Updating a Biomedical Database. Writing, Reading and Invisible Contribution

David Pontille

David Pontille. Updating a Biomedical Database. Writing, Reading and Invisible Contribution. Anthropology of Writing: Understanding Textually-Mediated Worlds, Continuum, pp.47-66, 2010. halshs-00575215

HAL Id: halshs-00575215
https://shs.hal.science/halshs-00575215

Submitted on 9 Mar 2011

Updating a Biomedical Database
Writing, Reading and Invisible Contribution[*]

David PONTILLE

chargé de recherche CNRS
pontille@ehess.fr

The development of information and communication technologies has multiplied our ability to produce, circulate and store large amounts of data. Whether for business use in managing stocks, suppliers and customers, or for administrative purposes of categorizing and classifying populations, databases are becoming increasingly important and their reliability is crucial. Over the last twenty years databases have become an essential part of biomedical research (Bowker 2000; Hine 2006). As a result of developments in molecular biology and informatics it is now possible to undertake genetic analysis of large amounts of data. For these databases to operate effectively, a link has to be made between very small amounts of biological material (only a few microlitres) and a wide range of personal data relating to the donors (age, sex, occupation, lifestyle, diet, etc.) and their state of health (clinical and biological data).

Yet most studies on bioinformatics databases take this link for granted, as if it emerged naturally and automatically from the data collection process. However, when we focus on the process whereby databases are made, the picture becomes more complicated. The relationship between samples and data does not emerge in and of itself. It is the result of the daily work of writing, and it is this point I wish to address. How, in actual concrete terms, does one produce a bioinformatics database? How does one forge a material link between different types of data? How does one ensure that this information is reliable and robust?

A previous interdisciplinary research project on several biomedical databases in France and the issues involved in different forms of organization was a first attempt to answer some of these questions (Pontille et al. 2007). I shall go further here by examining a specific case: writing practices devoted to the updating of a bioinformatics database that combines clinical, biological and radiological data sets. In these particular circumstances, the link between inscriptions and samples has to be reconstituted. Thus, I will analyze the day-to-day production of the link between several sorts of data as one action within a larger chain of activities.

This analysis gives me the opportunity to pursue several lines of investigation that have emerged from studies situated in the overlap between the anthropology of writing and the ethnography of work.

The first involves understanding how writing produces robust links between scientific propositions and data sets. Here I shall follow several laboratory studies carried out within the anthropology of sciences which have drawn particular attention to the major role played by graphic representations in scientific work (Lynch and Woolgar 1990). In this point of view, the notion of 'inscriptions', that 'summarises all traces, spots, points, histograms, recorded numbers, spectra, peaks and so on' (Latour and Woolgar 1979, p. 88), is particularly relevant to analyze the different writing activities that occur in scientific laboratories and many others workplaces.

The studies carried out by the *Language and Work* network have focused on the interweaving of productive work and language activities (Grosjean and Lacoste 1998; Borzeix and Fraenkel 2001), in reconsidering specific aspects of written performativity (Fraenkel 2006). The second line of investigation will be to consider written documents as a resource for action, but also to interrogate the specific actions of writing, and what they make it possible to do and to make happen.

Finally, several studies have emphasized the way in which the work of laboratory technicians who actually use the instruments has become devalued (Shapin 1989; Barley and Bechky 1994; Timmermans 2003). Such studies have generally shed light on the ways in which the technicians' contribution to scientific work may be represented or erased in the final published papers. I shall follow this third approach in order to stress out the active role played by these technicians in actually producing information through their daily involvement in writing practices.

I shall, first, briefly describe the present case study and the methodologies used, before looking specifically at three main issues:

1. Data management: What is done with the written data? How is it actually handled? And for what purposes?

2. Data processing: What is the graphic and textual information used for? What kinds of reading do they undergo? What combinations are made between them to produce other documents?

3. Writing and action: What links are made on the basis of these documents? What do they produce? What transformations do they undergo?

This will then make it possible to see how some of these documents are evaluated compared to others and the scientific weighting attributed to them.

The case study: constructing a patient cohort

At the beginning of the 1990s, a team of French doctors began to be interested in the predictive factors for a joint disease. To establish the precipitating factors for the disease and identify the most severe cases, they began with the premise that they had to gather a set of data (clinical, biological and radiological) at the outset of the disease in order to follow the course of its development in patients for a sufficiently long time. They therefore set about recruiting a cohort of individuals who were willing to be examined once a year.

Various types of data were collected between 1992 and 2002. First, at patients' annual health check X-rays of all painful joints were made and archived. Then patients underwent a clinical examination designed to elicit a set of data which was recorded into especially fashioned 'standardized research booklets' for each patient devoted to the cohort. These contained a detailed 16 pages questionnaire on patients' physical state (e.g. feelings of fatigue, severe pain, ability to move around) and mental state (e.g. feelings of being a burden to others, insomnia), as well as their daily activities (e.g. work activity, social life, family support, nervous tension). It also contained instances of graphic information (e.g. a cross placed on a scale of values, painful joints circled on a drawing of a limb, hand or foot). Finally, the participants gave a blood sample from which the medical team isolated two biological samples: blood serum itself and DNA.

The aim of setting up the cohort was two-fold. First, for clinical purposes, to enable doctors to make correct diagnoses and hence find therapeutic solutions. Secondly, the collection of data was a potentially powerful tool from the point of view of scientific research. Having access to a bioinformatics database held out the potential for formulating and testing a range of hypotheses using innovative techniques for detecting illness. It was a way of obtaining original results which could be published in major medical journals.

Clinical practice was thus linked to biological research[2]. This relationship involved a particular organization of work within the rheumatology unit of the hospital. First, the contribution of all personnel – the work of the head of department, the interns, the nurses, and the technicians – had to be coordinated. Then adequate procedures had to be established in the rheumatology unit for dealing with patients receiving a series of treatments. At the annual health checks, nurses accompanied the patients to X-ray and took their blood samples. For their part, the doctors carried out clinical examinations and wrote a range of data into the research booklets. Finally, a room in the rheumatology unit had to be fitted out for analyzing and storing biological samples, another made available for storing research booklets and another for setting up the bioinformatics database.

---

[2] See P. Keating et A. Cambrosio (2003) on the changes in clinical activities brought about by advances in molecular biology.

An ethnography of writing

At the time of my study, new patients were no longer being added. The stage of data collection from the 880 patients over the period 1992-2002 had come to an end. The biological samples were stored in secure fridges, the radiological data were archived in each patient file, and the clinical data were inscribed in research booklets according to a standard procedure. However, the database associated with this material had been in a state of limbo for a long time. The clinical staff's lack of time and the absence of specific funding for its upkeep had considerably delayed the project, even though an initial database had been started when the patients were enrolled[3].

The bioinformatics database had to be updated in order to be functional. The main aim was to construct links between the clinical data, the radiological data and the biological samples. The different data sets had to be organized according to the same criteria and brought together in a single material place (the database). In 2004, a laboratory technician called Kelly, who was trained in biology and had a complementary degree in informatics, was engaged by the medical team to devote herself exclusively to this work, and she set about keyboarding the data.

There were two imperatives that controlled the updating work. Kelly had to compile the bioinformatics database in conformity with regulations which had not been in place when it was originally set up. Since then regulation at stake had been pointed out and a number of increasingly detailed legal rules had been introduced (Pontille et al. 2007). Also, Kelly had to work relatively quickly to make the data available so that it could be used to produce innovative findings which could be published in international journals.

> 'The stage we're at now, we've got five years to make use of this cohort. Five years from now, things will have moved on; the therapies and the questions will have changed. So we've got five years. We have to hurry up and make use of it, publish and get our work known' (Database manager).

Here I shall focus particularly on the work carried out by Kelly. At the beginning of my research, in April 2005, the data on 600 patients was almost complete. I therefore observed the updating of data on 280 patients. This investigation was part of a larger interdisciplinary research project focused on five biomedical databases in France (Pontille et al. 2007). The ethnographic fieldwork was carried out at the same time to emphasize different forms of organization between these biomedical databases, especially selected in order to include a range of sizes, institutional contexts and developmental stages. Such a selection was made possible by the presence in the interdisciplinary project of a geneticist particularly informed with biobanking activities in France. She introduced me to the team of doctors interested in the genesis of joint diseases.

In this particular case, I started with an interview with the principal investigator of the patient cohort who finally showed me round his hospital unit. I then negotiated to stay near Kelly during her work in order to be familiar with the biodatabase updating process. Afterwards I followed Kelly during a three-weeks period to understand her different activities, which are partly shared with several people, and involve a variety of tools and locations. Concretely, I sat down during hours near Kelly while she was facing her computer screen, reading research booklets and compiling data from juxtaposed texts on her office desk. I also followed her in the several rooms of the rheumatology unit she went during the updating process to identify the range of documents and of devices she systematically relied on.

---

[3] This is the main preoccupation of most bioinformatics database managers. They have to constantly find grants to cover their use and maintenance: 'We need people to carry out these specific tasks. And this is not part of the clinical staff's daily work, so we have to get separate research funding' (Database manager).

Finally, I made a regular collection of written documents that she used in the course of her day-to-day activity.

Such an ethnography over time enabled me to study in detail the various elements that Kelly made use of in order to update the database, to note the constraints on her work as they emerged over time and to bring to light the range of writing activities that she engaged in. I supplemented my observations with photographs, interviews with several people involved in the clinical research project (the principal investigator and some of the doctors involved in setting up the cohort), and a sample of published articles based on this bioinformatics database.

Following Kelly through the course of her daily work showed how she contributed to producing the database. Apart from having her own space in the laboratory attached to the rheumatology unit and a computer dedicated to her work, she had priority access to the locked room where research booklets had been stored since 1992. She had everything she needed for rapid updating of the data. All she needed to do was take the research booklets one at a time, read the contents, and, using the software, inscribe the data into the appropriate tables of the bioinformatics database: one 'clinical datum', one 'biological datum', two 'radiological data', and one 'patient' whose description included personal data (e.g. sex, date of birth, marital status, educational level, number of screenings, presence of serum or DNA samples).

However, these facilities were far from sufficient. Circumstances had made it impossible to record all the data on the research booklets. The patient cohort was spread in time between the years 1992 and 2002, between different parts of the hospital rheumatology unit, and between a number of doctors and nurses some of whom had changed hospitals over the course of the project. Also the people who had seen the patients initially were not always the same ones who had written in the research booklets. The quality of the data varied markedly according to people's availability and commitment (some doctors had written their theses on this cohort). There was also variation in quantity: replies to the questionnaires were more complete for some patients than others; the data were inscribed more conscientiously by some doctors than others.

Contrarily to what one might think, updating the database therefore did not require just an office, a computer and the research booklets. As we shall see, Kelly regularly made use of other elements in order to carry out her task. In particular, she had one essential tool. Instead of updating the database directly into the computer as she went along, she first wrote the relevant data by hand on an 'intermediate form' designed by the team. We shall see the reason for this additional step later, but for the moment let us note merely that this form was part of the updating process. Each intermediate form, consisting of four A4 sheets, was a summary of the results of each annual health checks taken from the research booklets. This was a crucial first stage in stabilizing the information and preparing the updating of the bioinformatics database.

How did Kelly actually go about filling in the 'intermediate form'? How did she deal with the multiple inscriptions in the research booklets? How did she manage to make links between the different types of data?

Handling the documents

Kelly did not start by filling in the intermediate forms. She began by organizing the different spaces that constituted her work place. Ever since 1992, the research booklets had been stacked in one of the staff rooms. Every time a patient attended an annual health check, the nurses and doctors would fetch their research booklets in order to fill in them. While the cohort was being set up, the research booklets would be in daily circulation between the storage room and the rest of the rheumatology unit and so would be removed from their storage boxes on a regular basis. The nurses would replace them more or less promptly

depending on how busy they were and how many emergencies arose. Over the ten years of patient follow-up, an increasing disorder had encroached on the apparently orderly organization of boxes neatly aligned on the shelves (figure 1).

Figure 1. The stored research booklets room

The first thing Kelly did when she arrived was to spend several hours carefully sorting out the research booklets. This classification was twofold. First, she began by placing the research booklets for each patient in chronological order in box files. Then she arranged the box files in alphabetical order. This handling activity did not entail either reading or making sense of the research booklets, but merely noting the dates of the health checks and the patients' names – although the activity did require some concentration. By placing the research booklets in the box files and then arranging the latter, Kelly was imposing a spatial organization on the documents.

The importance of this classification was seen whenever research booklets mysteriously went missing. Doctors and nurses sometimes used them when following up former patients from the cohort. Kelly's first recourse was to look for the missing research booklet in the other box files in case it had been misfiled (some patients had similar surnames). If she found it in the wrong box she would sigh over her colleagues' failure to take due care with the documents they dealt with every day: 'I don't believe it! It was hiding in here!'

In these action sequences involving handling and manipulation, the documents are not texts to be read. Kelly saw them as sheer objects which she carried around, moved, put in piles, sorted and arranged. By regarding documents as written objects, Kelly took into account the range of their material supports, here specifically the research booklets and the box files[4]. It is important to emphasize this materiality of writing. Through her handling activities Kelly became increasingly familiar with the tools of her trade: updating the bioinformatics database presupposes an intimate knowledge of the various elements which constitute it. Hence the

---

[4] In other action sequences not discussed here, Kelly manipulated different written objects: she carried around patient files, handled X-rays, sorted out cardboard files, and put forms away in plastic cases.

documents are not simply textual resources. As material objects, they form an essential part of the working environment. In many work situations the fact that documents are objects which are produced and manipulated is just as important as their textual content (Latour and Woolgar 1979; Latour 1986; Grosjean and Lacoste 1998; Pontille 2006; Denis and Pontille 2008).

However, the handling of the written objects was not purely a manual task. By moving the research booklets and classifying box files, Kelly performed a stabilization of her working environment (Conein and Jacopin 1993). The shelves and the box files constituted a visual memory. Minimal as they were, these paper technologies produced an additional set of landmarks, enabling Kelly to put in order the different elements that formed the material infrastructure of the bioinformatics database.

The handling of the documents also required the ability to project oneself into the future. By arranging a spatial distribution of the data inscribed in different material supports, Kelly was preparing and anticipating future activities: the documents were organized in a particular way to produce a resource space rendered accessible through efficient routines (Kirsh 1995). Once every element was in its place, Kelly was able to devote herself to fill in the intermediate forms.

Forms of reading in order to write

The challenge for Kelly lay in condensing the data contained in the various material supports into a single form (the intermediate forms) in order to transfer it into the computer database. Although doctors assumed that the data gathering stage was complete, Kelly still had to engage in an active process of data collection. Filling in the intermediate forms involved searching, sorting and selecting the data contained in the research booklets. This selection therefore involved a change of attitude to the documents: Kelly was now fully focused on their textuality. In the case of both biological and clinical data, she needed to know what she was looking for and know how to read them. But let us be clear as to what this reading entails. It is far from being an obvious and unambiguous activity. On the contrary, Kelly adopted several ways of reading in order to identity the relevant data.

In some cases the reading was a rapid scan and a visual sorting to locate the essential data. Her gaze swept over the content and came to rest at precise points which formed landmarks thanks to the standardized presentation of the written notes. This reading could be done while standing up and continuing to pay attention to interactions with nurses and doctors who might be in the same room.

Kelly carried out this form of reading in order to pick out the results of biological test which might not always be written up in the research booklets. In such a case, she consulted the patient file where they were stored along with other documents[5]. Kelly rapidly scanned the biological results and picks out the relevant ones. Constant practice had given her a trained eye. For each patient she was able to rapidly distinguish the results that related to the cohort criteria from those that did not. The fact that the results were presented in the form of lists, tables or standardized formulas meant that the reading process could be structured and the contents scanned to decide which data need to be collected (figure 2).

The organization of the graphic space was also an important resource for action. Kelly held the biological results in one hand and fill in the requisite spaces on the intermediate form with the other. Her reading was inextricably linked to writing: while she scanned the graphic space of the documents, her attention was focused on the end of her pen. As soon as she identified a relevant biological result, she copied it down.

---

[5] See M. Berg and G.C. Bowker (1997) for a detailed analysis of the recording procedures that take place within patient files.

Figure 2. Reading in the office
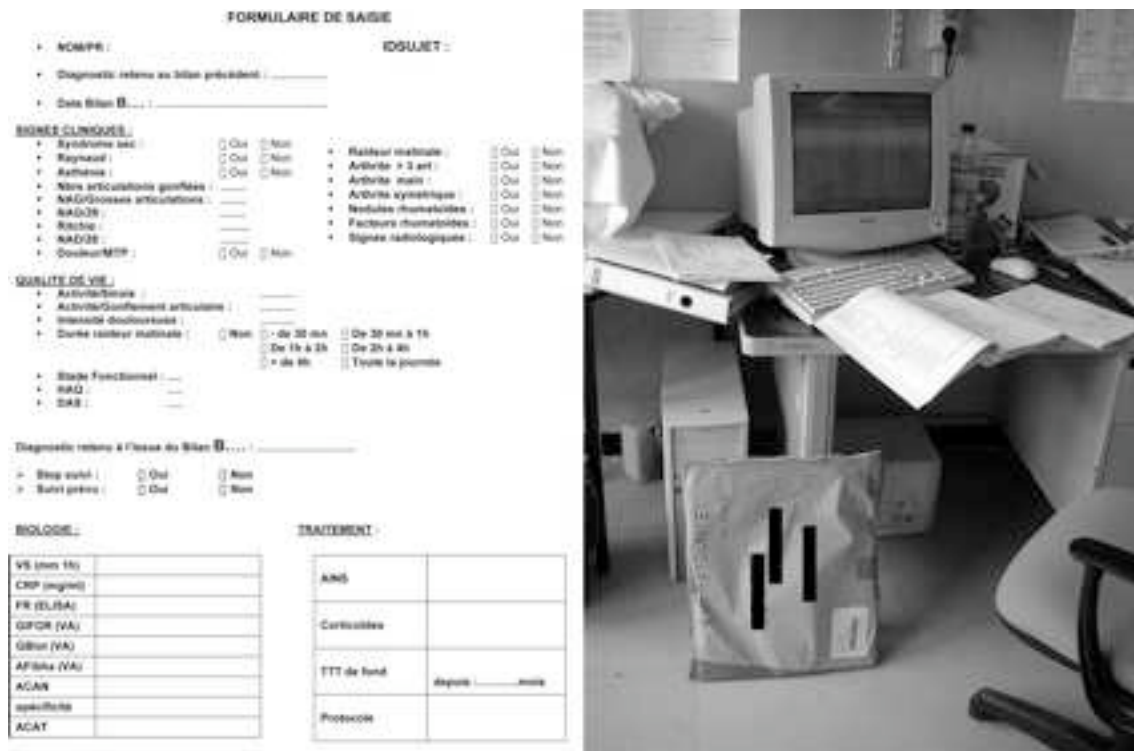
For other tasks, the reading was more detailed. It presupposed a studious concentration, a meticulous attention to each word, and might require several rereadings. This was the case when Kelly had to extract clinical data. How did she proceed? As figure 2 shows, she opened a research booklet at the annual health check page, placed the intermediate form beside it on the desk, and opened the computer database at the entries dealing with that patient while keeping the medical notes in their cardboard file nearby.

Here, reading involved first of all deciphering the doctors' handwriting. They tend to fill in the health check results rapidly and often use abbreviations which make reading more complicated. Kelly was sometimes unable to make them out. In order to update the data, she had to use her wits. She often turned to other written objects where the writing is more legible.

> 'Well, I have to read the letters sent to the patient [at the time of that health check] because it's written so illegibly, and the same for the treatment. Life would be a lot easier if they would all fill in the research booklet properly' (Kelly).

Here again the patient file was a resource: it contained all the letters sent to the patient. This correspondence provided data on how the pains were developing, the patient's state of health at the time of the consultation, the treatment prescribed and any test results.

But reading also involves interpreting. Without an understanding of what was written, the content would have been meaningless and Kelly would not have been able to process it correctly, that is, use it to fill in the intermediate form. The challenge was to study the data inscribed in the research booklets, evaluate its coherence and establish its meaning. This discernment is largely local: it is inscribed within the normative system of a group of professionals involved in the same clinical research. The shared writing objects, practices and common experiences are the basis for an understanding of the documents. The shared professional terminology provides specific terms for evaluating the patient's state of health, the progress of the illness and the medical treatment provided. Thus the ability to read is

closely linked to competencies based on medical knowledge situated in a specific configuration of several documents.

Finally, reading also involved checking that the data made sense. Kelly's practiced eye enabled her to recognize the note-writer (around ten of them between 1992 and 2002) and to attribute a relative weighting to the data contained in the research booklets. Kelly did not attach the same importance to all inscriptions to the research booklets, but evaluated them in terms of their enunciative authority.

> 'I think it must have been the technician before me who wrote that in the research booklet. I don't like it so much. I think it should be a doctor!' (Kelly).

Identifying the handwriting played an important role. The value of an inscription varied according to whether it was written by a technician, a junior doctor in training, a fully-fledged doctor or a noted mentor. Kelly attributed more or less weight to an inscription depending on the institutional hierarchy of the writers. Just as in medical diagnosis (Cicourel 1990), the structure of authority relations between the staff plays a central role.

Thus Kelly was engaged in an activity that involved different forms of reading that were closely linked to writing. Systematically, she brought the different elements together, compared the research booklets with parts of the patient files and compiled items of data from correspondence. All these activities were necessary in order to copy, report, transcribe or inscribe data in the intermediate forms. In carrying out her meticulous work, Kelly selected and interpreted material previously provided by the doctors in order to produce a new link. But the link she forged by writing had to be sufficiently robust to enable one to navigate between the clinical, radiological and biological data sets.

Producing reliable and durable information

In looking at the content of different forms of writing, Kelly was not merely transferring data from one support to another. Along the way, she was producing information with different kinds of validity. As we shall see, Kelly transformed the data into reliable scientific and medical information, she produced legal documents and she fashioned a material link by modifying the status of the elements which made up the database. The performativity of writing was crucial throughout this chain of transformations.

At several points, Kelly converted the inscriptions into figures and carried out calculations of certain biological parameters. That was the case when she was carrying out a recount of the number of swollen joints in order to convert it into an joint index: one of the sheets in the intermediate form contained a graphic representation of the human body showing the limbs, left and right, and including a calculation of the joint index (figure 3). Checking these calculations was not just a linear process, it involved going back and forth between different data sets. It involved numerous adjustments which were not solely reducible to copying or transcribing.

If she was in doubt about an annotation or a calculation, Kelly began reading, rereading, checking, and redoing calculations. As she went along, she tried to verify it from other sources. Her biomedical knowledge was essential for carrying out such calculations, but it was linked to skills which were heavily dependent on the ability to manipulate writings: gathering, compiling, sorting, checking, examining, and comparing (Latour 1986). During this cross-checking, Kelly was trying to ensure that the data were medically coherent and hence make sure that they were valid.

Updating the bioinformatics database was thus not an automatic process. Not only did Kelly gather inscriptions scattered among many different written objects, she also systematically made judgements of their value before writing them on the intermediate forms. The main issue here was to transform the data into reliable scientific and medical information.

Figure 3. The intermediate form: calculation of a polyarthritis index

The intermediate forms were actively involved in this production. They served to give spatial distribution to the set of data compiled and condensed by Kelly within a single graphic space. The intermediate forms became the sole support for the process of updating the computer database. The intermediate place of these forms within the chain of production can be simply stated here: they were a replacement for the written objects used previously (box files, research booklets, patient files, correspondence, the results of biological test, X-rays) and formed the starting point for the production of original information which could be published in specialized journals.

But the intermediate forms also had another purpose. By filling them out, Kelly was storing the data according to the requirements of the National Commission on Data Protection (Commission Nationale de l'Informatique et des Libertés). She gave each patient in the cohort a number, which appears on the intermediate form as well as in the bioinformatics database, in order to preserve anonymity. She then stored each intermediate form in a cardboard folder with all the other data on the same patient. Finally, she placed this folder in a locked metal filing cabinet reserved for the cohort.

This particular use of the intermediate forms was no longer concerned with the production of scientific and medical information. Kelly produced here a document with an explicitly legal function: if necessary, it would prove that the cohort set up for biomedical research purposes conformed to existing legislation and respected the patients' rights. In terms of the French legal system, it is absolutely necessary to have a paper version of the database. It is the only place where the biomedical data and the patient's name are found together, whereas the electronic version of the database, on which all subsequent treatment is based, must be completely anonymous.

In inscribing the information on the same physical record, the writing also transformed the status of the individuals. They were not only patients who circulated within the hospital service, they also became entities forming part of the bioinformatics database. The attribution of numerical codes and the systematic inscribing of data was not merely a writing gesture. It

was an act which impinged directly on persons (Bowker and Star 1999; Fraenkel and Pontille 2006). Here, it consisted of addressing them in a specific way by establishing their medical needs, and it influenced the course of their lives by obliging them to attend annual health checks. They were hospital patients, but they were also simultaneously categorized as individuals suffering from joint disease who featured in the bioinformatics database. The inscriptions made initially by the doctors and nurses, which Kelly then reworked, completed and archived, transformed their status permanently. It was through the intermediary of writing that they were constituted as members of the cohort and that their identity is newly organized through multiple inscriptions.

One last operation completed this process. After checking for missing data and correcting anomalies, Kelly decided to backup the first version of the database on a CD-ROM.

> 'It would be a good idea to make a back-up on CD so we have a clean and reliable copy. If we do this, if we have a problem at any time in the future, we will have a copy of the database that we can rely on' (Kelly).

The purpose of this saving, which was in addition to the backups made regularly by the hospital's computer server, was to have the same information available for all the analyses. The recording and storage were designed to give material form to the relationships that had been established between different pieces of information (Latour and Woolgar 1979; Berg 1996). As a transportable object, the CD-ROM played an active part in the construction of a material link between clinical, radiological and biological data sets. It formed the culmination of the process of updating the database.

In doing this, it also formed part of the rationalization of the process since it contained no trace of the many doubts, hesitations, and corrections which Kelly had laboriously overcome when making the database. It thus erased the whole set of day-to-day activities involved in updating it (figure 4).

Figure 4. The making of a bioinformatics database
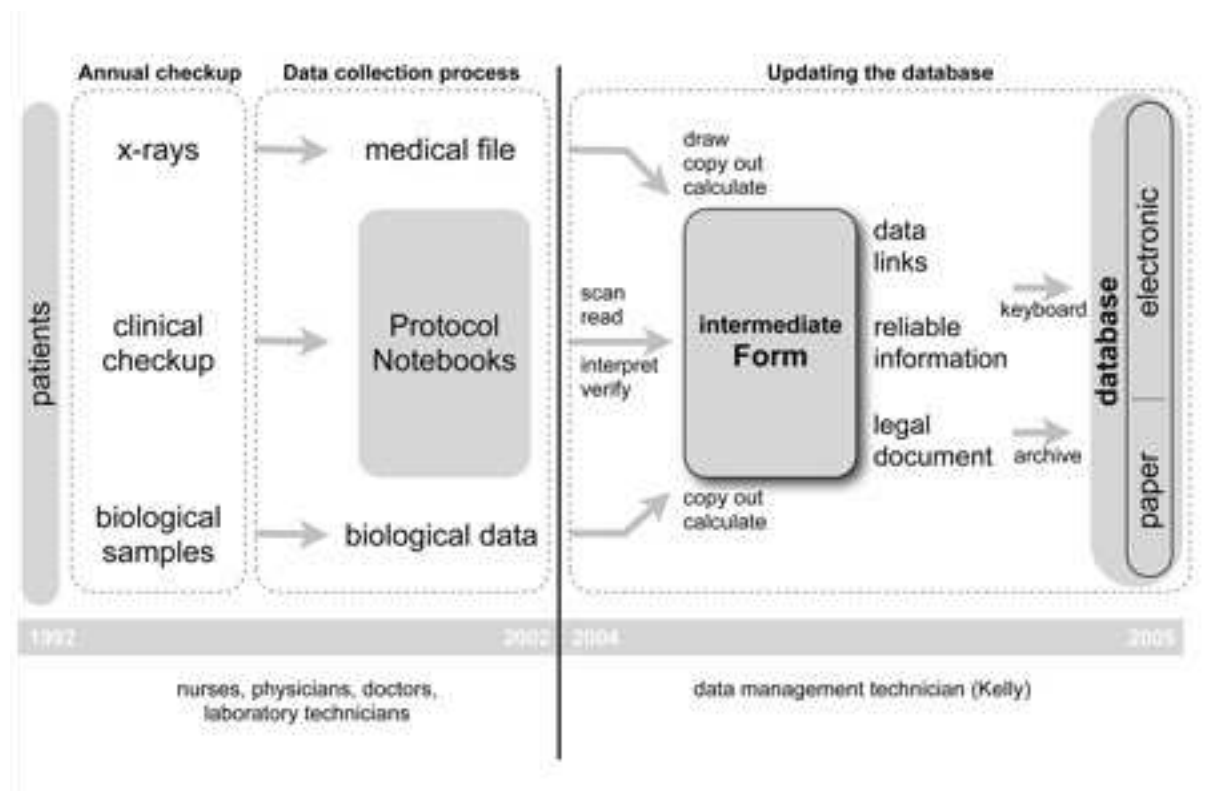
Invisible workers of writing and scientific contribution

This erasure of the conditions of production of the database should not, however, reduce invisibility to a one-dimensional phenomenon. On the contrary, observing Kelly go about her daily work enables to emphasize that there are various forms of invisibility.

Kelly was mainly visible during her handling activities. She regularly carried around research booklets, patient files and X-ray plates from place to place within the rheumatology unit of the hospital. This carrying around was of course part of her work. In the eyes of her colleagues, it drew on reserves of physical energy and needed few specific skills (Shapin 1989), although handling the documents was not reducible to this. Kelly used it to familiarize herself with her tools, organize her work environment and anticipate her future activities.

When she was in front of her computer, her work generally consisted of a simple and repetitive task: physically keyboarding the data in the database. In common with many other laboratory technicians (Barley and Bechky 1994), the importance of her activity remained largely invisible and unrecognized. However, Kelly participated actively in the process of data collection by virtue of different kinds of reading. Her expertise was closely linked to the way she used the graphic space of documents and to her ability to navigate between various written resources in order to act properly.

There was an even more invisible stage: Kelly contributed directly to the production of new knowledge by her daily writing work. Day after day, she filled in numerous gaps in the data contained in the research booklets which were supposedly her only source material. As she wrote, Kelly checked, corrected and filled in any missing cases. She forged a robust link between different items of clinical, biological and radiological data sets.

Of course her contribution was seen as essential: she was engaged to carry out work which no-one else within the rheumatology unit had sufficient time to do. Her employer also saw her as being necessary for the long term.

> 'At the moment I'm trying to think of a way of renewing my data management technician's temporary contract. This girl has had training in basic statistics, I need to keep her. So I'm now looking for funding to extend her contract' (Database manager).

However difficult, meticulous and crucial Kelly's writing work may have been, it was still reducible to making a stable bioinformatics database. All her writing activities were directed towards preparing the database so that it could be used by clinicians. As a data management technician Kelly was responsible for producing multiple *inscriptions* (gathering, keyboarding and manipulating data, carrying out statistical analyses, drawing up intermediate summaries…). In no way was she judged to be in a position to write *texts*, that is, articles publishable in scientific journals, like the doctors she worked with. The way the latter saw her contribution confirmed this point:

> 'The authors acknowledge the contribution of Kelly Whitehand as a clinical research data manager as well as the help of the Computational platform for Clinical Research and Analysis in Epidemiology & Public Health of the Beautiful University' ('Acknowledgments' section of an article published in a scientific journal in 2007).

When Kelly was thanked in published articles, it was in the same way as a technological platform. Although her work was essential, only its technical aspect was acknowledged. It was seen as inextricably linked to the 'inscription devices' (Latour and Woolgar, 1979, chap. 2), these laboratory instruments that can transform a material substance into a figure, a diagram or other inscriptions which form the starting point for scientific literature.

Conclusion

This case study shows that the solidity of the link between the different types of data which make up the bioinformatics database is based on writing. As well as favouring abstraction and making it easier to carry out mental operations (Goody 1977), writing is also a tangible object, easy to handle, to manage and to combine (Latour 1986). This is shown by the multiplicity of the physical records that Kelly manipulated: research booklets, intermediate forms, patient files, computer files, CD-ROMs etc. During her daily work, Kelly consulted various documents and committed herself in systematic writing practices which support the production of a strong material link between different data sets.

This study also shows that 'information' is not the starting point of Kelly's work. On the contrary, it is the result of all her actions devoted to the biodatabase update. The whole set of documents daily used by Kelly was precisely a way of making available information that has a polyvalent value: scientific, medical and legal. Classificatory tools and coding procedures were crucial to this differentiated information management. They had important political implications in terms of how individuals appearing on databases are identified (Bowker and Star 1999).

Finally, it shows that the work of laboratory technicians is not reducible to manipulating instruments. In some cases they are actively involved in collecting data and thus make a direct contribution to the production of knowledge through their daily writing work. Kelly produced an effective link between various pieces of data. Her involvement in different forms of reading and writing activities together culminated in a single operation: transforming data into information by giving them a specific and durable form. Yet, even if they effectively write some original results, laboratory technicians are rarely authorized to put their name on papers and few of them receive recognition for the importance of their scientific contribution (Timmermans 2003; Pontille 2004).

References

Barley, S.R. and Bechky, B.A. (1994), 'In the backrooms of science: the work of technicians in science labs'. *Work and Occupations*, 21(1), 85-126.

Berg, M. (1996) 'Practices of reading and writing: the constitutive role of the patient record in medical work'. *Sociology of Health & Illness*, 18(4), 499-524.

Berg, M. and Bowker, G.C. (1997) 'The multiple bodies of the medical record: towards sociology of an artefact'. *The Sociological Quarterly*, 38, 511-535.

Borzeix, A. and Fraenkel, B. (eds) (2001), *Langage et Travail. Communication, Cognition, Action*. Paris: CNRS Editions.

Bowker, G.C. (2000), 'Biodiversity data diversity'. *Social Studies of Science*, 30(5), 643-683.

Bowker, G.C. and Star, S.L. (1999), *Sorting Things Out. Classification and its Consequences*. Cambridge, MA: MIT Press.

Cicourel, A.V. (1990), 'The integration of distributed knowledge in collaborative medical diagnosis', in J. Galegher, R.E. Kraut and C. Egido (eds), *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 221-242.

Conein, B. and Jacopin, E. (1993), 'Les objets dans l'espace. La planification dans l'action', in B. Conein, N. Dodier and L. Thevenot (eds), *Les Objets dans l'Action: de la Maison au Laboratoire*. Paris: Editions de l'EHESS, Raisons Pratiques, n°4, pp. 59-84.

Denis, J. and Pontille, D. (2008), 'Placing subway signage: pragmatic properties of signs at work'. Working paper presented at the *58th Annual Conference of the International Communication Association 'Communicating for social impact'*, Montreal, Canada, 22-26 May.

Fraenkel, B. (2006), 'Actes écrits, actes oraux: la performativité à l'épreuve de l'écriture'. *Etudes de Communication*, 29, 69-93.

Fraenkel, B. and Pontille, D. (2006), 'La signature au temps de l'éléctronique'. *Politix*, 74, 103-121.

Goody, J. (1977), *The Domestication of the Savage Mind*. Cambridge: Cambridge University Press.

Grosjean, M. and Lacoste, M. (1998), 'L'oral et l'écrit dans les communications de travail ou les illusions du "tout ecrit"'. *Sociologie du Travail*, 40(4), 439-461.

Hine, C. (2006), 'Databases as scientific instruments and their role in the ordering of scientific work'. *Social Studies of Science*, 36(2), 269-298.

Keating, P. and Cambrosio, A. (2003), *Biomedical Platforms: Realigning the Normal and the Pathological in Late-Twentieth-Century Medicine*. Cambridge, Massachusetts: The MIT Press.

Kirsh, D. (1995), 'The intelligent use of space'. *Artificial Intelligence*, 73, 31-68.

Latour, B., (1986), 'Visualization and cognition. Thinking with eyes and hands'. *Knowledge and Society*, 6, 1-40.

Latour, B. and Woolgar, S. (1979), *Laboratory Life: the Social Construction of Scientific Facts*. Beverly Hills: Sage.

Lynch, M. and Woolgar, S. (eds) (1990), *Representation in Scientific Practice*. Cambridge, Massachusetts: The MIT Press.

Pontille, D. (2004), *La Signature Scientifique. Une Sociologie Pragmatique de l'Attribution.* Paris: CNRS Editions.

Pontille, D. (2006), 'Produire des actes juridiques', in A. Bidet, A. Borzeix, T. Pillon, G. Rot and F. Vatin (eds), *Sociologie du Travail et Activité*. Toulouse: Octares Editions, pp. 113-126.

Pontille, D., Milanovic, F. and Rial-Sebbag, E. (2007), Collectionner le vivant : régulation, marché, valeur'. *Revue d'Economie Industrielle*, 120, 195-212.

Shapin, S. (1989), 'The invisible technician'. *American Scientist*, 77, 554-563.

Timmermans, S. (2003), 'A black technician and blue babies'. *Social Studies of Science*, 33(2), 197-229.