Computers and Electronics in Agriculture 145 (2018) 130–141

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Open geospatial infrastructure for data management and analytics in interdisciplinary research

Jacob Høxbroe Jeppesen[a,*], Emad Ebeid[b], Rune Hylsberg Jacobsen[a], Thomas Skjødeberg Toftegaard[a]

[a] Department of Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N., Denmark
[b] Faculty of Engineering, University of Southern Denmark, Campusvej 55, 5230 Odense M., Denmark

A R T I C L E I N F O

Keywords:
Internet of Things
Remote sensing
Open software
Open data
Farm management information systems

A B S T R A C T

The terms Internet of Things and Big Data are currently subject to much attention, though the specific impact of these terms in our practical lives are difficult to apprehend. Data-driven approaches do lead to new possibilities, and significant improvements within a broad range of domains can be achieved through a cloud-based infrastructure. In the agricultural sector, data-driven precision agriculture shows great potential in facilitating the increase in food production demanded by the increasing world population. However, the adoption rate of precision agriculture technology has been slow, and information and communications technology needed to promote the implementation of precision agriculture is limited by proprietary integrations and non-standardized data formats and connections. In this paper, an open geospatial data infrastructure is presented, based on standards defined by the Open Geospatial Consortium (OGC). The emphasis in the design was on improved interoperability, with the capability of using sensors, performing cloud processing, carrying out regional statistics, and provide seamless connectivity to machine terminals. The infrastructure was implemented through open source software, and was complemented by open data from governmental offices along with ESA satellite imagery. Four use cases are presented, covering analysis of nearly 50 000 crop fields and providing seamless interaction with an emulated machine terminal. They act to showcase both for how the infrastructure enables modularity and interoperability, and for the new possibilities which arise from this new approach to data within the agricultural domain.

1. Introduction

The Internet of Things (IoT) and Big Data have gained tremendous attention in recent years, and we are currently witnessing information and knowledge from data becoming a critically important tradable asset. Most domains are becoming increasingly data-centric, however, it is difficult to gain an overview of the specific benefits one might achieve from this. Furthermore, interoperability has, and still is, an issue, and standardization is necessary for achieving modularity, where a broad range of software and hardware modules can be seamlessly connected. This calls for a standardized cloud-based infrastructure with an eco-system approach, resulting in third-party vendors becoming able to develop add-ons to existing systems, much like apps for smartphones. The agricultural sector is required to increase yield production to meet the expected doubling of crop demand from 2005 to 2050 (Tilman et al., 2011). Meanwhile, 70% of freshwater withdrawals are already devoted to irrigation and the agricultural sector is responsible for 30–35% of greenhouse gas emissions (Foley et al., 2011). This is partly caused by 50% of the global nitrogen applied not being absorbed by the crops they were distributed on (Stuart et al., 2014). Hence, there is a critical need for new methods for optimizing the agricultural domain. Precision agriculture employs technologies to manage spatial and temporal variability within fields to improve crop yield while decreasing the environmental impact. Hence, it facilitates site specific farming, such that e.g. irrigation and nitrogen fertilizer is distributed only where and when necessary. However, the adoption rate of precision agriculture has been slow, due to a range of factors, such as issues with interoperability, compatibility, and complexity (Aubert et al., 2012).

Recent advances in open source software, open standards, and open interfaces show potential for overcoming these issues. The standardization of interfaces and formats by The Open Geospatial Consortium (OGC) can ensure interoperability of geographic information systems (GIS) (Nash et al., 2009; Nikkilä et al., 2010), which is an essential part

* Corresponding author.
E-mail address: jhj@eng.au.dk (J.H. Jeppesen).

https://doi.org/10.1016/j.compag.2017.12.026
Received 6 July 2017; Received in revised form 8 November 2017; Accepted 18 December 2017
0168-1699/

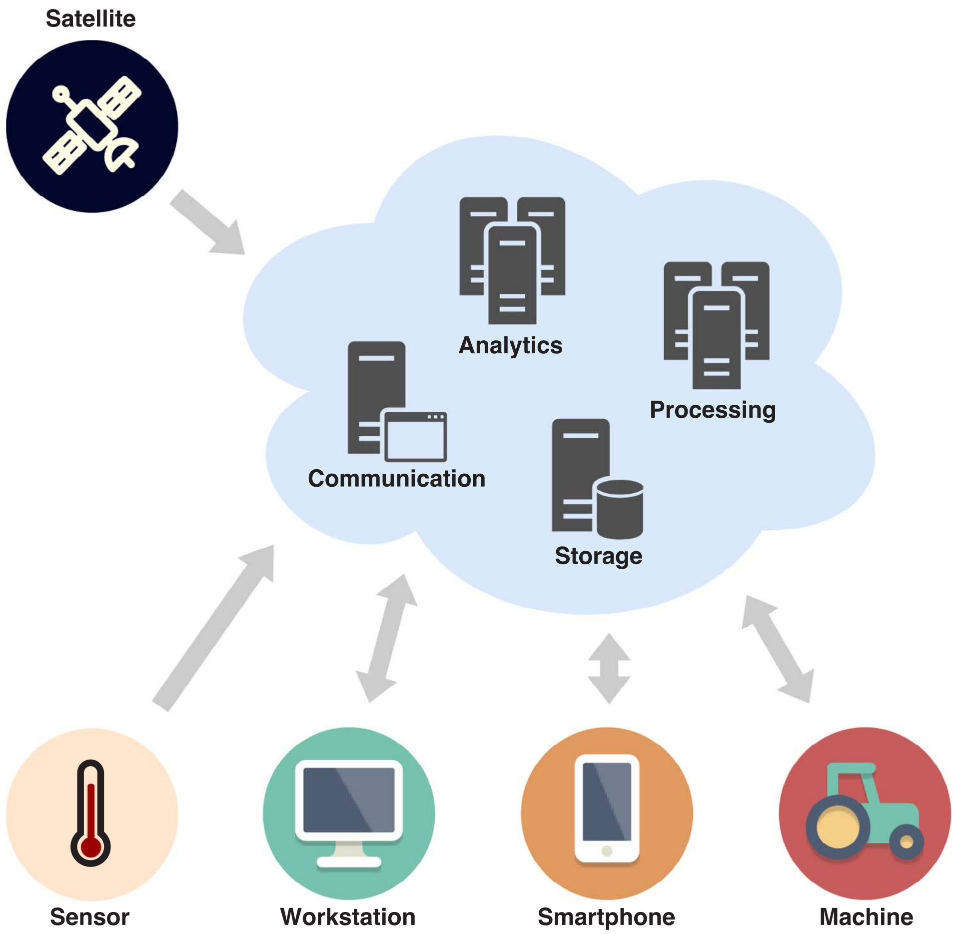

Fig. 1. An overview of a cloud-based data infrastructure, communicating with a broad range of devices.

of precision agriculture (Nikkilä et al., 2010). Furthermore, advances in cloud computing acts as an enabler for precision agriculture. The devices are interconnected, as seen in Fig. 1, and servers for storage, communication, processing, and analytics can use a variety of methods for extracting important knowledge from the vast amounts of data. Cloud computing capabilities are decreasing significantly in price, and high processing power along with large storage capabilities can be setup in an easy manner. Furthermore, open data provide valuable information for the agricultural practice. Although application maps based on satellite imagery has been found to improve profit while decreasing environmental impact (Seelan et al., 2003), the total cost of applying the system, alongside the implications regarding ease of use, made it infeasible. These issues have been significantly alleviated, due to the open multi-spectral satellite imagery from Sentinel-2 and Landsat satellites combined with the aforementioned advances in hardware and open software.

The work presented in this paper describes the geospatial data infrastructure required to facilitate data transfer, storage, processing, and analytics within precision agriculture. The contributions include an implementation of an OGC compliant infrastructure, including OGC standards for heterogeneous sensor webs. This covers important Information and Communications Technology (ICT) aspects of the precision agriculture practice, from satellite imagery or sensor data through cloud processing to tractor terminal, smartphone, or workstation. It is demonstrated how the infrastructure can be used to carry out large regional statistical analyses based on open data, which have not previously been viable. The infrastructure is implemented using open source software, and forms the foundation for interdisciplinary research with interoperability in mind, such that data can be shared in standardized formats, and visualized online. This results in the possibility for diverse research focusing on optimizing over the entire crop cycle. Finally, the infrastructure provides an easy transition to commercial products, hence coupling academic research with industrial objectives.

The remainder of this paper is organized as follows: Section 2 covers related work, including the advances in open data and the current state of Farm Management Information Systems (FMIS) with regard to precision agriculture. Section 3 presents the requirements and needs for the geospatial data infrastructure, before Section 4 describes the design and implementation of the proposed infrastructure. Section 5 validates the infrastructure through four use cases. Section 6 provides a discussion on the findings. Finally, Section 7 concludes the paper.

2. Related work

The Internet has provided an unprecedented platform for data sharing, where the cost of distribution to the public has become marginal. The availability of open data has progressed tremendously in recent years, and in 2013, the G8 countries decided to be trendsetters in open data policies (Castro and Korte, 2015). The financial benefits in 2008 coming directly and indirectly from open data was estimated to be about 140 billion euro in Europe (Vickery, 2011). NASA opened data access to the LANDSAT satellite programme in 2008, and the impact of this decision is discussed in Wulder et al. (2012). They describe how the use of satellite imagery have increased significantly, and address several future benefits when data from several satellites, such as LANDSAT and Sentinel, can be used together. The record of distributed Landsat images per year before the open access was 25 000 in 2001 (Wulder et al., 2012), compared to 1 million distributed images in the first 17 days of March 2016 (Landsat). In Janssen et al. (2012), the benefits and barriers of adopting open data are discussed. The major technical barriers involve the absence of standardized data along with the lack of