what kind of fairly basic research in software technology or possibly even deeper layers

of computer science, are needed to realize the “representation in an interpretative universe” model I have introduced in these pages. Many more could be added: at least for the concepts like “stop processing strings in arrays; understand them as branching graphs”, “frozen algorithms representing information” or “replacing embedded mark-up consistently” I consider it fairly simple to sketch research plans, if not solutions. This would certainly go beyond this autobiographical sketch, however. And their realization probably even beyond a next life.

The suspicion that just because things usually are done in a way makes that way not necessarily the right one is still holding strong. Strong enough for a sixth life, even a seventh.

List of Publications

- Thaller, Manfred. 1975. *Studien zum Europäischen Amerikabild. Darstellung und Beurteilung der Politik und inneren Entwicklung der Vereinigten Staaten von Amerika in Deutschland, Großbritannien und Österreich zwischen 1840 und 1941 im Vergleich*. Unpublished PhD thesis, Graz.
- Thaller, Manfred. 1980-1. DESCRIPTOR. Probleme der Entwicklung eines Programmsystems zur computerunterstützten Auswertung mittelalterlicher Bildquellen. In *Europäische Sachkultur des Mittelalters*. Wien (= Veröffentlichungen des Instituts für Mittelalterliche Realienkunde Österreichs 4 = Sitzungsberichte der Öster. Akademie der Wissenschaften, phil. hist. Kl. 374).
- Thaller, Manfred. 1980-3. Datenbankorientierte Verfahren bei der maschinenunterstützten Auswertung historischen Quellenmaterials. In *Datenverarbeitung in den Geisteswissenschaften*, ed. Thomas Kneser. Göttingen. (= GWDG Bericht 19).
- Thaller, Manfred. 1981-1. Praktische Probleme bei der interdisziplinären Untersuchung von Gemeinschaften 'Langer Dauer'. In *Aspekte der Historischen Forschung in Frankreich und Deutschland*, ed. Gerhard A. Ritter and Rudolf Vierhaus. Göttingen. (=Veröffentlichungen des Max-Planck-Instituts für Geschichte 69).
- Thaller, Manfred. 1981-2. Historical Software Issue 1: Introduction. *Historical Social Research* 6 (3): 75-82. doi: [10.12759/hsr.6.1981.3.75-82](https://doi.org/10.12759/hsr.6.1981.3.75-82).
- Thaller, Manfred. 1981-3. Historical Software Issue 2: Statistical Package for the Social Sciences (SPSS). *Historical Social Research* 6 (4): 87-92. doi: [10.12759/hsr.6.1981.4.87-92](https://doi.org/10.12759/hsr.6.1981.4.87-92).
- Thaller, Manfred. 1982-1. *Numerische Datenverarbeitung für Historiker*. Wien-Köln: Böhlau Verlag.
- Thaller, Manfred. 1982-2. Zur Formalisierbarkeit hermeneutischen Verstehens in der Historie. In *Mentalitäten und Lebensverhältnisse*. Göttingen.
- Thaller, Manfred. 1982-3. Klassifikatorische Voraussetzungen der Analyse historischer Quellen mittels formaler Verfahren. In *Numerische und Nicht-numerische Klassifikation zwischen Theorie und Praxis*, ed. Peter Ihm and Ingetraut Dahlberg. Frankfurt: Indeks-Verlag (= Studien zur Klassifikation 10).
- Thaller, Manfred. 1982-4. Historical Software Issue 3: MINITAB. *Historical Social Research* 7 (1): 70-77. doi: [10.12759/hsr.7.1982.1.70-77](https://doi.org/10.12759/hsr.7.1982.1.70-77).

- Thaller, Manfred. 1982-5. Historical Software Issue 4: Calculs et Analyses Sur Ordinateur Appliques aux Reconstitutions (CASOAR). *Historical Social Research* 7 (2): 80-6. doi: [10.12759/hsr.7.1982.2.80-86](https://doi.org/10.12759/hsr.7.1982.2.80-86).
- Thaller, Manfred. 1982-6. Historical Software Issue 5: Literary and Linguistic Computing - FAMULUS, OCP, COCOA, LEXICO, COBAPH. *Historical Social Research* 7 (3): 78-87. doi: [10.12759/hsr.7.1982.3.78-87](https://doi.org/10.12759/hsr.7.1982.3.78-87).
- Thaller, Manfred. 1982-7. Historical Software Issue 6: Graph Definition and Analysis Package (GRADAP). *Historical Social Research* 7 (4): 100-7. doi: [10.12759/hsr.7.1982.4.100-107](https://doi.org/10.12759/hsr.7.1982.4.100-107).
- Thaller, Manfred. 1983-1. Recycling the Drudgery. On the Integration of Software Supporting Secondary Analysis if Machine-Readable Texts into a DBMS. *Linguistica Computazionale* 3, Supplement.
- Thaller, Manfred. 1983-2. Ein neues Eingabegerät: Die Kurzweil Data Entry Machine. *Nachrichten für Dokumentation* 34 (6): 182-286.
- Thaller, Manfred. 1983-3. Historical Software Issue 7: Micro Computers. *Historical Social Research* 8 (1): 94-110. doi: [10.12759/hsr.8.1983.1.94-110](https://doi.org/10.12759/hsr.8.1983.1.94-110).
- Thaller, Manfred. 1983-4. Historical Software Issue 8: The Kurzweil Data Entry Machine (KDEM). *Historical Social Research* 8 (2): 88-94. doi: [10.12759/hsr.8.1983.2.88-94](https://doi.org/10.12759/hsr.8.1983.2.88-94).
- Thaller, Manfred. 1983-5. Historical Software Issue 9: Statistical Software in Historical Social Research. *Historical Social Research* 8 (3): 99-106. doi: [10.12759/hsr.8.1983.3.99-106](https://doi.org/10.12759/hsr.8.1983.3.99-106).
- Thaller, Manfred. 1983-6. Historical Software Issue 10: Statistical Analysis System (SAS). *Historical Social Research* 8 (4): 88-95. doi: [10.12759/hsr.8.1983.4.88-95](https://doi.org/10.12759/hsr.8.1983.4.88-95).
- Thaller, Manfred. 1984-1. Mittelalterliche Realienkunde und EDV. In *Die Erforschung von Alltag und Sachkultur des Mittelalters*, 219-28. Wien. (= Veröffentlichungen des Instituts für Mittelalterliche Realienkunde Österreichs 6 = Sitzungsberichte der Öster. Akademie der Wissenschaften, phil. hist. Kl. 433).
- Thaller, Manfred. 1984-2. Datenbankorientierte Verfahren in der Geschichtswissenschaft. In *Bericht über den 15. Öster. Historikertag in Salzburg 1981*. Wien (=Veröffentl. des Verbandes Österreichischer Geschichtsvereine 23).
- Thaller, Manfred. 1984-5. Historical Software Issue 11: Biomedical Computer Programs / BMDP. *Historical Social Research* 9 (1): 78-85. doi: [10.12759/hsr.9.1984.1.78-85](https://doi.org/10.12759/hsr.9.1984.1.78-85).
- Thaller, Manfred. 1984-6. Historical Software Issue 12: Statistical Package for the Social Sciences / SPSS X. *Historical Social Research* 9 (3): 96-104. doi: [10.12759/hsr.9.1984.3.96-104](https://doi.org/10.12759/hsr.9.1984.3.96-104).
- Thaller, Manfred. 1985-1. Beyond Collecting. On the Design and Implementation of CLIO, a DBMS for the Historical Sciences. In *Data Bases in the Humanities and the Social Sciences 2*, ed. Robert F. Allen. Osprey, Florida.
- Thaller, Manfred, Konrad Jarausach, and Gerhard Arminger. 1985-2. *Quantitative Methoden in der Geschichtswissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Thaller, Manfred. 1985-3. Possiamo permetterci di usare il computer? Possiamo permetterci di non usarlo? *Quaderni Storici* 60 (3): 871-89.
- Thaller, Manfred. 1986-1. Can We Afford to Use the Computer; Can We Afford not to Use it? In *Informatique et Prosopographie*, ed. Hélène Millet. Paris: CNRS.

- Thaller, Manfred. 1986-2. A Draft Proposal for the Coding of Machine Readable Sources. *Historical Social Research* 11 (4): 3-46. doi: [10.12759/hsr.11.1986.4.3-46](https://doi.org/10.12759/hsr.11.1986.4.3-46).
- Thaller, Manfred. 1987-1. Methods and Techniques of Historical Computation. In *History and Computing*, ed. Peter Denley and Deian Hopkin. Manchester: Manchester University Press.
- Thaller, Manfred. 1987-2. Auf dem Weg zu einem Standard für maschinenlesbare Quellen. In *Datennetze für die Historischen Wissenschaften?*, ed. Friedrich Hausmann et al. Graz: Leykam-Verlag.
- Thaller, Manfred. 1987-3. The Daily Life of the Middle Ages, Editions of Sources and Data Processing. *Medium Aevum Quotidianum* 10: 6-29.
- Thaller, Manfred. 1987-4. Secundum Manus. Zur Datenverarbeitung mehrschichtiger Editionen. In *Beiträge zur Geschichte und Ihren Grundlagen*, ed. Günther Cerwinka et al., Festschrift Friedrich Hausmann zum 70. Geburtstag. Graz.
- Thaller, Manfred. 1987-5. Clio – ein datenbankorientiertes System für die historischen Wissenschaften: Fortschreibungsbericht. *Historical Social Research* 12 (1): 88-91. doi: [10.12759/hsr.12.1987.1.88-91](https://doi.org/10.12759/hsr.12.1987.1.88-91).
- Thaller, Manfred. 1987-6. κλειω Ein fachspezifisches Datenbanksystem für die Historischen Wissenschaften Version 1.1.1. Göttingen: MPI für Geschichte.
- Thaller, Manfred. 1987-7. Κλειω. A Data Base System for Historical Research Version 1.1.1. Göttingen: MPI für Geschichte.
- Thaller, Manfred. 1988-1. Vom Beleg zum Begriff. Der Beitrag der Datenverarbeitung zur Lösung von Terminologieproblemen. In *Ut populus ad historiam trahatur*, ed. Gerhard Michael Dienes et al., 237-54. Graz: Leykam-Verlag.
- Thaller, Manfred. 1988-2. Gibt es eine fachspezifische Datenverarbeitung in den historischen Wissenschaften? Quellenbanktechniken in der Geschichtswissenschaft. In *Geschichtswissenschaft und elektronische Datenverarbeitung*, ed. Karl Heinrich Kaufhold and Jürgen Schneider, 45-83. Wiesbaden: Franz Steiner Verlag.
- Thaller, Manfred. 1988-3. A Draft Proposal for a Format Exchange Program. In *Standardisation et Vechange des bases de données historiques*, ed. Jean-Philippe Genet. Actes de la troisième Table Ronde internationale tenue au L.I.S.H. (C.N.R.S.). Paris.
- Thaller, Manfred. 1988-4. Was sind 'fortgeschrittene Kenntnisse' in formalen Verfahren für Historiker? In *Qualität und Quantität. Zur Praxis der Methoden der Historischen Sozialwissenschaft*, ed. Manfred Thaller, Gerhard Botz et al. Frankfurt / New York: Campus Verlag.
- Thaller, Manfred. 1988-5. Wider den Guru. In *Qualität und Quantität. Zur Praxis der Methoden der Historischen Sozialwissenschaft*, ed. Manfred Thaller, Gerhard Botz et al. Frankfurt / New York: Campus Verlag.
- Thaller, Manfred. 1988-6. Data Bases vs. Critical Editions. *Historical Social Research* 13 (3): 129-39. doi: [10.12759/hsr.13.1988.3.129-139](https://doi.org/10.12759/hsr.13.1988.3.129-139).
- Thaller, Manfred. 1988-7. κλειω Ein fachspezifisches Datenbanksystem für die Historischen Wissenschaften Version 2.1.1. Göttingen: MPI für Geschichte.

- Thaller, Manfred 1988-8. *κλειω Un système de base de données spécifique aux disciplines des sciences historiques Version 1.1.1*. Göttingen: MPI für Geschichte.
- Thaller, Manfred. 1989-1. *κλειω. Ein Datenbanksystem* St. Katharinen and later editions (=Halbgraue Reihe zur Historischen Fachinformatik B 1).
- Thaller, Manfred. 1989-2. *Query Net I/O*. St. Katharinen (= Halbgraue Reihe zur Historischen Fachinformatik B 2).
- Thaller, Manfred. 1989-3 Have Very Large Data Bases Methodological Relevance? In *Conceptual and Numerical Analysis of Data*, ed. Otto Opitz, 311-26. Berlin: Springer.
- Thaller, Manfred. 1989-4. Warum brauchen die Geschichtswissenschaften fachspezifische datentechnische Lösungen? Das Beispiel kontextsensitiver Datenbanken. In *Computer in den Geisteswissenschaften. Konzepte und Berichte*, ed. Manfred Thaller and Albert Müller, 237-64. Frankfurt a. M. (= Studien zur Historischen Sozialwissenschaft 7).
- Thaller, Manfred. 1990-1. Geographische Angaben in einer Historischen Datenbank. *Eratosthene-Sphragide 2*.
- Thaller, Manfred. 1990-2. Databases and Expert Systems as Complementary Tools for Historical research. In *Tijdschrift voor Geschiedenis*, 103, 233-47.
- Thaller, Manfred. 1991-2. The Historical Workstation Project. In *Historical Social Research* 16 (4): 51-61. doi: [10.12759/hsr.16.1991.4.51-61](https://doi.org/10.12759/hsr.16.1991.4.51-61).
- Thaller, Manfred. 1991-3. The Historical Workstation Project. In *Computers and the Humanities* 25 (2-3): 149-62.
- Thaller, Manfred, and Heinrich Best. 1991-4. Technik als Hilfsmittel in den Geschichtswissenschaften. In *Technik und Wissenschaft*, ed. Armin Hermann and Charlotte Schönbeck. Heidelberg: Springer.
- Thaller, Manfred. 1992-1. The Historical Workstation Project. In *Histoire et Informatique*, ed. Josef Smets. Montpellier.
- Thaller, Manfred, ed. 1992-2. The Processing of Manuscripts. In *Images and Manuscripts in Historical Computing*. St. Katharinen (=Halbgraue Reihe zur Historischen Fachinformatik A 14).
- Thaller, Manfred. 1992-3. Bildanalyse in der Geschichtswissenschaft. In *Deutscher Dokumentartag 1991. Proceedings*, ed. Wolfram Neubauer and Karl-Heinz Meier, 265-301. Frankfurt a. M.: Deutsche Gesellschaft für Dokumentation.
- Thaller, Manfred. 1992-4. On the Conception, Training and Employment of Historical Data and Knowledge Daemons. In *Eden or Babylon?*, ed. Jan Oldervoll. St. Katharinen.
- Thaller, Manfred. 1993-1. *κλειω. A Database System*. St. Katharinen (= Halbgraue Reihe zur Historischen Fachinformatik B 11).
- Thaller, Manfred. 1993-2. The Archive on the Top of your Desk? On Self-Documenting Image Files. In *Image Processing in History: towards Open Systems*, ed. Jurij Fikfak and Gerhart Jaritz. St. Katharinen (= Halbgraue Reihe zur Historischen Fachinformatik A 16).
- Thaller, Manfred. 1994-1. Source Oriented Data Processing. In *Informatik Forum* 8.
- Thaller, Manfred. 1994-2. Die Herausforderung großer Korpora unstrukturierter Texte. In *Max-Planck-Gesellschaft Berichte und Mitteilungen* 1, 32-44.
- Thaller, 1994-3. Bilder und Manuskripte als Gegenstand der rechnergestützten Bearbeitung. In *EDV-Tage Theuern 1993. Kolloquiumsbericht*, 77-88. München.

- Thaller, Manfred. 1994-4. Computerunterstützte Bildanalyse in der Geschichtswissenschaft. In *Informations- und Wissensverarbeitung in den Sozialwissenschaften*, ed. Heinrich Best, Brigitte Endres-Niggemeyer, Matthisa Herfurth and Peter Ohly, 173-84. Heidelberg: Springer.
- Thaller, Manfred. 1996-1. Inventare und Forschungssysteme: Zwei Seiten einer Münze oder unterschiedliche Währungen? In *EDV-Tage Theuern 1995*. Kolloquiumsbericht. München.
- Thaller, Manfred. 1997-1. L'immagine del passato: archiviare ed accedere a fonti culturali visive. *Schede Umanistiche* 11: 163-88.
- Thaller, Manfred. 1998-1. Virtuelle (Zeit-)geschichte? Eine Disziplin zwischen Popularität, Postmoderne und dem Post-Post-Positivismus. In *Zeitgeschichte im Wandel*, ed. Gertraud Diendorfer, Gerhard Jagschitz and Oliver Rathkolb. Innsbruck: Studien Verlag.
- Thaller, Manfred. 1998-2. Bilddatenbanken im Internet. In *EDV-Tage Theuern 1997*. Kolloquiumsbericht. München.
- Thaller, Manfred. 1999-1. Digitale Archive: Technik und Methode. In *Digitale Archive*, ed. Hans-Heinrich Ebeling and Manfred Thaller. Göttingen.
- Thaller, Manfred, Stefan Aumann, Hans-Heinrich Ebeling, Hans-Reinhard Fricke, Peter Hoheisel, and Malte Rehbein. 1999-2. From digital archive to digital edition. *Historical Social Research* 24 (1): 101-44. doi: [10.12759/hsr.24.1999.1.101-144](https://doi.org/10.12759/hsr.24.1999.1.101-144).
- Thaller, Manfred. 2000-1. Historische Datenbanken. Vorteile und Probleme. *Geschichte und Informatik* 11: 7-25.
- Thaller, Manfred. 2000-2. *Vom verschwindenden Unterschied zwischen Datenbanken und Texten: Konsequenzen neuerer WWW Technologien am Beispiel von museumsnahen Datenbanken*. EDV-Tage Theuern 2000. Kolloquiumsbericht. München.
- Thaller, Manfred, ed. 2001-1. Die Handschriftenbibliothek des Kölner Doms im Internet. In *Codices Electronici Ecclesiae Coloniensis. Eine mittelalterliche Kathedralbibliothek in digitaler Form*, 21-39. Göttingen: Duehrkohp und Radicke.
- Thaller, Manfred. 2002-1 Von der Bibliothek zur Objektsammlung In *ABI Technik* 22 (1): 25.
- Thaller, Manfred. 2003-1. Historische Fachinformation & Historische Fachinformatik schwierige Beziehungen. *Information in Wissenschaft und Praxis* 6: 325-9.
- Thaller, Manfred. 2003-2. Kultur und Informatik Spannung und Symbiose. In *Perspektiven der Wissensgesellschaft*, ed. Eva Moser and Rainer S. Elkar. München: Kastner.
- Thaller, Manfred. 2003-3. Bemerkungen zu kunsthistorischen Informationssystemen; vornehmlich aus der Sicht der Informatik. In *Zeitenblicke – Online-Journal für die Geschichtswissenschaften* 2 (1). <<http://www.zeitenblicke.de/2003/01/thaller/>> (Accessed July 19, 2017).
- Thaller, Manfred. 2004-1. Texts, Databases, Kleio: A Note on the Architecture of Computer Systems for the Humanities. In *Digital Tools for the History of Ideas*, ed. Dino Buzzetti, Giuliano Pancaldi and Harold Short, 49-76. London: Office for Humanities Communication.

- Thaller, Manfred. 2004-2. Digital Manuscripts as base line for dynamic editions. In *Digital Technology and Philological Disciplines*, ed. Andrea Bozzi, Laura Cignoni and Jean-Louis Lebrave. In *Linguistica Computazionale XX-XXI*, 489-511.
- Thaller, Manfred. 2005-2. „Historische Fachinformatik“ – Ein Kölner Modell. In *hist 2003 Geschichte und Neue Medien*, ed. Daniel Burckhardt, Rüdiger Hohls and Vera Ziegeldorf. Berlin (= *Historisches Forum 7*, 2005) Teilband I. Persistent <http://edoc.hu-berlin.de/e_histfor/7_1/> bzw. http://edoc.hu-berlin.de/e_histfor/7_1/PDF/HistFor_7-2005-1.pdf>.
- Thaller, Manfred, and Claudia Loebbecke. 2005-3 Preserving Europe's Cultural Heritage in the Digital World. In *ECIS 2005 Proceedings*. 31.
- Thaller, Manfred. 2006-1. Wie ist es eigentlich gewesen, wenn das Gedächtnis virtuell wird? In *Forschung in der digitalen Welt* (= *Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt Hamburg 20*), ed. Rainer Hering, Jürgen Sarnowsky, Christoph Schäfer and Udo Schäfer. Hamburg.
- Thaller, Manfred. 2007-1. Was macht einen Quellentext für die Informatik ‚historisch‘? In *De litteris, manuscriptis, inscriptionibus*, Festschrift zum 65. Geburtstag, ed. Walter Koch, Theo Kölzer, Christian Friedl and Georg Vogeler. Wien / Köln / Weimar.
- Thaller, Manfred, Robert Neumayer, Hannes Kulovits, Eleonora Nicchiarelli, Michael Day, Hans Hofmann, and Seamus Ross. 2007-2. On the need for benchmark corpora in digital preservation. Proc. Of the 2nd DELOS Conference on Digital Libraries.
- Thaller, Manfred, Volker Heydegger, Jan Schnasse, Sebastian Beyl, and Elona Chudobkaite. 2008-1. Significant Characteristics to Abstract Content: Long term Preservation of Information. In *Research and Advanced Technology for Digital Libraries. ECDL 2008. Lecture Notes in Computer Science*, vol 5173, ed. Brite Christensen-Dalsgaard, Donatella Castelli, Bolette Ammitzbøll Jurik and Joan Lippincott. Heidelberg: Springer.
- Thaller, Manfred, Volker Heydegger, Jan Schnasse, Sebastian Beyl, and Elona Chudobkaite. 2008-2. XCL: The Extensible Characterisation Language – One Step towards an Automatic Evaluation of Format Conversions. In *Research and Advanced Technology for Digital Libraries. ECDL 2008. Lecture Notes in Computer Science*, vol 5173, ed. Brite Christensen-Dalsgaard, Donatella Castelli, Bolette Ammitzbøll Jurik and Joan Lippincott. Heidelberg: Springer.
- Thaller, Manfred, Christoph Becker, Andreas Rauber, Volker Heydegger, and Jan Schnasse. 2008-3. Systematic Characterisation of Objects in Digital Preservation: The eXtensible Characterisation Languages. *Journal of Universal Computer Science* 14 (18): 2936-52
- Thaller, Manfred, Christoph Becker, Andreas Rauber, Volker Heydegger, and Jan Schnasse. 2008-4. A generic XML language for characterising objects to support digital preservation. In *Proceedings of the 23rd Annual ACM Symposium on Applied Computing (SAC 2008)* 1, 402-6.
- Thaller, Manfred, Thomas Bürger, Claudia Fabian, Rupert Schaab, Barbara Schneider Kempf, and Heiner Schnellling. 2008-5. Das VD 18: eine Einladung ins 18. Jahrhundert. *Bibliothek Forschung und Praxis* 32 (2): 195.
- Thaller, Manfred, Marc Guttenbrunner, Jan Wieners, and Andres Rauber. 2010-1. Same Same But Different – Comparing Rendering Environments for Interactive

- Digital Objects. In *EuroMed 2010. Lecture Notes in Computer Science*, vol. 6436. Heidelberg: Springer.
- Thaller, Manfred, and Claudia Loebbecke. 2011-1. Digitization as an IT Response to the Preservation of Europe's Cultural Heritage. In *Emerging Themes in Information Systems and Organization Studies*, 359-72. Heidelberg: Springer.
- Thaller, Manfred, Daniel Ebner, and Jochen Graf. 2011-2. Virtual German Charter Network: *A Virtual Research Environment for the Handling of Medieval Charters*, presented at the SDH 2011: Supporting the Digital Humanities conference. <https://www.researchgate.net/publication/228533838_Virtual_German_Charter_Network_A_Virtual_Research_Environment_for_the_handling_of_medieval_charters> (Accessed July 19, 2017).
- Thaller, Manfred, and Dino Buzzetti. 2012-1. Beyond Embedded Markup. In *Digital Humanities 2012: Conference Abstracts*, ed. Jan Christoph Meister, 142-7. Hamburg: Hamburg University Press.
- Thaller, Manfred, Ana Nika, Tizianma Catarci, Yannis Ioannidis, Akrivi Katifori, Georgia Koutrika, Natalia Manola, and Andreas Nürnberger. 2012-2. A Survey of Context-Aware Cross-Digital Library Personalization. In *Adaptive Multimedia Retrieval. Context, Exploration, and Fusion AMR 2010. Lecture Notes in Computer Science*, vol. 6817, ed. Marcin Detyniecki et al. Heidelberg: Springer.
- Thaller, Manfred, ed. 2012-3. Controversies around the Digital Humanities: An Agenda. Special Issue of *Historical Social Research* 37 (3). ISSN 0172-6404.
- Thaller, Manfred, Sebastian Cuy, Martin Fischer, Daniel de Oliveira, Jens Peters, Johanna Puhl, and Lisa Rau. 2013-1. DA-NRW: Eine verteilte Architektur für die digitale Langzeitarchivierung. In *Nestor Edition Sonderheft 1 - Beiträge des Workshops „Digitale Langzeitarchivierung“ auf der Informatik 2013*, <http://nbn-resolving.de/urn:nbn:de:0008-2014012419>.
- Thaller, Manfred. 2014-1. The Humanities are about research, first and foremost; their interaction with Computer Science should be too. In *Computational Humanities - bridging the gap between Computer Science and Digital Humanities*, ed. Chris Biemann, Gregory R. Crane, Christiane D. Fellbaum and Alexander Mehler (Dagstuhl Seminar 14301) (= Dagstuhl Reports 4/7). <<http://drops.dagstuhl.de/opus/volltexte/2014/4792>>.
- Thaller, Manfred. 2015-1. Are the Humanities and Endangered or a Dominant Species in the Digital Ecosystem? In *AIUCD '14 Proceedings of the Third AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem*, ACM.
- Thaller, Manfred. 2016-1. Wege zu einer Informatik der Geisteswissenschaften. In *Digitalität. Theorien und Praktiken des Digitalen in den Geisteswissenschaften*. <<https://digigeist.hypothesen.org/77>> (Accessed July 19, 2017).
- Thaller, Manfred. 2016-2. Digital Humanities – eine Suche. *Forum Wissenschaft* 33 (4): 33-6.
- Thaller, Manfred. 2016-3. Was sind (keine) methodischen Implikationen der Digital Humanities? *Rechtsgeschichte – Legal History* 24: 335-6. <<http://dx.doi.org/10.12946/rg24/335-336>> (Accessed July 19, 2017).
- Thaller, Manfred. 2017-1. Geschichte der Digital Humanities / Digital Humanities als Wissenschaft. In *Digital Humanities*, ed. Fotis Jannidis, Hubertus Kohle and Malte Rehbein, 3-18. Stuttgart: Springer.

As Editor

- Thaller, Manfred, ed. 1984E. *Die Praxis der Quantifizierung in der österreichischen Geschichtsforschung*. Unterlagen der Arbeitssitzung des Arbeitskreises „Quantifizierung in der österreichischen Geschichtsforschung“ auf dem österreichischen Historikertag 1984 in Krems a. d. Donau. Göttingen: Ludwig-Boltzmann-Institut für Historische Sozialwissenschaften.
- Thaller, Manfred, ed. 1986E. *Datenbanken und Datenverwaltungssysteme als Werkzeuge historischer Forschung* (= *Historisch-sozialwissenschaftliche Forschungen*. Bd. 20). St. Katharinen: Scripta-Mercaturae.
- Thaller, Manfred, Ursula Klenk, and Peter Scherber, eds. 1987E. *Computerlinguistik und philologische Datenverarbeitung*. (= *Linguistische Datenverarbeitung*. 7). Beiträge der Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung 1986 in Göttingen. Olms u. a., Hildesheim u. a.
- Thaller, Manfred, Gerhard Botz, Christian Fleck, and Albert Müller, eds. 1988E. „Qualität und Quantität“. *Zur Praxis der Methoden der Historischen Sozialwissenschaft* (= *Studien zur Historischen Sozialwissenschaft*. Bd. 10). Frankfurt a. M. u. a: Campus.
- Thaller, Manfred, and Albert Müller, eds. 1989E. *Computer in den Geisteswissenschaften. Konzepte und Berichte* (= *Studien zur Historischen Sozialwissenschaft*. Bd. 7). Frankfurt a. M. u. a.: Campus.
- Thaller, Manfred, Heino Best and Ekkehard Mochmann, eds. 1991E. *Computers in the Humanities and the Social Sciences. Achievements of the 1980s, Prospects for the 1990s*. Proceedings of the Cologne Computer Conference 1988 Uses of the Computer in the Humanities and Social Sciences held at the University of Cologne, September 1988. München u. a.: Saur.
- Thaller, Manfred, ed. 1992E. *Images and Manuscripts in Historical Computing*. (= Halbgraue Reihe zur Historischen Fachinformatik A14). St. Katharinen.
- Thaller, Manfred, Leonid Borodkin, and John Turner, eds. 1995E. *Statistics for Historians: Standard Packages and Specific Historical Software*. (= Halbgraue Reihe zur Historischen Fachinformatik A26), St. Katharinen: Scripta-Mercaturae.
- Thaller, Manfred, and Hans-Heinrich Ebeling, eds. 1999E. *Digitale Archive. Die Erschließung und Digitalisierung des Stadtarchivs Duderstadt*. Göttingen: Max-Planck-Institut für Geschichte.
- Thaller, Manfred, ed. 2001E. *Codices Electronici Ecclesiae Coloniensis. Eine mittelalterliche Kathedralbibliothek in digitaler Form* (= *Fundus. Forum für Geschichte und ihre Quellen*. Beiheft 1). Göttingen: Duehrkohp und Radicke.
- Thaller, Manfred, ed. 2003E. *Digitale Bausteine für die geisteswissenschaftliche Forschung* (= *Fundus. Forum für Geschichte und ihre Quellen*. Beiheft 5). Göttingen: Duehrkohp und Radicke.
- Thaller, Manfred, ed. 2005E. *Retrospektive Digitalisierung von Bibliotheksbeständen. Evaluierungsbericht über einen Förderschwerpunkt der DFG*. Köln: Universität zu Köln.
- Thaller, Manfred, ed. 2009E. *The eXtensible Characterisation Languages – XCL*. Hamburg: Verlag Dr. Kovac.
- Thaller, Manfred, Simone Görl, and Johanna Puhl, eds. 2011E. *Empfehlungen für die weitere Entwicklung der Wissenschaftlichen Informationsversorgung des Landes NRW*. Berlin: epubli.

Thaller, Manfred, ed. 2012E. Controversies around the Digital Humanities Special Issue of *Historical Social Research* 37 (3). ISSN 0172-6404.

Thaller, Manfred, ed. 2013E. *Das Digitale Archiv NRW in der Praxis*. Hamburg: Verlag Dr. Kovac

Thaller, Manfred, ed. 2016E. Computer Science in Humanities. *it – Information Technology* 58 (2): 61-3.

As Series Editor

Halbgraue Reihe zur Historischen Fachinformatik

[HGR A1] BAND A1

Peter Becker, and Thomas Werner. 1991. *Kleio. Ein Tutorial*. St. Katharinen: Scripta Mercaturae.

[HGR A2] BAND A2

Thomas Werner, and Thomas Grotum. 1989. *Sämtlich Hab und Gut ... Die Analyse von Besitzstandslisten*. St. Katharinen: Scripta Mercaturae.

[HGR A3] BAND A3

Jürgen Nemitz. 1989. *Die historische Analyse städtischer Wohn- und Gewerbelagen: die Auswertung sozialtopographischer Quellen*. St. Katharinen: Scripta Mercaturae.

[HGR A4] BAND A4

Barbara Schuh. 1989. *"Von vilen und mancherlay seltsamen Wunderzeichen": die Analyse von Mirakelbüchern und Wallfahrtsquellen*. St. Katharinen: Scripta Mercaturae.

[HGR A5] BAND A5

Peter Becker. 1989. *Leben, Lieben, Sterben: die Analyse von Kirchenbüchern*. St. Katharinen: Scripta Mercaturae.

[HGR A6] BAND A6

Martin Gierl, Thomas Grotum, and Thomas Werner. 1990. *Der Schritt von der Quelle zur historischen Datenbank. StanFEP: Ein Arbeitsbuch*. St. Katharinen: Scripta Mercaturae.

[HGR A7] BAND A7

Josef Smets. 1990. *Créer une base de données historiques avec Kleio. (Avec un contribution de Nathalie Gardiol.)*. St. Katharinen: Scripta Mercaturae.

[HGR A8] BAND A8

Thomas Engelke, Jürgen Nemitz, and Carolin Trenkler. 1990. *Historische Forschung mit Kleio. Ergebnisse des ersten Treffens der Kleio-User-Group im Stadtarchiv Regensburg vom 18. bis 21. März 1990*. St. Katharinen: Scripta Mercaturae.

[HGR A9] BAND A9

Steffen Wernicke, and Martin Hoernes. 1990. *„Umb die unzucht die ich handelt han ...“: Quellen zum Urfehdedwesen*. St. Katharinen: Scripta Mercaturae.

[HGR A11] BAND A11

Daniel Greenstein, ed. 1991. *Modelling Historical Data. Towards a Standard for Encoding and Exchanging Machine-Readable Texts*. St. Katharinen: Scripta Mercaturae.

[HGR A12] BAND A12

- Donal Spaeth, Peter Denley, Virginia Davis, and Richard Trainor, eds. 1992. *Towards an International Curriculum for History and Computing. A Workshop of the International Association for History and Computing, University of Glasgow, 15-17 May 1992*. St. Katharinen: Scripta Mercaturae.
- [HGR A13] BAND A13
Jan Oldervoll, ed. 1992. *Eden or Babylon? On Future Software for Highly Structured Historical Sources*. St. Katharinen: Scripta Mercaturae.
- [HGR A14] BAND A14
Manfred Thaller, ed. 1992. *Images and Manuscripts in Historical Computing*. St. Katharinen: Scripta Mercaturae.
- [HGR A15] BAND A15
Leonid I. Borodkin, and Wolfgang Levermann, eds. 1993. *Istorija i komp'juter: Nove informacionnye tehnologii v istoriceskich issledovanijach I obrazovanii*. St. Katharinen: Scripta Mercaturae.
- [HGR A16] BAND A16
Jurij Fikfak, and Gerhard Jaritz, eds. 1993. *Image Processing in History: Towards Open Systems*. St. Katharinen: Scripta Mercaturae.
- [HGR A17] BAND A17
Virginia Davis, Peter Denley, Donald Spaeth, and Richard Trainor, eds. 1993. *The Teaching of Historical Computing: An International Framework. A Workshop of the International Association for History and Computing, University of London, 26--28 February 1993*. St. Katharinen: Scripta Mercaturae.
- [HGR A18] BAND A18
Optical Character Recognition in the Historical Discipline. Proceedings of an International Workgroup organized by: Netherlands Historical Data Archive, Nijmegen Institute for Cognition and Information. iv und 165 S., St. Katharinen 1993.
- [HGR A19] BAND A19
Kevin Schürer, and Herman Diederiks, eds. 1993. *The Use of Occupations in Historical Analysis*. St. Katharinen: Scripta Mercaturae.
- [HGR A20] BAND A20
Seamus Ross, and Edward Higgs, eds. 1993. *Electronic Information Resources and Historians: European Perspectives*. St. Katharinen: Scripta Mercaturae.
- [HGR A21] BAND A21
Leonid I. Borodkin, and Wolfgang Levermann, eds. 1993. *History and Computing in Eastern Europe*. St. Katharinen: Scripta Mercaturae.
- [HGR A22] BAND A22
Gerhard Jaritz. 1993. *Images. A Primer of Computer-Supported Analysis with Kleio IAS*. St. Katharinen: Scripta Mercaturae.
- [HGR A23] BAND A23
Matthew Woollard, and Peter Denley. 1993. *Source-Oriented Data Processing for Historians: a Tutorial for Kleio*. St. Katharinen: Scripta Mercaturae.
- [HGR A24] BAND A24
Laboratoire d'Etudes et de Recherches sur l'Information et la Documentation: Histoire et Informatique 1993. Une Bibliographie Internationale / History and Computing 1993. An International Bibliography. 170 S., St. Katharinen 1994.
- [HGR A25] BAND A25

- Michael Goerke. 1994. *Coordinates for Historical Maps*. St. Katharinen: Scripta Mercaturae.
 [HGR A26] BAND A26
- Leonid Borodkin, Manfred Thaller, and John Turner, eds. 1995. *Statistics for Historians: Standard Packages and Specific Historical Software*. St. Katharinen: Scripta Mercaturae.
 [HGR A27] BAND A27
- Herman Diederiks, and Marjan Balkestein, eds. 1995. *Occupational Titles an their Classification*. St. Katharinen: Scripta Mercaturae.
 [HGR A28] BAND A28
- Laboratoire d'Etudes et de Recherches sur l'Information et la Documentation: Histoire et Informatique 1994. *Une Bibliographie Internationale / History and Computing 1994. An International Bibliography*. 130 S., St. Katharinen 1995.
 [HGR A29] BAND A29
- Matthew Woollard, and Peter Denley, eds. 1996. *The Sorcerer's Apprentice: Kleio Case Studies*. St. Katharinen: Scripta Mercaturae.
 [HGR A30] BAND A30
- Peter Denley, ed. 1996. *Computing Techniques and the History of Universities*. St. Katharinen: Scripta Mercaturae.

Series B – Software Descriptions

- [HGR B1] BAND B1
 Manfred Thaller. 1992. *Kleio 4. Ein Datenbanksystem*. St. Katharinen: Scripta Mercaturae.
- [HGR B2] BAND B2
 Manfred Thaller. 1989. *Query Net I/O*. St. Katharinen: Scripta Mercaturae.
- [HGR B4] BAND B4
 Wolfgang Levermann. 1989. *CMATCH: Mustererkennung in Zeichenketten*. St. Katharinen: Scripta Mercaturae.
- [HGR B5] BAND B5
 Ingo H. Kropac, and Ursula Leiter-Köhler. 1991. *Kleio. Der Datenbankeditor*. St. Katharinen: Scripta Mercaturae.
- [HGR B6] BAND B6
 Kathrin Homann. 1990. *StanFEP. Ein Programm zur freien Konvertierung von Daten*. St. Katharinen: Scripta Mercaturae.
- [HGR B8] BAND B8
 Wolfgang Levermann. 1991. *Kontextsensitive Datenverwaltung*. St. Katharinen: Scripta Mercaturae.
- [HGR B9] BAND B9
 Jan Oldervoll. 1992. *CensSys. Eine Softwareumwelt für die Analyse historischer Daten*. St. Katharinen: Scripta Mercaturae.
- [HGR B10] BAND B10
 Gabriele Gross. 1992. *Kleio. Eine Einführung in die Menüsteuerung*. St. Katharinen: Scripta Mercaturae.
- [HGR B11] BAND B11
 Manfred Thaller. 1993. *Kleio. A Database System*. St. Katharinen: Scripta Mercaturae.

Series C – Databases as Editions

[HGRC 1] Band C1

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Die Vorauer Bücher Mosis: machine-readable texts*. St. Katharinen: Scripta Mercaturae.

[HGRC 2] Band C2

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Das St. Trudperter Hohe Lied: a machine-readable text*. St. Katharinen: Scripta Mercaturae.

[HGRC 3] Band C3

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Notker Labeo: Psalmen nach der Wiener Handschrift*. St. Katharinen: Scripta Mercaturae.

[HGRC 4] Band C4

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1990. *Deutsche Predigten des XII. und XIII. Jahrhunderts: maschinenlesbare Texte*. St. Katharinen: Scripta Mercaturae.

[HGRC 5] Band C5

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Altdeutsche Predigten II: die Oberaltaicher Sammlung: machine-readable texts*. St. Katharinen: Scripta Mercaturae.

[HGRC 6] Band C6

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Konrad der Pfaffe: Altdeutsche Predigten III: die deutschen Predigten des Priesters Konrad: machine-readable texts*. St. Katharinen: Scripta Mercaturae.

[HGRC 7] Band C7

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Altdeutsche Predigten und Gebete: machine-readable texts*. St. Katharinen: Scripta Mercaturae.

[HGRC 8] Band C8

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Speculum ecclesiae: eine frühmittelhochdeutsche Predigtsammlung: machine-readable texts*. St. Katharinen: Scripta Mercaturae.

[HGRC 9] Band 9

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Eilhart von Oberg: Tristrant*. St. Katharinen: Scripta Mercaturae.

[HGRC 10] Band 10

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *König Rother: nach der Heidelberger Handschrift: a machine-readable text / prepared by Roy A. Wisbey; revised by Astrid Reinecke*. St. Katharinen: Scripta Mercaturae.

[HGRC 11] Band 11

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1990. *Konrad der Pfaffe: Das Rolandslied*. St. Katharinen: Scripta Mercaturae.

[HGRC 12] Band 12

Wisbey, Roy Albert, and Astrid Reinecke, eds. 1990. *Lamprecht der Pfaffe: Das Alexanderlied*. St. Katharinen: Scripta Mercaturae.

[HGRC 13] Band 13

Jones, George Fenwick and Astrid Reinecke, eds. 1991. *Die Weingartner-Stuttgarter Liederhandschrift: machine-readable texts*. St. Katharinen: Scripta Mercaturae.

[HGRC 14] Band 14

- Bennewitz, Ingrid. Ulrich Müller, and Franz Viktor Spechtler, eds 1991. *Verskonkordanz zur Berliner Neidhart-Handschrift* [Staatsbibliothek Preussischer Kulturbesitz, Berlin, mgf 779], St. Katharinen: Scripta Mercaturae [HGRC 15] Band 15
- Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Otte: Eraclius*, St. Katharinen: Scripta Mercaturae. [HGRC 16] Band 16
- Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Rudolf von Ems: Barlaam und Josaphat*. St. Katharinen: Scripta Mercaturae. [HGRC 17] Band 17
- Wisbey, Roy Albert, and Astrid Reinecke, eds. 1991. *Wolfram von Eschenbach: Lieder, Parzival und Titurel*. St. Katharinen: Scripta Mercaturae. [HGRC 18] Band 18
- Kropač, Ingo H., 1990. *Das Urkundenbuch des Herzogtums Steiermark: eine Textdatenbank*. St. Katharinen: Scripta Mercaturae.

Cologne Contributions to Humanistic Specialist Informatics

- [KBGF 1] Band 1
Schlarb, Sven. 2008. *Unschärfe Validierung strukturierter Daten. Ein Modell auf der Basis unscharfer Logik*. Hamburg: Verlag Dr. Kovac.
- [KBGF 2] Band 2
Margulies, Simon B. 2009. *Digitale Daten als Quelle der Geschichtswissenschaft. Eine Einführung*. Hamburg: Verlag Dr. Kovac.
- [KBGF 3] Band 3
Thaller, Manfred, ed. 2009. *The eXtensible Characterisation Languages – XCL*. Hamburg: Verlag Dr. Kovac.
- [KBGF 4] Band 4
Delve, Janet, and David Anderson. 2012. *The Trustworthy Online Technical Environment Metadata Database – TOTEM*. Hamburg: Verlag Dr. Kovac.
- [KBGF 5] Band 5
Thaller, Manfred, ed. 2013. *Das Digitale Archiv NRW in der Praxis. Eine Softwarelösung zur digitalen Langzeitarchivierung*. Hamburg: Verlag Dr. Kovac.

Special References

Contributions within this HSR Supplement 29:
From History to Applied Computer Science in the Humanities

- Thaller, Manfred. 2017a [1980]. Automation on Parnassus. CLIO – A Databank Oriented System for Historians. *Historical Social Research Supplement 29*: 113-137. doi: [10.12759/hsr.suppl.29.2017.113-137](https://doi.org/10.12759/hsr.suppl.29.2017.113-137)
- Thaller, Manfred. 2017b [1984]. Ungefähre Exaktheit. Theoretische Grundlagen und praktische Möglichkeiten einer Formulierung historischer Quellen als Produkte ‚unscharfer‘ Systeme. *Historical Social Research Supplement 29*: 138-159. doi: [10.12759/hsr.suppl.29.2017.138-159](https://doi.org/10.12759/hsr.suppl.29.2017.138-159).

- Thaller, Manfred, ed. 2017c [1986]. Vorüberlegungen für einen internationalen Workshop über die Schaffung, Verbindung und Nutzung großer interdisziplinärer Quellenbanken in den historischen Wissenschaften. *Historical Social Research Supplement* 29: 160-177. doi: [10.12759/hsr.suppl.29.2017.160-177](https://doi.org/10.12759/hsr.suppl.29.2017.160-177).
- Thaller, Manfred. 2017d [1990]. Entzauberungen: Die Entwicklung einer fachspezifischen historischen Datenverarbeitung in der Bundesrepublik. *Historical Social Research Supplement* 29: 178-192. doi: [10.12759/hsr.suppl.29.2017.178-192](https://doi.org/10.12759/hsr.suppl.29.2017.178-192).
- Thaller, Manfred. 2017e [1991]. The Need for a Theory of Historical Computing. *Historical Social Research Supplement* 29: 193-202. doi: [10.12759/hsr.suppl.29.2017.193-202](https://doi.org/10.12759/hsr.suppl.29.2017.193-202).
- Thaller, Manfred. 2017f [1991]. The Need for Standards: Data Modelling and Exchange. *Historical Social Research Supplement* 29: 203-220. doi: [10.12759/hsr.suppl.29.2017.203-220](https://doi.org/10.12759/hsr.suppl.29.2017.203-220).
- Thaller, Manfred. 2017g [1992]. Von der Mißverständlichkeit des Selbstverständlichen. Beobachtungen zur Diskussion über die Nützlichkeit formaler Verfahren in der Geschichtswissenschaft. *Historical Social Research Supplement* 29: 221-242. doi: [10.12759/hsr.suppl.29.2017.221-242](https://doi.org/10.12759/hsr.suppl.29.2017.221-242).
- Thaller, Manfred. 2017h [1993]. The Archive on Top of your Desk. An Introduction to Self-Documenting Image Files. *Historical Social Research Supplement* 29: 243-259. doi: [10.12759/hsr.suppl.29.2017.243-259](https://doi.org/10.12759/hsr.suppl.29.2017.243-259).
- Thaller, Manfred. 2017i [1993]. Historical Information Science: Is there such a Thing? New Comments on an old Idea. *Historical Social Research Supplement* 29: 260-286. doi: [10.12759/hsr.suppl.29.2017.260-286](https://doi.org/10.12759/hsr.suppl.29.2017.260-286).
- Thaller, Manfred. 2017j [1995]. Source Oriented Data Processing and Quantification: Distrustful Brothers. *Historical Social Research Supplement* 29: 287-306. doi: [10.12759/hsr.suppl.29.2017.287-306](https://doi.org/10.12759/hsr.suppl.29.2017.287-306).
- Thaller, Manfred. 2017k [2001]. From the Digitized to the Digital Library *Historical Social Research Supplement* 29: 307-319. doi: [10.12759/hsr.suppl.29.2017.307-319](https://doi.org/10.12759/hsr.suppl.29.2017.307-319).
- Thaller, Manfred. 2017l [2005]. Reproduktion, Erschließung, Edition, Interpretation: Ihre Beziehungen in einer digitalen Welt. *Historical Social Research Supplement* 29: 320-343. doi: [10.12759/hsr.suppl.29.2017.320-343](https://doi.org/10.12759/hsr.suppl.29.2017.320-343).
- Thaller, Manfred. 2017m [2009]. The Cologne Information Model: Representing Information Persistently. *Historical Social Research Supplement* 29: 344-356. doi: [10.12759/hsr.suppl.29.2017.344-356](https://doi.org/10.12759/hsr.suppl.29.2017.344-356).

Curriculum Vitae:

Prof. em. Dr. Manfred Thaller

Personal Information

Born March 18th, 1950, in Feldbach, Austria
1978 – 1987 married with Ilse Klotzinger
since 1993 with Sigrid Amedick
no children

Address: Rosenhof 16b, D 50321 Brühl, Germany
manfred.thaller@uni-koeln.de

Education

1960 – 1964 Abteigymnasium Seckau
1964 – 1969 Bundesrealgymnasium Gleisdorf (Abitur)
[1969 – 1970 Military service]
1970 – 1975 PhD in Modern History at Karl-Franzens Universität Graz, Austria
1976 – 1978 Postdoctoral diploma in Sociology at Institute for Advanced Studies, Vienna, Austria

Employment

1978 – 1999 Research fellow (*Wissenschaftlicher Referent*) at *Max-Planck-Institut für Geschichte*, Göttingen, Germany
1997 – 2000 Professor *historisk informatik* and director of Humanities Information Technology Center at *Universitetet i Bergen*, Norway
2000 – 2015 Professor *Historisch Kulturwissenschaftliche Informationsverarbeitung* (Humanities Computer Science) at *Universität zu Köln*, Germany

Honors

1989 "*Univ. Prof. Dr. Harry Kühnel-Förderungspreis*" of the Johann Wolfgang Goethe-Stiftung, Basel
1987 Visiting professor Hebrew University, Jerusalem, Israel
1993 Visiting professorial fellow, Queen Mary and Westfield College, London
1993 Visiting professor Istituto Universitario Europeo di Firenze, Italy
2011 Google European Digital Humanities Award

Teaching

Between 1982 and 1997	Seminars and lectures at the Universities of Köln, Germany (1982), Graz, Austria (1983/84, 1985); Göttingen, Germany (1984/85, 1985/86, 1986, 1986/87, 1987/88, 1988, 1988/89, 1989, 1989/90, 1990, 1990/91, 1991, 1991/92, 1992, 1993, 1994, 1994/95, 1995, 1996, 1997), Siegen, Germany (1984/85), München, Germany (1985, 1988, 1988/89, 1990, 1990/91, 1994, 1994/95, 1996, 1996/97), Hamburg, Germany (1985), Wien, Austria (1985/86, 1987), Salzburg, Austria (1986, 1988), Jerusalem, Israel (1987), Queen Mary and Westfield, London (1992/93), Firenze, Italy (1993/94), Bergen, Norway (1995/96, 1996, 1996/97, 1997)
1979 – 1999	Courses taught at the Linz / Salzburg / Bergen summer school
1987 – 1992	Organized and taught at a summer school in Göttingen
1992 – 1996	Organized and taught at a summer school in Moscow
2005 – 2008	Courses taught at digital preservation summer schools in Sophia Antipolis, France (2005), San Miniato, Italy (2006), Pisa, Italy (2007), Tirrenia, Italy (2008)

Grants Received

Received substantial 3rd party funding for approximately two dozen research projects of 50.000 € or – sometimes substantially – more. Main sources of funding received: German National Research Council (DFG), *Volkswagenstiftung*, European Commission, various German ministries.

Professional Organisations

1980 – 1992	QUANTUM – Association for Quantification and Methods in Historical and Social Research, Member of the Executive Board
1989 –	Journal <i>Historical Social Research</i> (HSR), Cooperating Editor
1990 – 1994	Association for History and Computing, president
2012 – 2015	Digital Humanities im deutschsprachigen Raum, Vorstandsmitglied

Advisory Functions Funding Programs

Regular reviewing for *Volkswagenstiftung*, *Deutsche Forschungsgemeinschaft* (DFG), Alexander von Humboldt Foundation, European Commission funding calls, Austrian Science Fund (FWF), Royal Netherlands Academy of Arts and Sciences (KNAW), Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), Swiss National Science Foundation, Social Sciences and Humanities Research Council of Canada

Funding Advisory Board Memberships

2000 – 2008	Member library committee (<i>Bibliotheksausschuss</i> later AWBI) of the DFG (German National Research Council)
2001 – 2013	Continuous reviewing for German Academic Exchange Service (DAAD)

- 2010 – 2017 Strategic Advisory Board German Ministry for Education and Research (BMBF)
- 2013 – 2016 Evaluation committees for the Belgian Research Action through Interdisciplinary Networks of the Belgian Science Policy Office
- 2013 – 2017 Advisory board for Horizon Societal Challenge 6 "Europe in a changing world – inclusive, innovative and reflective societies" of the European Commission
- 2014 – 2016 Jury for Digital Humanities proposals Austrian Academy of Science
- 2014 – 2018 *Evaluationspanel für geisteswissenschaftliche Forschungsinfrastrukturen* of the Swiss National Science Foundation