



The need for a RDC network and for sharing data in an RDC approach

Decentralised Access to European
Microdata

Kloster Johannisberg, June, 2010

Stefan Bender

German developments in data access

- **80's: no data access**
- **Constant pressure of the scientific community**
- **1999: Commission for the improvement of informational infrastructure between science and statistics**
- **2001/2004: German Data Forum (RatSWD)**
 - **Development of the German data infrastructure**
 - **Establishment of Research Data Centers (RDC)**
- **2001: First RDC was established**

The RDC of the BA at the IAB - Basics

- **Started in 2004; 100% financed by the German Federal Employment Agency (BA)**
- **Data based on:**
 - **Administrative processes**
 - **Data from IAB surveys**
- **No costs for data access**

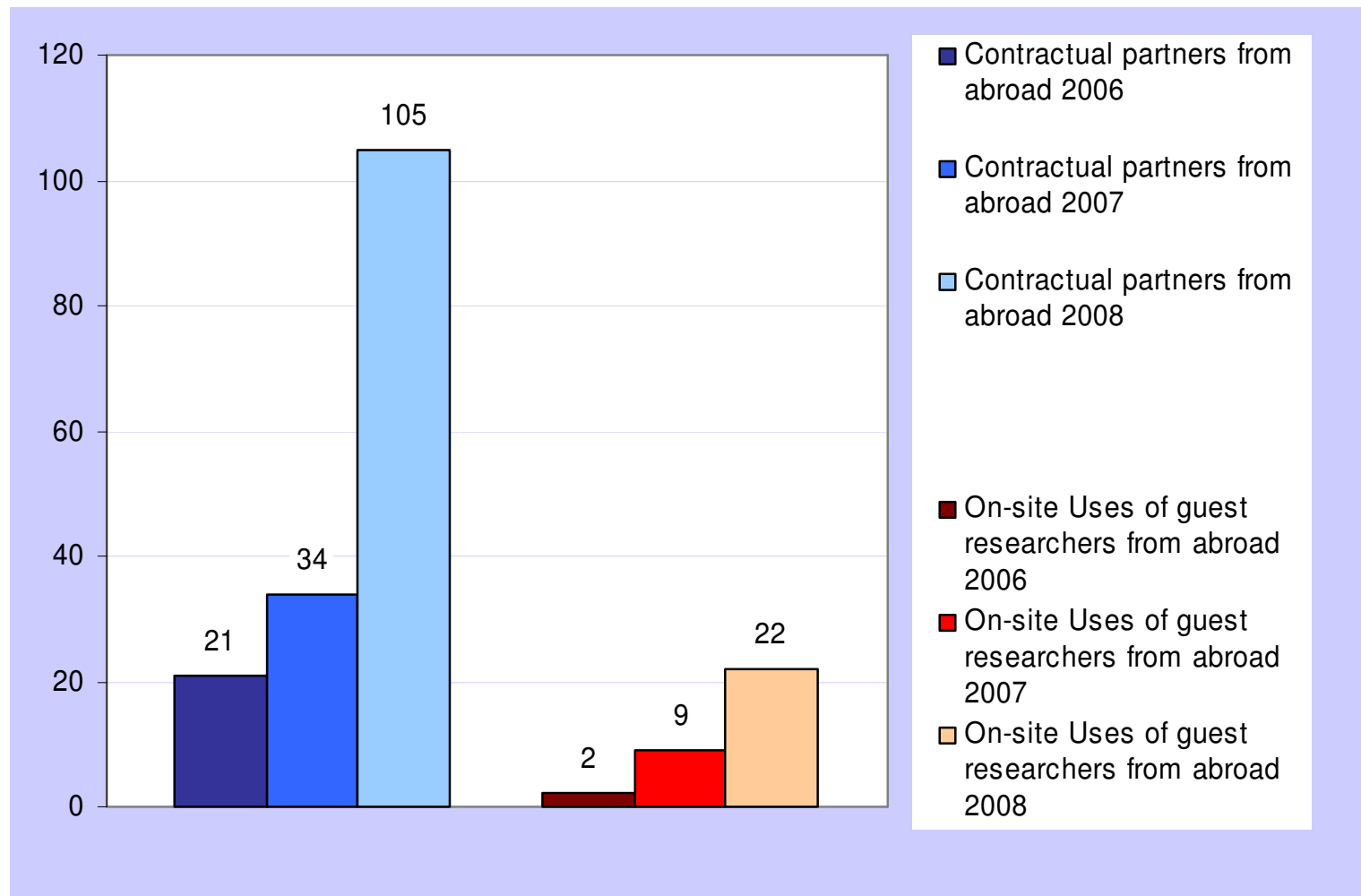
RDC of the BA in the IAB – General Idea

- **Development of data infrastructure**
 - **National access to IAB/BA microdata**
 - **International access to IAB/BA microdata**
 - **Access to combined German micro data (other RDC, patent data and commercial micro/transaction data)**
 - **Access to combined and harmonized international microdata**
- **New data: “double” linked employer employee or combined data (Danish-German admin data, patent or tax data with IAB data)**

RDC of the BA in the IAB – International Activities

- **At the beginning: mostly users from Germany**
- **Translation of website and documentation,**
- **Financial support for user from abroad**
 - ⇒ **Growing number of users from abroad**
- **Development in the number of users and data access**
- **RDC-in-RDC**
- **Participation in international activities (international conferences and cooperations, COST, FP7, WDA)**

Number of users outside Germany



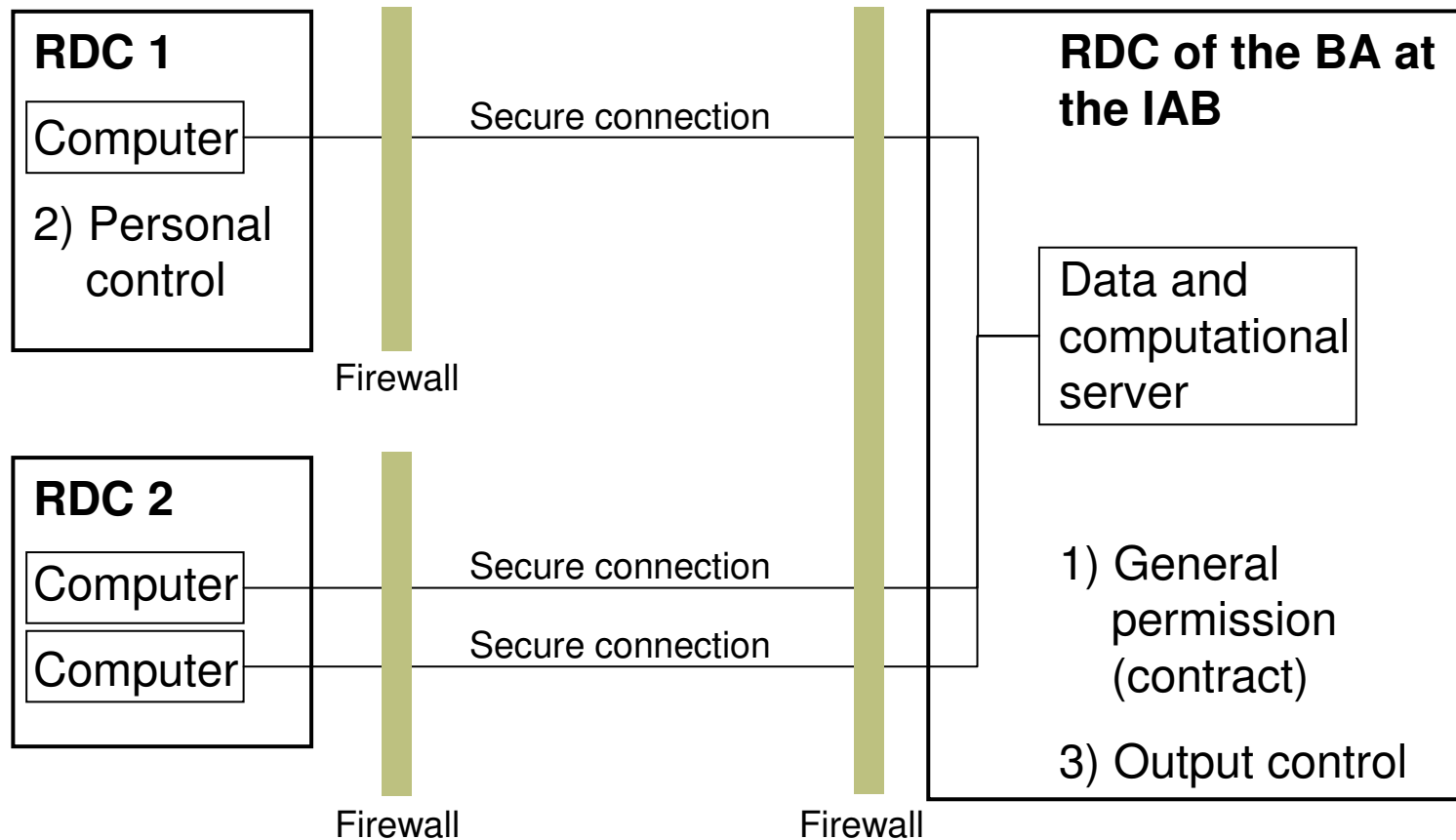
RDC-in-RDC I

- Ideal perception of data access: remote data access
 - Due to legal and technical restrictions not implemented in Germany so far
- ⇒ Costly guest stays
- Step towards remote data access: RDC-in-RDC:

RDC-in-RDC II

- Data protection standards/regulations/laws vary over countries
- But: Nearly comparable standards in RDCs all over the world
- Access to BA/IAB data from other locations/other RDCs
⇒ RDC-in-RDC

RDC-in-RDC III



RDC-in-RDC IV

- **Fall 2010: 4 locations of the Statistical Offices of the Länder in Germany**
- **Fall 2010/Spring 2011: Institute for Social Research, Ann Arbor**
- **Planned: Italy (Blue-ETS);
UK, France, The Netherlands (DwB)**

WDA – General idea I

- **All statistical agencies face the dilemma of how to provide microdata access to scientists while at the same time guaranteeing the confidentiality of data providers**
- **New innovations have recently emerged (multiple imputation techniques, technological solutions to provide remote access)**

WDA – General idea II

The aims of the Workshop on Data Access (WDA) are

- **to review access to confidential government data for research purposes**
- **to bring together researchers, share knowledge, outline a future research agenda and build up a network for future activities.**

WDA – General idea III

“Bottom Up” Approach

- **Self-organized by (a group of) RDCs/Archives**
- **Topics are generated by the Archives/RDCs**
- **“Working base”, highly interactive**
- **Asia, Australia/NZ, Europe, North America**

1st and 2nd WDA is documented: <http://fdz.iab.de/en>

WDA – 1nd Workshop: Basics

- **August 2007 organized by IAB in Nuremberg**
- **UK, Germany, the Netherlands, Canada, US, Denmark, New Zealand, Serbia.**
- **Developments in technology, the legal and statistical environments, and management issues.**
- **Multiple Imputation, metadata.**
- **Workshop should be done a second time.**

WDA – 2nd Workshop: Basics

- **February 2009 organized by Felix Ritchie (ONS) in Newport (supported by IAB and RatSWD).**
- **UK, Germany, the Netherlands, Slovenia, France, Canada, US, Japan, China and Australia.**
- **Same topics (without imputation).**

WDA – 2nd Workshop: Results I

- **Increasing convergence of statistical legislation in respect of research use of data (still large variation across countries).**
- **Many countries have developed, or are on the way to develop, access for researchers to microdata via research data centers or remote data access.**
- **Growing importance of metadata, particularly as a source of analysis in its own right.**

WDA – 2nd Workshop: Results II

- **Expand the network to underrepresented regions**
- **Continue with the conference delivery programme**
- **Build a ‘portal’ website to support the network, identify current models of operation and circulate best practice ideas**

WDA – 3rd Workshop

- **May 2009 organized by ICPSR (Felicia LeClere), ISR (Maggie Levenstein) in Ann Arbor**
- **UK, Germany, Netherlands, Canada, US, Slovenia, Japan, Sweden and Australia**
- **Same topics, but with a more international (European) perspective**

WDA – 3rd Workshop: Results

- **Four Working groups**, which define the subject matter and develop an agenda for the next meetings
 - **Metadata:** Chuck Humphrey (chuck.humphrey@ualberta.ca) and Bill Block (block@cornell.edu)
 - **Disclosure:** Jan Mol (j.mol@cbs.nl)
 - **Data Access:** Lynn Riggs (tammy.l.riggs@census.gov)
 - **International Data Sharing:** Maggie Levenstein (MaggieL@umich.edu)
- **Web page** done by ICPSR
- Workshops should be connected to **IASSIST** (2011: Canada).

Developments
are
POSITIVE

My View I: data

- **Linked data of all kind of sources**
- **Micro data are getting more complex (metadata, training)**
- **Surveys will no longer play the key role (admin data, tax data, transaction data)**

My View II: metadata

- **RDC/NSI need solutions for metadata**
- **Need for “standard” solutions (no need for a 100% solution).**

My View III: researcher's role

- **Swedish researchers are the best trusted, US researchers are the less trusted in the world (could be the same researcher). There is room for improve data access**
- **EU-Access-Standards have been developed (without researchers?)**
- **Need for a voice of researchers (German Data Forum)**

My View IV: data access and dissemination

- **Cost of not publishing data compared to disclosure risk (Ritchie)**
- **More knowledge how to use sensitive data on researcher's side (naive data user model)**
- **Interdisciplinary discussion on disclosure risk**
- **New ideas of data access (imputation, RDC in RDC, International Remote Access)**
- **Standards: positive and negative (Sweden vs. US)**

My View V: last slide!!!

- **European Developments: German Data Forum in Europe, FP7 project “Data without Boundaries”, ESSnet, OECD, WDA**
- **Strong need for a broad international network of**
 - **Researchers, Data Archives, National Statistical Institutes**
 - **International Institutions (Eurostat, OECD)**
- **Strong need for a (international) division of work (DwB)**

Developments
are
POSITIVE
!!!

Institute for Employment
Research

The Research Institute of the
Federal Employment Agency



Contact:
stefan.bender@iab.de

<http://fdz.iab.de>

