Crowdsourcing for agricultural applications : a review of uses and opportunities for a farmsourcing approach

Authors order : Minet Julien, Curnel Yannick, Gobin Anne, Goffart Jean-Pierre, Mélard François, Tychon Bernard, Wellens Joost, Defourny Pierre.

- Julien Minet, University of Liège, Arlon Campus Environnement, Avenue de Longwy, 185, B6700 Arlon
- Yannick Curnel, Walloon Agricultural Research Center, Farming Systems, Territories and Information Technologies Unit, Rue de Liroux, 9, B5030 Gembloux, y.curnel@cra.wallonie.be
- Anne Gobin, Flemish Institute for Technological Research (VITO), Boeretang 200, B2400 Mol, anne.gobin@vito.be
- Jean-Pierre Goffart, Walloon Agricultural Research Center, Bâtiment A. Petermann, Rue du Bordia, 4, B5030 Gembloux. j.goffart@cra.wallonie.be
- François Mélard, University of Liège, Arlon Campus Environnement, Avenue de Longwy, 185, B6700 Arlon, f.melard@ulg.ac.be
- Bernard Tychon, University of Liège, Arlon Campus Environnement, Avenue de Longwy, 185, B6700 Arlon, bernard.tychon@ulg.ac.be
- Wellens Joost, University of Liège, Arlon Campus Environnement, Avenue de Longwy, 185, B6700 Arlon, joost.wellens@ulg.ac.be
- Pierre Defourny, Université catholique de Louvain, Earth and Life Institute, Croix du Sud, B1348 Louvain-la-Neuve, pierre.defourny@uclouvain.be

Keywords: Crowdsourcing, Citizen Science, Smart Farming, Participatory Approaches, Big Data, ICT, Data Collection

Highlights:

- We reviewed crowdsourcing activities in agriculture and classified them in 4 categories
- We identified 8 types of inputs that can be collected by crowdsourcing for agricultural applications
- We discussed data quality and contributors and farmers’ participation
- We introduced the concept of farmsourcing as a professional crowdsourcing strategy in farming activities

Abstract

Crowdsourcing, understood as outsourcing tasks or data collection by a large group of non-professionals, is increasingly used in scientific research and operational applications. In this paper, we reviewed crowdsourcing initiatives in agricultural science and farming activities and further discussed the particular characteristics of this approach in the field of agriculture. On-going crowdsourcing initiatives in agriculture were analysed and categorised according to their crowdsourcing component. We identified eight types of agricultural data and information that can be generated from crowdsourcing initiatives. Subsequently we described existing methods of quality control of the crowdsourced data. We analysed the profiles of potential contributors in crowdsourcing initiatives in agriculture, suggested ways for increasing farmers' participation, and discussed the on-going initiatives in the light of their target beneficiaries. While crowdsourcing is reported to be an efficient way of collecting observations relevant to environmental monitoring and contributing to science in general, we pointed out that crowdsourcing applications in agriculture may be hampered by privacy issues and other barriers to participation. Close connections with the farming sector, including extension services and farm advisory companies, could leverage the potential of crowdsourcing for both agricultural research and farming applications. This paper coins the term of farmsourcing as a professional crowdsourcing strategy in farming activities and provides a source of recommendations and inspirations for future collaborative actions in agricultural crowdsourcing.

1 Introduction

First coined in 2006 by J. Howe, the editor of the *Wired* magazine (Howe, 2006), the term crowdsourcing rapidly gained public uses in the social milieux of the internet and blogospheres. Crowdsourcing was associated (but not defined) by Howe as a new organisational form inspired by online firms such as Amazon.com using the crowd or online communities as a way to outsource several tasks. David Brabham, after a comprehensive work of literature review (Brabham, 2008; Brabham, 2009; Brabham, 2013) proposed to define crowdsourcing scientifically as… "(…) an online, distributed problem-solving and production model that leverages the collective intelligence of online communities to serve specific organizational goals" (Brabham, 2013, p. xix). According to Brabham, the specificity of crowdsourcing lies in the topical sharing of responsibilities between an organization (typically a firm) and an online community; between "a bottom-up, open, creative process [and] a top-down organizational goals" (Brabham, 2013, p. xv). In this sense, crowdsourcing practices are situated under the control of the institution (being an administration, an academic community, a corporate firm, …), i.e., the one that manages the activity and defines its objectives, or purposes. As a result, according to Brabham, open sources software or common-based peer productions such as Wikipedia should not be labelled crowdsourcing initiatives as the locus of control is in the hand of the online community.

The term "crowdsourcing" was progressively assigned to many scientific and operational initiatives aimed at collecting contributions from a large group of people. In scientific research, outstanding initiatives based on crowdsourcing managed to yield significant scientific outputs (Franzoni & Sauermann, 2014) such as the project "Foldit", where contributors can help, through the gamification of a scientific task, to improve the understanding of the structure of proteins. This project currently gathers nearly 200,000 contributors, has resulted in a number of publications in top journals and has inspired dozens of similar crowdsourcing initiatives in biomedical research (Belden et al., 2015). Often denominated as community-based monitoring, citizen sensing, or citizen monitoring, the majority of crowdsourcing initiatives aim at collecting environmental and wildlife observations by volunteers (Roy et al., 2012). Besides this major field of application, crowdsourcing initiatives were reported in the fields of astronomy (Raddick et al., 2010), meteorology (Muller et al., 2015), cartography (Heipke, 2010), mathematics (Cranshaw & Kittur, 2011) and human health (Ranard et al., 2014). These initiatives, in relation with the concept of citizen science, have gained an increasing interest in the scientific community not only for the potential outcomes that crowdsourcing-based projects may bring to the researcher's field of interest but also for studying crowdsourcing as a scientific object per se (Wiggins & Crowston, 2011; Franzoni & Sauermann, 2014). Although the use of volunteering contributions in the scientific research area originates well before the internet era (Koerten & van den Besselaar, 2014), current crowdsourcing initiatives are always mediated by internet platforms. Other ICT tools such as mobile phones considerably foster the development of citizen sensing initiatives. The quality of the inputs collected through crowdsourcing is a major point of discussions in several projects (e.g., Muller et al., 2015), as well as the data quality procedures that are needed to improve the quality of the inputs (Allahbakhsh et al., 2013). Some authors claim that "higher quality information can be derived from vast amounts of low quality data" (De Longueville, 2016), which is related to the so-called "big data" paradigm. Several studies further investigated the profiles of the contributors to crowdsourcing initiatives (e.g., Newman et al., 2012; Neis & Zielstra, 2014; Ranard et al., 2014) and their motivations (e.g., Raddick et al., 2010; Reed et al., 2013; Koerten & van den Besselaar, 2014; Nov et al., 2014).

Large-scale, successful projects such as the ones developed in environmental monitoring are still lacking in the agriculture sector. It is sometimes argued that farmers may be reluctant to use new ICT tools such as crowdsourcing applications. However, specific applications are increasingly adopted when the tools are relevant and meet their current practices, e.g., weather forecasts on a mobile application. More complex ICT tools such as precision agriculture applications are also increasingly used (GNSS, 2015), in both industrialized and developing countries (USAID, 2013). This trend is supported by the facts that mobile phones are used worldwide and mobile connectivity is increasing to reach complete spatial coverage in many rural areas.

Although not always denominated as crowdsourcing, there is a long tradition of setting participatory approaches in research and development projects in agriculture, attempting to facilitate the farmers-researchers interactions or to simply collect and aggregate agricultural information from farmers (van Etten, 2011). Dissemination of research and development knowledge in agriculture is often organized by national or regional agricultural agencies or structures, also known as extension services, or by farm consultants from private companies, which all aim to transfer scientific knowledge and new technologies to farmers. However, a gap remains between scientists and farmers. Scientists may not understand or even know the farmer needs. In addition, many project outputs fail to meet the farmers' fields or needs, even if research outputs are pertinent. More recently, there was a receding investment in agricultural extension services in some countries due to a decrease of public funding, and/or their missions had to be largely reformulated, which delayed research and technology dissemination and transfer. Some participatory approaches such as participatory learning were successfully applied to agricultural research and development projects and helped to bridge the gap between scientists and farmers (Pretty, 1995). Recently, Beza et al. (2017) identified crowdsourcing of farmers' data as an alternative way of getting field observations to conduct yield gap analysis, alongside with remote sensing and sensor networks. Crowdsourcing applications in agriculture cannot only provide inputs that meet the agricultural researchers' needs, but also help closing the knowledge dissemination loop between researchers and practitioners and foster farmer-to-farmer interactions. Therefore, there are huge opportunities for scientists and practitioners in developing crowdsourcing applications in agriculture.

We reviewed crowdsourcing projects in agricultural applications and classified them according to the type of inputs (data, information or knowledge) provided through crowdsourcing (section 2). We reported crowdsourcing applications mentioned in the scientific literature and websites, especially citizen science platforms, and also built on the cumulated experience of the authors of this manuscript in several years of participatory research projects in agriculture with close connections to extension services and farmers (e.g., Lebrun et al., 2014). Subsequently we identified 8 types of inputs that could be collected through crowdsourcing for agricultural applications (section 3) and reported data quality control methods (section 4). Finally, we identified the profiles of contributors, discussed farmers' participation and contributors' motivations, and explored the potential benefits for science and/or farming activities of these crowdsourcing initiatives in agriculture (section 5). We coin the term of farmsourcing as a crowdsourcing strategy involving professionals working in the field of agriculture.

2 Crowdsourcing applications in agriculture

We defined crowdsourcing as (1) the realisation of specific tasks or (2) the collection of data, information or knowledge, by a network of persons (the contributors) that are not doing so for their normal professional activities. Crowdsourcing contributors may receive monetary retribution for their work, or not (Schenk & Guittard, 2011). Following the first component of this definition, crowdsourcing resorts to the externalisation of repetitive tasks at no or low cost by volunteers or low-paid contributors, or in a broader perspective to more complex collaborative activities where the expertise of highly-skilled experts is needed. Following the second component, crowdsourcing is applied to the collection of data or information by volunteers, which is also known as citizen sensing (Boulos et al., 2011), participatory sensing, or community-based monitoring (Conrad & Hilchey, 2011). Crowdsourcing is often associated with the longer-standing concepts of citizen science (Wiggins & Crowston, 2011) or participatory science. Crowdsourcing and citizen science initiatives in the last decade were strongly supported by the development of the internet. In agriculture research, the use of crowdsourcing can particularly be filiated with the participatory approaches in research and development projects (van Etten, 2011). In the remainder of this manuscript, we propose the term of "farmsourcing" for crowdsourcing applications involving professional stakeholders in the agriculture sector working on a voluntary basis to exchange information. Unlike common crowdsourcing, information timeliness and information beyond observation, i.e., which cannot be derived from earth imagery or ground observation, are particularly important in the farmsourcing approach.

Note that throughout this paper, we distinguish between data, information, and knowledge, as proposed by Ackoff (1989). Crowdsourced data can be raw measurements of environmental variables, geographical features such as the coordinates of a measurement point or field boundaries, visual observations (notes or photographs) or any inputs that are provided without interpretation. Information is interpreted data that becomes useful, as for example, the processing of an image leading to the identification of a plant. Finally, knowledge is understood as organized information that is held by the contributors, based on their experience and empirical observations.

We reviewed recent applications of crowdsourcing use in agricultural research and development activities in a non-exhaustive list (Table 1) and propose the following categorization of crowdsourcing developments in the agricultural sector, with a few examples reported in the field of agricultural research or applications. This categorization applies to the reviewed projects but may be also used for future projects. Note that the initiatives reported here are mainly from the research domain, but not only. The beneficiaries of these initiatives are either scientists, farmers, other beneficiaries, or a mix of these categories, as discussed in the last section. An exhaustive list of crowdsourcing-based projects in the field of agriculture is out of the scope of this work and it worth noting that many initiatives are not reported in the scientific literature.

2.1 Crowdsourcing of tasks

According to the first component of our definition, crowdsourcing is based on the externalisation of simple or complex tasks. Numerous scientific projects using crowdsourcing are based on this approach, such as human interpretation of galaxy images (Raddick et al., 2010), mathematical problem solving (Cranshaw & Kittur, 2011) and classification of land cover based on easily accessible earth imagery (Fritz et al., 2012; See et al., 2015).

In the field of agriculture and environment, projects such as Pl@ntNet (Goëau et al., 2013) aim at identifying plant species by a combination of computer-automated image analysis based on machine learning and plant image collection through crowdsourcing. Using a mobile application, the contributor takes a picture of a plant and attempts to identify the plant using the automated image classification. However, if the contributor can identify the plant based on his own knowledge, the correctly classified image is used to feed the image database and so contributes to the efficiency of the machine learning algorithm. More directly applied to the agricultural sector is a project about plant disease image identification, PlantVillage Image (Hughes & Salathé 2015), which is dedicated to help farmers identify pests and diseases that affect their crops. In late 2015, more than 50,000 crop disease images of 16 crops were available. Most images originated from experimental research stations and were provided by skilled technicians who took high-quality photographs of the leaves affected by several diseases following a rigorous protocol. Similarly, Rahman et al. (2015) set up a methodology for weed identification based on a combination of computer-automated image analysis and two levels of crowdsourcing. First, a non-expert crowd from a web-based crowdsourcing platform attempts to perform weed identification, if not already performed by the computer. Secondly, a set of experienced professionals working in agricultural extension services validates the identification or contribute to the identification of non-classified images. Weed images are mostly taken by farmers who in turn benefit from the identification of the weed but also other related information about weed management and control.

2.2 Crowdsourcing of local visual observations

In situ data collection is a cumbersome task and data limitation is often encountered in environmental modelling, especially for highly complex environment systems or with a high spatio-temporal variability. One of the main issues of field officers that collect data or make observations is the time it takes to go from one place to another. Crowdsourcing of visual observations made by local people allows that a large amount of data can be more easily collected through many operators. A large number of examples of crowdsourcing of local visual observations in environmental sciences exists in the field of biodiversity monitoring, often denominated as community-based environmental monitoring (Conrad & Hilchey, 2011). Visual observations are recorded and communicated by short notes and/or by photographs (Roy et al., 2012) that are uploaded on a web platform. Dedicated mobile applications can support the monitoring of environmental observations and dozens of mobile phone applications are available (e.g., see http://brunalab.org/apps/ for a list). In particular, a generic tool is the Open Data Kit framework that allows to quickly deploy a customized mobile phone application for environmental data collection. This framework was used in several research applications (Chaudhri et al., 2012) and has the prominent advantage of enabling the field data collection without internet connection.

However, while several crowdsourcing of local visual observations initiatives took off in environmental sciences, there are only few examples of widely-used initiatives in the specific field of agriculture. In the field of crop breeding, van Etten (2011) proposed the design of a large scale crowdsourcing system for crop improvement and implemented it through pilot studies in Africa, India and Central America (van Etten et al., 2016). In this approach, denominated as participatory variety selection, farmers are asked to evaluate and report crop growth performance for different crop varieties, in farming conditions. For a long time, farmers are asked to provide observations to extension services and scientists, with or without monetary compensation, and at relatively small scale. In a sense this is a very early form of crowdsourcing.

2.3 Crowdsourcing of data from disseminated sensor measurements

A lot of environmental and agricultural variables can be measured in situ using permanent or portable measuring devices. Using a large network of distributed sensors has been successful in numerous crowdsourcing projects. In the field of hydrology, Lowry & Fienen (2013) compared stream gauge measurements from volunteers to official measurements. Based on the potential of every mobile phone as an environmental sensor, Overeem et al. (2013) demonstrated that temperature sensors of the mobile phone batteries could be used to retrieve outside air temperature in 8 major cities, where the concentration of mobile devices is particularly high.

GPS-driven agriculture machinery, low-cost environmental sensors and mobile devices equipped with basic sensors have considerably increased the amount of data that could be exploited for agricultural applications, which together are at the origin of the concept of big data in agriculture (Sonka, 2014; Wolfert et al., 2017). In the field of agricultural research, Francone et al. (2014) developed a mobile application, PocketLAI, that enables measuring the leaf area index of a crop cover using the accelerometer and camera of common mobile devices. The leaf area index, which is defined as the total one-sided area of leaf tissue per unit ground surface area, is a commonly-used biophysical variable in the crop and remote sensing scientific community for assessing the biomass of a crop cover and monitoring vegetation growth (Bréda et al., 2003). However, this variable per se is not familiar to farmers. While the leaf area index is correlated with biomass or grain yield, empirical relationships that translate leaf area index into other variables of interest are necessary to meet the farmers' needs and to further leverage the potential of the application to be used outside of the research area. Another project is PhotosynQ (Kramer, 2016), which aims at collecting a large amount of plant health data using a web-based platform together with disseminated plant sensors. The sensor measures plant fluorescence parameters as well as basic environmental variables (air temperature, relative humidity, barometric pressure and altitude). Contributors can submit on the web platform their own project based on this sensor, resulting in a large number of projects targeting different plant species in various environments. Another example is the participatory experiment reported by Marx et al. (2016), where GPS devices and cameras embedded in mobile phones were used for measuring crop height in a maize field. The experiment concluded that measuring maize height using a simple ruler was more robust than a method based on the automated processing of plant images taken by the phone camera.

Although rarely reported in the scientific literature, applications based on the crowdsourcing of data that are more dedicated to the farmers' needs exist. For instance, the YieldCheck application from the Potato Crop Management project in the UK (www.potatocropmanagement.com) combines a web platform and mobile application allowing potato growers to collect data about their crops (field location, crop variety, planting and emergence dates) with a mobile phone and gives back to growers yield forecasting information. In the same way, the Akkerweb platform developed by Wageningen University and Research in the Netherlands (www.akkerweb.nl) allows the centralisation of field information combined with satellite and soil data to provide an integrated cropping plan to farmers. These farm-sourced information can then be shared with consultants to optimize crop production at the field scale. However, both projects so far do not gather the collected data for further research or operational applications, meaning that data are just collected for a single use.

2.4 Crowdsourcing of knowledge

Although rarely viewed as crowdsourcing applications, user-generated content web platforms, such as knowledge portals, Questions & Answers (Q&A) forums and wikis have a great potential for gathering information and knowledge. A knowledge portal can be defined as a web platform that provides access to an ensemble of knowledge resources. Q&A forums are specific knowledge portals where the information is generated by the contributors of the forums through questions and answers, potentially moderated by the administrator of the forum or other contributors. Readers of the forum can build their own knowledge about a specific topic by gathering the user-generated information that is archived in the forum (passive use) or by contributing with questions and answers (active use). In some fields of expertise such as computer programming, Q&A forums such as www.stackoverflow.com have become a primary source of information that have rendered the existing official software documentation obsolete (Treude et al., 2011). The wiki platforms are based on a collaborative content-management system that allows to organise the user-generated content, ultimately generating online encyclopedias, such as Wikipedia.

In the field of agriculture, a recent platform is Croprotech (Bruce, 2016) that aims to provide information about weeds, pests and diseases to farmers in the UK while seeking to establish a two-way relationship between scientists (who designed the platform) and farmers. Several Q&A forums relate to agriculture in different languages and regional contexts (e.g., Hansen et al., 2014; Hughes & Salathé, 2015, Table 1). Q&A forums can help practitioners getting know-how, disseminating new practices and technologies, and validating and legitimating informal knowledge, as the backbone of a community of practice (Hansen et al., 2014). The topics addressed in these Q&A forums are diverse and the forums are often subdivided in specific sections, such as agricultural machinery with advice for choice, hints for maintenance and repair, crop and animal productions, pests and diseases (identification and interventions), trade and market, information about agricultural regulations, job offers in the agriculture sector, informal discussions, and sections about specific crops or agricultural activities (e.g., vineyard production). In line with the open-source knowledge movement, some web platforms propose to elaborate innovative prototypes of tools designed for a particular purpose, which are shared under an open source licence. In the farming sector, FarmHack (www.farmhack.org) proposes hundreds of prototypes of agricultural machines, often tailored for small-scale farming, as well as software and mobile applications. Contributors are able to submit their ideas and co-construct an open-source licensed prototype, potentially with the help of other contributors. A large number of blogs exist on agriculture, often maintained by farmers who want to share specific information to other farmers or to popularize their work to the general public. These knowledge portals and web-based platforms help to disseminate new agricultural practices worldwide, such as the use of new technologies, or alternative ways of conducting specific operations such as organic farming.

Name	**Website**	**Short description**	**Crowdsourcing component**	**Reference**
Pl@ntNet	www.plantnet-project.org	Plant identification by image analysis	Task	Goëau et al., 2013
PlantVillage Image	www.plantvillage.org	Plant disease identification by image analysis	Task	Hughes & Salathé, 2015
none	*none*	Weed	Task	Rahman et

		identification by image analysis		al., 2015
GeoWIKI	http://geo-wiki.org	Land-use mapping based on satellite imagery	Task	Fritz et al., 2009
DIYlandcover	http://mappingafrica.princeton.edu	Land-use mapping based on satellite imagery	Task	Estes et al., 2016
none	*none*	Reporting of on-farm trial of crop varieties	Local visual observations	van Etten et al., 2016
PocketLAI	www.cassandralab.com	Mobile application for enabling leaf area index measurements	Data from disseminated sensor measurements	Francone et al., 2014
PhotosynQ	www.photosynq.org	Web platform for crowdsourcing projects based on plant measurements	Data from disseminated sensor measurements	Kramer, 2016
Akkerweb	www.akkerweb.nl	Web platform for farm-sourcing information and private/public-supported applications	Data from disseminated sensor measurements	*none*
Potato Crop Management	www.potatocropmanagement.com	Web platform for data collection and yield forecasting	Data from disseminated sensor measurements	*none*
LandPKS	http://landpotential.org	Integrating scientific and local knowledge and improving farmer-to-farmer interactions	Knowledge	Herrick et al., 2013
FarmHack	http://farmhack.org	Web platform for sharing plans of prototypes for agricultural applications	Knowledge	*none*
Croprotech	https://croprotect.com	Web platform for sharing scientific information about weeds, pests and diseases.	Knowledge	Bruce, 2016
PlantVillage	www.plantvillage.org	Q&A forum on plant culture and phytopathology	Knowledge	*none*
AgTalk	http://talk.newagtalk.com	General Q&A forum on agriculture	Knowledge	Hansen et al., 2014

Table 1: Review of crowdsourcing projects related to agricultural applications

3 What information to collect?

We established a list of data, information and knowledge inputs that can be provided by crowdsourcing in agriculture (Table 2). These items are either inputs that are collected in existing projects, or inputs for which potential crowdsourcing projects could be designed in the future.

	Short description	Source	Target	Projects / References
Agricultural land-use /land cover data	Delineation of agricultural parcels and description of land use and land cover (crop sequencing)	From satellite, airborne or UAV imagery digitalisation and visual observations	For environmental and crop modelling, yield forecasting, …	OpenStreetMap (Minet et al., 2015); Geo-Wiki (Fritz et al., 2012); Collect Earth (Bey et al., 2016); DIYlandcover (Estes et al., 2016)
Soil data	Soil parameters useful for agricultural applications: texture, structure, organic matter content, pH, nutrient content (N, P, K), water content	From ground or near remote sensing, from soil surveys, from laboratories database	For farmers and recommendation systems	Rossiter et al., 2015
Weather data	Records of weather variables (temperature, precipitation, relative humidity)	From meteorological stations network	For farmers, warning and recommendation systems	Muller et al., 2015, Overeem et al., 2013
Crop phenology and calendar information	Records of phenological events and of field interventions	From UAV or close range remote sensing, from farmers	For crop modelling, yield forecasting, legal aspects	*none*
Pests and diseases	Observations of pests and diseases, photographs	From farmers, from technical expert	For pests and diseases monitoring, warning systems, time scheduling for farmers	PlantVillage Image (Hugues & Salathé, 2015); Rahman et al., 2015
Yield and vegetation status	Yield data per field, vegetation status measurements, fractional cover, biomass, leaf area index, …	From UAV or close range remote sensing, from farmers	For crop modelling, yield forecasting, crop monitoring	PocketLAI, Francone et al., 2014
Prices	Prices of agricultural products	From farmers and marketers	For farmers and marketers	Pommak
General agricultural knowledge	General knowledge and know-how about agriculture: can be information about crop calendar, farming practices, agricultural machinery issues, crop and animal productions, pests and diseases, stocks and market information, information about regulations, etc.	From farmers	For farmers	Agtalk; PlantVillage; Hansen et al., 2014

Table 2 : Type of inputs to collect in farmsourcing projects

3.1 Agricultural land-use data

Agricultural land-use data is supported by points or agricultural parcels limits, namely the geolocalised shapes of the parcels and their land-use / land cover (LULC) features. The LULC of agricultural parcels may be stable over the years in case of permanent grasslands or for specific crops (e.g., rice, sugar cane) or may change from one growing season to another where fields are cultivated under crop rotations. In the latter case, the information should be frequently updated. Crowdsourcing projects such as OpenStreetMap (www.openstreetmap.org), Geo-Wiki (Fritz et al., 2012), Collect Earth (Bey et al., 2016) or DIYlandcover (Estes et al., 2016) have proven that a tremendous amount of geographical land-use information can be efficiently collected through crowdsourcing, i.e., in so-called volunteered geographical information systems. The two projects Geo-Wiki and DIYlandcover produced maps of the cropping area and of the parcel sizes, at the global scale for Geo-Wiki (Fritz et al., 2015) and in South Africa for DIYlandcover. However, these crowdsourcing-based projects do not inform about the specific crop cover, namely the crop species that are usually or currently grown in the agricultural lands. In the case of OpenStreetMap, a number of tags (i.e., classes) were defined to map agricultural areas, i.e., farmland, meadow, orchard, vineyard and horticulture, mostly under the commonly-used "landuse" tag, and to a lesser extent, under the "landcover" tag. From these two tags, only basic distinctions between crop land and grassland could be made, and, up to now, precise delineations of individual parcels is mostly missing. However, specific crop types are mapped in OpenStreetMap under the tag "crop" that is preferably used for permanent cropping systems, with rice, grass and maize being the top used values. OpenStreetMap therefore has the potential of providing crop types map (Minet et al., 2015) but this is still not fully exploitable due to the lack of completeness and the poor update of the information.

3.2 Soil data

There are several kind of soil data that can be of interest for agricultural applications: textural classes, structure, organic matter content, pH, nutrient content (particularly mineral nitrogen submitted to quick change). These data may be of direct interest for farmers or for pedologists in order to improve current soil maps. Rossiter et al. (2015) reviewed existing applications of crowdsourcing projects targeting soil data and listed soil properties that could be collected in order to improve soil maps. Crowdsourcing platforms for soil data could benefit from regular soil analysis made by extension services and private laboratories, which may communicate their soil survey results to a centralized web platform.

3.3 Weather data

Accurate local weather information is very important for farmers since the farming calendar and several farm operations but also agricultural warning or recommendation systems heavily depend on the weather. A lot of mobile applications providing weather services exists, a majority of them using data from official and/or amateur weather stations (Muller et al., 2015). Crowdsourcing precipitation data is particularly appealing as rainfall, especially from convective storms, is largely uncorrelated in space. Not only farmers but also farm planners, operators of large machinery or processing factories largely depend on good weather information. While waterlogging and flood can be very localised, drought or heat conditions are more widespread in a region. In exchange for providing data on potential damages, farmers might benefit from specific forecasts that may help adaptation for instance on irrigation scheduling or crop disease control.

3.4 Phenology and crop calendar information

Phenology refers to the study of the seasonal life cycle events, which are dependent to the variations of climate from year to year. Crowdsourcing projects for the collection of environmental phenology observations are the most developed crowdsourcing projects globally (e.g., National Phenology Network, US, Betancourt et al., 2007). These projects report phenological observations of plant or animal life cycles, such as bud bursting, flowering or animal migration, with contributions mostly from citizen scientists interested in wildlife observations. Though specific professional applications to collect phenological data exist (McEwan & McCarthy, 2005), we did not find similar, large-scale projects for agricultural applications. In cultivated lands, phenology is not only driven by climate, but also by crop and field management. In that respect, the crop calendar, comprising field preparation dates (ploughing, fine-ploughing), sowing/planting dates, emergence, flowering, maturity and harvest dates are key data to collect. There is a long tradition of reporting crop phenology stages from agronomists and crop scientists (e.g., Demarée & Curnel, 2008), who may report not only the aforementioned events but also more precise plant phenological stages that are specific to each crop, such as the Zadoks (1974) or the BBCH (Bleiholder et al., 1989; Hack et al., 1992) scales. These data are of a primary importance for vegetation monitoring and crop modelling, at times supported by earth observation data. Despite the ease of phenological observations, specific protocols need to be developed to make reliable observations that can be compared across geographical regions, such as for projects on wildlife phenology data collection (Wiggins & Crowston, 2011). For instance, the emergence of a crop (i.e., the time when first green tissues emerge from the soil after sowing or planting) at the field level is usually defined when emergence can be observed within 90% of the field area.

3.5 Weeds, pests and diseases

A large amount of crowdsourcing projects are related to the observations of wildlife in general through visual observations and photographs (Franzoni & Sauermann, 2014). These projects usually apply to a global or local watch of a particular taxon or species, sometimes providing the contributor with automated or semi-automated recognition by image processing. However, projects that applied the same principles for farming applications are few, e.g., PlantVillage Image (Hugues & Salathé, 2015) or the weed identification set up by Rahman et al. (2015). While the primary goal of Croprotech (Bruce, 2016) is to provide information toward farmers, they can also contribute by reporting weed, pest, and disease observations. Since these projects were recently designed, it is too early to evaluate the success of these initiatives in terms of data collection and farmers' satisfaction. Pest and diseases infections on crops can be difficult to identify, and require expert contributions. Farmers may be particularly motivated to provide observations as they are highly concerned with the emergence of pests and diseases and their control. From the scientific side, there is a huge demand for regional or global observations of pests and diseases in cropping systems for crop disease modelling and forecasting, particularly within the context of climate change and agriculture extensification. For extension services in agriculture, such information is of prime importance to identify risks and supply in return recommendation to farmers for diseases and pests control.

3.6 Yield and vegetation status

Agricultural yield, namely, the quantity of crop or animal production per farming area or agricultural inputs, is a commonly collected variable that is important for many stakeholders in the field of agriculture: extension services, government officers, market analysts, researchers and the agri-food industry. In crop research and development, crop yield is usually the top variable of interest when comparing crop varieties, different crop management options (e.g., regarding the date of sowing, soil preparation, fertilization, etc.), as well as for crop model calibration. Crop production can be estimated using crop models or earth observation data. In that respect, vegetation status can be approximated by several variables, such as biomass (aerial or ground), fractional cover, or leaf area index (LAI). These variables are often used for monitoring the vegetation growth or detecting diseases during the growing season before being able to measure the yield by harvesting. The mobile application PocketLAI (Francone et al., 2014), which aims at measuring the LAI in the field, was not so far coupled with a crowdsourcing system for gathering crop development information. Actually, although crop and animal production (meat or dairy production) is often well-known by the farmers, this information is confidential and is rarely shared. Therefore, the sensitivity of this information precludes the development of crowdsourcing projects aimed at gathering yield data.

3.7 Prices

Prices of agricultural products is a key information for producers and marketers. Individual stakeholders of production chain may strengthen their commercial position through exchange of information on prices. An example is the Pommak initiative (www.pommak.be) that has been developed recently in Belgium by extension services of the potato sector to gather prices information on potato aiming to mutualise the access to the free market prices.

3.8 General agriculture knowledge

General agriculture knowledge concerns agricultural activities, such as agricultural machinery, crop and animal productions, pests and diseases, trade and market, information about regulations, etc. Several web-based knowledge portals gather general knowledge and know-how concerning agricultural activities (see section "Crowdsourcing of knowledge"). Most of this knowledge is provided by the farmers based on their own experience. Examples of knowledge that is shared are mechanical problems arising with agricultural machinery and how to fix it, advices for purchasing new machinery or agricultural inputs, tips and tricks about conducting specific agricultural operations, animal health issues, pests and diseases identification and how or when to intervene, etc. A major issue in knowledge portals, Q&A forums or blogs is that they are often mono-linguistic, and that minor language communities have access to a lower amount of information than widely-spread languages.

4 Data quality

The question of the data quality obtained through crowdsourcing or citizen science projects has received specific attention, especially since many crowdsourcing-based studies are published in the scientific literature. Data quality is reported to increase with the number of contributions (e.g., as cited in Muller et al., 2015), and it is even reported that high quality data can emerge from a vast amount of low quality data (De Longueville, 2015). It does not mean that more contributions automatically result in better quality outputs, but rather that more contributions allow for a stronger quality check of the crowdsourced data.

Data quality is particularly sensitive for agricultural applications that aim to give recommendations to practitioners. For instance, information on crop diseases observation or pests detection coming from a network of operators is often used to feed warning systems to help control the development of pests and diseases and to give pesticides application recommendations. Therefore, systems of validation of the collected information are strongly recommended to avoid mismatch in the delivered recommendations towards farmers.

4.1 Crowdsourcing data model approaches

The goal of many crowdsourcing initiatives is to collect inputs at a lesser cost or in a greater extent than traditional, professional data collection methods. It can be expected that crowdsourced contributions are of lesser quality than professionally-collected data. This issue can be partly circumvented by designing a strict crowdsourcing data model, e.g., based on pre-defined classes with a certain number of properties to fill in and by setting data quality control methods.

However, crowdsourcing approaches based on more flexible data models have shown a greater potential to 1) support unanticipated evolution of the crowdsourcing initiatives and 2) attract and maintain a larger group of contributors (Lukyanenko et al., 2014). For instance, the specific data model of OpenStreetMap allows the contributors to map any kind of features that exist on the ground and to describe them with properties left to the user choice. First, this allows unexpected geographical features to be digitalized in the database, while a minimum of standardisation is ensured by the community that edicts rules of mapping for the most common features. Second, this fosters the participation, as contributors with varying levels of expertise can apply. Lukyanenko et al. (2014) empirically demonstrated that letting the contributors to define the classes of the model results in an overall better accuracy and minimises information losses.

4.2 Recommendations for data quality control

Following Allahbakhsh et al. (2013), we distinguish between design and run-time approaches for improving data quality in crowdsourcing systems. Prior to its deployment, the crowdsourcing platform must be properly designed to ensure that high quality contributions will be recorded (design approach). In the run-time approach, a number of quality control checks may include the identification of high quality contributions and the subsequent evaluation of the quality of the inputs. Note that the approaches of validation and quality control traditionally applied in science also apply to crowdsourcing-based research projects, such as the ultimate peer-reviewing process (Franzoni & Sauermann, 2014). Hereafter, we identified several ways of improving and evaluating the quality of data obtained through crowdsourcing initiatives:

1. Replication

A common way of ensuring high quality data obtained by crowdsourcing is to replicate the task or the data collection by different contributors. This is easily applicable in projects based on the crowdsourcing of tasks, such as the human interpretation of weeds images in Rahman et al. (2015).

2. Community checking

In some crowdsourcing initiatives, the community of contributors can correct low-quality inputs or detect vandalism. Similarly to the collaborative encyclopedia Wikipedia, this is observed in the OpenStreetMap project, where some contributors declare that they are busier with checking the inputs of other contributors than with providing new inputs (Heipke, 2010). Of course, community checking mostly applies to objectively collected data such as geographic information or to specific tasks that can be replicated.

3. Contributors training, profiling and reputation

Some crowdsourcing initiatives impose a specific training of the contributors prior to their effective contributions. Contributors can also be evaluated during their work using specific quality check tasks, e.g., the DIYlandcover project (see below), which allows evaluating the accuracy of the contributions. Online forums increasingly rely on user reputation to benchmark the answers to questions or even to allow performing specific tasks, such as in the www.stackoverflow.com fora where users can vote or comment on a question only if they have a sufficient reputation score.

4. Discard outliers

Outliers can be discarded according to physically-plausible bounds or based on expert-knowledge. This is an obvious way of performing a first quality check of the data, although good/bad contributions may be discarded/kept depending on the threshold that is used.

5. Comparison to reference data

Crowdsourced inputs may be compared to reference data or other source of information, if available. This approach was applied to volunteered geographic information initiatives, where reference geographic information exists (e.g., Flanagin & Metzger, 2008; Heipke, 2010; Neis & Zielstra, 2014; Senaratne et al., 2016).

6. Checking coherence of spatial/temporal information

Since environmental variables such as weather records, soil properties or ecological data are often spatially and/or temporally correlated, they can be easily validated against a regression model or interpolation method. For instance, temperature records can be cross-validated against a simulated temperature field obtained by spatial interpolation (Muller et al., 2015).

In the DIYlandcover project, the quality of the classification of land cover is assured by a complex procedure of contributors' training and evaluation (Estes et al., 2016). The contributors must pass a qualification test in order to be able to participate. Once qualified, the quality of their work is continuously monitored and assessed by making them partly work on data sets that were already mapped before by experts, the so-called quality assessment sites. DIYlandcover contributors do not know if they are working on quality assessment or on unmapped sites. This quality assessment method allows to score the contributions from specific contributors and also to attribute bonus payments to the most qualified contributors, since DIYlandcover contributors are remunerated for their work. In an application of the mapping of land cover in South Africa (Estes et al., 2016), the overall classification quality was 91%, but varied largely depending on the contributors. In a similar image interpretation task based on a non-expert crowd, the overall accuracy of weeds classification was about 80% (Rahman et al., 2015), with a minimum of 10 answers to reach this accuracy.

5 Contributors and beneficiaries of agricultural crowdsourcing

5.1 Contributors profiles

In a similar study about the potential of citizen science for soil mapping, Rossiter et al. (2015) identified 9 groups of persons, sharing common work and/or interests, which may contribute to soil data and information collection. In the particular case of farmsourcing, a large part of the inputs should ideally be provided by contributors who belong to the farming sector. Farmers are the first contributors in the crowdsourcing of knowledge through Q&A forums dedicated to agriculture. However, we could not assess the participation rate of farmers in the other farmsourcing initiatives. Extension agents and agricultural scientists were reported to be strongly involved in some initiatives, such as in the collection of weeds (Rahman et al., 2015) and plant diseases (Hugues & Salathé, 2015) images. Lastly, citizen scientists, such as outdoor enthusiasts (i.e., people spending leisure time in hiking, bird watching, etc.), actively contribute to crowdsourcing projects in community-based environmental monitoring activities. Contributors from other crowdsourcing projects may also join farmsourcing projects, especially if those are supported by large citizen science platforms such as www.zooniverse.org and www.scistarter.com (Franzoni & Sauermann, 2014). Similar to platforms of open-source software development, new projects that are hosted in Zooniverse benefit from the expertise of the network of crowdsourcing projects for designing the project and the ways to foster the participation and also potentially attract contributors from other projects.

Regarding the skills of the contributors, the assumption that volunteers in crowdsourcing projects are always non-experts can be faulty (Wiggins & Crowston, 2011). Even if made on a voluntary basis, numerous crowdsourcing contributions are actually performed by professionals working in the field of the crowdsourcing application. For instance, a study about the collaborative mapping project OpenStreetMap showed that 50% of the respondents of a survey about the profile of contributors had degrees or worked in the fields of computer sciences or geography-related disciplines (Neis & Zielstra, 2014). In the application proposed by Rahman et al. (2015), weeds identification is performed by a combination of contributions made by non-expert, inexpensive crowdsourcing and expert, professional extension agents. In the PlantVillage Q&A forum on plant diseases (Hughes & Salathé, 2015), the top contributors are plant pathologists working in scientific research.

5.2 Farmers' participation and motivations

Farmers' involvements in participatory research can take different forms: providing useful information to scientists, collecting data themselves, or helping design the research questions. However, farmers may be reluctant to participate in research projects because of potential mistrust between farmers and scientists. This mistrust partly originates from the fear that research results may be used to put in place burdening regulations against some farming activities. Therefore, a trustful relationship should be established between practitioners and researchers, which may be more difficult in large-scale projects such as those reported in this review.

The intrinsic motivation of contributing to science, or to tackle intellectual challenges are sufficient reasons for a lot of volunteered contributors to join crowdsourcing initiatives in science (Reed et al., 2013). Nevertheless, it is worth noting that crowdsourcing projects do not always mean that the contributors do not receive remuneration for their work (Schenk & Guittard, 2011). In particular, farmers may need a monetary compensation when the tasks are time-consuming or when potential loss of production is foreseen due to modifications in the agricultural operations. In any case, establishing a crowdsourcing platform requires infrastructure and does not result in a costless operation (Franzoni & Sauermann, 2014).

An obvious way of increasing farmers' participation would be to provide them with useful and cost-saving agricultural services. For instance, the success of PlantVillage Image (Hugues & Salathé, 2015) is based on the fact that farmers are looking for support about plant disease identification and management. Crowdsourcing projects could also be inspired by examples from agricultural cooperatives and companies that closely work with farmers. For instance, in yield forecasting using crop modelling, knowing crucial data such as the date of sowing and observations of phenological stages largely increases the reliability of the yield estimates (Curnel et al., 2011). As a result, farmers have to provide this information if they want to obtain reliable yield estimates and predictions. Similarly, near-real time satellite remote sensing data processing could be improved to deliver more accurate results for all farmers if some of them provide appropriate data for training in a timely manner.

It is also worth noting that farmers already collect a lot of information in a digital format for their own purpose in farm management software or for complying with agricultural regulations. For instance, beneficiaries of the Common Agricultural Policy in Europe must digitize their field boundaries in a Land Parcel Information System and report the crop cover in order to get subsidies. As a result, a large number of digital information already exists but would need to be gathered in order to be further exploited for research or supplemental services to farmers' purposes.

5.3 Beneficiairies

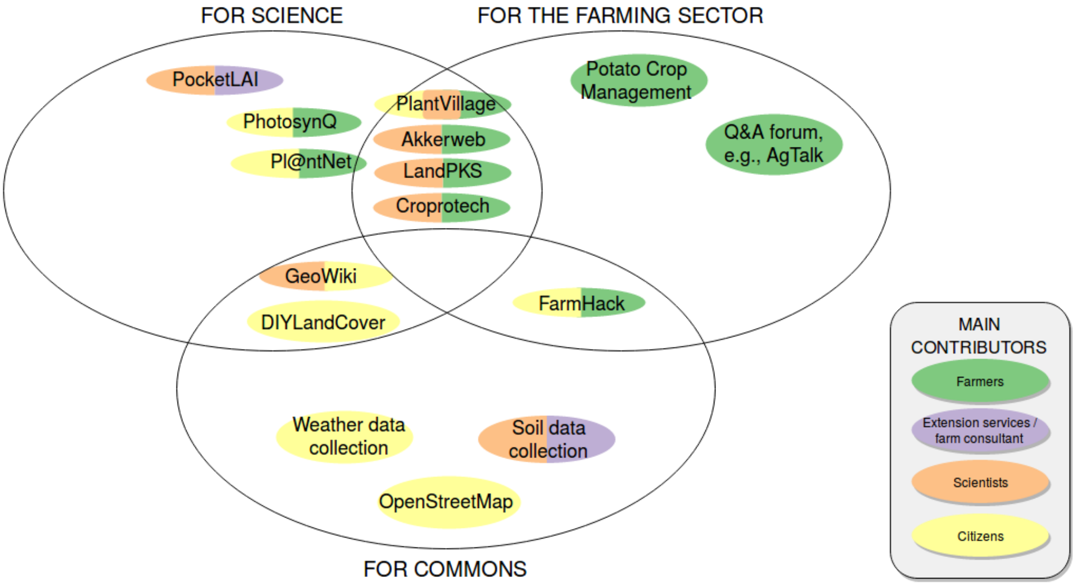


Figure 1: Classification of farm/crowdsourcing initiatives. The initiatives are grouped according to the main beneficiaries they are designed for, and highlighted according to the main contributors participating in the crowdsourcing of inputs.

The farmsourcing projects reviewed in this paper are highly diverse in terms of contributors' profiles and beneficiaries. We summarized this diversity in a diagram (Fig. 1) and presented a classification of crowdsourcing initiatives. This classification in terms of beneficiaries and contributors was made based on our review of the initiatives but, of course, it may change following the development of the initiatives, with unforeseen beneficiaries or contributors, or differ according to specific appropriations of the initiatives by beneficiaries or contributors. Initiatives devoted to the building of commons, such as weather, soil or geographic data infrastructure, have actually many different potential end-users (including farmers or scientists), but these are usually not necessarily known by contributors at the moment of their contribution. However, initiatives designed for the farming sector directly provide inputs to farming activities, such as an advice taken from a Q&A forum that helps a practitioner for taking a decision. Other initiatives for the farming sector such as PlantVillage, Croprotech and Potato Crop Management were specifically designed as practical advisory tools for farmers, with a crowdsourcing component. Initiatives that benefit to science aim at providing scientists some inputs for scientific advancements. For instance, it is foreseen that useful information about agricultural management, e.g., crop calendar, will be reported to crop scientists through the Akkerweb platform.

The initiatives that were reviewed in this manuscript were mainly designed by scientists. Some were specifically dedicated to farmers (e.g., PlantVillage, Potato Crop Management, Pommak) and attempted to reach them and to engage a win-win relationship between scientists and practitioners. In this kind of approach, scientists or consultants aimed for contributions by setting an open call to contributions and providing outputs as a payback for the contributors. Some crowdsourcing initiatives are based on contributions that were traditionally performed by experimented professionals, such as field technicians or scientists from extension services collecting field data. If successful, these farmsourcing initiatives could threaten specific professions, such as those dedicated to the collection of data in remote areas. However, data collected from field technicians or scientists from extension services can still remain useful as reference, authoritative data for validation that could be used to validate larger set of data supplied from practitioners.

An opposition exists between initiatives that aim to benefit to a “commons” and those that benefit the farming sector, since farming is mostly a private sector activity. The outstanding motivations of contributors to crowdsourcing of environmental observations that are often reported are the desire to contribute to environmental conservation achievements, as well as the pleasure to be outdoors while performing useful observations (Roy et al., 2012; Koerten & van den Besselaar, 2014). The fact that a farmsourcing approach could rarely be dedicated to the improvement of a “commons” is a serious drawback for the implication of a potentially large group of contributors (Reed et al., 2013). This might be the cause of the poorer development of farmsourcing initiatives, together with the privacy issues discussed hereafter. Nevertheless, there might be initiatives that reconcile these two goals, such as projects that aim to build an “agricultural commons” (e.g., FarmHack) or to contribute to agricultural science. Some contributors might be interested in the development of national or global agricultural data infrastructure, by providing useful inputs for farmers.

5.4 Economic benefits for farmers

We did not find comprehensive studies on the economic benefits of current farmsourcing initiatives. The rise of so-called big data applications in agriculture, with the collection of data from high-tech machinery, crowdsourced environmental and remote sensing data shows a huge potential in the coming future. The collection of environmental and management information is the basis of the smart farming paradigm (Wolfert et al., 2017) and supports the development of precision agriculture, which aims at increasing the profitability of farming operations by optimising the use of agricultural production factors (soil, inputs, machinery, people). However, these new applications may mostly benefit to large-scale farming and/or to specific crops. For instance, satellite remote sensing data can be better exploited in areas with large, homogeneous and flat agricultural parcels, and may not benefit small-scale parcels and mixed crop cover. However, small hand-held or embedded sensors can still find their place in developing countries where lower field size still allows non costly equipment to reach the goal of precision agriculture, based on sharing information through farmsourcing linked to local extension service expertise.

Exchanges of agricultural practices and knowledge are particularly appealing for new farmers or early-adopters of cutting-edge technologies. Farmers usually make decisions based on their local knowledge supported by external information. Even though they must be confronted with the local farming conditions, crowdsourcing-based initiatives such as a forum giving farming advice can support the decision-making process (Bruce, 2016).

Crowdsourcing may change the selling power in particular of smallholder or specialised farms. For example, horticultural farmers may benefit from selling their produce to local customers and markets which may be facilitated by a web-service. When fulfilling a market place role crowdsourcing works well: middlemen or retailers may be avoided and consumers may have a role in the way their food is produced. Prices may be set through consumers' demands, and consumers may participate in setting the quality of the produce they want, e.g. "la marque du consommateur" (https://lamarqueduconsommateur.com/).

5.5 Barriers to participation

5.5.1 Technical barriers

It is widely recognized that successful crowdsourcing projects are designed so that minimal barriers to participation occur (Franzoni & Sauermann, 2014). The participation in terms of subscription to the platform, understanding the protocol, and the way of making the contributions must be fast and easy, while ensuring a high quality of the contributions. For instance, while user subscription and login on a web platform is convenient for tracking user contributions, this is often perceived as a barrier to entry for new contributors. Some projects added "gamification" features to increase the entertainment of the contributors and also to enable a stimulating competition between contributors. An outstanding example of a citizen science project that successfully relies on gamification is FoldIt (Franzoni & Sauermann, 2014), where contributors are engaged in a friendly competition to model the structure of proteins. Contributors' achievements and performances are listed as in an online computer game, with user ranking and the possibility of creating teams of contributors. Nevertheless, technological barriers, such as the absence of internet connectivity, the lack of access to computer devices or, more simply, the aversion or the disinterest to use ICT tools are clearly large constraints to the participation in farmsourcing projects.

5.5.2 Privacy issues

Compared to other environmental crowdsourcing projects, privacy issues may limit the expansion of crowdsourcing initiatives to the agriculture domain, which may partly explain the bigger success of environmental monitoring projects with regards to farmsourcing projects. Indeed, there are no or few privacy restrictions in collecting environmental data such as wildlife observations, even on private lands. However, when applied to farmsourcing, the contributor might be discouraged to share data or information from his own private agricultural fields. Similarly, farmers may be troubled if they know that collected information about their field or agricultural practices (such as crop type, application of fertilizers, …) are available on an open platform. In addition, even if information is returned anonymously, identification of contributors and/or farmers in geolocalised datasets from sparsely populated rural areas is easier than in urban environments (Li & Goodchild, 2013).

As discussed by Franzoni & Sauermann (2014), one may distinguish between the openness of the participation and that of intermediate inputs. Open participation is a prerequisite for most crowdsourcing systems and it potentially maximises the number of contributors participating in the project. The openness of the intermediate inputs (such as the crowdsourced raw data) often benefits the project, as for instance it allows experienced contributors to correct potential mistakes or errors made by other contributors. Increasing the degree of openness of intermediate inputs also encourages the contributors to participate, because they can immediately see their contributions. However, in some situations, the openness of intermediate inputs may be in conflict with some parties, such as the farmers that may not want personal information about their agricultural production or practices being disclosed. Therefore, in general, the privacy of sensitive data will need to be ensured, otherwise farmers will be reluctant to contribute.

As a professional activity, farming generates a lot of sensitive information, such as the crop or animal production that is often kept secret by both the producer and the company that bought the product. Since some companies or public-supported extension services are active in farming advices and decision support, they may enter in conflict with new crowdsourcing projects aimed at providing free information about farming activities. For instance, a project relying on the community-based monitoring of crop disease may break the market of current crop disease early-warning systems based on the monitoring of some fields by experienced agents. At the same time, if companies or extension services could join the project, they could bring in considerable expertise as well as their network of observations. As a result, contractual agreements should be conceived between all parties to avoid potential conflicts (Franzoni & Sauermann, 2014). These agreements may include the degree of openness of the intermediate inputs and ensure that the privacy of the farmers will be respected. Data collected from companies or extension services can also find a place in the mutualisation of their data as reference data aiming to validate large set of lower accurate data as mentioned earlier.

5.6 Final recommendations for participation

Beyond all these pioneer initiatives currently exploring the field potential, a farmsourcing initiative would probably require the following elements of a decreasing importance to be attractive to many:

- Open participation;
- Convenience of a fully intuitive interface already proposing all existing data about a given location or field;
- Data quality policy;
- Salience of the service or information provided to farmers, which also include the service timeliness and its reliability;
- Trust in data management and security insuring full confidentiality wherever needed;

- Legitimacy of the entity supporting the platform;
- Credibility of the provided information to the farmers or agricultural professionals.

The relative importance of these criteria will surely evolve when it comes to a regular sustainable service. The sustainability is also very dependent on both the information return from the farmer's perspective and the sense of belonging to a large community doing better together than alone. For that reason sustainable farmsourcing initiatives should rather not consider to rely only a one-to-one relationship even with two-way information exchanges (one data provider – one data user better and the return) but should rather build on one-to-many relationship in addition to the one-to-one relationship. Such a more collaborative approach balances the asymmetric situation of the farmer and matches the increasing interest of the young farmer generation for ICT.

6 Conclusions and perspectives

We analysed crowdsourcing initiatives and discussed major issues related to the use of crowdsourced contributions in the specific field of agriculture. We defined crowdsourcing as (1) the realisation of specific tasks or (2) the collection of data, information or knowledge, by a network of contributors that are not doing so for their normal professional activities. We did not restrict our search to the term of "crowdsourcing", as many other concepts such as citizen science, community-based monitoring, citizen sensing, etc., are used to describe initiatives that are related to our definition. We coined the term of "farmsourcing" as a professional crowdsourcing strategy in farming activities, in relation to the professional nature of the farming sector. Current crowdsourcing initiatives were reviewed and sorted in four categories: crowdsourcing of tasks, of local observations, of data disseminated by sensors and of knowledge. We subsequently identified eight types of inputs that are or could be generated by farmsourcing and further discussed the ways of ensuring and controlling the quality of crowdsourced inputs. We reviewed the profiles of contributors that are active in farmsourcing initiatives and stressed that farmsourcing contributors are by definition professionals in the farming sector and/or agricultural experts. We addressed the degree of participation and the different motivations of contributors. The initiatives reviewed in this paper were highly diverse and could be highlighted according to the beneficiaries who can benefit from the crowdsourced contributions: they may benefit scientists or the farming sector, improve a "commons" or they may be a combination of these three categories. This synthesis could be used as a source of recommendations for elaborating collaborative projects using crowdsourcing in agricultural research and development.

We now summarize the key discussions about the development of farmsourcing initiatives. Some crowdsourcing and citizen science initiatives outside of the agricultural research domain have proven to be successful in terms of scientific achievements. In particular, community-based monitoring of the environment, e.g., crowdsourcing of wildlife observations, was reported to be highly developed with several projects existing for a long time. We found less successful initiatives in the field of agriculture. Our analysis pointed to privacy issues in relation to participation, since farming is a professional activity in the private sector. This prevents the collection of inputs by contributors on private lands and decreases the attractiveness for potential users interested in the contribution to a commons. Nevertheless, some crowdsourcing projects not directly related to agriculture could benefit to specific agricultural applications, such as weather forecasting systems based on crowdsourced meteorological observations. The development of specific regulations for the use of private data will also help to decrease the fear of contributors to share data on web–based platforms that will be made open-access for the development of collaborative initiatives in agriculture. A first perspective is a technological farmsourcing approach, related to the big data paradigm, where applications are built upon crowdsourced database (e.g., soil, weather database) and the gathering of inputs from farm management software. A second perspective is to develop a more collaborative farmsourcing approach, where farmers share information and knowledge to improve their farming practices, in close connection with extension services and farm advisory private companies. There is a long tradition of participatory approaches in agricultural science, aimed at bridging the gap between scientists and farmers. Benefiting from the increase of ICT in agriculture, innovative farmsourcing approaches could be designed to increase the participation of farmers in scientific research.

Acknowledgements

This work was funded in the framework of the BELgian collaborative Agriculture Monitoring at parcel level for sustainable cropping systems (BELCAM) project within the STEREOIII program of the Belgian Federal Science Policy. We thank two anonymous reviewers who help us improving our manuscript.

References

Ackoff, R. L. (1989), 'From data to wisdom', *Journal of applied systems analysis* **16**(1), 3--9.

Allahbakhsh, M.; Benatallah, B.; Ignjatovic, A.; Motahari-Nezhad, H.; Bertino, E. & Dustdar, S. (2013), 'Quality control in crowdsourcing systems', *IEEE Internet Computing* **17**(2), 76--81.

Belden, O. S.; Baker, S. C. & Baker, B. M. (2015), 'Citizens unite for computational immunology!', *Trends in immunology* **36**(7), 385--387.

Betancourt, J. L.; Schwartz, M. D.; Breshears, D. D.; Brewer, C. A.; Frazer, G.; Gross, J. E.; Mazer, S. J.; Reed, B. C. & Wilson, B. E. (2007), 'Evolving plans for the USA National Phenology Network', *Eos, Transactions American Geophysical Union* **88**(19), 211-211.

Bey, A.; Sánchez-Paus Díaz, A.; Maniatis, D.; Marchi, G.; Mollicone, D.; Ricci, S.; Bastin, J.-F.; Moore, R.; Federici, S.; Rezende, M.; Patriarca, C.; Turia, R.; Gamoga, G.; Abe, H.; Kaidong, E. & Miceli, G.

(2016), 'Collect Earth: Land Use and Land Cover Assessment through Augmented Visual Interpretation', *Remote Sensing* **8**(10), 807.

Beza, E.; Silva, J.V.; Kooistra, L. & Reidsma, P. (2017), ‘Review of yield gap explaining factors and opportunities for alternative data collection approaches’, *European Journal of Agronomy* **82B**, 206-222.

Bleiholder, H.; van den Boom, T.; Langeluddeke, P. & Stauss, R. (1989), 'Einheitliche Codierung der phanologischen Stadien bei Kultur- und Schadpfanzen', *Gesunde Pflanzen* **41**, 381-384.

Boulos, M. N. K.; Resch, B.; Crowley, D. N.; Breslin, J. G.; Sohn, G.; Burtner, R.; Pike, W. A.; Jezierski, E. & Chuang, K.-Y. S. (2011), 'Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples', *International journal of health geographics* **10**(1), 1.

Bréda, N. J. (2003), 'Ground-based measurements of leaf area index: a review of methods, instruments and current controversies', *Journal of experimental botany* **54**(392), 2403--2417.

Brabham, D. (2013), *Crowdsourcing*, Mit Press.

Brabham, D. (2012), 'Crowdsourcing: A Model for Leveraging Online Communities,[w:] Delwiche A', *Henderson J.(red.), The Participatory Cultures Handbook, New York*, 120-129.

Brabham, D. (2008), 'Crowdsourcing as a model for problem solving an introduction and cases', *Convergence: the international journal of research into new media technologies* **14**(1), 75--90.

Bruce, T. J. (2016), 'The CROPROTECT project and wider opportunities to improve farm productivity through web-based knowledge exchange', *Food and Energy Security* **5**(2), 89--96.

Chaudhri, R.; Brunette, W.; Goel, M.; Sodt, R.; VanOrden, J.; Falcone, M. & Borriello, G. (2012), Open data kit sensors: mobile data collection with wired and wireless sensors, *in* 'Proceedings of the 2nd ACM Symposium on Computing for Development', pp. 9.

Conrad, C. C. & Hilchey, K. G. (2011), 'A review of citizen science and community-based environmental monitoring: issues and opportunities', *Environmental monitoring and assessment* **176**(1-4), 273--291.

Cranshaw, J. & Kittur, A. (2011), The polymath project: lessons from a successful online collaboration in mathematics, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', pp. 1865--1874.

Curnel, Y.; de Wit, A. J.; Duveiller, G. & Defourny, P. (2011), 'Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS Experiment', *Agricultural and Forest Meteorology* **151**(12), 1843--1855.

De Longueville, B. (2016), 'Citizens as Earth Observation Sources: a Workflow for Volunteered Information Sensing', PhD thesis, Université catholique de Louvain.

Demarée, G. & Curnel, Y. (2008), Plant phenology in Belgium, *in* 'Cost action 725: The history and current status of plant phenology in Europe.', pp. 29-33.

Estes, L.; McRitchie, D.; Choi, J.; Debats, S. R.; Evans, T.; Guthe, W.; Luo, D.; Gagazzo, G.; Zempleni, R. & Caylor, K. (2016), 'Diylandcover: Crowdsourcing the creation of systematic, accurate landcover maps', *Environmental Modelling and Software* **40**, 41-53.

van Etten, J. (2011), 'Crowdsourcing Crop Improvement in Sub-Saharan Africa: A Proposal for a Scalable and Inclusive Approach to Food Security', *IDS bulletin* **42**(4), 102--110.

van Etten, J.; Beza, E.; Calderer, L.; van Duijvendijk, K.; Fadda, C.; Fantahun, B.; Kidane, Y. G.; van de Gevel, J.; Gupta, A.; Mengistu, D. K.; Kiambi, D.; Mathur, P. N.; Mercado, L.; Mittra, S.; Mollel, M. J.; Rosas, J. C.; Steinke, J.; Suchini, J. G. & Zimmerer, K. S. (2016), 'First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot)', Experimental Agriculture, 1--22.

Flanagin, A. J. & Metzger, M. J. (2008), 'The credibility of volunteered geographic information', *GeoJournal* **72**(3-4), 137--148.

Francone, C.; Pagani, V.; Foi, M.; Cappelli, G. & Confalonieri, R. (2014), 'Comparison of leaf area index estimates by ceptometer and PocketLAI smart app in canopies with different structures', *Field Crops Research* **155**, 38--41.

Franzoni, C. & Sauermann, H. (2014), 'Crowd science: The organization of scientific research in open collaborative projects', *Research Policy* **43**(1), 1--20.

Fritz, S.; McCallum, I.; Schill, C.; Perger, C.; Grillmayer, R.; Achard, F.; Kraxner, F. & Obersteiner, M. (2009), 'Geo-Wiki. Org: The use of crowdsourcing to improve global land cover', *Remote Sensing* **1**(3), 345--354.

Fritz, S.; McCallum, I.; Schill, C.; Perger, C.; See, L.; Schepaschenko, D.; Van der Velde, M.; Kraxner, F. & Obersteiner, M. (2012), 'Geo-Wiki: An online platform for improving global land cover', *Environmental Modelling & Software* **31**, 110--123.

Fritz, S.; See, L.; McCallum, I.; You, L.; Bun, A.; Moltchanova, E.; Duerauer, M.; Albrecht, F.; Schill, C.; Perger, C.; Havlik, P.; Mosnier, A.; Thornton, P.; Wood-Sichra, U.; Herrero, M.; Beckerreshef, I.;

Justice, C.; Hansen, M.; Gong, P.; Abdel, Aziz, S.; Cipriani, A.; Cumani, R.; Cecchi, G.; Conchedda, G.; Ferreira, S.; Gomez, A.; Haffani, M.; Kayitakire, F.; Malanding, J.; Mueller, R.; Newby, T.; Nonguierma, A.; Olusegun, A.; Ortner, S.; Ram, Rajak, D.; Rocha, J.; Schepaschenko, D.; Schepaschenko, S.; Terekhov, A.; Tiangwa, A.; Vancutsem, C.; Vintrou, E.; Wenbin, W.; Van Der Velde, M.; Dunwoody, A.; Kraxner, F. & Obersteiner, M. (2015), 'Mapping global cropland and field size', *Global change biology* **21**(5), 1980–1992.

-, 'GNSS Market Report, Issue 4, Publications Office of the European Union, ISBN 978-92-9206-013-8', 84 p.

Goëau, H.; Bonnet, P.; Joly, A.; Bakić, V.; Barbe, J.; Yahiaoui, I.; Selmi, S.; Carré, J.; Barthélémy, D.; Boujemaa, N.; Molino, J.-F.; Duché, G. & Perronet, A. (2013), Pl@ntnet mobile app, *in* 'Proceedings of the 21st ACM international conference on Multimedia', pp. 423--424.

Hack, H.; Bleiholder, H.; Buhr, L.; Meier, U.; Schnock-Fricke, U.; Weber, E. & Witzenberger, A. (1992), 'Einheitliche Codierung der phänologischen Entwicklungsstadien mono-und dikotyler pflanzen–Erweiterte BBCH-Skala, Allgemein', *Nachrichtenbl. Deut. Pflanzenschutzd* **44**(12), 265--270.

Hansen, J. P.; Melby Jespersen, L.; Leck Jensen, A.; Holst, K.; Mathiesen, C.; Brunori, G.; Halberg, N. & Ankjær Rasmussen, I. (2014), 'ICT and social media as drivers of multi-actor innovation in agriculture', *CIGR Proceedings* **1**(1).

Heipke, C. (2010), 'Crowdsourcing geospatial data', *ISPRS Journal of Photogrammetry and Remote Sensing* **65**(6), 550--557.

Herrick, J. E.; Urama, K. C.; Karl, J. W.; Boos, J.; Johnson, M.-V. V.; Shepherd, K. D.; Hempel, J.; Bestelmeyer, B. T.; Davies, J.; Guerra, J. L.; Kosnik, C.; Kimiti, D. W.; Ekai, A. L.; Muller, K.; Norfleet, L.; Ozor, N.; Reinsch, T.; Sarukhan, J. & West, L. T. (2013), 'The global Land-Potential Knowledge System (LandPKS): Supporting evidence-based, site-specific land use and management through cloud computing, mobile applications, and crowdsourcing', *Journal of soil and water conservation* **68**(1), 5A--12A.

Howe, J. (2006), 'The rise of crowdsourcing', *Wired magazine* **14**(6), 1--4.

Hughes, D. & Salathé, M. (2015), 'An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing', *arXiv*, 1511.08060.

Koerten, H. & van den Besselaar, P. (2014), Citizen science and Crowd science in biodiversity research, *in* 'Proceedings of the Internet, Policy & Politics conference'.

Kramer, D. (2016), PhotosynQ: Community-driven plant phenotyping for understanding plant responses to climate change, *in* 'Plant and Animal Genome XXIV Conference'.

Lebrun, P.; Goffart, J.-P. & Glorvigen, B. (2014), 'Report on Workshop “Connecting Research to Practice in the Potato Sector—Top-Down and Bottom-Up Approaches”', *Potato Research* **57**(3-4), 359--364.

Li, L. & Goodchild, M. F. (2013), Is privacy still an issue in the era of big data?—Location disclosure in spatial footprints, *in* '2013 21st International Conference on Geoinformatics', pp. 1--4.

Lowry, C. S. & Fienen, M. N. (2013), 'CrowdHydrology: crowdsourcing hydrologic data and engaging citizen scientists', *GroundWater* **51**(1), 151--156.

Lukyanenko, R.; Parsons, J. & Wiersma, Y. F. (2014) ‘The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-Generated Content’, *Information Systems Research* **25**(4), 669--689.

Marx, S.; Hämmerle, M.; Klonner, C. & Höfle, B. (2016), '3D Participatory Sensing with Low-Cost Mobile Devices for Crop Height Assessment–A Comparison with Terrestrial Laser Scanning Data', *PloS one* **11**(4), e0152839.

McEwan, R. W., & McCarthy, B. C. (2005). Phenology: An integrative environmental science.

Minet, J.; Robert, B. & Tychon, B. (2015), 'The potential of OpenStreetMap for land use/land cover mapping', FOSS4G.be, Brussels, Belgium.

Muller, C.; Chapman, L.; Johnston, S.; Kidd, C.; Illingworth, S.; Foody, G.; Overeem, A. & Leigh, R. (2015), 'Crowdsourcing for climate and atmospheric sciences: current status and future potential', *International Journal of Climatology* **35**(11), 3185--3203.

Neis, P. & Zielstra, D. (2014), 'Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap', *Future Internet* **6**(1), 76--106.

Newman, G.; Wiggins, A.; Crall, A.; Graham, E.; Newman, S. & Crowston, K. (2012), 'The future of citizen science: emerging technologies and shifting paradigms', *Frontiers in Ecology and the Environment* **10**(6), 298--304.

Nov, O.; Arazy, O. & Anderson, D. (2014), 'Scientists@ Home: what drives the quantity and quality of online citizen science participation?', *PloS one* **9**(4), e90375.

Overeem, A.; R Robinson, J.; Leijnse, H.; Steeneveld, G.-J.; P Horn, B. & Uijlenhoet, R. (2013), 'Crowdsourcing urban air temperatures from smartphone battery temperatures', *Geophysical Research Letters* **40**(15), 4081--4085.

Pretty, J. N. (1995), 'Participatory learning for sustainable agriculture', *World development* **23**(8), 1247--1263.

Raddick, M. J.; Bracey, G.; Gay, P. L.; Lintott, C. J.; Murray, P.; Schawinski, K.; Szalay, A. S. & Vandenberg, J. (2010), 'Galaxy zoo: Exploring the motivations of citizen science volunteers', *Astronomy Education Review* **9**(1), 010103.

Rahman, M.; Blackwell, B.; Banerjee, N. & Saraswat, D. (2015), 'Smartphone-based hierarchical crowdsourcing for weed identification', *Computers and Electronics in Agriculture* **113**, 14--23.

Ranard, B. L.; Ha, Y. P.; Meisel, Z. F.; Asch, D. A.; Hill, S. S.; Becker, L. B.; Seymour, A. K. & Merchant, R. M. (2014), 'Crowdsourcing—harnessing the masses to advance health and medicine, a systematic review', *Journal of general internal medicine* **29**(1), 187--203.

Reed, J.; Raddick, M. J.; Lardner, A. & Carney, K. (2013), An exploratory factor analysis of motivations for participating in Zooniverse, a collection of virtual citizen science projects, *in* 'System Sciences (HICSS), 2013 46th Hawaii International Conference on', pp. 610--619.

Rossiter, D. G.; Liu, J.; Carlisle, S. & Zhu, A.-X. (2015), 'Can citizen science assist digital soil mapping? ', *Geoderma* **259-260**, 71 - 80.

Roy, H.; Pocock, M.; Preston, C.; Roy, D.; Savage, J.; Tweddle, J. & Robinson, L. (2012), 'Understanding Citizen Science & Environmental Monitoring', Technical report, NERC Centre for Ecology & Hydrology and Natural History Museum, Final Report on behalf of UK-EOF.

Schenk, E. & Guittard, C. (2011), 'Towards a characterization of crowdsourcing practices', *Journal of Innovation Economics & Management* **7**(1), 93--107.

See, L.; Fritz, S.; Perger, C.; Schill, C.; McCallum, I.; Schepaschenko, D.; Duerauer, M.; Sturn, T.; Karner, M.; Kraxner, F. & Obersteiner, M. (2015), 'Harnessing the power of volunteers, the internet and Google Earth to collect and validate global spatial information using Geo-Wiki', *Technological Forecasting and Social Change* **98**, 324--335.

Senaratne, H.; Mobasheri, A.; Ali, A. L.; Capineri, C. & Haklay, M. (M. (2016), 'A review of volunteered geographic information quality assessment methods', *International Journal of Geographical Information Science* **31**(1), 139-167.

Sonka, S. (2014), 'Big data and the ag sector: more than lots of numbers', *International Food and Agribusiness Management Review* **17**(1), 1.

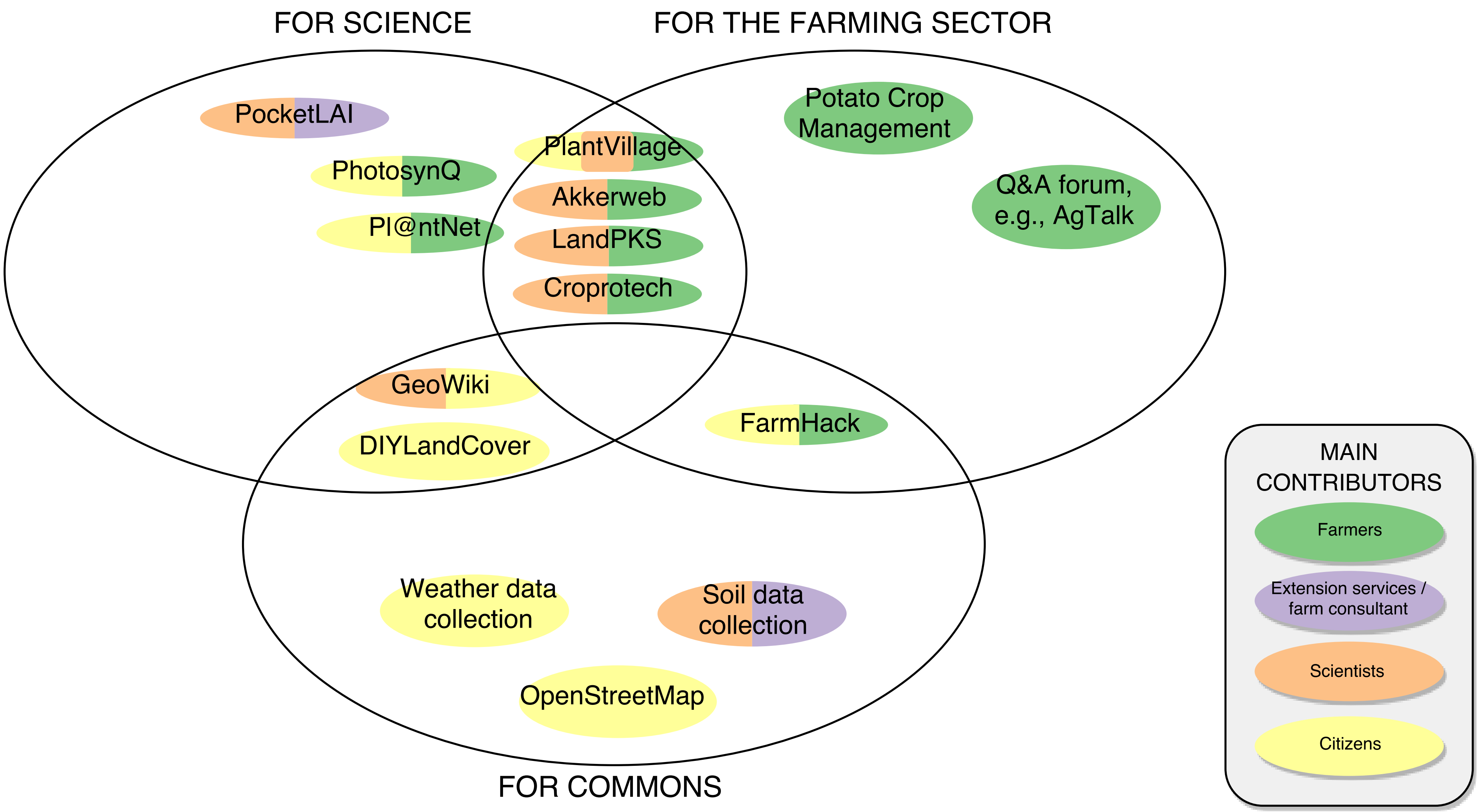
Treude, C.; Barzilay, O. & Storey, M.-A. (2011), How do programmers ask and answer questions on the web?: Nier track, *in* 'Software Engineering (ICSE), 2011 33rd International Conference on', pp. 804--807.

USAID (2013), 'Crowdsourcing applications for agricultural development in Africa', http://pdf.usaid.gov/pdf_docs/PA00J7P7.pdf, 1--6.

Wiggins, A. & Crowston, K. (2011), From conservation to crowdsourcing: A typology of citizen science, *in* 'System Sciences (HICSS), 2011 44th Hawaii international conference on', pp. 1--10.

Wolfert, S.; Ge, L.; Verdouw, C. & Bogaardt, M.J. (2017), ‘Big Data in Smart Farming – A review’, *Agricultural Systems*, **153**, 69 – 80.

Zadoks, J. C.; Chang, T. T. & Konzal, C. F. (1974), 'A decimal code for the growth stages of cereals', *Weed research* **14**(6), 415--421.


FOR SCIENCE
FOR THE FARMING SECTOR
PocketLAI
PhotosynQ
Pl@ntNet
PlantVillage
Akkerweb
LandPKS
Croprotech
Potato Crop Management
Q&A forum, e.g., AgTalk
GeoWiki
DIYLandCover
FarmHack
Weather data collection
Soil data collection
OpenStreetMap
FOR COMMONS
MAIN CONTRIBUTORS
Farmers
Extension services / farm consultant
Scientists
Citizens