Sharing data through an API platform - API AGRO

E. Rehben[1], B. Balvay[2], and T.P.Haezebrouck[3]

[1]IDELE - Institut de l'Elevage, 149 rue de Bercy, 75012 Paris, France
[2]IDELE - Institut de l'Elevage, 23 rue Jean Baldassini, 69007 Lyon, France
[3]ACTA, 149 rue de Bercy, 75 012 Paris, France

E-mail of the corresponding author: erik.rehben@idele.fr

Abstract

Sharing data is becoming a hot topic in agriculture and for livestock in particular. This paper outlines the main features of the organization which has been recently established in France to facilitate data sharing through Application Programming Interface (API) for agriculture and breeding industry. The different issues which have been addressed to operate the services are described in detail: the need analysis, the implementation strategy, the establishment of an API service provider, the description of the first set of services which are already available, the development plan for the future and the IT architecture. The conclusion underlines the interest of that technology to share data for an industry which needs a lot of update reference data and sophisticated calculation which may not be performed locally because of the complexity of the calculation or because of the size of the data bases and for privacy reason.

Keywords: API, livestock, breeding, application, cloud, web services.

Introduction

Remote Process Call (RPC) is an Information Technology (IT) which is existing for more than twenty years. The lack of standards and the lack of widely used and standardized network have limited the uptake of that technology for many years.

Now, the extensive use of connected applications through the internet protocols and Hyper Text Transfer Protocol (http) pave the way to opportunities to create new services which may enhance the value delivered by connected applications in a transparent way for their end users at a reasonable price.

This paper deals with the case of the implementation of Application Programming Interface (API) for breeding industry starting from the business requirements to the IT architecture including implementation strategy and service description.

Goal

For animal breeding, a majority of farmers and almost all the technicians are using connected applications to optimize their activities.

The uptake of update reference data and of sophisticated calculation requiring large data bases remains complicated.

The driving idea is to meet the needs of the breeding industry, by providing the missing resources by remote services through the cloud using API.

The target is the players of the French breeding industry: farmers, breeding companies, breed societies…

Implementation strategy

Operating API services implies to consider several stakeholders:

- The end users whose applications are using API services. These applications may operate either on line or off line as long as to be connected permanently with the cloud of the company which has developed the software. The purpose of this last provision which mainly addresses farm management applications is to avoid too many customers and the problems resulting in poor network performance. Two main types of end users are considered, the farmers who are using farm management applications, more than the half of the French dairy farmers, and the technicians who are providing farmers with advice for mating plan with artificial insemination sires.
- The technology providers which are providing the end users with applications integrating the API services in their software. These companies play a critical role to channel API services to the end users. In case of off line farm management applications, the company should also reroute the requests and the responses between the end user applications and the API service provider through its cloud. In France, it means approximatively ten companies.
- The API service provider which is operating API services based on data provided by data owners. The API service provider do not need to be the data owner, it should only have got the authorization to use data from the owner. The value of the services should not be linked to data ownership but based on the service availability, the data reliability, the neutrality of the calculations and the seriousness of the methodology.
- The data owners which are providing the API service provider with data either to be distributed or for calculation.

Figure 1 provides the value chain for API services.

The critical issue is sharing the value with the stakeholders through appropriate contracts with the additional problem when data ownership is not clearly defined.

The API services should be developed step by step, starting with a Minimum Viable Product (MVP) which consists in as set of services starting by those having the highest value for the end users. This first step is critical, since, in addition to MVP, it needs to address simultaneously organization and IT arrangements which will be used for the other steps.

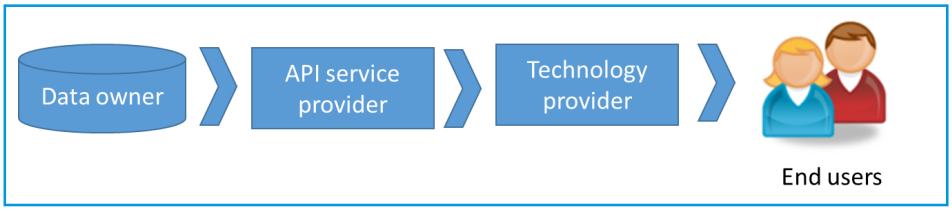

Figure 1. API services value chain.

Establishing the API service provider

Considering the potential of API for agriculture, a consortium that includes IDELE - Institut de l'Elevage, and the French key players of research and development and of technology providers established and invested in 2016 in a common subsidiary, API AGRO® (api-agro.fr) to promote API services for all agriculture: crops, livestock, poultry, vegetable…

Currently, about 40 API are already available dealing with different topics such as localizations, weather, soil, phytosanitary products, economics, livestock…

API services for breeding industry are available through that company whose mission is to provide its users with resources to operate API services under satisfactory economic conditions:

- A central repository to register the API with standardized tags to make them findable through a search engine.
- Monitoring the uptake of each API.
- Tools to allow the data owners to make by themselves their data available through API.
- Control of authorizations.
- License management.
- Data for automatic invoicing.
- Basic data visualization.
- API version management.
- …

Most of the above services are delivered through the cloud by an API platform which is operated by a company which has contracted with API AGRO®: Open data soft®.

Description of the services

Business requirement specifications

For the first step, the below needs have been considered:

- All user applications need at any time a unique update list of breed code.
- All user applications need at any time a unique update list of sires which are available through artificial insemination.
- To optimize mating plans, farmer and technician applications need at any time to test different options through the on line calculation of inbreeding coefficients for a limited set of parents. These coefficients should be calculated with a unique method, accurately described and widely used with all the data, the end user of the application is authorized to use.
- When network is missing or when network performance is too poor, the application of the technicians needs to calculate in advance, the inbreeding coefficient for the farms which will be visited.
- The application developers need streamlined services, easy to implement, with working principles which must be easily understood.

Service description

The first release includes four services:

1. "Breed code": that service is free. It can be used through a simple request by anybody according an open data license which provides that commercial use is possible but that no derivative is allowed in compliance with the "Creative common" framework.
2. "Sire list": the service is free. It can be used by anybody through a simple request according an open data license which provides that commercial use is possible but that no derivative is allowed in compliance with the "Creative common" framework.
3. "On line inbreeding coefficient": a paid service which requires a contract. It is available every hour, every day through requests which allow to submit a set of a maximum of 5 males and 150 females.
4. "Off line inbreeding coefficient": a paid service which requires a contract. It is available on working days through requests which allow to submit a set of a maximum of 80 males and 100 000 females with the guarantee to get the results in less than eight hours.

Table 1 below reported gives the distribution of the service and of their users.

Table 1. Services and users.

	Name of the service			
End user	Breed code	Sire list	On line inbreeding coefficient	Offline inbreeding coefficient
Farmers	X	X	X	
Technicians / mating advice	X	X	X	X

Development plan

For 2017, we have planned to sell services to at least three major breeding organizations representing several hundred of technicians who are providing farmers with mating advice.

By the end of 2018, it is expected to sell services to technology providers which are developing farm management software in order to address several thousands of farmers.

IT architecture

The API services should be designed in order to facilitate their uptake by people who are developing the applications. It implies at least:

- Fine grained services whose working principle may be easily understood.
- Few data to be exchanged.
- Using a widely spread language.
- Clear documentation.
- A test environment to test the integration of the API services in the software.

Because it is widely used, easy to implement and relevant for simple data structure, it was decided that the API services will operate as REST FUL web services.

Figure 2, provided below, gives the general architecture.

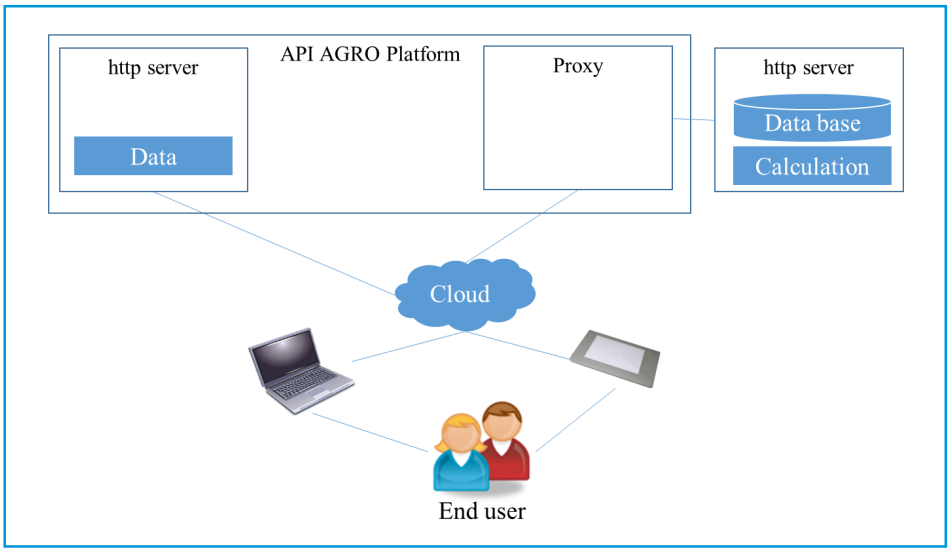

Figure 2. General architecture.

Two "http" servers are hosting the services: one for data, breed code and sire list, one for inbreeding coefficient calculation.

The data server is provided by the API AGRO platform which is only a proxy the calculation server which is located on an other platform.

End user PC or smartphones are connected to the servers through the cloud using internet protocols.

The communication between end user application and API services is performed by http protocol:

- Each service which has a specific URL (Uniform Resource Locator), for instance /api/consang/syncDemande for the service "On line inbreeding coefficient".
- The service is invoked by the client application through a specified "htpp" method. For the service 'Online inbreeding coefficient' the method 'POST' should be used in conjunction with the transmission of parameters. The request looks like "POST /api/consang/syncDemande PARAMETERS"
- PARAMETERS includes two lists of animal identification codes, on for the males and the other for the females.
- " The format used for exchanged data is J son. PARAMETERS looks like {population1": ["string"] , "population 2:" ["string"]} where "population1" is the list of males and "population2" the list of females.
- The response is transmitted by a J son file which contains for each couple of male and female the inbreeding coefficient.
- Because of the size of the files, for the service "Offline inbreeding coefficient", FTP (File Transfer Protocol) is used in combination with http.

Upcoming developments

For breeding industry we have planned to develop an improved release including some new services which would be :

- High precision inbreeding calculation for technicians who are in charge to prepare mating for the next generation of sire.
- Inbreeding calculation for sheep and goats.
- Inbreeding calculation including foreign pedigrees.
- Risk assessment of genetic abnormalities and mortality.

Conclusion

API have a huge interest for agriculture where a lot of applications for farmers and for technicians need update reference data and sophisticated calculation from large data bases. Both, may be delivered through the cloud by remote API services.

However, operating professional API requires several pre requisites to make the users confident in the services: availability of the service, quality of the service, result neutrality and privacy.

Operating API at a reasonable cost requires also resources which should be easily provided by one of the numerous marketed API platform.

The impact of research and development should strengthened when the results and the methods will be available through API.