

GIS Technologie News

SOGI-Informationsblatt Nr 3/2015

Big Data

Immer mehr Daten fallen in kürzerer Zeit an. Big Data wird als Wirtschaftsmotor der digitalen Welt angesehen. Was wird als Big Data bezeichnet und welche Bedeutung hat Big Data für GIS?

Big Data wird allgemein mit grosse (big) Daten (data) übersetzt. Der Begriff steht aber auch für die Technologien, die zum Sammeln und Analysieren dieser Datenmengen verwendet werden. Die Kosten und die Zeitintensität sind in den vergangenen Jahren stark gesunken. Dadurch können grosse Datenmengen aus vielfältigen Quellen, mit hoher Verarbeitungsgeschwindigkeit erfasst und gespeichert werden. Die Aufbewahrung erfolgt über eine lange Zeit und die Daten können für beliebige und noch unbestimmte Zwecke verwendet werden. Es erfolgt ein Sammeln auf Vorrat.

Big Data werden durch vier V's beziehungsweise Merkmale beschrieben:

- **Volume:** Big Data sind grosse Datenmengen
- **Velocity:** Big Data ändern bzw. vergrössern sich schnell
- **Varietät:** Big Data zeichnet sich durch Vielfalt und unterschiedlicher Beschaffenheit aus
- **Value:** Durch die Datenanalyse entsteht eine neue Information, welche Mehrwert schafft

Daten aus unterschiedlichen Quellen, welche sich bisher nicht aufeinander bezogen haben, werden miteinander kombiniert. Unstrukturierte und heterogene Informationen werden verknüpft und ausgewertet. In der Analyse wird nach Mustern gesucht, um daraus neue Erkenntnisse zu gewinnen. Bisherige Methoden und Speichersysteme (wie Relationale Datenbanksysteme) sind dafür wenig performant und geeignet. Für die Analyse der verschiedenen Informationstypen (Video, Bild, Text, Zahlen, Koordinaten) sind gleichzeitige und zeitnahe Verarbeitungen gefordert.

Viele Anwendungen erfolgen zum Beispiel in der Werbung und Marktforschung, in der Elektrobranche (Smart Metering) und Telekommunikation (Billingsysteme), bei Finanztransaktionen (Fraud-Detection, Kreditkartenmissbrauch), bei der Vorhersage von Wetter und Epidemien, bei Rückversicherungen oder in der geheimdienstlichen oder polizeilichen Arbeit (Erstellung von Bewegungsprofilen, Früherkennung von Verbrechen). Zu erwähnen ist auch der Datenzugriff und die -Analyse auf raumzeitlichen Rasterdaten (Web Coverage Service von OGC).

GIS beschäftigt sich hauptsächlich mit Small Data. Mit geografischer Analyse von Social Media und beispielsweise Hadoop, ein Java basiertes Framework von Apache, stossen auch heutige GIS in Big Data vor. Big Data bietet neue Chancen insbesondere für soziale oder wissenschaftliche Erkenntnisse und bildet eine veränderte Form der Wertschöpfung für die Unternehmen. Die neu geschaffenen Informationsinhalte sind nicht richtig oder falsch, sie stellen eine Wahrscheinlichkeit oder Interpretation dar!

Für eine in Betracht zu ziehende Anwendung Big Data sind folgende Fragen zu beantworten:

- Sind grössere und heterogene Daten die qualitativ besseren Daten und sind alle Daten gleich wertvoll?
- Ist es ethisch vertretbar die verfügbaren Daten zu verwenden (Wahrung der Privatsphäre und Schutz von personenbezogenen Daten)?

Big Data ist bereits gestern. Weiter wird von Fast Data und Smart Data gesprochen, aber davon vielleicht ein anderes Mal.

Fachgruppe GIS Technologie
technologie@sogi.ch
Urs Flückiger

Quellen:

- Wikipedia (https://de.wikipedia.org/wiki/Big_Data)
- SRF Einstein (<http://www.srf.ch/wissen/big-data-das-grosse-vermessen>)
- EDÖB (<http://www.edoeb.admin.ch/datenschutz/00683/01169/index.html?lang=de>)