MULTIPASS: managing the consents of access to farm data in a chain of trust to make new services emerge for farmers

B. Lauga[1], B. Balvay[2], L. Topart[3], J. Leclaire[4], A. Clenet[5], F. Brun[6], F. Pinet[7], C. Roussey[7] and M. Sine[1]

[1]ARVALIS, Institut du Végétal, Agricultural Technical Institute for Arable Crops, France
[2]IDELE, The French Livestock Institute, France
[3]ORANGE, Telecommunication Company, France
[4]FIEA, France Informatique Elevage et Agriculture, France
[5]SMAG, Farm Software Provider, France
[6]ACTA, Agricultural Technical Coordination Association, France
[7]Université Clermont Auvergne, IRSTEA, UR TSCF, Clermont-Ferrand, France

Abstract

With the emergence of digital technologies, farms become a relevant source of data to meet the challenges of multi-performance agriculture. Beyond the services provided, access to farmers' data depends on a clear understanding of their use, which must be done in a transparent way. Several codes of conduct at a national or international level push for a voluntary commitment to respect some good practices in the use of agricultural data. To provide a tool and answer farmer's questions on the control of their data and the transparency of the data processing, the partners of the MULTIPASS project, have imagined an interoperable ecosystem of farmer consents management, protecting farmers from no consented uses of their data.

Farmers' expectations of such an ecosystem have been expressed during workshops. They want to better identify existing data flows, including actors, data processes, and data clusters. Based on the farmers' expectations, the MULTIPASS project stakeholders have proposed the architecture of an ecosystem integrating two consent management tools as "pilots". This ecosystem should take in charge the interoperability between each consent management tools or with future tools.

This solution is based on a shared typology of data and data processes as well as on the specifications of the consent message content. All these elements should be easily accessible to meet the interoperability need of the ecosystem. It is also based on a router, which provides unified access to consent management tools (using API). In particular, it provides the farmer (beneficiary) with an exhaustive view of his/her consents (which can be distributed on several consent management systems), meeting farmers' expectations for transparency. It is also the point where a data provider can check whether the consent required to provide data exists, without needing to know which consent management system is concerned.

In this project, the stakeholders want to demonstrate to agricultural professional organizations the benefits and feasibility of a consent management ecosystem. By strengthening the confidence of farmers to share data, the project will allow the emergence of new knowledge and new services.

Keywords: farm data, data management, consent, transparency, chain of trust.

Introduction

Farmers are engaged in a progress for sustainable and productive agriculture. With the emergence of digital technologies, farms become a relevant source of data to meet the challenges of multi-performance agriculture. These data are the basis of the decision-making process. There is a data-driven agriculture based on the data transfer within the farm. These data also make it possible to create new knowledge or tools that improve the precision and relevance of agricultural operations in order to increase yields without negative impact on the environment.

Beyond the services provided, access to farmers' data depends on a clear understanding of their use, which must be done in a transparent way (Brun et al., 2016). This is a real concern for both farmers, who cannot control the uses of their data, and also data providers who have difficulties in determining the access and reuse permissions they can provide on the farmers' data they host. Access rights must be properly managed, as well as the farmer's consent for the uses of her/his data. The conditions related to this consent must be easily accessible and modifiable by the farmer. It is this chain of trust that the MULTIPASS project wants to implement.

Through the MULTIPASS project, the partners want to make available to farmers and data producers an interoperable farmers' consent management ecosystem, protecting data exchanges improving confidence to share their data with other organizations.

Challenges for a consent management ecosystem

Towards a widespread use of consents

Consents are the adherence of one party to the request made by another. In the case of personal data, consent is one of the 6 legal bases provided by the EU General Data Protection Regulation (GDPR, 2016) which authorizes the implementation of data processing. The law does not require the systematic collection of consents before processing personal data, because other legal bases can be invoked to process these data such as a mission of general interest or a contractual commitment (CNIL, 2018). Nevertheless, consents will enable the management of agricultural data exchanges not specified in the contracts.

To authorize an agricultural data processing and to reinforce the transparency of these uses, the farmer must be able to express her/his consent as shown in Figure 1.
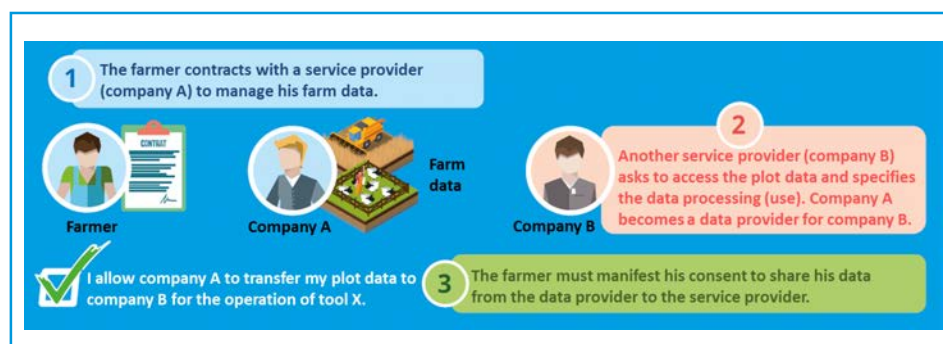

Figure 1. Example of a consent use for data exchanges.

Our goal was to build an ecosystem of stakeholders to manage consents and to create the engagement rules of these actors. We defined the typology of stakeholders presented in Table 1 involved in any farm data exchanges and consents management.

In this chain of trust, each actor has a responsibility and must satisfy good practices related to the use of agricultural data and consents.

Build a chain of trust

Table 1. Typology of actors in a consent management ecosystem.

Term	Definition
Right holder	The person who has the rights on the data. The consent of this person is needed to exchange data. In the MULTIPASS project, she/he is a farmer or breeder.
Delegatee	The right holder has delegated to a person or an organization (i.e., a delegatee) the right to give consents on her/his behalf.
Consent manager	The manager in charge of a consent management system.
Service provider	The organization that sells service to farmers and that needs an access to data. It is the beneficiary of the consent.
Data provider	The manager of the service (database) in charge of providing the data to the service provider.
Consent recorder	The organization that registers consents in the consent management system.

Jurists seem to think that, in the absence of a specific legal regulation, the control of agricultural data is ensured only by contracts with the farmer (Douville, 2019). The control will not come from the law but from a voluntary commitment made by the parties to respect some good practices in data uses. The French DataAgri code of conduct (FNSEA, 2018) leaded by the "Fédération Nationale des Syndicats d'Exploitants Agricoles" (FNSEA) and the "Jeunes Agriculteurs" (JA), and the European CODE OF CONDUCT (EU code of conduct, 2018) clearly goes in this direction.

Respecting good practices

In this context, farmers' expectations of such an ecosystem have been expressed in various workshops. Farmers regret that so far they had not been consulted much when the service providers processed their data. They expressed a need for transparency and want to better identify existing flows, including stakeholders, data uses and associated data categories. Based on farmers' expectations, the MULTIPASS project stakeholders have proposed one architecture of an ecosystem integrating two consent management tools as "pilots" and the conditions for their interoperability with each other or with future tools.

There are already existing consent management solutions dedicated to agriculture. These systems are often designed for particular needs. These different consent management systems can be freely chosen by the ecosystem stakeholders. Consents are stored in the consent management systems with the only constraint to register the information expected in the MULTIPASS ecosystem interfaces.

Implementation of the multipass ecosystem

Proposed architecture

The main tool defined in the MULTIPASS ecosystem is a router that guarantees the interoperability of the different consent management systems. It allows a unified access to consents to provide a list of them (by right holders, service providers, etc.) or to

verify the existence of consents before data exchanges. For this, it knows and can query the various consent management systems which will have interfaces (APIs) similar to those of the MULTIPASS router.

In particular, it provides the right holder with an exhaustive view of her/his consents (which can be distributed across several consent management systems), meeting farmers' transparency needs. The router also allows a data provider to check if the consent required for a data exchange exists, without needing to know in which consent management system it is managed. There is also a traceability of these controls. The use case diagram presented in Figure 2 shows the expected roles of each of the actors as well as the functional scope of the MULTIPASS router.
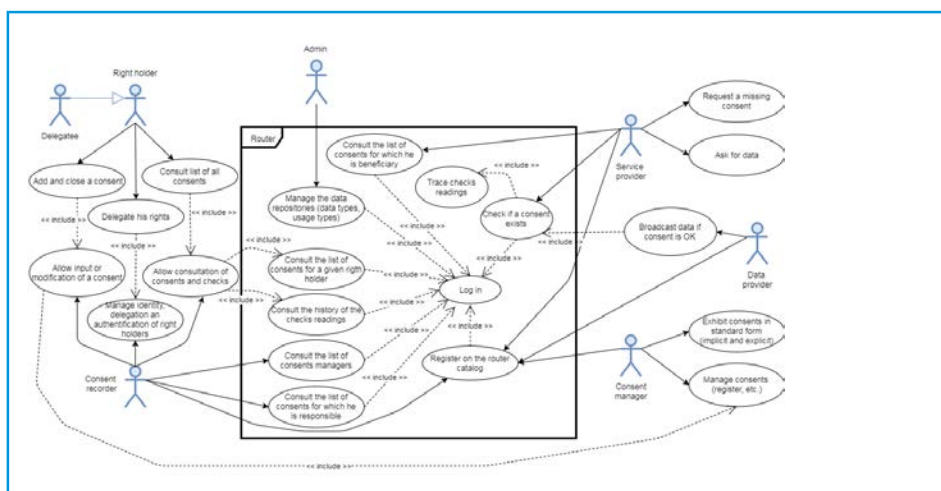

Figure 2. Use case diagram of MULTIPASS router.

All actors other than right holders must register on the router before they can use it. Their registration is validated by a router's administrator. A right holder (or her/his delegate) does not interact with the router. It is the role of the consent recorder to allow the input or the modification of consents. Only the right holder can see her/his consents once she/he is authenticated. The consent recorder cannot see them. This security is especially needed when the consent recorder is also a service provider (it must not see if the farmer works with its competitors). For this, either it will have made a contractual commitment in its contract with the farmer, or it will be committed by adhering to a charter or it will be obliged by GDPR in the case of personal data.

The management of the data repositories is the responsibility of a router's administrator.

Technical architecture

The router has a Java REST API that exposes business and administrative services. As consents are by nature sensitive data that must be secured, HTTPS is used for the exchanges. The OAuth protocol is used for authentication. A signature mechanism guarantees the API that the token issued during authentication process has been generated by the system. The passwords of the different users are stored in a SSHA

hashed form in an LDAP server. A Java human-machine interface allows system administrators to manage the different users and data repositories (data categories and uses) of the router.

A reverse proxy "HA Proxy" is used to secure the application upstream. This system will also be used for load balancing between different downstream application servers. The PostgreSQL database that registers actors, data repositories and logs could eventually be transferred into an elastic stack.

The router is an important part of the ecosystem interoperability. It is based on the main concept of consent. Consents are not managed in the database of the router, but only in the interfaces. The identification of the companies (farm, data or service provider) is done by the French SIRET identifier but the system allows the use of another identifier.

Table 2. Description of the concept of consent.

SIRET number of the farm exploitation (data producer)	WHO: actors of the data exchange
SIRET number of the service provider (beneficiary)	
SIRET number of the data provider	
SIRET number of the consent recorder	
Data categories	WHAT: What is the data exchange about
Use case (codes)	
Use case description (free wording)	
Consent beginning	Scope of consent
Consent end	
Restrictions on consent (and data): (optional free wording)	
Anonymisation	Constraints
Contract (explicit, implicit ...) : If yes, contract reference or terms of use	
Reversibility of the consent (not possible if based on a contract)	

Ontologies are one of the possible solutions for solving data interoperability issues. The word ontology covers a large number of different data sources ranging from thesauri to schemas shared on the Web through semantic Web technologies (Roussey *et al.*, 2011). In the MULTIPASS project, we studied different agricultural data exchange schemes, and in particular GIEA ("Gestion des Informations de l'Exploitation Agricole" - a model for Farm Information Management), a model created in France for data sharing (Pinet *et al.*, 2009). These schemes propose a vocabulary dedicated to agriculture, but too complex and not suitable for the uses in the context of consent management. The definition of consents will be associated with a typology of data and a typology of uses that remains to be defined. We recommend that these lists will be organized (hierarchies of category) and shared on the Web to meet the interoperability need of the ecosystem.

Discussion

A Blockchain could constitute the ecosystem on its own, but the challenge at this stage is to explore its promises in terms of trust decentralization. For this, in the second phase of the project, two consent management tools will be compared within use cases. The first one is based on a trusted third party (France Génétique Elevage, 2016) and the second one will be based on Blockchain technology.

MULTIPASS does not have the ability to interfere with consent management systems. They have to verify that the person who registers a consent is the one for whom the consent is given. It is therefore recommended to clearly identify the users with the creation of identity providers for agriculture, as there are elsewhere (French administration, Google or Facebook). Finally, it is the responsibility of the consent manager to ensure the legal value of the consents collected. The participants of the MULTIPASS workshop held on Sept 27th, 2018 (bringing together socio-economic partners of the farmer) highlighted the overlap in the regulatory bases of contracts and consents. There may be a risk of contradiction between a consent and a pre-existing contract.

Conclusions

The project aims to demonstrate to agricultural professional organizations the benefits and feasibility of a consent management ecosystem through limited but concrete use cases in France. The Blockchain technology will be evaluated to explore its promises in terms of trust decentralization. The router designed by the partners will implement a proof of concept for interoperability between existing and future consent management systems. It provides a solution ("data passport") to farmers for the control on their data and on the transparency of the data uses.

By strengthening the confidence of farmers to share their data, the project will bring new knowledge and new services. It promotes open innovation, i.e. the emergence of agricultural applications coupled with farmers' data from any data source or connected object. In this context, the goals are (1) to avoid the risk of concentration of innovation, and (2) the creation of knowledge by the analysis of massive farm data, in a chain of trust.

Acknowledgments

List of references

Douville, T. (2019) 'Contrat et données agricoles', Droit rural, n° 469, 7p.

CNIL (2018), 'Conformité RGPD : comment recueillir le consentement des personnes?' *https://www.cnil.fr/fr/conformite-rgpd-comment-recueillir-le-consentement-des-personnes*

GDPR (2016) 'General Data Protection RegulationEU 2016/679, April 27th, 2016.

FNSEA and JA (2018) 'DataAgri Charte sur l'utilisation des données agricoles', 12p.

EU Code of Conduct (2018) 'EU Code of conduct on agricultural data sharing by contractual agreement', 20p.

Roussey, C., Chanet, J-P., Soulignac, V. and Bernard, S. (2011) 'Les ontologies en agriculture', journal'Ingénierie des systèmes d'information (ISI)', specialissue 'Systèmes d'informations pour l'environnement' Vol 16, n°3, pp55-84.

Brun, F., Siné, M., Gallot, S., Lauga, B., Colinet, J., Cimino, M., Haezebrouck, T-P., Besnard, J., (2016) 'ACTA - Les Instituts Techniques Agricoles: L'accès aux données pour la Recherche et l'Innovation en Agriculture. Position des Instituts Techniques Agricoles'.

France Genetique Elevage. (2016) 'Le consentement des éleveurs pour l'accès à leurs données - Comment le gérer? Comment le faire respecter?', note proposed by the Information Systems Strategy Commission.

Pinet, F., Roussey, C., Brun, T., Vigier, F (2009) 'The use of UML as a tool for the formalisation of standards and the design of ontologies in agriculture'. Chapter in: Advances in Modelling Agricultural Systems, Springer, p.131-147.