

Streiflicht

Big Data in der Landwirtschaft – Perspektiven eines Daten- dienstleisters

**Dr. Martin Kunisch**

Foto: privat

Das Thema Big Data ist allgegenwärtig und auch für die Landwirtschaft von höchster Aktualität. So ist die Zahl der Veranstaltungen, die sich um dieses Thema ranken, sprunghaft in die Höhe geschneit. Zunächst stellt sich die Frage, was sich hinter Big Data verbirgt und welche Rolle Big Data heute in der Landwirtschaft spielt:

Big Data bezeichnet Datenmengen, die zu groß oder zu komplex sind oder sich zu schnell ändern, um sie manuell oder mit klassischen Methoden der Datenverarbeitung auszuwerten. Dementsprechend setzt man für Big-Data-Analysen eigens entwickelte Methoden der Datenspeicherung und -auswertung ein. Big Data wird anhand von vier Kriterien, die auch als die vier „V“ bezeichnet werden, charakterisiert: Volume, Velocity, Variety und Veracity.

- **Volume:** Relevante Datenmengen beginnen im Terabyte-Bereich. Einzelbetriebliche Aufzeichnungen aus dem ISOBUS oder aus dem Stall liegen volumenmäßig noch nicht im Bereich von Big Data.
- **Velocity:** Spielt in der Landwirtschaft derzeit eine untergeordnete Rolle. Bei der Fortentwicklung von Analysen von Pflanzenbeständen während der Überfahrt oder des Überfluges oder von Telemetrie-Plattformen können künftig aber höhere Verarbeitungsgeschwindigkeiten gefordert sein.
- **Variety:** Damit sind Daten in vielen verschiedenen Formaten und sehr unterschiedlichem Strukturierungsgrad gemeint. Dieses Kriterium kommt dann zum Tragen, wenn in der Landwirtschaft damit begonnen wird, vielfältige Datenquellen zu nutzen, um komplexe Szenarien abzubilden.
- **Veracity:** Notwendigkeit, Daten zu verarbeiten, deren Aussagekraft nicht endgültig gesichert ist. Hier fehlen entsprechende Erfahrungswerte. Es ist zu erwarten, dass die Akzeptanz in der Landwirtschaft zumindest anfänglich vergleichsweise gering sein wird.

Demnach ist die Landwirtschaft von heute nur bedingt ein Einsatzgebiet für Big-Data-Technologien.

Häufig wird Big Data in einem Atemzug oder gar synonym mit Entwicklungen genannt, die unter dem Begriff „Digitalisierung der Landwirtschaft“ zusammengefasst werden können. Dahinter verbergen sich Entwicklungen der Automatisierung und Robotik einerseits sowie des Datenaustauschs und der Vernetzung andererseits. Auch wenn die Digitalisierung boomt und aktuell mehr denn je neue Dienstleistungen rund um Daten angeboten werden: der Zusammenhang mit Big Data ist nicht zwingend. Allerdings bleibt festzuhalten, dass bei Fortschreibung der genannten Entwicklungen über die nächsten Jahre auch in der Landwirtschaft ein Datenfundus entsteht, der sich für die Auswertung mit Big-Data-Methoden nahezu aufdrängt. Die Landwirtschaft ist eine sehr komplexe, multifaktoriell beeinflusste Produktionslandschaft, die zu einem großen Teil unter freiem Himmel wirtschaftet und

dadurch sehr vielfältig und schwer planbar ist. Das heißt, sie ist intensiv mit der Umwelt und zusätzlich mit der Wirtschaft und der Gesellschaft vernetzt.

Es ist eine Chance, dass sich mit Big-Data-Technologien Zusammenhänge aufdecken und Fragen beantworten lassen, deren Erforschung bislang an der Komplexität des Systems und der Interdisziplinarität der Forschenden gescheitert ist. Es wäre fatal, wenn diese Erkenntnisse einer kleinen Zahl weltumspannender Datenkonzerne überlassen blieben, die zwar hochkompetent in Big Data, aber noch nicht – zumindest nicht erkennbar – kompetent in Landwirtschaft sind.

In diesem Sinne wäre es hilfreich, die Datenbestände zur landwirtschaftlichen Produktion und Forschung im Rahmen von Public-Private-Partnership und im Sinne Open Source und Open Access verfügbar zu machen. Es ist ein vielleicht ideales, aber doch schönes Bild, wenn Daten aus Behörden, der Zuliefer- und verarbeitenden Industrie, aus der landwirtschaftlichen Produktion sowie den Forschungsdisziplinen, die mit Landwirtschaft zu tun haben, zusammengeführt würden und einer gemeinsamen Auswertung zur Verfügung stünden. Und wenn wir zögern oder uns gar verweigern, wird es nach den Zuboff'schen Gesetzen letztendlich doch gemacht, und zwar von Branchenfremden, ohne dass wir, die Landwirtschaft, eine Chance haben, daraus Nutzen zu generieren.

Als Konsequenz sollte die landwirtschaftliche Forschung dem Anstoß der Deutschen Agrarforschungsallianz (DAFA) folgen und sich zügig damit beschäftigen, welche Fragen mit Big-Data-Methoden zu beantworten und welche Maßnahmen zu ergreifen sind, um dies zu ermöglichen. Big Data bietet neue Perspektiven für die Simulation bestimmter Szenarien anhand komplexer Modelle auf einer umfassenden Datenbasis. Vielleicht kann man dies als eine Art „Virtuelles On-Farm Research“ bezeichnen. Dazu müssen wir die Infrastruktur für ein gemeinsames öffentliches Forschungsdatenmanagement schaffen und die Bereitschaft mitbringen, individuelle Datenbestände und Erkenntnisse in diesen öffentlichen Pool einzubringen. Auch muss die Entwicklung gemeinsamer komplexer Modelle im Vordergrund stehen, an die untergeordnete Teilmodelle andocken können. Und am meisten Überwindung wird es kosten, Akzeptanz dafür zu entwickeln, dass möglicherweise nicht vollständig belastbare Aussagen zu einzelnen Sachverhalten getroffen werden. Bei Lichte betrachtet, haben jedoch die klassischen Methoden auch ihre Unzulänglichkeiten. Big-Data-Methoden hingegen bieten die Chance, die Unsicherheiten in Modellen transparent zu machen.

Das KTBL beschäftigt sich in seiner Funktion als Datendienstleister intensiv mit diesen Entwicklungen. In der von Präsidium und Hauptausschuss getragenen KTBL-Strategie 2014 wurde ein Arbeitsschwerpunkt „Datenbanken und Wissenstechnologien“ etabliert, in dem Projekte zur Digitalisierung in der Landwirtschaft insbesondere im Hinblick auf die Rolle des KTBL als Datendienstleister gebündelt werden. Dies basiert auf der Erkenntnis, dass Daten aus öffentlich finanzierten Projekten auch öffentlich zugänglich sein sollten sowie dem Gedanken, als Teil der zuvor genannten Forschungsdateninfrastruktur Vorreiter zu sein.

Diese Entwicklung hatte 2004 mit agroXML ihren Anfang genommen und findet in agroRDF ihre Fortsetzung. Derzeit sind die ersten Planungsdaten als Linked Open Data in maschinenlesbarer Form – ergänzt um eine formale, strukturierte Beschreibung – im Internet verfügbar. Hierdurch wird die KTBL-Datenbank Teil eines weltweiten Netzes von untereinander verbundenen, frei zugänglichen Datenbeständen. KTBL-Daten können so beispielsweise in Farmmanagement-Informationssystemen oder Smartphone-Apps von Drittanbietern integriert werden. Zudem ist das KTBL in der Interest Group on Agricultural Data aktiv, die Teil der Research Data Alliance ist und von der FAO getragen wird. Deren Ziel ist es, Forschungsdaten aus dem Agrarbereich mit innovativen Methoden internati-

onal zugänglich zu machen. Vergleichbare Ziele auf technischer Ebene verfolgt das KTBL auch über die Mitarbeit in der Agricultural Electronic Foundation (AEF).

Nach diesem fachlichen Blick auf Big Data darf nicht ignoriert werden, dass die wertvollsten gesellschaftlichen Errungenschaften unserer Kultur, die persönliche Freiheit und Selbstbestimmtheit auf dem Spiel stehen. Heute erfolgt Big Data meist ohne Wissen oder Bewusstsein der Gesellschaft durch Strukturen, die unter dem Deckmantel einer Dienstleistung oder der inneren Sicherheit Daten sammeln und mit Big-Data-Technologien auswerten. Wenn wir diese Situation nicht neu denken und global regeln, geht nicht nur der Nutzen an der Gesellschaft vorbei, sondern auch unsere individuelle Selbstbestimmtheit verloren.



Dr. Martin Kunisch

Hauptgeschäftsführer des KTBL