

# Data and Data Access at the Research Data Center of the German Federal Employment Agency at the Institute for Employment Research

INFO7470

Understanding Social and Economic Data

Session 8: Census Geography – Restricted Access Data

03 / 21 / 2016

# Agenda

1. Administrative Data Sources
2. Administrative Research Data
3. Data Products Available at FDZ (selection)
4. Data Access

# 1. Administrative Data Sources

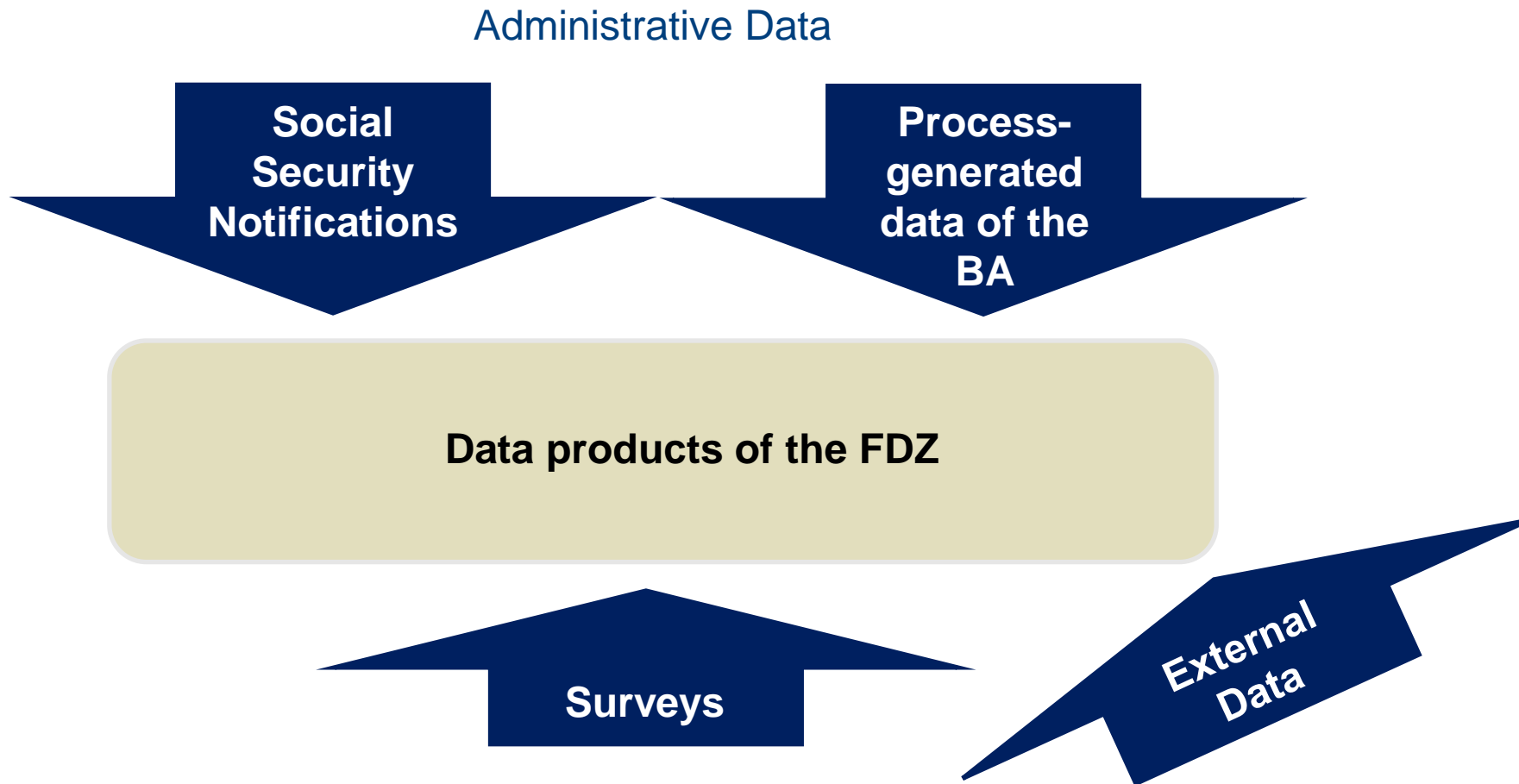
# The Data Producers (BA, IAB and FDZ)

- BA: Federal Employment Agency
  - Labor administration
  - Calculation of social security contributions and unemployment benefits
- IAB: Institute for Employment Research
  - Research department of BA
  - Preparation of admin data for academic purposes
  - Linkage to (own) survey data
- FDZ: Research Data Center of BA at IAB
  - Established in 2004
  - Provides access to confidential micro data of BA/IAB for academic purposes
  - Standardized data products
  - Since 2011: Remote access centers in Germany and USA



## Data Available at FDZ

- Micro labor market data on individuals, households, and establishments



# Notification Procedure to the Social Security Institutions



- Employers are required by law to notify the social security institutions on:
  - Every employee and marginal worker covered by the social security system
  - At the beginning or end of employment, employment interruption, change of health insurance, etc.
  - At least once a year

PRO: Highly reliable and detailed information on employment spells

CON: No information on civil servants, freelancers, or self-employed

# Notification Procedure to the Social Security Institutions cont'd

The screenshot shows a web browser window with the title 'sv.net/classic 14.0.0 - [Meldung zur Sozialversicherung - Ende der Beschäftigung]'. The browser's address bar and menu bar are visible. The main content area displays a form with the following sections:

- Personalangaben:** Includes fields for 'Personalauswahl', 'Versicherungsnummer', 'Personalnummer (freiwillige Angabe)', 'Name', 'Vorsatz', 'Zusatz', 'Titel', 'Vorname', 'Straße und Hausnummer', 'Land', 'Postleitzahl', and 'Wohnort'.
- Grund der Abgabe:** Includes 'Grund der Abgabe' (with a dropdown set to '30'), 'Entgelt in Gleitzone', and 'Namensänderung'.
- Beschäftigungszeit:** Includes 'von' and 'bis' dates, 'Betriebsnummer des Arbeitgebers', 'Personengruppe', 'Mehrfachbeschäftigung', and 'Betriebsstätte'.
- Betragsgruppen:** Includes 'KV', 'RV', 'ALV', 'PV', 'Angaben zur Tätigkeit', and 'Aktuelle Staatsangehörigkeit'.
- Beitragspflichtiges Bruttoarbeitsentgelt:** Includes 'DM', 'Euro', and 'Statuskennzeichen'.
- Unfallversicherung:** Includes 'Betriebsnummer des UV-Trägers', 'Mitgliedsnummer beim UV-Träger', 'Grund', 'Betriebsnummer Gefährtarbeitsstelle', 'Gefährtarbeitsstelle', 'UV-betrpfl. Arbeitsentgelt', and 'Arbeitsstunden'.
- Wenn keine Versicherungsnummer angegeben werden kann:** Includes 'Geburtsname', 'Vorsatz', 'Zusatz', 'Geburtsort', 'Geburtsdatum', and 'Geschlecht'.
- Name der Einzugsstelle (Geschäftsstelle):** Includes 'Datum' (07.02.2014), 'Firma', 'Straße', 'PLZ', and 'Ort'.

- Social Security Number
- Establishment Number
- Last Name
- First Name
- Address
- Reason for Notification
- Period of Employment (on a daily basis)
- Nationality
- School Education
- Vocational Training
- Type of Employment
- Wages
- Occupation (since 2011: new classification)
- Industry code

## Process-Generated Data of BA

- Data from operative processes of BA:
  - Information on earnings replacement/unemployment benefit receipt
  - Participation in labor market programs
  - Registered job search
- Daily information on benefit receipt and amounts at the individual level
- Social security number or BA customer ID, and several other variables



## 2. Administrative Research Data

# Main Database – The Integrated Employment Biographies (IEB)

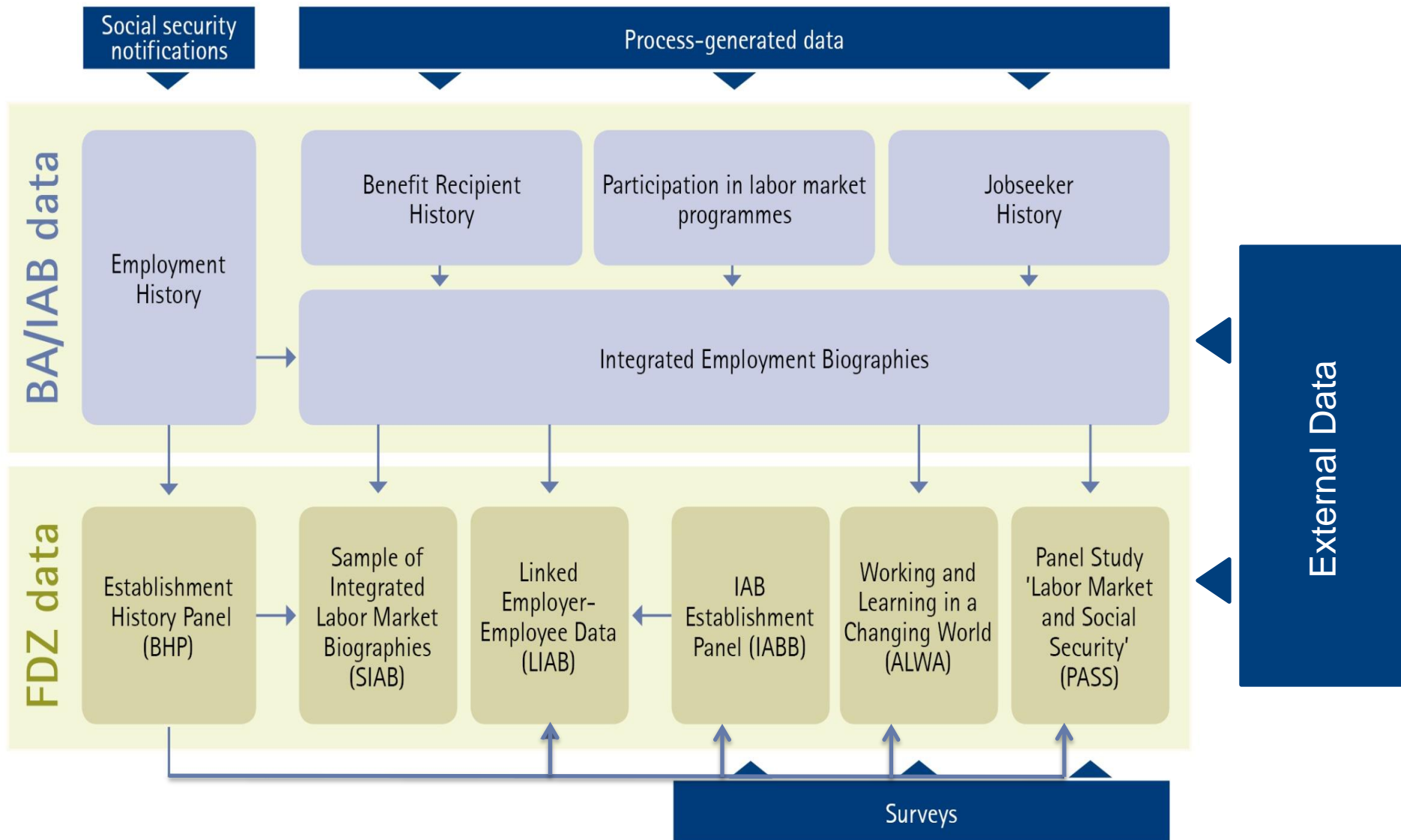
- IAB merges social security records and BA data ⇒ (complete) individual employment biographies



- Employment history covered by social security system (since 1975)
- Unemployment benefit receipt (SGB III since 1975, SGB II since 2005)
  - Registered job search (since 2000)
  - Participation in labor market programs (since 2000)
- Current version: Information on 86.090.505 individuals

### 3. Data Products Available at FDZ (selection)

# Data Available at FDZ - Overview



# Establishment History Panel (BHP)

- Repeated cross sectional dataset on all establishments in Germany with
  - at least one employee liable to social security (until 1998)
  - and/or at least one marginal worker (since 1999)
- Aggregation of the individual admin data by establishment IDs and year
- Two versions:
  - 50% random sample
  - Random sample stratified by eight establishment sizes
- Information on
  - employment structure of the establishment
  - worker in- and out-flows
  - indicators of establishment entries and exits

## Establishment History Panel (BHP) cont'd

|                 |   |
|-----------------|---|
| Period covered  | West-Germany: 1975 - 2010<br>East-Germany: 1992 – 2010  |
| Time reference  | June 30   |
| Number of cases | Between approx. 650,000 and 1.3 million establishments p.a.<br>(in the samples)   |
| Topics          | <b>Establishment information (examples):</b> <ul style="list-style-type: none"><li>• Total number of employees</li><li>• Number of employees by main occupation</li><li>• Employee age structure</li><li>• Division of employees by educational and vocational qualifications</li><li>• Wage structure of full-time employees</li><li>• Branch of industry</li><li>• Number of inflows and outflows</li><li>• Classification of establishment entry and exit</li><li>• Etc.</li></ul> |

# Sample of Integrated Labor Market Biographies (SIAB)



- 2 percent random sample drawn from the IEB database
- May be linked to the Establishment History Panel (BHP) data

|                 |  |
|-----------------|--|
| Period covered  | 1975 – 2010 (2011) (East-Germany since 1992)   |
| Time reference  | Employment biographies on a day-to-day basis   |
| Number of cases | 1,639,325 individuals or 1,594,466 individuals   |
| Topics          | <b>Socio-demographic characteristics (examples):</b><br><br>Gender, year of birth, nationality, marital status, number of children, school leaving qualification, etc. |

# Sample of Integrated Labor Market Biographies (SIAB) cont'd



Topics  
(continued)

**Information on employment, benefit receipt, job search (examples):**

Beginning and end of employment, benefit receipt and job search, daily wage, daily benefit rate, occupation, termination of last job, etc.

**Information on location (examples):**

Place of residence, place of work (federal state, county, regional directorate, employment agency)

**Information on establishments (examples):**

Branch of industry, total number of employees, median wage of the establishment , etc. ⇒ **Establishment History Panel**



# Linked-Employer-Employee Data (LIAB)

- Worker information from the IEB database are linked to the **IAB Establishment Panel** survey data
- IAB Establishment Panel:
  - Annual establishment survey
  - West-Germany since 1993, East-Germany since 1996
- Two versions of the LIAB data are available:
  - **Cross-sectional model:** *all* establishment surveys from the IAB Establishment Panel combined with administrative worker data from IEB database *on June 30<sup>th</sup>* (linked establishments: 4,188 – 14,981 per year; workers: 1,629,542 – 2,584,520 per year)
  - **Longitudinal model:** *selection of repeatedly interviewed* establishments from the IAB Establishment Panel combined with *longitudinal* worker biographies from the IEB database (linked establishments: 2,702 – 11,117 per year; workers: 1,090,728 – 1,536,665 per year)

## Panel ‘Labor Market and Social Security’ (PASS)

- Annual household survey started in 2006/2007: Labor market and poverty situation in Germany
- Situation of recipients of benefits in accordance with SGB II (recipients of Unemployment Benefit II (“Hartz IV”))
- Two random samples from two partial populations:
  - Persons and households in receipt of Unemployment II Benefits
  - All households with persons registered as residents of Germany
- 7 waves are currently available
- Combined data: PASS – IEB (PASS-ADIAB)

# Biographical data of selected Insurance Agencies in Germany (BASiD)



- 1% sample of the population of the German Pension insurance related to the sample of insured persons and their insurance accounts 2007 (Versichertenkontenstichprobe - VSKT)
  - Employment liable to social security (in the data since **1951**)
  - Marginal employment (in the data since 1999)
  - Self-employment with voluntary pension insurance
  - Receipt of unemployment benefits according to the German Social Code Book III (in the data since 1975) or German Social Code Book II (in the data since 2005)
  
- Further information from the IEB database are added:
  - Times of job-seeking registered at BA (in the data since 2000)
  - (planned) participation in labor market programs

## IAB-SOEP Migration Sample (IAB-SOEP MIG)

- Household survey conducted by IAB and German Socio-Economic Panel (SOEP) at DIW Berlin
- First wave 2013: direct interviews with 4,964 persons in 2,723 households
- Individual and household questionnaire
- Survey address various aspects of immigration including
  - Migration history
  - Education history
  - Employment history
  - Labor market background
- Combination of IAB-SOEP MIG with IEB data will be published in the near future

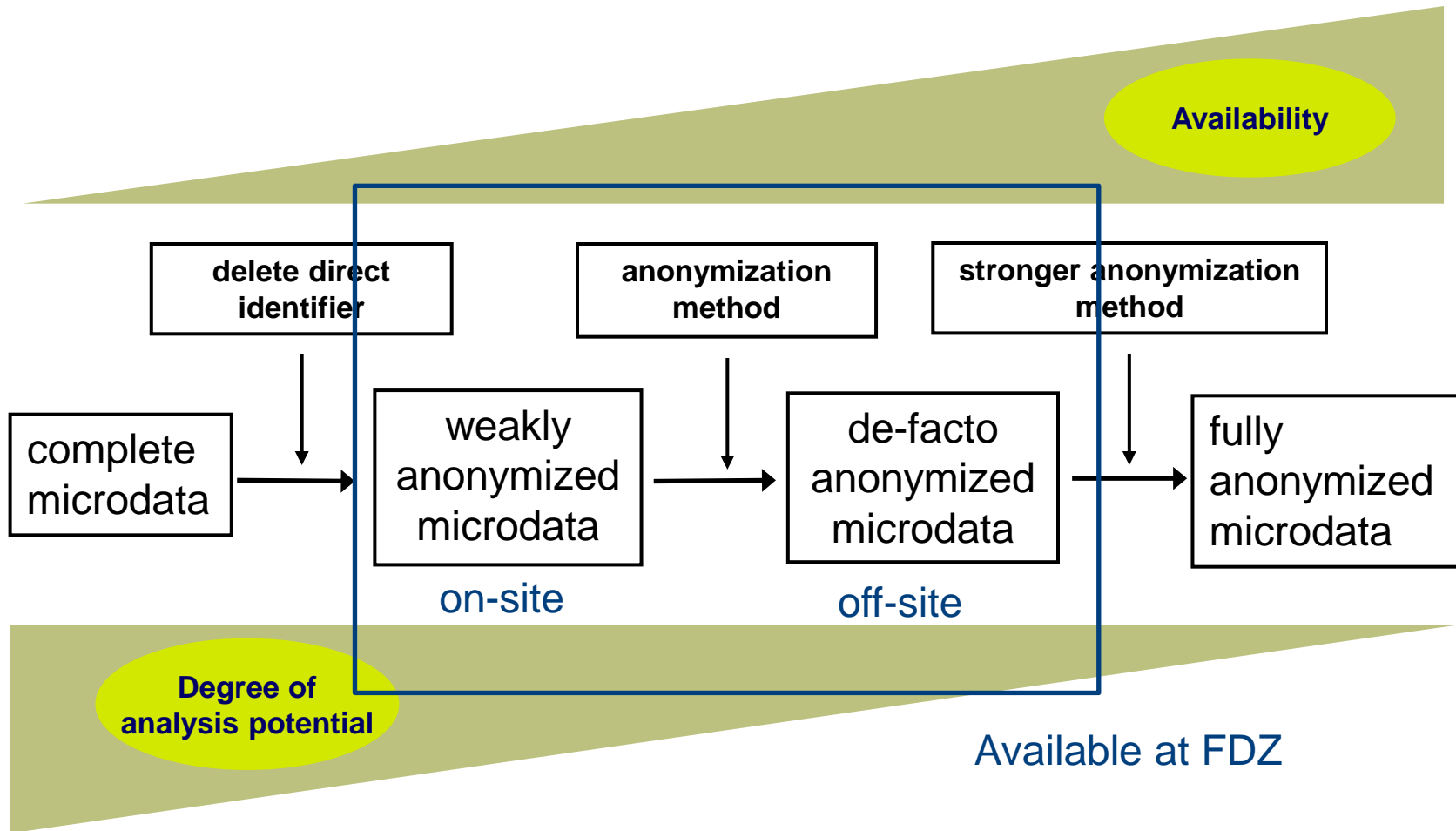
## New Data Developments (further examples)

Most recent examples:

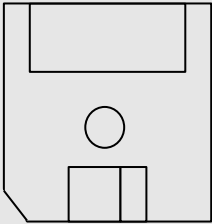
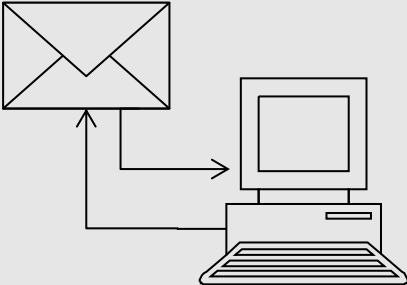
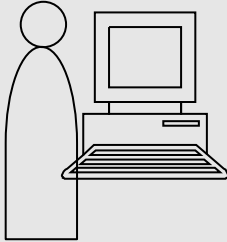
- Geocoded data
- Patent data
- Commercial business data dating back to 2006  
(source: Bureau van Dijk, BvD): Combined BvD-IEB data

## 4. Data Access

# Data Access - Types of Data



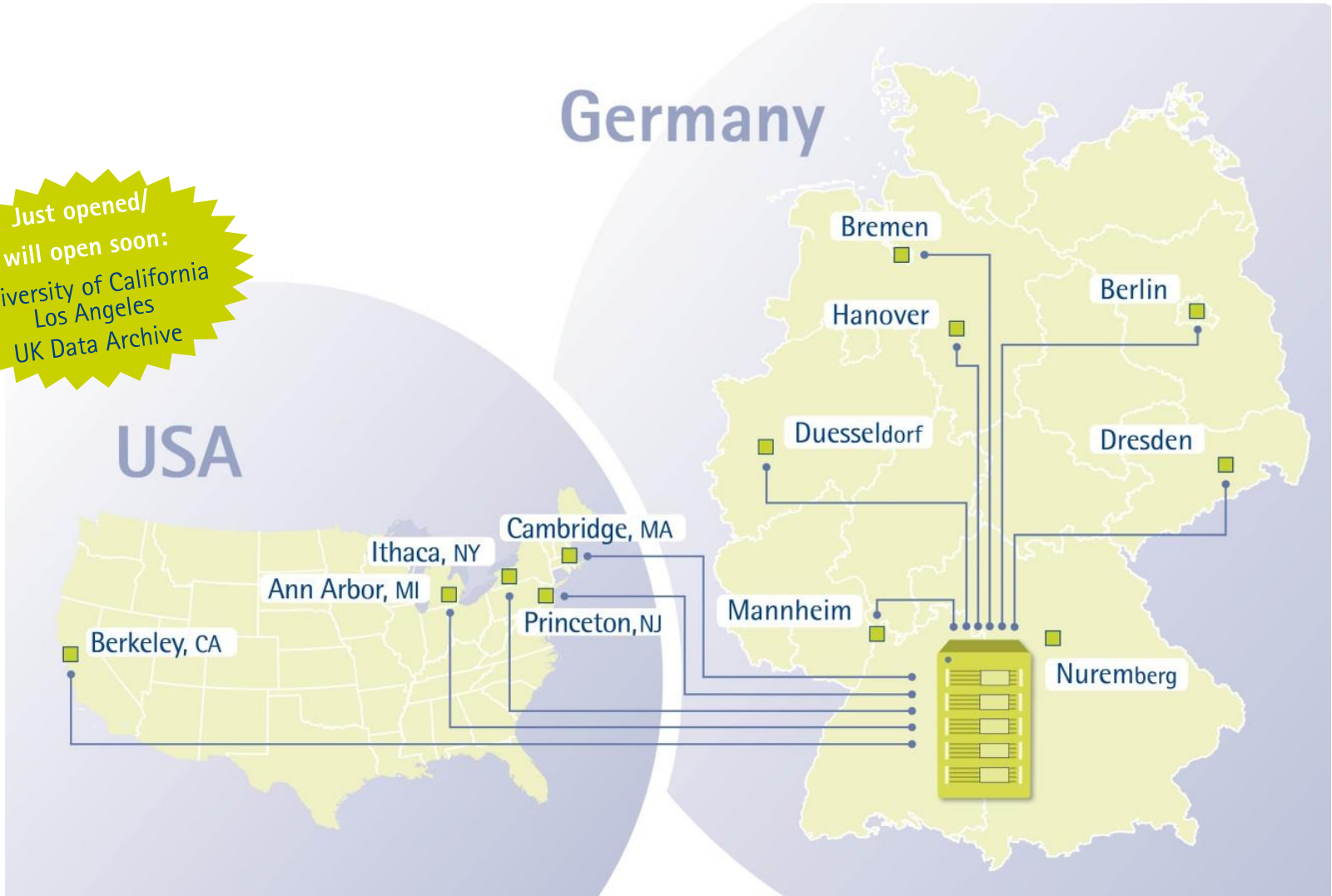
# Access Modes

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

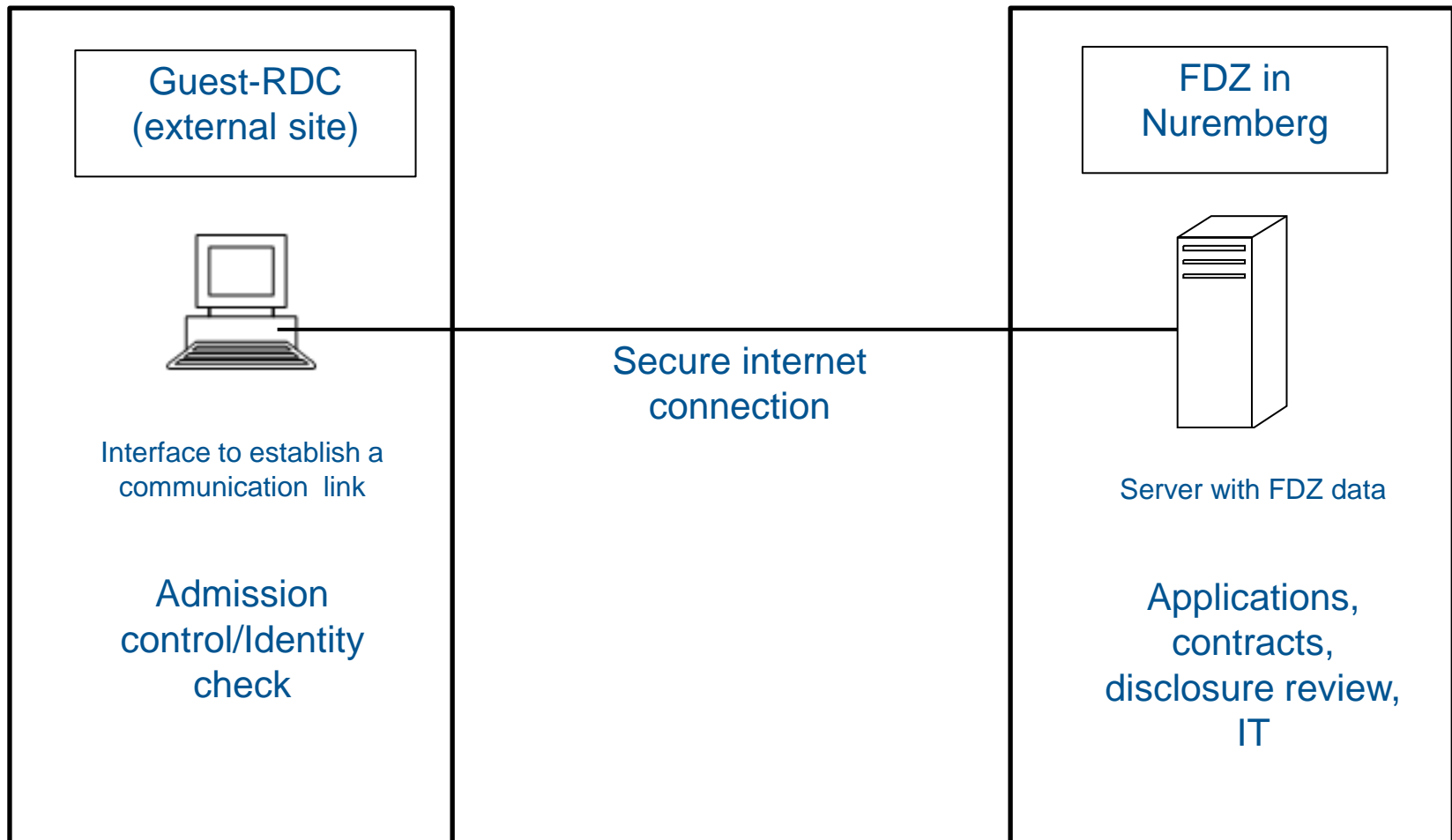


# Remote Access Centers of FDZ

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



# On-site Access



# Remote Execution with JoSuA

- JoSuA is a *Job Submission Application* developed by the Institute for the Study of Labor (IZA).
- Users upload programs in JoSuA from outside the FDZ online.
- These programs must follow the guidelines for the remote data access and must have been tested in advance using the test data (“Blacklist”).
- The results are available in JoSuA after verification of compliance with data protection legislation.
- Two types of output:
  - „internal use“; output automatically reviewed; appears directly on your screen; **must** not be extracted.
  - „publication/presentation“: output is reviewed by FDZ staff; can be extracted after review.

# Easy, Quick and Cheap Access

## ■ Easy

- Non-technical project proposal ⇒ <http://fdz.iab.de/en.aspx>
- Approval by FDZ (off-site access) or BMAS (on-site access)
- Use agreement with the institution of the researcher

## ■ Quick

- (Estimated) Time until user/institution receives agreement:
  - 2 weeks for off-site access (scientific use file)
  - 6 to 8 weeks for on-site access

## ■ Cheap

- Access to standard FDZ data is free of charge
- No lab fees
- No restrictions on hours/visits of on-site use facilities or runs of remote executions

**Thank you!**

**Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit  
(BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB)**

Regensburger Straße 100  
90487 Nuremberg, Germany

+49-(0)911-179-1752

+49-(0)911-179-1728

[iab.fdz@iab.de](mailto:iab.fdz@iab.de)

<http://fdz.iab.de/en.aspx>

Matthias Umkehrer