

EAIR: Special Interest Group (SIG) on Exploiting Data Repositories

Marcel Herbst & Mantz Yorke

November 12, 2007

Contents

1 Introduction	1
2 Session at Innsbruck	2
3 New Steering Group	2
4 Aims	3
5 Major Issues	4
6 Projects, Publications, and the Future of our SIG	5
A Questionnaire to SIG Participants	5

1 Introduction

This note comes a bit late, unfortunately (for reasons that are, we presume, all too familiar). We had, once again, a good session and a fine second meeting at the 29th Annual EAIR Forum in Innsbruck (August 26, 2007).

For those of you who attended our SIG-session for the first time, our previous notes are to be found under the following address:

<http://www.eair.nl/SIGS/sigdata.asp>

2 Session at Innsbruck

During our session, we had contributions of the following people (in alphabetical order):

- Vic Borden (Indiana University)
- Urs Hugentobler (ETH Zürich)
- Roz Collins and Helena Lim (Southampton Solent University)
- Bernard Longden (Liverpool Hope University)
- Mantz Yorke (Lancaster University)

Session presentations by Urs (“Data Repositories and Management Information”) and by Roz and Helena (“Risotto, Paella, Rice Pudding!”) will be mailed together with this note.

3 New Steering Group

We (Mantz and Marcel) have been active in EAIR for some years now, and with our retirement we are looking for ways to reduce our various commitments. One of us (Mantz) has stepped down from the Editorial Board of TEAM, and the other (Marcel) will follow suit in due time. In this situation, we decided to hand over the torch to run this SIG to a new group of able members, some of whom had been helping us in the past year. We are grateful that

- Vic Borden (vborden@indiana.edu)
- Roz Collins (Roz.Collins@solent.ac.uk)
- Urs Hugentobler (urs.hugentobler@fc.ethz.ch)
- Helena Lim (Helena.Lim@solent.ac.uk) and
- Bernard Longden (longdeb@hope.ac.uk)

have agreed to serve in the Steering Group. They will organize themselves and will contact you regarding upcoming activities and events, and regarding a SIG-meeting at the upcoming 30th EAIR Forum in Copenhagen (August 24-27, 2008).

Marcel Herbst & Mantz Yorke

November 12, 2007

4 Aims

In this last note of ours, we shall like to summarize the points mentioned at our Innsbruck session. The first major point refers to the aims pursued by this SIG:

- to exploit available data (guide the buildup of — existing or new — repositories?);
- to enhance comparative research (inter-institutional — and international — comparisons);
- to strengthen quantitative research (help or coach researchers? enhance arguments? fight abuses? broaden the methodological spectrum?);
- to stay in touch (communicate).

Perhaps, we might reiterate our general feeling that lots of data resources are “out there”, but few people use them. The reasons for this situation are manifold: we all have our own agenda, and we presumably just lack the time to do this work (because we have other things to do). Nonetheless, higher education research requires arguments, and arguments ought to be based on observations and on data.

A second observation in this context is that comparative analyses are not that common, and they rarely cross the boundaries of one’s own respective higher education culture or Nation. Without comparative analyses no proper benchmarking is possible, and without benchmarking, improvement is severely hampered.

A third observation, perhaps, is more controversial: to foster quantitative research. This is a difficult task in the social sciences (and the humanities). Youngsters with a quantitative bent will likely pursue studies and careers in engineering or the natural sciences, and those who are more verbally inclined will flock toward the humanities and the social sciences. This self-selection mechanism helps to cement a division which C.P. Snow, in his famous lecture on the “Two Cultures” in 1959, had decried. There is no intrinsic reason why quantitative research ought to be outside the realm of the social sciences (or the humanities) and, sooner or later, the social sciences (and the humanities) will become increasingly quantitatively oriented. Presumably, no real progress is possible in these sciences without further quantification and the modelling of social phenomena.

And, fourth, staying in touch with people of similar professional interests is important, and if this SIG can make a contribution in this direction, we shall be happy!

5 Major Issues

The second major point of our Innsbruck meeting refers to major issues (in this context):

- MIS versus Data Repositories (micro data versus macro data?);
- Cross-National (inter-institutional) Classification Schemes;
- Possibilities of Abuse of Data (league tables, competence centers, excellence initiatives; performance budgeting);
- Forms of Data Analysis and Use (description, simulation, prognosis, exploration).

Urs Hugentobler addressed the relationship between Management Information Systems (MIS) and data repositories (see his presentation). It is, in fact, a delicate issue, and it touches on various concepts of governance and management. Normally, MIS are used by the institutions themselves (or by their governing bodies) in their attempt to do institutional research (for their own benefit) and to manage (or govern). Data Repositories are normally open (to the public) and their data are used to do research (for various ends). If data repositories are not open to the public, if their access is restricted, research findings will have to be accepted with the utmost skepticism only.

A related issue refers to attempts to do inter-institutional — and cross-national — comparative research (benchmarking, et cetera). Here, the problem of cross-national (inter-institutional) classification schemes will have to be addressed: definitions will have to refer to the same type of population or strata, to the same bookkeeping categories, and occasionally mapping issues arise. This type of work requires researchers familiar with more than one higher education system, and it normally requires team work where the team members represent the whole spectrum of higher education systems addressed: a worthy task for members of our SIG!

It has been our observation that data abuse is not recognized often and not fought against to the extent necessary. This, of course, would be the task of professionals — such as those assembled by this SIG. Data abuse is

so frequent because there are interests behind “using” and abusing data. To expose these interests, and to show the form of abuse, is a worthy task of all professionals in the broader field of institutional research. In fact, to be able to recognize data abuse is a sign of, and a test for, being a professional.

The last point under this heading concerns forms of data analysis and use. Observations over the past decade indicate that data analysis (within our disciplinary fields) has a tendency toward simplicity. Normally, rather straightforward analyses and models are used that are taught in an introductory social sciences statistics class. The use of more advanced models is very rare. Perhaps, in due time, this SIG could make a contribution regarding the adoption of, and familiarity with, useful methodologies that are not that common.

6 Projects, Publications, and the Future of our SIG

At the Innsbruck session, we were also musing about possible common projects or publications some of us could focus on. A new possibility is presented by EAIR’s new Monograph Series whose Editor is Bernard Longden, a member of our Steering Group (see 3).

The future of this SIG will depend on at least a few members who have a long-term interest regarding issues of data analysis. Sessions at the annual meetings of EAIR might continue, or we might hold intensive seminars to introduce software or to discuss issues of methodologies.

A Questionnaire to SIG Participants

The following are a few questions Mantz Yorke had assembled, and which you might want to answer; try to send the answers back to

- mantzyorke@mantzyorke.plus.com

by Friday, November 30, 2007:

QUESTION 1: What is the focus of your interest regarding data repositories (e.g. analyzing data at HE system level; analyzing data at institutional level; et cetera)?

QUESTION 2: Very briefly, what work do you do, or have you done in the past, with data repositories?

QUESTION 3: What data analysis package(s) do you use, or have you used?

QUESTION 4.1: If you are interested in collaborating with others, would you like to work with others who are members of the same higher education system? (Yes / No)

QUESTION 4.2: Work with colleagues who come from different higher education systems? (Yes / No)