
Data From the Central Bureau of Statistics to the Social Science Community: The Norwegian Experience

by Bjørn Henrichsen¹, Director
Norwegian Social Science Data Services

The other presentations in this session are from official producers of statistics. I am representing an organization, the Norwegian Social Science Data Services (NSD), established by the Norwegian Research Council for Science and the Humanities (NAVF) and with strong links to the Norwegian Research Council for Applied Social Research (NORAS).

Due to the composition of the panel, I will, however, first give the background for the Social Science Data Archive movement, and then I will describe briefly our organization and the way we have organized access to data for the research community on Norway. In the presentation I am leaning heavily on the article Stein Rokkan and I wrote about 10 years ago.

NSD was set up by NAVF in 1971 and is to-day financed by the two research councils funding social science research in Norway, NAVF and NORAS. The set-up was a result of two parallel, but still distinctive developments; the one intellectual, the other technological. **Intellectually**, the decisive factor was the increasing prestige of quantitative methods and statistical testing procedures across the entire range of the social sciences; the demand for solid data across large populations spread rapidly from demography and economics, across the entire spectrum of the generalizing disciplines, far into the traditional fields of history and geography.

Technologically, the very notion of a data service was unthinkable before the computer revolution.

In fact the standard definition contrasts the man-readable information stored in the traditional document archive with the machine-readable information stored in the data archive. But the technological conditioning of developments goes much further: the data archive is only one component in the broader set of facilities offered by a data service and all these facilities were developed in one way or another in response to the opportunities opened up by computer technology. A fully-fledged data service

would not only offer files of machine-readable information for various fields of social science research, it would also place at the disposal of the research community a variety of other computer-based tools: facilities for ready conversion of textual data machine-readable form, systems for easy retrieval of information stored in all the files, teaching packages, and workbooks for computer analysis at classroom level, inventories of programs, of on-going research projects, and software systems.

These two sets of developments, the intellectual and the technological ones, interacted at different speeds in different national settings and generated markedly different conditions for institutional innovations.

To gain some insight into the sources of such differences we have to start out from an analysis of the characteristics of what we might call the national "information establishments" on the eve of the computer revolution.

What agencies were there for the collection, storage, classification, and distribution of basic information about social structures and processes?

Institutions for storage and display of information-bearing artifacts came first: the museum, the library, the document archive. Later came the institutionalization of standardized data-gathering operations: tax registers, census, registers of births, marriages and deaths, police records - all elements of a growing system of national book-keeping.

Before the arrival of the computer, these giant data-collection agencies cooperated without great difficulty with the data-storage institutions: the tables and the analyses published by the statistical bureaus were stored in libraries and the original data sheets (census forms, register protocols, and the like) were with some regularity transferred to the established archives.

The computer upset these well-established links: once the central statistical agencies had started to transfer all the new information to magnetic tape or other machine-readable media, there was no stopping the demand for similar technological changes at the storage-display end. The statistical agencies could store all their tapes themselves but this raised serious problems of access for academic social scientist working outside the agencies. The agencies might of course also send the tapes for storage to libraries or archives but this made little sense

as long as these institutions were still firmly wedded to man-readable media: given the predominance of textual material in libraries and archives, it took much longer to introduce even rudimentary computer technology in their operations.

The advent of the computer created a gap in the established network for information flow: at the production end, data were regularly transferred to machine-readable form; at the consumption end, the agencies responsible for storage and distribution to the wider community of social scientists could only handle man-readable data and had no facilities for meeting the new demands.

All this happened just during the period when the methodological revolution in the social sciences had created an increasing demand for mass data and widespread dissatisfaction with procedures of analysis: more and more social scientists had become aware of the potentialities of the mass of information collected by governmental agencies and wanted to subject them to reanalysis.

Theoretically, this problem of data access could have been solved in two ways: the governmental statistical agencies could have set up special divisions for servicing the academic community, or enterprising academics could undertake to reorganize statistics from the central services into manageable data banks for clients in the universities. The first type of solution suited the economists best: they normally tended to get very good service from Central Bureaus. It proved much more difficult to satisfy the academic clientele in the other social sciences: the Central Bureaus were not so easily ready to link up data from censuses, school statistics, elections, and so on to meet the demands of sociologists and political scientists.

This interaction between governmental book-keeping and academic research was heavily influenced by events on another front: the emergence of a methodology offering an alternative to the total enumerations so characteristic of official data collection activities. The demand for facilities for secondary analysis increased markedly in the wake of the methodologies for market research firm and the survey agency proved distinctly more flexible than the established governmental bureaus and were soon able to produce vast quantities of data for the social science community. At first, arrangements for access to such data were irregular and haphazard, but the pressure soon built up for some form of institutionalization.

The result was the establishment of organizationally distinct archives for raw data from polls and surveys, the ICPSR in Ann Arbor, the Roper Center, now in Connecticut and the Zentralarchiv in Cologne are typical examples of this trend.

This development on the narrow "poll-survey" front proved catalytic: it triggered a series of efforts to build up broader-range agencies for the reorganization and transmission of data for research in the social sciences. During the 1960s more and more of the academic data

archives took an interest in the great masses of data available from governmental agencies and started to link up information from many sources into broad-gauged systems for rapid computer retrieval. The Inter-university Consortium in Ann Arbor took an early lead in this direction but a number of groups in Europe had also seen the opportunities. One of these was a group of sociologists and political scientists in Norway: this group had built up its own private files in the early 1960s and was later persuaded to take the step that led to the establishment of the Norwegian Social Science Data Services under the Research Council. This Service covers probably a wider range of different types of data than most of the other archives and deserves some scrutiny just for this reason.

Norwegian social scientists had taken part in the early discussions that led to the establishment of survey archives and had taken steps to acquire the raw data of a number of Gallup polls for reanalysis. But the decisive development took place within the programme of electoral studies launched at the end of the 1950s. Within this programme it proved essential to build up an archive of information for all the communes of Norway: this archive linked up data from election statistics, party membership records, censuses, tax returns and a great variety of other official book-keeping activities. This archive was later expanded into an impressive Commune Data Bank that was to become the core of the Data Service finally established under the Research Council in 1971.

Curiously, the decision to establish a National Data Service was not triggered by increased pressures for access to machine-readable information by mounting dissatisfaction among social scientists with the services they were offered at the University Computing Centers. The transition from mechanical sorting-counting machinery to electronic computers had not improved the conditions for the social science research: on the contrary the many promises of great improvements had caused frustration and anger and the difficulties experienced in writing software specifically for social science users had produced large backlogs of analysis jobs.

By comparison with most of the other data facilities built up over the last two decades, NSD is probably the one giving highest priority to book-keeping and "process produced" data. It is multi-sectoral and sees it as its primary task to link up and to systemize data of different types. We have built close links with the different user groups and have established local offices at all the Norwegian universities and also established close links with data producers like the Central Bureau of Statistics (SSB) and the private polling agencies. From the start the SSB has appointed a member to the NSD Board.

Up until 1976 the SSB responded to requests from individual scholars and from institutions like NSD. Often the same requests came from many users. The procedure was very time consuming both for the SSB and the research community.

In 1976 the SSB was looking for new ways of organizing their transfer of data to the Norwegian research community.

After negotiations between the SSB and NSD, an agreement was signed. The agreement gives NSD a broker function between the SSB and the Social Science Community and it specifies what sorts of data that can be transferred to NSD for further distribution to the individual users.

The agreement secures transfer of both survey data and data from registrations and enumerations. Services to the NSD is in principle free of any charge. Data are transferred to the NSD as soon as data has been processed by the SSB and usually the only restriction on our side is that data shall not be presented before the official publications has been sent out by the SSB.

The purpose of the agreement has been to give the research community as easy access as possible to data from the Central Bureau. Data has been refined by the NSD before they are serviced to the NSD User Community. Both the SSB and the NSD are in their procedures for data transfer following strict rules concerning data protection. The agreement has now been in operation for about 12 years and for the social sciences it has been a success. It has secured easy access to SSB data for the social sciences and stimulated use of such data. In the years since the agreement was signed, the cooperation between the SSB and the NSD has been strengthened and data transfer is to-day secured for most of the areas covered by the Central Bureau. To secure close cooperation the SSB is not only represented on the NSD Board but also on various groups discussing new priorities for the NSD. NSD is also represented in similar groups set up by the SSB to discuss needs and priorities. We are also reporting twice a year to the SSB on the usage of their data, and are sending them publications where their data have been used. As far as I know it is the only agreement of its kind giving free access to such a broad specter of services.

Seen from both the SSB and the NSD I think it is fair to say that the agreement has secured an improvement in the efficiency of the distribution of statistical information in Norway and access has stimulated use of data from the Census Bureau in research at the universities.□

¹Presented at the IFDO/IASSIST 89 Conference held in Jerusalem, Israel, May 15-18, 1989.