Dear ERIM colleagues,

Given recent events, we wish to inform the entire ERIM community of our on-going initiatives and further plans for debating, clarifying and upholding scientific integrity values.

In November 2011, ERIM established an initiative to address integrity issues within our research institute. Following an analysis of then recent cases of scientific misconduct, and a first analysis of the particular risks associated with management research, it was decided that this initiative would first address the management of research data including collection, analysis and archiving of data.

Five working groups then started working to identify a set of minimum and best practices with respect to these different phases of data management, where each group dealt with a specific research strategy (surveys; behavioural experiments; existing or collectible data; qualitative studies; and quantitative modelling). On 21 May 2012, ERIM hosted a symposium during which the preliminary recommendations of the working groups were reported. The final reports of these working groups were completed by the end of June. These reports contain specific recommendations, for example regarding data collection and the storage of raw and processed data.

In the second half of this Summer, ERIM will integrate these recommendations in the form of research protocols, in consultation with the ERIM research community. We will start implementing and following up these protocols in September. ERIM is dedicated to realizing these improvements in research quality and transparency, together with its members, and aims to make the necessary investments and support available to enable these improvements (such as facilities for data curation and storage).

In parallel, a new course on academic integrity has been developed and added to the training programme for doctoral students; it starts this autumn. This course will consist of a combination of online tutorials, panel discussions and interactive discussions. Elements of the course will most likely also be deployed in other courses for senior faculty, such as the Academic Leadership course offered by the EUR. Recently, these initiatives regarding the doctoral programme have been extended to the entire Erasmus University, where a Task Force has been installed that reports to the Rector Magnificus, with representatives from different schools and institutes.
In addition to initiatives for training and better awareness for all researchers, including senior faculty, this EUR taskforce is detailing procedures for better management of research data in co-operation with the university library, and for more explicit discussions of integrity issues during career counselling and performance appraisals. Implementation and follow-up will start during coming academic year.

While finalizing the more detailed and specific recommendations for the different research strategies, ERIM finds it appropriate already at this stage to provide a number of general recommendations for data storage (see attachment).

Please note that the most recent ERIM Membership Charter (2012) already states that “ERIM expects all its members and fellows to respect academic integrity, and to adhere to and uphold the general principles of professional academic practice at all times. The Netherlands Code of Conduct for Scientific Practice (VSNU 2005) further details these principles.” The code states, among others:

“The quality of data collection, data input, data storage and data processing is guarded closely. All steps must be properly reported and their execution must be properly monitored. Raw research data are archived in such a way that they can be consulted at a minimum expense of time and effort.”

While the code provides a good start, recent events have illustrated that in management research, there are still quite some divergent and risky practices on storing data. Hence, we offer these somewhat more detailed recommendations to our ERIM community primarily to support the quality and reputation of our research—one of ERIM’s key strategic priorities. Also, following these recommendations can help to avoid simple disasters, such as loss of data due to computer crashes and theft. Note that the rules below explicitly do not seek to make all research data publicly available; this will remain the decision of the individual researcher, in consultation with data owners/sponsors.

Consider this as a first recommendation; in the Fall of 2012, these recommendations will be further specified for the different research strategies, and the (IT) infrastructure and other support required for this will be explicated further. This will be done in consultation with the Tinbergen Institute, among others.

Naturally, ERIM will fully support researchers in actually complying with the research protocols under development; we certainly are not planning to create impractical and unnecessary bureaucratic rules.

With best regards,
on behalf of the ERIM Management Team,

Marno Verbeek

Attachment:
General recommendations for storing research data
General recommendations for storing research data

1. Always maintain copies of the original, “raw” research data. In case of paper and pencil questionnaires, this means storing the actual forms. In case of electronic data, it means the original completed electronic forms. In case of qualitative research, it means the original audio files or transcripts of interviews, or field notes. In case of secondary data or data collected by others, it means the originally obtained data (data ownership issues permitting). Thus, while the nature and form of the actual raw data may vary, the basic principle applies that the researcher should be able to convincingly demonstrate that this original version of the raw research data has not yet undergone any selection, purification or transformation steps.

2. We recommend tying the original data to the identity of the research informant/participant, even in the case of confidential data. Confidentiality can be maintained by separately storing a key, controlled by the (lead) researcher. At a minimum, the identity of individual research informants/participants should be recorded (without necessarily relating this to a specific response). The key principle here is that anonymity can be guaranteed (if necessary) with respect to published data, without sacrificing the identity of research participants for the original data collected.

3. The data collection process should be clearly described. This includes the names and roles of the researchers involved and/or the organisations providing the data (such as research agencies). The descriptions should be detailed to the extent that the process can fully be traced back.

4. The data input and analysis procedure should be documented in detail, so that the analysis can be replicated exactly. This includes major analysis steps that may in the end not be reported in further publications, but which have been instrumental in steering the analysis process. All substantial files should be stored, including for instance specific software syntax, diagrams, graphical presentations, etcetera. Again, the names and roles of the researchers involved should be provided.

5. For each crucial data compilation, purification or transformation step, it is recommended that clearly identifiable and described data sets are stored. (Crucial steps transform data such that it is impossible to revert to the rawer data when only the transformed data is available.)

6. All original, “raw” data and the documentation of the data collection and analysis process should be stored for a minimum of five years after publication of the most recent publication using this data. This applies as long as specific professional or journal policies do not require a longer storage period.

7. All (electronic) “raw” data and the documentation of the data collection, input and analysis process should be stored in duplicate. At least one set of data should be stored on a university or external network, with appropriate safeguards regarding anonymity and data ownership. Data collected on paper should be transferred to an electronic medium in its entirety and stored electronically (where possible by scanning the entire documents).
8. In the case of co-authored papers where another person is executing data collection, input and/or analysis, we recommend storing a copy of the raw data yourself as well (confidentiality and data ownership issues permitting), and storing the documentation of data collection, input and data analysis procedures.