

# **Big Data in der amtlichen Statistik - Möglichkeiten und Grenzen**

**Dr. Susanne Schnorr-Bäcker**

*Statistisches Bundesamt*

Big Data im Sinne von Massendaten, die elektronisch und zeitnah in so großen Mengen vorliegen, dass sie mit herkömmlichen Methoden und Verfahren nicht aufbereitet und ausgewertet werden können, sind vor allem ein „big issue“ für die amtliche Statistik.

Traditionelle Datenquellen der amtlichen Statistik sind primärstatistische Erhebungen, vor allem Stichprobenerhebungen, Verwaltungsdaten und Register. Bislang bilden diese die Basis für die rund 400 statistischen Produkte der Bundesstatistik, die darüber hinaus zu mehr als zwei Dritteln europaweit harmonisiert sind. Mehr und mehr gewinnen deutschlandweit vergleichbare kleinräumige Daten an Bedeutung in Politik, Wirtschaft, Wissenschaft und Gesellschaft. Die notwendigen Schritte für die Bereitstellung derartig hochwertiger statistischer Informationen sind etabliert und werden ständig verbessert. Für Big Data als Datenquelle sind diese traditionellen Vorgehensweisen nach dem derzeitigen Kenntnisstand höchstens bedingt anwendbar.

Bei Big Data handelt es sich um ein breites Spektrum von Datenquellen. Es reicht von frei zugänglichen Informationen aus dem Internet über Informationen und Online-Transaktionen, die überwiegend bei privaten Firmen entstehen, Sensordaten aufgrund von elektronischen Hilfsmitteln wie RFID oder Satellitenbildern bis hin zu Verhaltensdaten aus sozialen Medien (wie Facebook, Twitter) und elektronischen Suchmaschinen wie Google oder Mobiltelefonen oder anderen elektronischen Kommunikationsmitteln. Diese Datenquellen könnten für verschiedene Statistikbereiche genutzt werden, wobei eine einzelne Datenquelle auch für verschiedene Statistikbereiche in Betracht kommen kann. Aufgrund der Vielfalt von Big Data und dem Umgang damit für statistische Zwecke gibt es bislang nur wenige Pilotprojekte in Ländern mit einem gut ausgebauten amtlichen statistischen Informationssystem wie in Deutschland.

Die Arbeiten im Statistischen Bundesamt konzentrieren sich derzeit auf ausgewählte Anwendungsbeispiele für die Preisstatistiken. Eine bereichsübergreifende Beschäftigung mit Big Data in der Bundesstatistik ist ein strategischer Arbeitsschwerpunkt für 2015. Darüber hinaus arbeitet das Statistische Bundesamt an Projekten auf europäischer und internationaler Ebene zur Weiterentwicklung des statistischen Systems in Hinblick auf Big Data mit. Über den aktuellen Stand soll berichtet werden.