Editor's Notes

Torture, Numbers, and Digital Tape

Welcome to the IASSIST Quarterly (IQ) volume 33 number 3. One third of 100 volumes! We have reached the autumn of 2009 in our IQ chronology. This issue is very much about quantitative approaches, and without getting into a discussion on precision and accuracy here, we can report from our own world that most other activities presently experience themselves as being in autumn 2010. By now, you have probably learned that it is unwise to set your clock by the IQ.

With this issue concentrating on quantitative investigations and the use of statistics, I came to think of the Mark Twain citation "There are three kinds of lies: lies, damned lies, and statistics." Mark Twain did not take credit for the remark; he attributed it to Disraeli but that has been questioned and that is another story. However, Mark Twain showed his goodwill by making the reference and demonstrating free dissemination of knowledge.

With the violent title of the first article (see below), the article on numbers and statistics, and the report from a national statistics agency, the "Torture, Numbers, and Digital Tape" title of this editorial surfaced. A fourth article bends and uses a famous film title, so I used and bent another film title for the heading for this introduction. (A remark to the non-cineasts: Steven Soderbergh directed in 1989 a motion picture called "Sex, Lies, and Videotape"). With that I welcome you to an issue of the IQ that is filled with tales about data and their uses.

"Torturing Nurses with Data". Now, that's a title to remember! Maybe it could be a song too! Kristi Thompson, Data Librarian at the Leddy Library of the University of Windsor, presented this paper at the IASSIST 2008 conference at Stanford in the session "Numeracy, Quantitative Reasoning and Teaching about Data". She describes and discusses two iterations in the creation of a short module in quantitative research in the programme for a Masters in Nursing. The module included a lecture as well as hands-on practice. In the second version of the module the hands-on part was more extensive and more structured including an analysis assignment using a pre-selected data set. The paper explains what worked and what didn't work. We can reveal that the article includes an analysis of feedback from an anonymous questionnaire filled out by the students following the second unit. As a data provider for eager students Kristi Thompson speaks of the "Impossible Dataset" - the dream dataset searched for by students that is very unlikely to exist as it would break nearly all confidentiality codes. This was remedied by spending more time on the realities of data collection. The second iteration had learned from the first but was also

faced with its own quantitative problems as the number of students in the class was now much higher. For both of these reasons the student task was changed into choosing from a selection of prepared datasets and accompanying research questions. However, the module then was experienced more as a class in statistics than in quantitative research. This wisdom is what triggered the "torturing nurses with data" title. There needs to be more direct interest from the students. This can probably be achieved by again letting the students formulate their own research questions and thus fulfilling the subtitle of the article: "Building a Successful Quantitative Research Module".

The second article is also within the quantitative area. With the short title of "Numbers", Flavio Bonifacio takes us on a tour of "numbers" from the representation of reality, through processing as a system of guarantee and numbers as model parts, to what is called a path between numbers and society. Flavio Bonifacio works at Metis Ricerche in Torino, a company which does data collection and processing, forecasting, and analysis, and has used SAS software in several of its projects. He starts by addressing the issue of objectivity and gives what looks as the concluding insight: "how numbers are not more true than other representations of reality, that every statement about reality must be responsibly supported, that it is necessary to support this responsibility with an explicit agreement between the producers of the data, that this agreement must be institutionally granted, that only this agreement can make it possible not to surrender the objectivity of the measurements to the tastes of the moment." The article is on "the misuse of numbers" and shows with some citations how the same event in a public bus takes many different forms in Queneau's "Exercises of style". The obvious differences in texts can lead to the assumption – and a jump to a false conclusion - that numbers are "more true".

In the session on "Building on Data: Resources, Tools and Applications" at the 2009 IASSIST conference in Tampere Chiu-Chuang (Lu) Chou, a senior special librarian at the Data and Information Service at the University of Wisconsin-Madison presented her "The Good, the Bad and the Ugly of Playing a Data Custodian". The article concerns the National Survey of Families and Households (NSFH) that is a longitudinal study on family life in the US. The survey has been carried out in three waves in 1987-1988, 1992-1994, and 2001-2003. It is a very expensive data collection that is heavily used; an ICPSR database shows that so far the data has been used in more than a thousand publications. The NSFH project has ended but researchers continue to use the data. So data librarians face the job of documentation though they were not involved in the actual project; the people who were involved are now retired or unavailable. The paper describes the content of the three NSFH waves that are concentrated on family living arrangement, marriage, cohabitation, fertility, parenting relations, kin contact and economic and psychological well-

being. It also includes some measure of the requirement for user support, as the design and structure of the NSFH waves are complex. The article shows that some of this support is quickly done but I noted that more than a quarter of the users' questions took two hours or "substantially" more (my statistics). So there is a great need for support even though most of the contact support takes place through email. The article also includes examples showing that acting as a data custodian for such complicated studies also involves making corrections in the data based upon users' reports.

The last article is a presentation from the session "Building Data Archives and User Communities: Greece, Estonia and Ethiopia" also at the 2009 IASSIST conference in Tampere. Yacob Mudesir Seid from the Central Statistical Agency of Ethiopia (CSA) describes the agency as being responsible for providing accurate and timely statistical information for development planning and monitoring purposes. The paper describes the history of the use of Information Communication Technology in the agency's data processing, archiving and dissemination efforts. The CSA is considered as one of the leading institutions in Ethiopia in utilizing ICT and has developed through generations of hardware and software. On the archiving and dissemination side the CSA has taken advantage of DDI (Data Documentation Initiative) for improvement of the metadata documentation to an international standard. It is interesting how younger agencies as "late movers" very quickly move to the highest levels of standards. The article shows many graphic examples of ICT use at the CSA, among them Geographical Information Systems (GIS) that are providing easy access for decision makers.

Articles for the IASSIST Quarterly are very welcome. Articles can be papers from IASSIST conferences, from other conferences, from local presentations, discussion input, etcetera. Authors are very welcome to contact me. If you don't have anything to offer right now, then please prepare yourselves for the coming IASSIST conference. You can start planning for participation in a session there. Should you be interested in compiling special issues for the IQ as guest editor(s) please also contact me. Chairing a conference session with the purpose of aggregating and integrating papers for a special issue IQ is much appreciated as the information reaches many more people than the session participants and will be readily available on the IASSIST website. By the way, if you have not experienced the new website then take a look at http://www.iassistdata.org. Also the IASSIST blog - the IASSIST Communiqué – is found at http://www.iassistblog.org.

Contact the editor via e-mail: kbr@sam.sdu.dk.

Karsten Boye Rasmussen   -   October 2010

.