The Culture of the Search Society
Data Management as a signifying practice

David Gugerli, ETH Zurich

November 13, 2009, Amsterdam

Abstract

Databases are operationally essential to the search society. Since the 1960's, they have been developed, installed, and maintained by software engineers in view of a particular future user, and they have been applied and adapted by different user communities for the production of their own futures. Database systems, which, since their inception, offer powerful means for shaping and managing society, have since developed into the primary resource for search-centered signifying practice. The paper will present insights into the genesis of a society which depends on the possibility to search, find, (re-)arrange and (re-)interpret of vast amounts of data.

I am aware of the fact that the title of my talk is both very ambitious and theoretically subversive. The "Culture of the Search Society" undermines the distinction Gilles Deleuze once made between the operating principles of the Foucauldian societies of discipline on one hand, and the operating principles of late capitalist societies on the other hand, i.e. societies, which seem to replace earlier disciplinary surveillance techniques of inclusion and exclusion with a diverse set of juxtaposed rules that rather serve to control "input / output relations", i.e. societies that are tightly linked to the notion of management and the allocation of resources. In a somewhat paradoxical sense, Deleuze's control society is a society which is characterized by a high degree of flexibility, by distributed, rather than hierarchical, networks, by stochastic processes, and by an increased level of tolerance with regard to norms. And it is a society which is flourishing on a both infrastructural and cultural seedbed of search practices or search technologies. It is, I want to argue, not so much a control society, but rather a search society.

The world as a database: CSI

Let me start with something probably familiar, something which you actually might have seen on television. "CSI: Crime Scene Investigation" , one of the most popular, Emmy Award-winning, CBS television series, trails the investigations of a team of Las Vegas forensic scientists as they unveil the circumstances behind mysterious and unusual deaths and crimes.

Most episodes conform to the traditional detective story whodunit-structure, and depict the work of two forensic teams which are usually analyzing two different cases of murder at a time, using, in both cases, the most sophisticated technological and scientific means for their forensic laboratory work.

One reviewer stated that the series' techno-scientific orientation has an astonishing effect both for the role of the murderer as well for the figure of the victim. In fact, both of these figures are, dramaturgically speaking, only interesting as carriers of evidence. There is no attempt at understanding the social dynamics between murderer and victim. The motives of the suspect are almost irrelevant, the tragedy of the victim is not really taken into account. There are a dead and a living body whose encounter in the past has produced forensically relevant evidence. In addition to the crime scene, the two bodies involved in the crime are, to put it bluntly, a mere repository of traces, a hub of evidential markers, a base of pieces of information that can be retrieved, technically stabilized and scientifically analyzed in order to – and this is crucial - recombine them in such a way that the whole set of aggregated data might verify or falsify parts of an ever more differentiated hypothetical narrative of the crime under investigation. These scenarios are shown in a somewhat blurred view of the investigating agents' imaginations as they indicate where they should look for more data at the evidence-providing crime scene. One can say without exaggeration that CSI is depicting the world as a database; its dramatic development is about the excitement of search and query.

Global Software Business 2007

The popularity of the database as a general model for search and query processes is not only evident if one is wasting his or her time in front of a TV set. Database management systems are also shaping our "real" world to a great degree. Most importantly, DBMS are big business. A glance at one of many market reports can give us a sufficient orientation. Carl W. Olofson wrote in a study released last year by IDC: "The RDBMS market is estimated to have grown by more than 12% from $16.7 billion in 2006 to $18.8 billion in 2007." Moreover, it is noteworthy that the top three vendors control 84% of the global market (IBM, Oracle, and Microsoft) and that customers face a high client lock-in due to the costs of migrating data and data models." They are, so to speak, highly dependent on their data management systems."

A signifying practice

Popular culture, globalized markets, and the path dependencies of enterprise are certainly the most promising entry points for a historian's account of this emerging computer-based field of practice, especially if we remember the crucial fact that DBMS - like all other technologies - are mainly instruments both for their producers and for their users to shape society and to manage social change.

Probably the most important feature of a state-of-the-art database is its virtually unlimited possibility to fine-tune, to enlarge, to aggregate, to export and to recombine datasets or even entire databases. Current database management systems have enormous flexibility. However, this flexibility has been developed over a considerable period of time. They have, so to speak, a history of their own. This becomes clear if you have a look at the computer and software landscape of the late 1950s and the 1960s, where computers were dealt with as exceptionally powerful calculators, whose programs were run in a linear, batch-processing mode.

At that time, computers were the obscure, powerful, tireless and efficient assembly-line workers that could sort, count and even filter all sorts of almost endless lists of data. Complex data, however, could not be treated by such a computer in an efficient way – its analysis was a very cumbersome task. No wonder, people started to dream of another kind of data retrieval and search than the one that was available.

The main problem

Take for instance, the very fundamental limits faced in the 1960s by people trying to realize the dream of a Management Information System which could provide an endless pool of information ready to be parsed by decision makers. It would be too simple to blame the insufficiencies of the hardware for these limitations. Of course, storage and memory were expensive, and processing power was nothing compared to even a present-day handheld device. However, the more important limitations I am referring to were of a conceptual nature.

At that time, data was usually stored in a hierarchical, tree-structured array. In order to retrieve a specific item, the programmer had to know exactly its logical position, sometimes even its physical address on a disk or on a magnetic tape. A manager would not be able to actually use a computer other than through the review of standardized output on endless reams of paper. The programmer, on the other hand, was a highly specialized navigator who could find a particular path to a particular item within a specific set of variables. Given these circumstances, even for a specialist it became a very troublesome if not an impossible task to rearrange data which had previously been stored within a particular expectation of how the information would be used.

This is why a database of the 1960s could only answer questions that were already foreseen when the database was constructed: The structure of the database defined what kind of questions it could answer. Nevertheless, the tempting idea of an all embracing pool of information, which allowed for the recombination of data and the aggregation even of whole databases, a database which was open for future problems, questions, and interpretations, had already occurred to some people by the early 1960s.

The dream of a pool of data, from which relevant pieces of information could be selected in any desired combination, became theoretically feasible in the 1970s. This new form of database was gradually implemented in the 1980s and has had a huge practical and economic consequences throughout the last one and a half decades. The fundamental question to be solved was how to separate storage from retrieval. This problem was dealt with for the first time in a seminal paper presented by Ted Codd in 1970, which can be seen as a pivotal moment of transition from hierarchically structured databases to a relational model of databases.

Ted Codd and “Large Shared Data Banks”

In terms of intellectual achievement, Ted Codd’s contribution to the development of Large Shared Data Banks, as he called them, was highly original. He was absolutely convinced that “future users of large data banks must be protected from having to know how the data is organized in the machine (the internal representation). (..). Activities of users at terminals, and most application programs, should remain unaffected when the internal representation of data is changed, and even when some aspects of the external representation are changed.” (Codd 1970, 379)

Two principal steps led Codd to his relational model. Firstly, he decided to organize all data in tables consisting of records in rows (these are also called tuples) and attributes in columns. Additionally, each record was given its own key which uniquely identified the record. This key eventually allowed for the combination of records in two different tables.

Secondly, Codd and others developed a search and query language which allowed for a mathematically consistent retrieval of data stored in related tables. Whoever knows how to use a SQ-Language, is able to interpret any possible combination of records in a database, without being forced to study the physical address or the logical position of an entry.

Recombination of Data

Thus, recombination and relation became the most powerful mode of query and interpretation of data. Since the interpretation of data suddenly became independent of its form of storage, database managers did not have to anticipate future searches which might be done on their database. Therefore, an existing database could be confronted with new unexpected questions.

Tables certainly have been the most prevalent means of representation used by statisticians and political scientists ever since the late 18th century. The statistical bureaus of the nation-state have developed this technology to near-perfection during the 19th and early 20th centuries.

The database-driven views and reports of the late 20th century, however, produced a new category of tables. These are characterized by a configuration which is, in principle, dynamic in its composition, produced in real time and able to be reconfigured at practically no marginal costs. This leads to a remarkable discursive shift, from a narrative based on a standardized table, to a narrative based on a flexible tabular view.

Interpretative Flexibility and Critical Theory

The reason why I think that this discursive turning-point is historically so exciting is to be found in a temporal coincidence with a development, which was completely disconnected with what happened in the realm of relational databases. This coincidentaly simultaneous development allows for a very similar emancipation of the signifying practice to that of the report-generating user.

In this parallel development I will describe presently, the user is usually called a reader. In the second half of the 1960s and the early 1970s, French critical theorists elaborated a concept of the literary work which takes into account the interpretative flexibility with which a reader may approach a text. In the theoretical reasoning of Barthes, Derrida, and Foucault, a literary work like a novel or a poem, are seen as a machine which enables its readers to produce interpretations with a remarkable degree of flexibility.

For Roland Barthes, for instance, an ideal text "is a galaxy of signifiers, not a structure of signifieds; it has no beginning; it is reversible; we gain access to it by several entrances, none of which can be authoritatively declared to be the main one; the codes it mobilizes extend as far as the eye can reach, they are indeterminable." As Barthes argues in S/Z, published in 1970, (the same year Codd published his paper on the relational database model) the interpretation of a text cannot be determined by its author. There is no direct connection between the author's intentionally structured text and the views or interpretations a reader produces while querying the same galaxy of signifiers.

Patterns of the Control Society

These are, I should add, not just signs of mere shifts of intellectual vogue. Both in critical theory and in the development of database concept, we find an astonishing coincidence of change towards interpretative flexibility, genuine fascination with the open work (to quote Umberto Eco), and attempts at separating the structure of the input from the structure of its derivative output. The cultural consequences of this change in operative preconditions are enormous. They obviously affect not only the relation between author and text on one hand, and the text and its readers on the other, they have profound bearing on information processing in any form. This paradigmatic shift in information processing has had tremendous organizational and societal consequences. Newly emerging organizational structures, new administrative and productive processes, and new forms of resource allocation have been some of the most important consequences of this change in information processing.

The most prominent consequences of changes in software architecture and social organization can be observed in the large industrial and financial corporations of the late 20th century. Real time production, lean production, logistics, supply chain management, human resource management, all these tasks have been considered since the middle of the 1970s as the most important challenges faced by late modern enterprise. All these processes are determined by, or at least supported by, database management systems.

As a consequence of the implementation of relational databases and other enterprise application software which reside upon such databases, the entire corporate world of the late 20th century has become subject to the manager's disposal and command. The database-driven enterprise is represented and available in all its operational processes and structures. Public administrations and even universities are joining the club of Oracle and SAP customers. The operations and logistics of a supermarket, the accounting of a bank, including its grid of automatic tellers, as well as large technological systems in the telecommunications or transport sector, depend on the consistent interaction of databases, the recombination of data which is often stored in distributed data servers.

The search society and its culture of recombination had started to emerge even before Gilles Deleuze was able to coin a confusing term for its adaptive and distributed mode of operation. It is a society which operates in real-time mode, a culture which thinks of itself as a project with continuous change management, a society which has to be understood in terms of the permanent fluctuation of its conditions and relations, of its writing and reading, of its calculations and decisions, and indeed, of its practices of search and query.

Further reading including bibliographical notes

· David Gugerli: Die Welt als Datenbank. Zur Relation von Softwareentwicklung, Abfragetechnik und Deutungsautonomie, in: Gugerli, David et al. (eds.): Daten, (Nach Feierabend. Zürcher Jahrbuch für Wissensgeschichte, Bd. 3), diaphanes: Zürich - Berlin, 2007, pp. 11-36

· David Gugerli, Suchmaschinen. Die Welt als Datenbank. Suhrkamp Verlag, Frankfurt am Main, 2009

The Culture of the Search Society, Amsterdam November 13, 2009