LANDTECHNIK 71(1), 2016, 1–3

DOI: 10.15150/lt.2016.3117

LANDTECHNIK
AGRICULTURAL ENGINEERING

Spotlight

Big Data in agriculture – perspectives for a service organisation

Dr. Martin Kunisch

Photo: private

The theme big data is omnipresent nowadays, and very topical for agriculture too. There is a very rapid increase in the number of events concerning big data. But just what does this term mean? And what kind of role can big data play in today's agriculture?

First of all, big data means an amount of information so massive, complex, or changing so rapidly that analysis is not possible manually or via classical data processing methods. Applied for storage and processing of big data is specifically developed software. Big data is characterised by four criteria, known as the four 'Vs': volume, velocity, variety and veracity.

- Volume: Relevant data amounts start in the terabyte range. Volumes of single farm ISOBUS or livestock information are not yet in this range.
- Velocity: To date, this plays a subordinate role in agriculture. But higher processing speeds may be required in further development of crop analyses while driving through crops, flying over them, or from telemetry platforms.
- Variety: This refers to data in many different formats and very varied degrees of complexity. Such criteria are then useful when multiple data sources are utilised in agriculture to depict complex scenarios.
- Veracity: Where data is processed without definite establishment of its validity. Missing are appropriate empirical values. Acceptance in agriculture may be comparatively limited – at least initially.

Based on the above, current agriculture is only in some cases an application area for big data technologies.

Often, big data is mentioned in the same breath, or even as synonym for, the general term "agricultural digitalisation". Behind this term lie developments in automation and robotics, but also in data exchange and networking. Even where digitalisation is now booming with increasing numbers of new data services on offer, this still does not mean that big data is necessarily involved. However, we have to accept that this will certainly change. Even in agriculture, the steady progress means information is piling up at such a rate that big data processing methods will almost certainly have to be applied. Agriculture is a very complex, multi-factorial production sector, mainly conducted under open skies and therefore very diverse and difficult to plan. This means intensive interactions with the environment, as well as with the business world and society in general.

Big data technologies allow us the opportunity to discover relationships and to answer questions, the research for which up until now has foundered on the complexity of the system and on the in-

terdisciplinary of the researchers. It would be fatal if this knowledge was left to a small number of worldwide data concerns which may be highly competent in big data but not yet – at least apparently – reliably knowledgeable as far as agriculture is concerned.

In this sense, it would be helpful if the stored data for agriculture production and research were made available within the framework of public-private partnership and in terms of open source and open access. It is perhaps an idealised concept. But it is still a beautiful one: of data from government sources, from the supply and process sectors, from agricultural production as well as the research disciplines, were all brought together for processing. And should we hesitate, or even completely baulk at such a development, then Zuboff's Laws mean it will eventually take place anyway – but carried out by people that are not acquainted with the agricultural sector, perhaps resulting in the exclusion of farming experts that could have exploited it best for agriculture.

Consequently, agricultural research should follow the urging of the German Agricultural Research Alliance (DAFA) and rapidly involve itself, deciding on the questions to be answered through big data methods and the actions needed to this end. Big data offers new perspectives for the simulation of certain scenarios with the help of complex models on a comprehensive database. Perhaps one can describe all this as a form of "virtual on-farm research". Hereby, we have to create the infrastructure for a common public form of research data management. And we have to bring a willingness to contribute individual data collections and knowledge into this pool. Also prioritised here should be the development of publically shared complex models onto which subordinate component models can be docked. Maybe the most effort will be needed to come to terms with the fact that answers arrived at concerning individual issues will possibly not be completely reliable. Seen in perspective, however, classical methods also have their inadequacies. Big data methods, on the other hand, offer the opportunity to make transparent the uncertainties that exist in models.

The KTBL is intensively involved in such developments within its function as data service organisation. In the KTBL strategy 2014 supported by the presidium and central committee, a focal point "Databases and knowledge technologies" is established which includes a project on digitalisation of agriculture, in particular with regard to the role of the KTBL as data service organisation. This is based on the acceptance that data from publically financed projects should also be publically accessible. Also attractive is the thought of establishing the organisation as a leader within the previously mentioned research data infrastructure.

This development started 2004 with agroXML and continued in agroRDF. Currently, the first planning data are accessible in Internet as linked open data in machine-readable form supplemented with a formal, structured description. Hereby, the KTBL database is part of a worldwide network of intermeshed, freely accessible databases. Already, KTBL data can be integrated in farm management information systems or smartphone apps from third-person suppliers. Additionally, the KTBL is active in the Interest Group on Agricultural Data which is part of the FAO supported Research Data Alliance. The group target is to apply innovative methods to aid international availability of research data from the agricultural sector. KTBL has comparable targets on a technical level concerning cooperation within the Agricultural Electronic Foundation (AEF).

Following this specialist view of big data, we still must not ignore its inherent risk to the most crucial social values of our culture: personal freedom and self-determination. Nowadays, big data is usually applied without society's knowledge or understanding via structures that collect data and analyse them under cover of service operations or security systems. Should we fail to think out the

implications of this situation and establish appropriate safety rules to contain developments on a global scale, then not only might we miss the target of social usefulness for big data but also lose our individual independence and freedom.

Dr. Martin Kunisch
KTBL Managing Director