



Research Data Centre (FDZ)
of the German Federal
Employment Agency (BA)
at the Institute for
Employment Research (IAB)

Neue Entwicklungen beim Datenangebot im Forschungsdatenzentrum der BA im IAB

Institute for Employment Research
Research Data Centre

Peter Jacobebbinghaus

Agenda

Long-term perspective

News about:

1. Administrative data
2. Survey data
3. Data linkage
4. Data access

Long-term perspective

- Micro data for German researchers
- Micro data for researchers from abroad
- Linkage of BA/IAB data
 - with one another
 - with other German data
 - with international data

1. News about administrative data

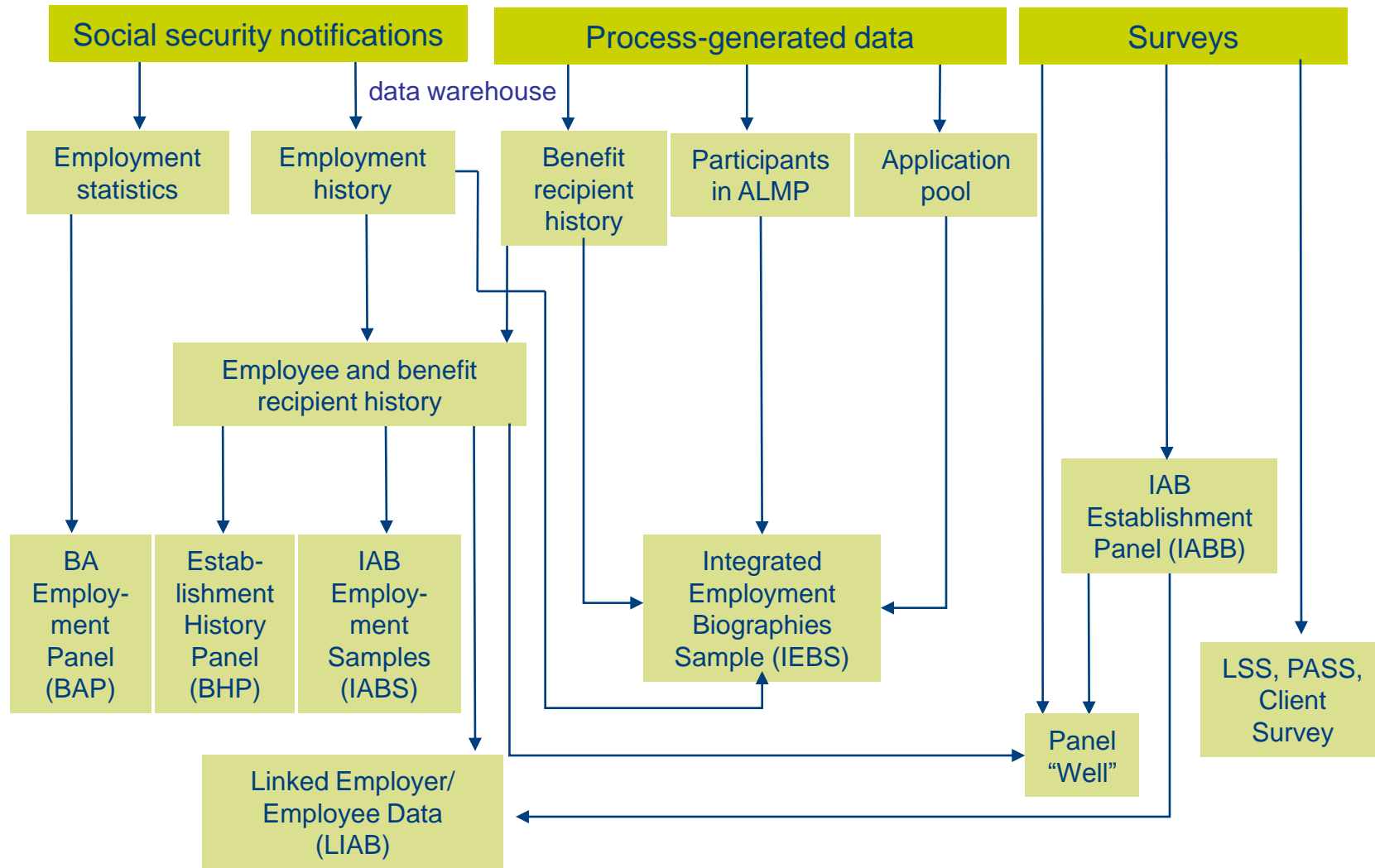
- Harmonization of administrative datasets
 - all data based on the same data base (IEB)
 - standardized naming and coding of variables

 - data portfolio gets clear structure
 - easier to switch between different datasets

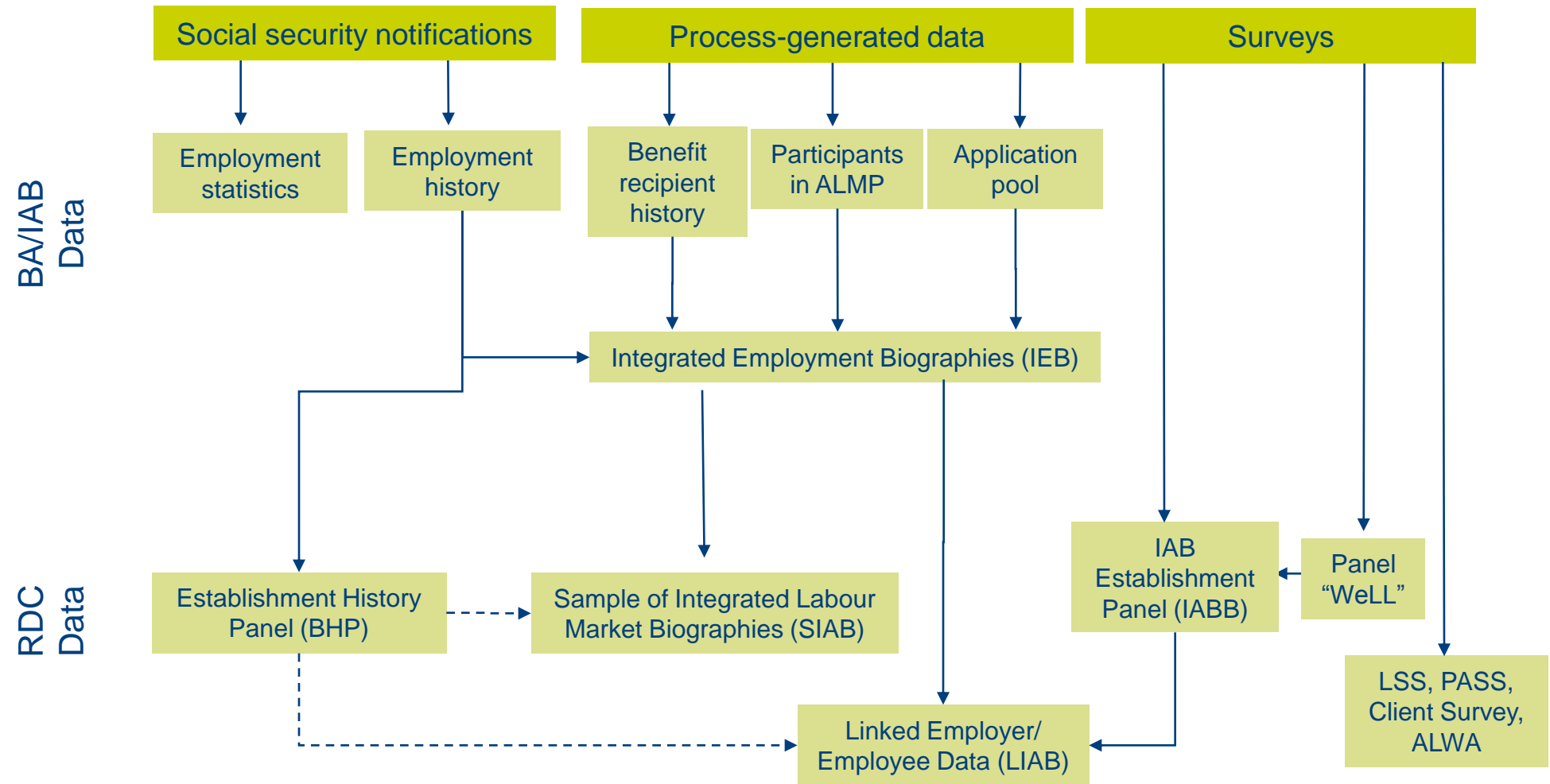
BA/IAB data structure old

BA/IAB
Data

RDC
Data



BA/IAB data structure new



1. News about administrative data

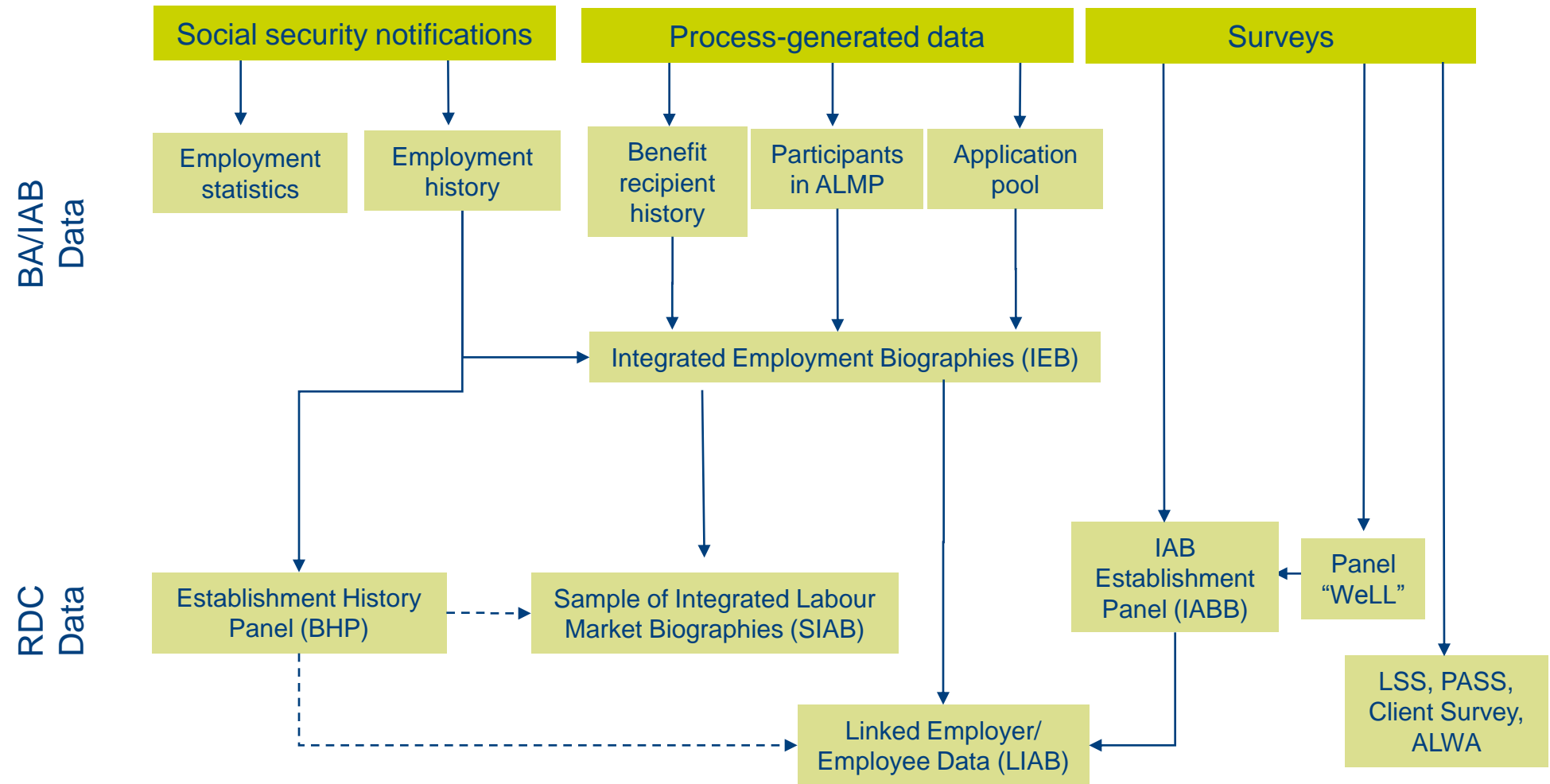
- Sample of Integrated Labour Market Biographies (SIAB)
 - successor of IABS and IEBS
 - 1975-2008
 - employment, benefits, job search, programme participation
 - wide range of establishment information based on BHP

1. News about administrative data

- Extension of Establishment History Panel (BHP)
 - 25% sample of German establishments with employees liable to social security
 - information on employment structure on 30 June.

 - 1975-2008
 - employment flow data
 - classification of id changes
 - openings, closings, spin-offs
 - based on employment flows

BA/IAB data structure new



2. News about survey data

- Working and Learning in a Changing World (ALWA - Arbeiten und Lernen im Wandel)
 - cohorts 1956-1988, more than 10.400 respondents
 - contains information about
 - education and employment trajectories
 - residential, partnership and parenthood histories
 - longitudinal analysis of
 - schooling and training decisions
 - labour market behaviour
 - processes of family formation and regional mobility.

2. News about survey data

- **Panel WeLL – Further Training as a Part of Lifelong Learning: 2nd wave**
 - further vocational training since January 2006
 - socio-demographic information, retrospective employment biography since January 2006, job satisfaction, expectations, questions about the respondent and his/her household

2. News about survey data

- **Panel WeLL – Further Training as a Part of Lifelong Learning: 2nd wave**
 - 149 establishments
 - 6404 employees 2007, 4894 employees 2008
 - RWI-Essen (Rheinisch-Westfälisches Institut für Wirtschaftsforschung e.V.), DIE (German Institute for Adult Education), IAB (Institute for Employment Research) and infas GmbH (Institute for Applied Social Sciences)

2. News about survey data

- **PASS – Panel Labour Market and Social Security: 2nd wave**

- labour market, welfare state and poverty research
- two partial populations:
 - persons and households in receipt of Unemployment Benefit II
 - persons and households registered as residents of Germany

2. News about survey data

- **PASS – Panel Labour Market and Social Security: 2nd wave**
- **Wave 1:**
 - Sample I: 9,386 individuals (6,804 households)
 - Sample II: 9,568 individuals (5,990 households)
- **Wave 2:**
 - Sample I: 4,753 individuals (3,491 households)
 - Sample II: 6,392 individuals (3,897 households)
 - Sample III: 1,342 individuals (1,041 households)

3. Linkage of data

- In all new IAB surveys respondents are asked to consent to linkage with administrative data (WeLL, ALWA, PASS)
- In addition: projects on linkage with external admin. data
 - Patent data (DPMA)
 - KombiFiD
 - BasiD
- Commercial data (Bureau von Dyke)
- Multinational data
 - Commuters Denmark-Germany
 - RDC-in-RDC approach

3. Linkage of data: patent data

German Patent and Trademark Office data (LMU Munich)

- all entries for 2002 by inventors name and address
- education information very reliable
- over 90% of inventors are in the social-security system
- 63,278 observations in 2002

3. Linkage of data: patent data

- Linkage
 - no identifier to link the data
 - record linkage by name+address (MTB 2004)
 - exact linkage: 60% declared matches
 - different Jaro-techniques: 12% matches
 - 45,782 (72%) inventors found

- Result: linked employer-employee-inventors data

3. Linkage of data:

KombiFiD – Combined Firm Data for Germany

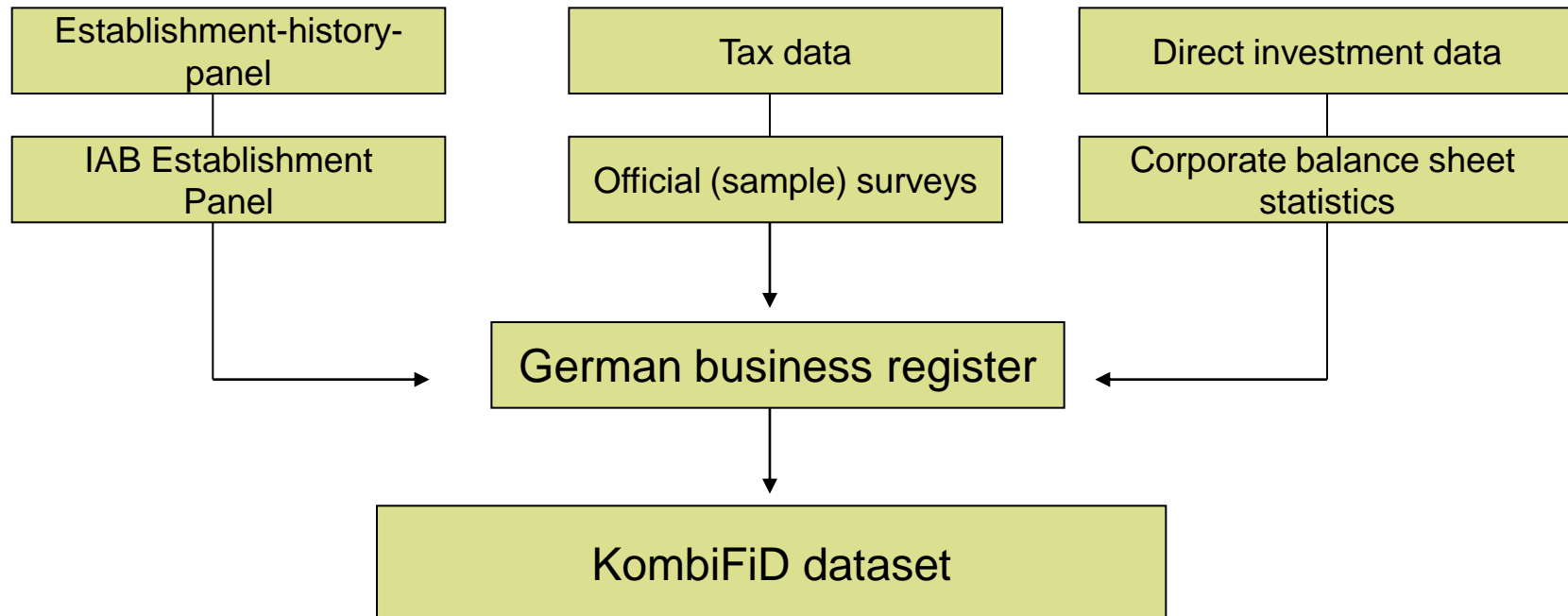
Combining firm datasets of different data producers for the first time in Germany; time horizon 1993 - 2006



New dataset for researchers with:

- a variety of variables,
- a variety of data generating processes
- a variety of economic sectors

Reduce respondent burden for the firms



Matching Results: BHP and Business register (URS)

Year	Identified firms in BHP (%)
------	--------------------------------

2003	96.20
------	-------

2004	96.10
------	-------

2005	95.70
------	-------

2006	95.00
------	-------

2007	94.40
------	-------

2008	94.00
------	-------

- 84% - 88% of URS firms consist of one establishment

- around 99% of those identified in BHP

Source: FDZ-Methodenreport 01/2010, Anja Spengler, Kerstin Lorek (only in German)

4. Data access

- Meta data
 - DDI 3 standard
 - web based search across datasets
 - tools suited for administrative and survey data

- JoSuA - Job Submission Application
 - web tool to make the submission easier
 - provided by IDSC

4. Data access: RDC-in-RDC

- Ways of data access at the FDZ:
 - on-site use
 - remote data execution
 - scientific use files
- Due to legal restrictions: **no** remote data access
- Problems:
 - on-site use: guest stays are costly and time consuming
 - remote data execution: researcher has to „program at random“
- Step towards remote data access: RDC-in-RDC approach

4. Data access: Conception of RDC-in-RDC

- Idea:

Comparable standards in nearly all RDCs over the world

⇒ Access to sensitive BA/IAB data from other RDCs or comparable institutions

- Locations:

- Germany:

Statistische Landesämter (Statistical offices of the German states)

Berlin, Bremen, Düsseldorf, Dresden

- United States:

MiCDa Enclave, University of Michigan/Institute for Social Research, Ann Arbor

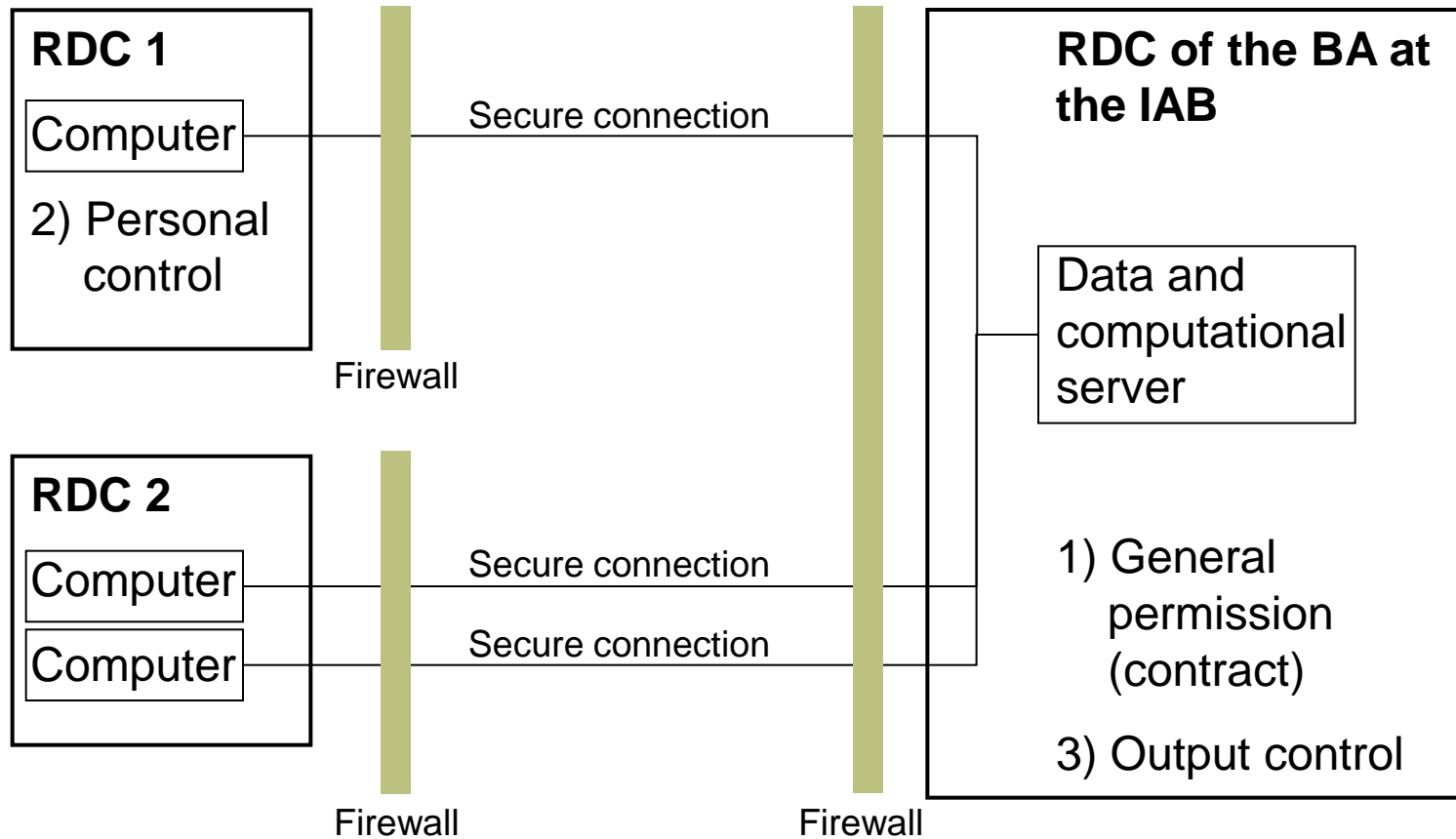
4. Data access: Conception of RDC-in-RDC

- Requirement: physical access control
 - Germany: comparable standards at the RDCs of the Statistische Landesämter ⇒ no on-site staff is needed
 - Sites abroad: (German) trained on-site staff is required

- Funding of the project: Grant from the Federal Ministry of Education and Research for three years

- Evaluation after two years

4. Data access: RDC-in-RDC who does what?



4. Data access: RDC-in-RDC outlook

- Time schedule:
 - Fall 2010: start of technical implementation in Germany and Ann Arbor
 - Trial period as of the beginning of 2011
 - Regular operation after successful trial period

- Further sites are planned:
 - Germany
 - Italy (Rome)
 - UK (Colchester, University of Essex)

Summary

- Micro data for German and foreign researchers
 - harmonization of administrative data
 - new surveys: ALWA, waves of WeLL and PASS

- Better access for researchers from abroad
 - RDC-in-RDC

- Linkage of BA/IAB data
 - Patent data
 - KombiFiD
 - ...



Research Data Centre (FDZ)
of the German Federal
Employment Agency (BA)
at the Institute for
Employment Research (IAB)

Contact:

peter.jacobebbinghaus@iab.de

<http://fdz.iab.de>

Selected datasets



1. Federal german business register (URS)

- Implemented in late 90s
- Regulated by european law (Council regulation No 2186/93)
- Master file

- Identifier:
 - Unique business identifier
 - All corresponding establishment identifiers
 - All corresponding tax numbers

Selected datasets



2. Mix of official (sample) surveys

- Cost structure surveys
- Information about added value, turnover, subsidies,...
- Salary and wage-earning set up surveys
Information about wages, salaries, working hours,
vacation entitlement

3. Tax data

- Turnover tax
- Corporate income tax

Selected datasets

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit



1. Process generated data:

Personal data from the employment notification
annual aggregation on establishment level
(establishment number)

Establishment-History-Panel (BHP)

2. Survey data:

IAB Establishment Panel

- Annual panel (10.000 establishments)
- Conducted by the IAB since 1993
- Information about various and changing topics (e.g. labour demand)

Selected datasets

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit



1. Process generated data:

Personal data from the employment notification
annual aggregation on establishment level
(establishment number)

Establishment-History-Panel (BHP)

2. Survey data:

IAB Establishment Panel

- Annual panel (10.000 establishments)
- Conducted by the IAB since 1993
- Information about various and changing topics (e.g. labour demand)

Selected datasets



Process generated data:

- **Microdatabase Direct Investment (MiDi)**
 - Annual firm data on foreign direct investment stocks
 - Due to german trade and payment regulation
- **Corporate Balance Sheet Statistics (Ustan)**
 - Annual balance sheet statistics of non-financial enterprises
 - main economic sectors: manufacturing, construction, wholesale and retail trade

First Matching Results: BHP - URS

Years of appearance	Rate of establishments in %	Rate of firms in %
1	10.89	3.06
2	9.06	3.34
3	6.13	3.11
4	6.45	4.55
5	7.15	6.30
6	60.32	79.65

⇒ High potential for longitudinal analysis

Source: FDZ-Methodenreport 01/2010, Anja Spengler , Kerstin Lorek (only in German)