

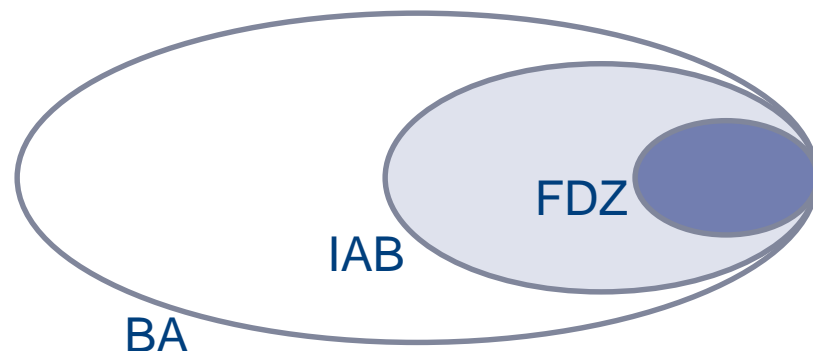
Tools for Enhancing Access for Researchers - An Example from Germany

Association of Public Data Users
2015 Annual Conference
Arlington, VA
September 1, 2015

Joerg Heining

BA, IAB and FDZ

- BA: Federal Employment Agency
 - Labor administration
 - Nuremberg
- IAB: Institute for Employment Research
 - Research department of BA
 - Staff: about 300, 150 researchers
- FDZ: Research Data Center of BA at IAB
 - Provides access to confidential micro data of BA/IAB for academic purposes
 - Established in 2004
 - Since 2011: Remote access centers in Germany, UK and USA



Some Facts about FDZ

- 25 staff members, 20 researchers (25% working time dedicated to research)
- approx. 500 active projects of approx. 700 users
- 160+ new projects in 2014
- 1800+ job submissions for remote execution in 2014
- 2200+ access days at on-site visits in 2014
- 15 standardized data products
- Bilingual (German and English) documentation
- Comparatively easy access rules: no background checks, no prior scientific evaluation of the project, etc.
- Access to standardized data products is free of charge

Data available at FDZ

- Micro labor market data on individuals/households and establishments

Administrative Data

**Social Security
Notifications**

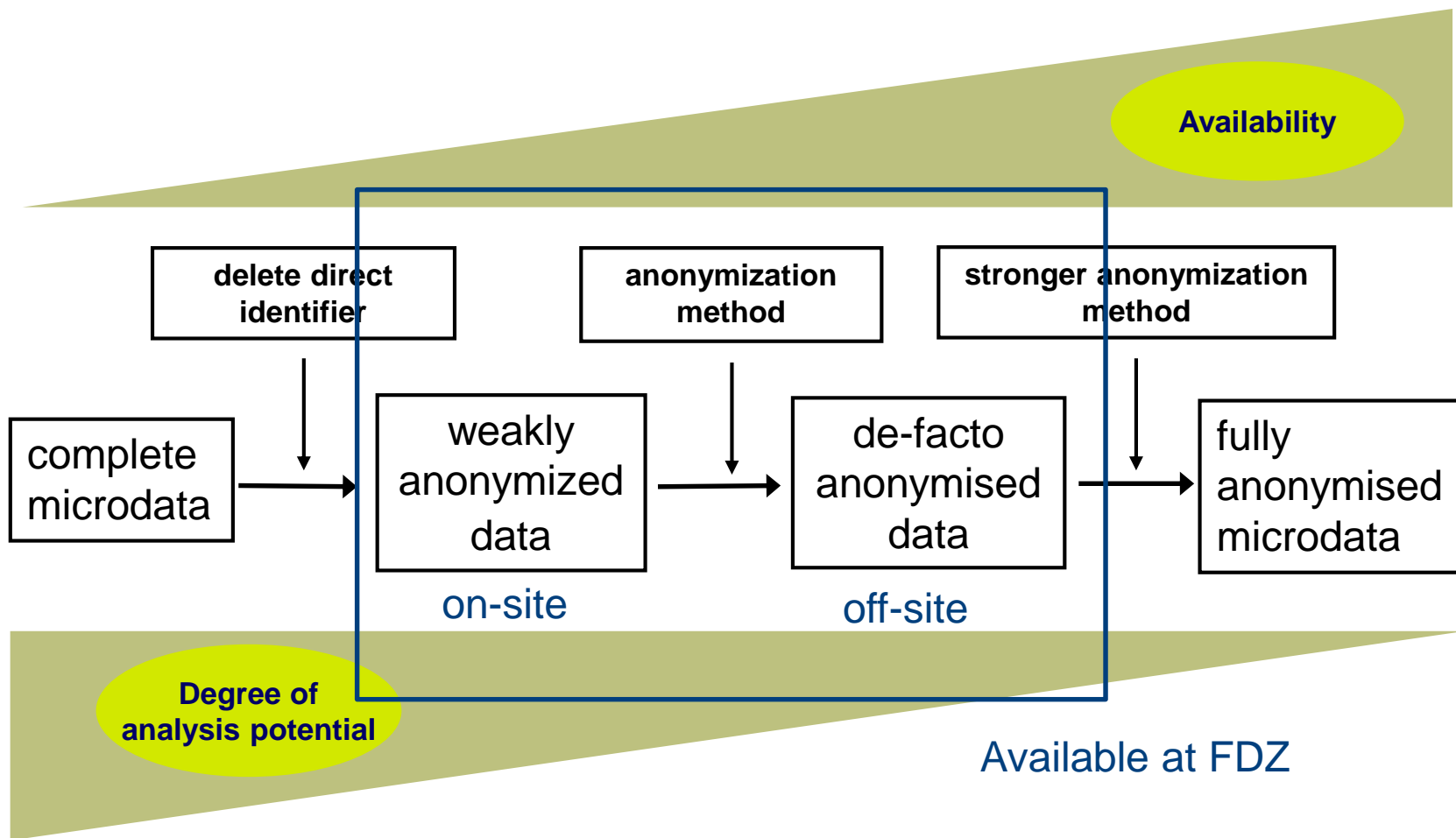
**Process-
generated data
of the BA**

Data available at the FDZ

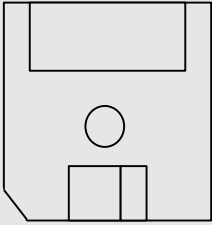
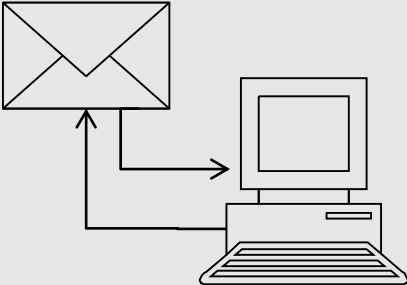
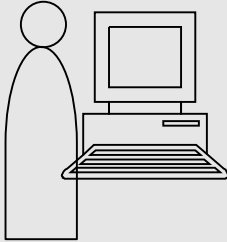
Surveys

**External
Data**

Types of Data and Anonymity



Access Modes

Off-Site Access		On-Site Access
		
<p>Download (Scientific Use File)</p>	<p>Remote Execution</p>	<p>Visit to FDZ or Remote Access Centers</p>
<p>De-facto anonymized data</p>	<p>Weakly anonymized data</p>	

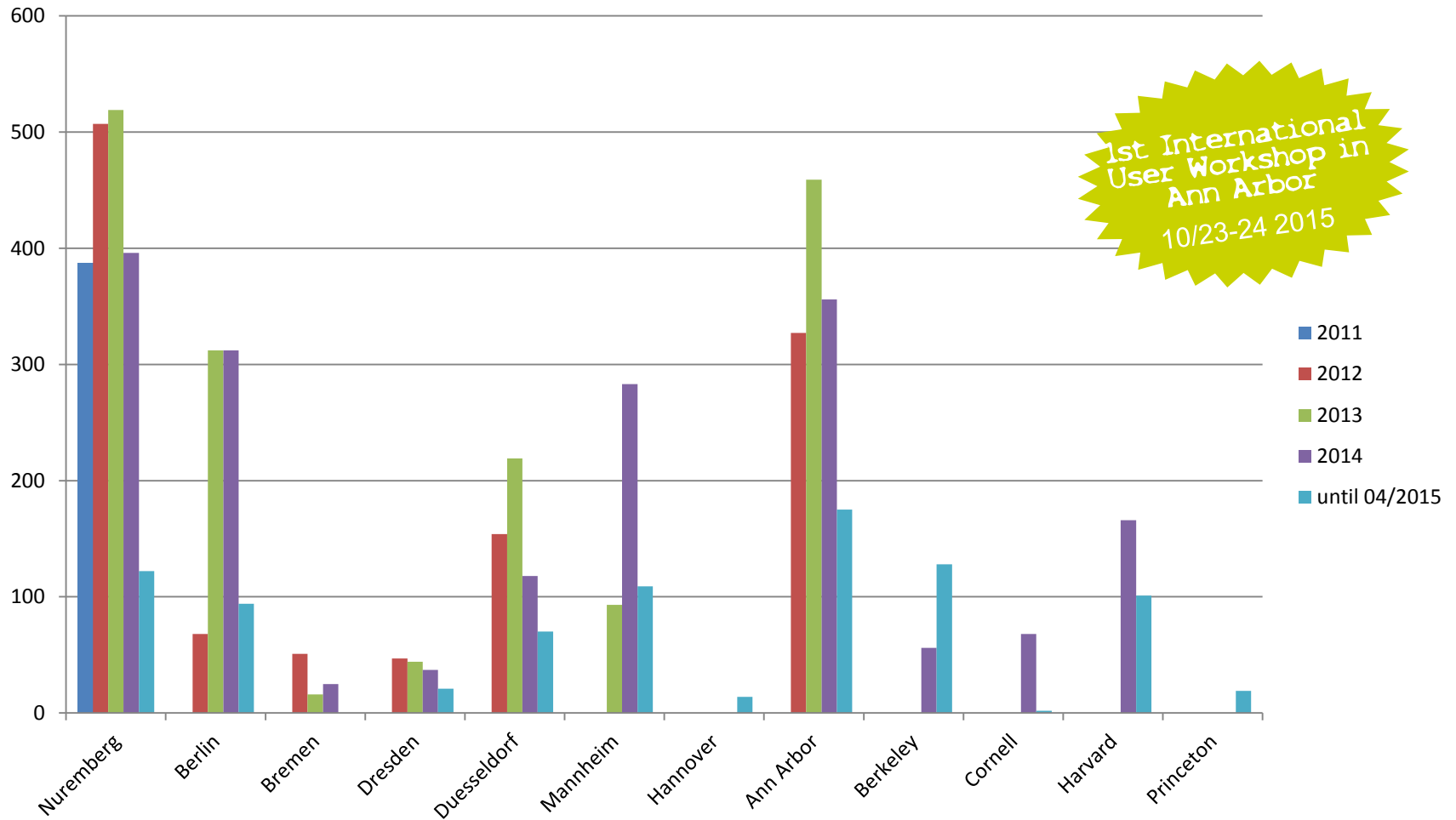
On-Site Access

- Until 2012: On-site Access was only possible in Nuremberg
- Since 2012: Several on-site access points in Germany and the US
RDC-in- RDC approach ⇨ Citrix-thin client technology

Just opened/
will open soon:
University of California
Los Angeles
UK Data Archive



Access Days



Scientific Use Files

- FDZ prepares and disseminates so-called Scientific Use Files (SUF) for off-site access
 - Worldwide accessible (more or less)
 - Data can be downloaded from an secure exchange server after a use agreement is concluded
 - Deletion of the data by the user after the project has expired (use agreement)

- Security Problem:
 - Approved users download SUF data from secure server to install locally
 - Potential abuse of illegally keeping local copy of the data after expiration of the use agreement

Scientific Use Files II

- Problem Solved:
 - Joint project between IAB and Cornell Institute for Social and Economic Research (CISER)
 - IAB SUF data will be stored in the Cornell Restricted Access Data Center (CRADC)
 - Users connect to CRADC after concluding use agreement with IAB
 - Advantage: No local copies of SUF data

Remote Execution

- A “hand-made” system was used in the past
 - Now: JoSuA (Job Submission Application)
 - Software tool developed by the Institute for the Study of Labor (IZA) in Bonn, Germany
 - Researchers log to a secure computing environment
 - Upload of scripts
 - Scripts are processed with the data on the servers of FDZ
 - Download of outputs
- ⇒ More efficient workflow and less administrative work

Review Services and Disclosure Control within JoSuA



- Currently: 1800+ jobs for remote execution \Rightarrow 1800+ files of output need to be reviewed
- Review policy: absolute anonymity
- Manual inspection after “AutoDafe” script (in-house development)
- Only 5% (rough estimate) of the reviewed output is published by the user
- New concept for disclosure control: **internal vs. external output**
- Both concepts are embedded within the JoSuA environment
- User indicate what kind of output is requested

Review Services and Disclosure Control within JoSuA II

	Internal Output	External Output
Characterization	work material	used for publications
Download or Printing	no	yes
Review Policy	de-facto anonymous	absolute anonymous
AutoDafe script	yes	yes
Manual inspection	no	yes

- Additional means to preserve confidentiality within internal output:
 - Explicit contractual stipulations and severe penalties
 - Black list (of Stata) commands in JoSuA
 - Advanced version of the AutoDafe script
 - Several warnings within the JoSuA environment

Benefits

- **Access to a Unique Data Resource for Cutting Edge Research**
 - Several high ranked peer review journal articles by data users
 - Benefit for (German) policy-makers: more knowledge and insights on the German labor market
 - Value added to the University research centers hosting on-site access
- **Training and Education**
 - Students and scholars learn how to work with administrative data and deal with problems
 - Good and transferable experiences/skills when they want to work with comparable data
- **Positive Externalities in data sharing**
 - Knowledge transfer between international experts on data access and dissemination
 - Basis for the further development of transnational data access, dissemination strategies and international data sharing

Conclusion

- FDZ continuously further developed the access modes to confidential microdata of IAB and BA
 - Transnational access
 - International data sharing
 - More efficient tool for remote execution
 - Benefits for all sides involved
 - Students and scholars
 - (German) Policy-makers
 - FDZ and its partners
- ⇒ FDZ will continue to engage in the further development of data access in the future!

Thank You!

Joerg Heining
Acting Head of FDZ of BA at IAB

joerg.heining@iab.de