We describe the linkage of commercial firm data from Bureau van Dijk (BvD) with administrative employment data of the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB). BvD is a commercial provider of firm data and business intelligence. Its databases have mainly been used for analyzing business intelligence and are well renowned with companies researching competitors, suppliers or business partners. Meanwhile, the FDZ has been providing cost-free access to administrative and survey data to researchers for over a decade.

To combine the research potential of both data sources, the FDZ has performed a record linkage of companies (independent unit) given in BvD's database orbis with the establishments (dependent subunits) given in the Establishment History Panel (BHP) of the FDZ. Thus far, no systematic linkage between BvD data and administrative data had been successful for Germany (or worldwide for that matter). The major obstacle was that the two data sources do not contain a common identifier that would allow a direct linkage. The FDZ thus performed the linkage by comparing the names of companies and establishments given in the original databases. We therefore first present the challenges of the linkage process (e.g. preprocessing, comparison, classification) and the methods the FDZ has applied overcome them.

As the FDZ's administrative establishment and employee data both contain a unique establishment number, the research potential can be increased even further. Therefore, the final data set includes longitudinal information on companies, their dependent establishments and all of their employees. Creating and providing access to this innovative and unique database entails additional challenges. The presentation will thus elaborate on the measures necessary to prepare this complex dataset for researchers and to ensure data quality and anonymity of all observational units.