Summary brief of the case *in itinere* "Pl@ntNet"

Supervisor/case team:

Pierre Bonnet,

Timothée Morin (CERDI, intern), Benoit Bertrand (CIRAD, Ipme), A. Joly (INRIA, Zenith), S. Dufour-Kowalski (INRA, Amap), Julien Champ (INRIA, Zenith), Hervé Goëau (CIRAD, Amap), Antoine Affouard (INRIA, Zenith), J. Carré (Tela Botanica), J.-F. Molino (IRD, Amap), N. Boujemaa (INRIA Saclay), D. Barthélémy (CIRAD, Bios).

I- The innovation story

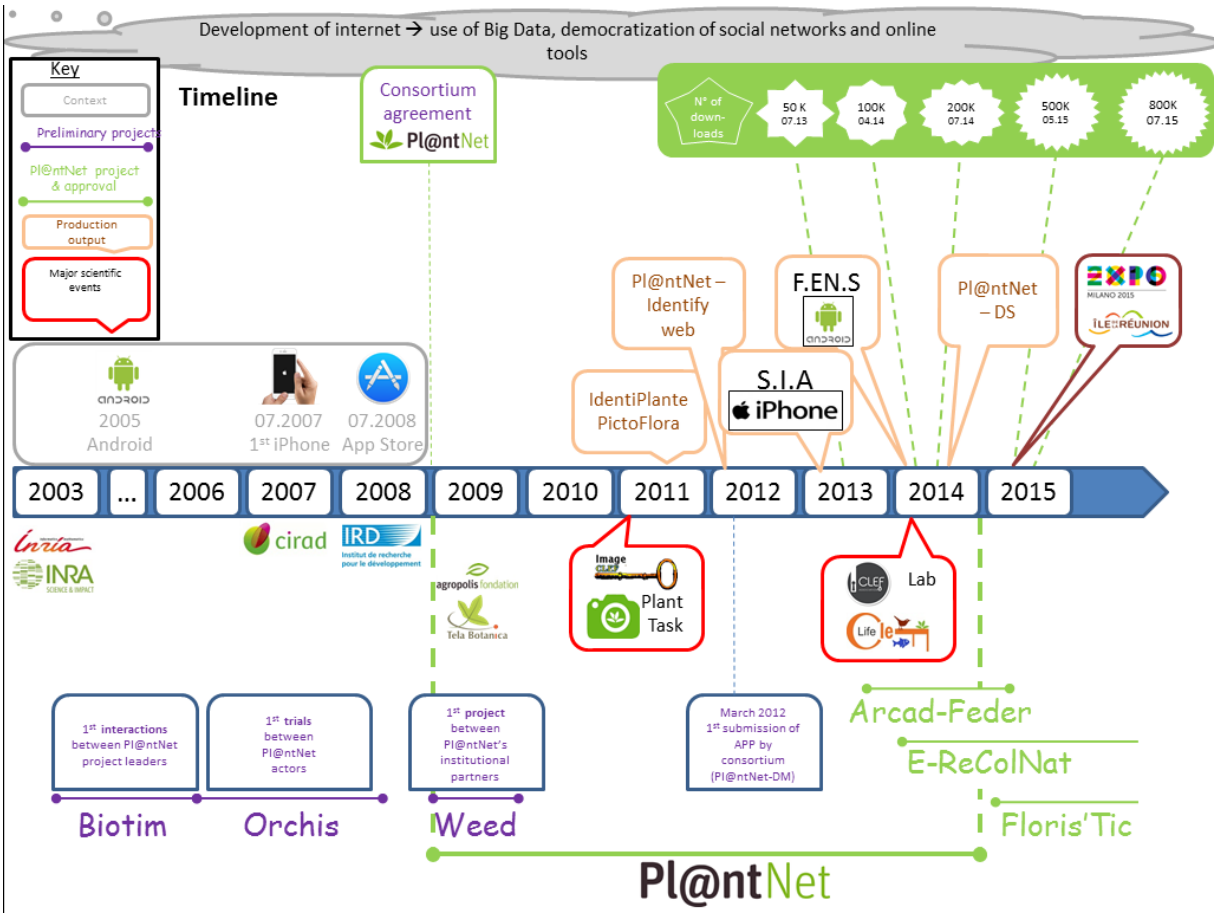


Figure 1 - Timeline

One of the major disincentives for studying botany is plant identification, which requires (i) considerable experience and (ii) learning specialized vocabulary. This first stage constitutes a major barrier that dissuades many beginners who find it difficult to overcome. For years, the mixed research unit Amap has been striving to make taxonomic identification easier. This work has led to the development of IDAO, an original approach to plant identification that uses a identikit robot. In addition, the work has focused on finding new and pertinent criteria for plant identification based, preferably, on the use of architectural and vegetative criteria. The two future project leaders for the Pl@ntNet project (Nozhaa Boujemaa and Daniel Barthélémy) actually met within the framework of the Biotim project (2003-2006), which focused on the use of large text/image datasets in biodiversity. Pl@ntNet was designed in 2007-2008 and officially launched in 2009. The pl@ntNet initiative focused on the development of innovative digital tools (i) visual aids for taxonomic

identification, (ii) collaborative revision of data quality, (iii) management of large volumes of botanical observations. The pl@ntNet initiative draws on the expertise, methods and complementary data records that the teams involved have been developing for nearly 15 years. Organizing the network of people involved in the project made it possible to put together unique structured datasets, which are among the most extensive and diversified in the world. The data pool on the flora of Western Europe was used via the Pl@ntNet-mobile system. This is a mobile interactive navigational tool and visual aid for plant identification. The tool can be used to browse all the jointly aggregated data via different systems. However, only the data that has been revised collaboratively is visible. It includes a system that aids automatic plant identification using photographs and a novel visual search engine. Thanks to contributions to the project, the application covers an ever-increasing number of species and has a growing number of images. Since 2011, the visual search engine that was developed was assessed via the task of identifying plants for the international campaign ImageCLEF (2011-2013) and then LifeCLEF (2014-2015), organized by the project teams. By working on (i) building up a structured dataset, (ii) developing innovating research tools, (iii) building up a community of volunteers, the Pl@ntNet initiative made it possible to aggregate a huge volume of botanical observations (over 2 million observations are currently being analysed) from the user community's identification requests. The infrastructure set up has been used by a wide typology of users (from beginners to experienced researchers) in over 150 countries in the world and by more than 1 400 000 people. This participative scientific initiative has led to the development of a new form of access to botanical knowledge that can be consulted directly in the field by a vast typology of actors.

When it comes to mapping the actors involved (shown below), we identify (i) the actors affected in red, (ii) the actors that influence the innovation in yellow, (iii) and those who are central to the Pl@ntNet project in purple. The key indicates the types of actors present. We distinguished between the actors who are from the world of research (in blue) from those from the associative sector (green), and between the users of Pl@ntNet (orange) from the media who help diffuse the innovation (grey). The arrows indicate the existing links between the different actors.

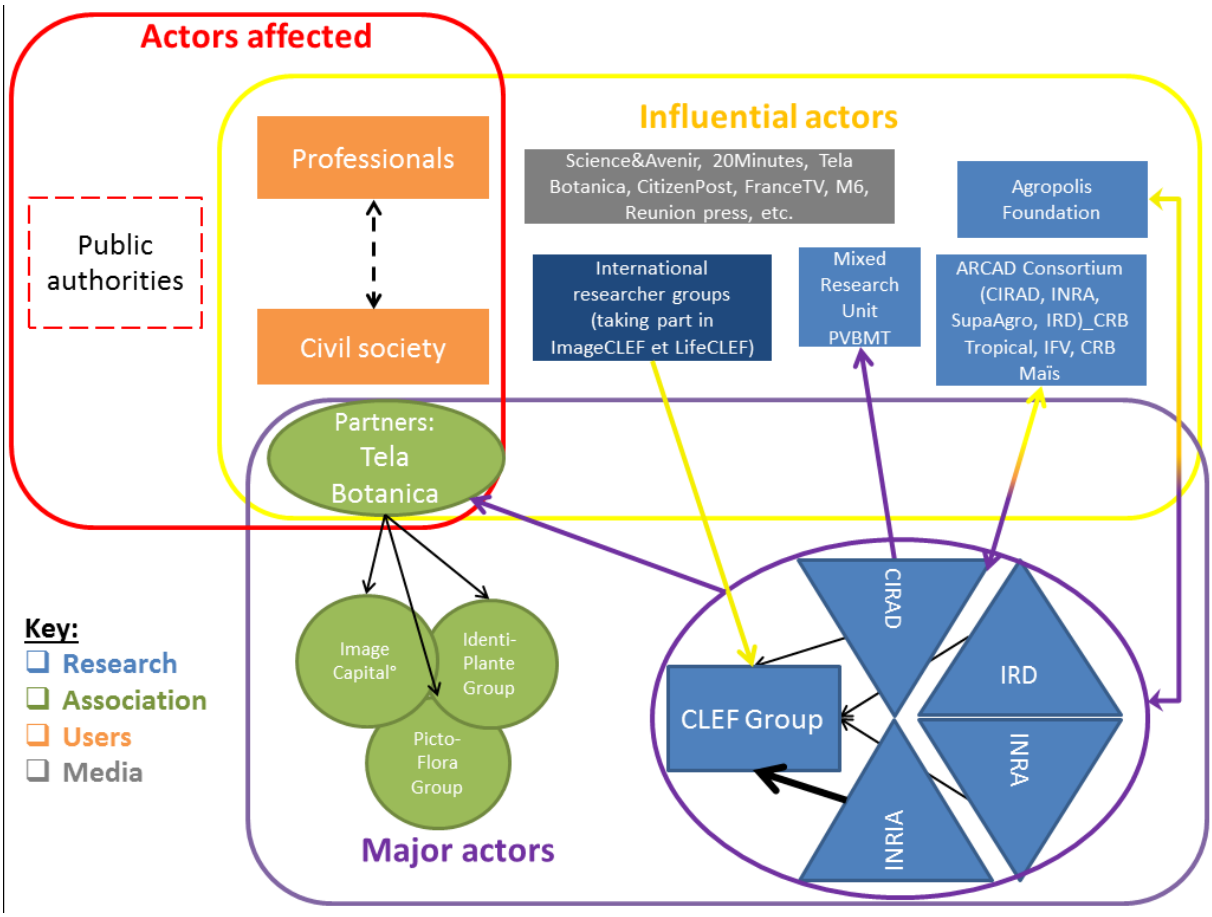


Figure 2- Map of actors

II- Lessons drawn from the analysis of the impact pathway

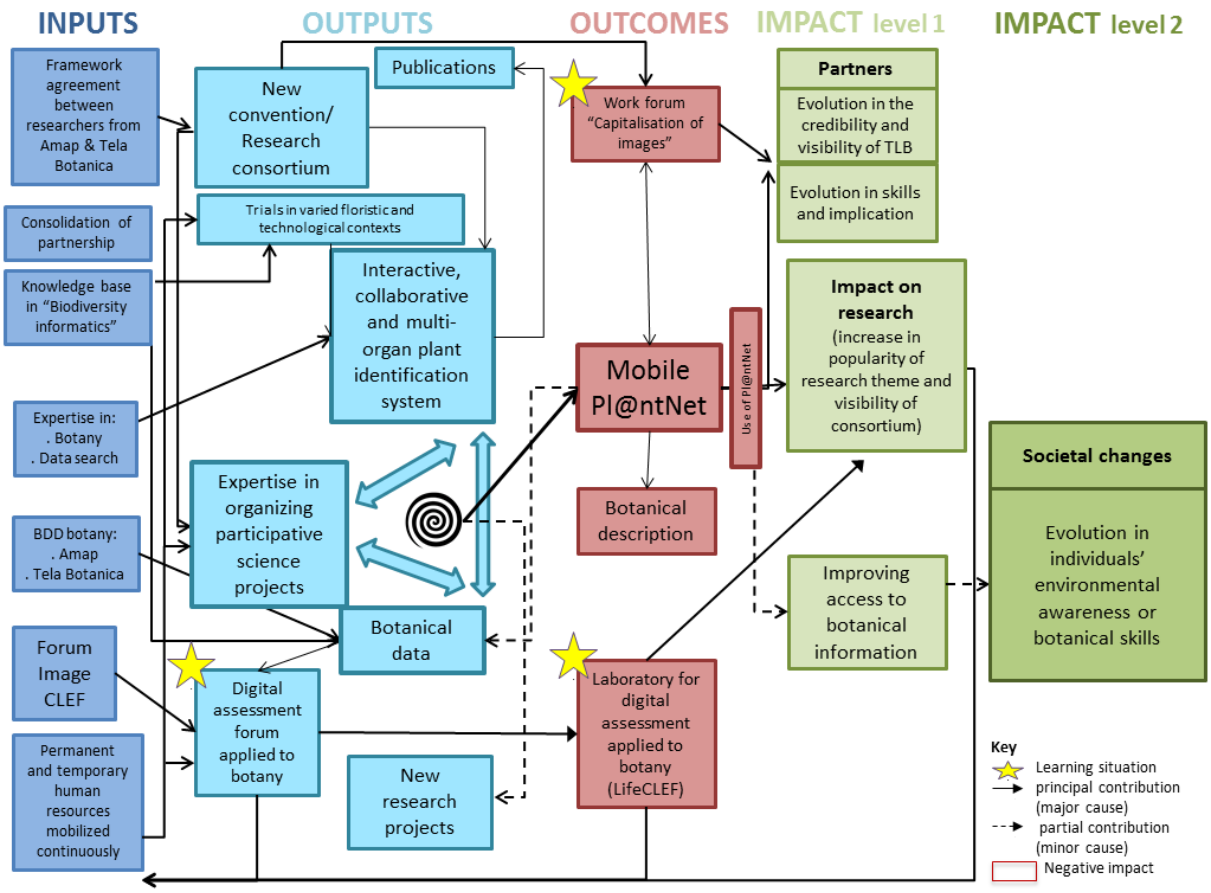


Figure 3 – The impact pathway

The work provided the opportunity to formalize a structured analysis of the perceptions of the different actors involved and affected, and to discover the reported and potential future impacts. The analysis was well received by actors beyond the realms of research. They gave us honest feedback of their perception and expectations. In our case, two important points facilitated the work: (i) the fact that we had access to a pool of several tens of thousands of people already interested in the Pl@ntNet initiative and who we could question (which meant we could collect more than 700 responses to the questionnaire sent out), (ii) the fact that Pl@ntNet benefitted from considerable media support meant that the people encountered had a better vision of the existing tool and its potential.

On the contrary, one major constraint for our case study was the fact that the innovation analysed was only diffused recently (Feb. 2013) and it is located on the boundary between shifting disciplinary fields (digital nomadism, citizens' sciences, open data, visual data mining, etc.). This makes it quite difficult to identify the impact scenarios because they depend on numerous factors in these disciplinary fields.

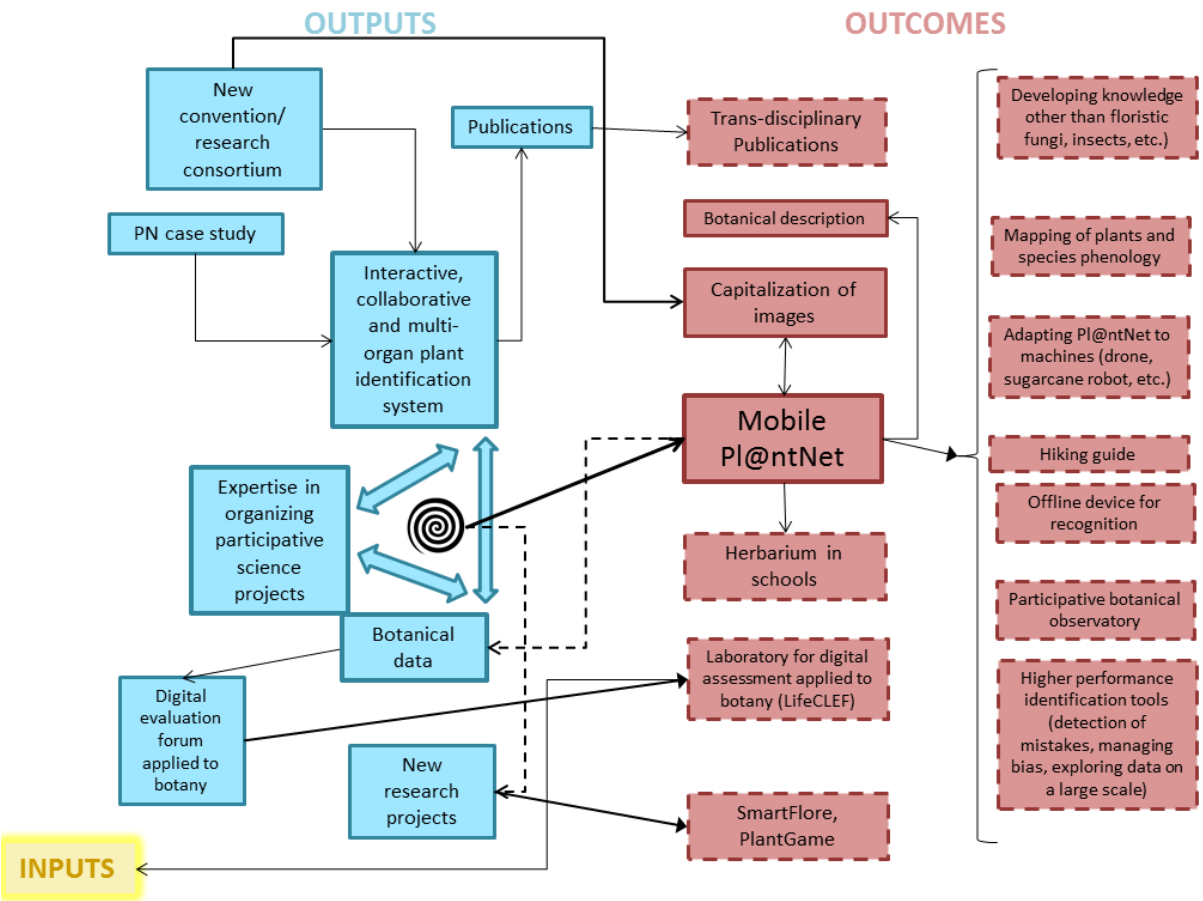
We regret not having been able to examine the negative impacts in sufficient detail, because they could probably have provided many elements to help the case team (or other teams working in similar disciplinary fields) to conduct work of this type in the future.

This work is helpful for identifying the expectations of the communities that may potentially be affected by the innovation developed. Thus, it provides elements that are more concrete than those

that were previously available to us. This will help steer the development of the innovation in new contexts (geographic, thematic).

III – Which scenarios could achieve the desired impacts? What type of monitoring (role of indicators?)

In our methodological framework, we focused on the description of hypothetical outcomes (shown by a dotted line) in the diagram below.

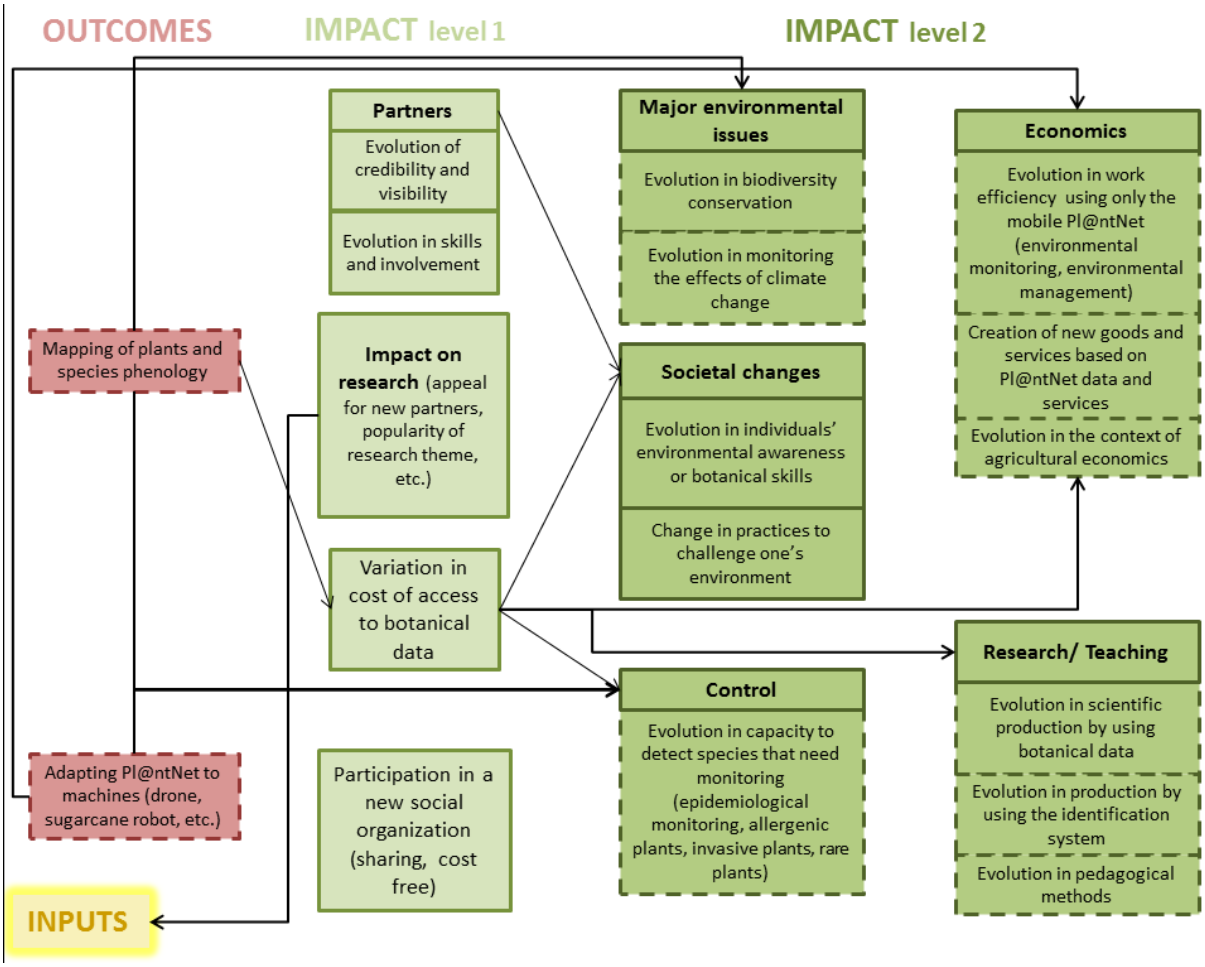


The impacts were identified in two ways. Firstly, responses to an online questionnaire helped us identify the domains where Pl@ntNet is used and the changes that it generated among users. Secondly, the numerous interviews with people more or less close to the project meant that we could collect the views of each one with regard to the proven and potential impacts. The ongoing measure of impacts involved the collection of quantitative data from different sources. As far as the presumed impacts are concerned, we tried to determine the indicators so that we could observe their future progression. Given their hypothetical nature, we chose to assess them during focus groups by using a common rating grid. We asked about the maximum future intensity of the impact, as well as the probability of its occurrence ("what chance would have this intensity to achieve?"), its proximity in time ("how long before it occurs?") and, lastly, the conditions required to realize it.

The results of the work on the likely outcomes are shown below:

Supposed outcomes	Probability of realization (1 to 5)	Time scale
Herbarium in schools	4	1 to 3 years
Development of recognittion other than floristic (fungi, insects, etc.)	3	4 to 6 years
Mapping of plants + species phenology	4	4 to 6 years
Adapting Pl@ntNet to machines (drone, sugarcane robot, etc.)	4	7 to 10 years
Hiking guide	5	1 to 3 years
Offline recognition device	5	1 to 3 years
Participative botanical observatory	3	4 to 6 years
Higher performance identification tools (detection of mistakes, bias management, exploring data on a large scale)	4	4 to 6 years

III-1 Scenarios envisaged to link outcomes to impacts



This work was conducted on the basis of interviews and focus groups so that the responses provided by the people interviewed could be taken into account properly. Initially, we did not consider the

constraint of monitoring indicators in terms of type of impact. Later, we chose solely the impacts that could be monitored and measured over time.

III-2 Which monitoring methods are possible that include indicators per type of impact?

Type of impact	Identified impacts	Indicators	Results	Sources used
It is more of an outcome given the indicators	Use of the plantnet application? (Involved in the development of non-commercial	Evolution in n° of users per annual peak use	Evolution in peak use per year from 2013 to 2015: 24 536 - 303 297	Google Analytics
		Evolution in n° of contributions to the pl@ntnet database	Evolution from Feb. 2014 to June 2015: 1000 - 2500	Pl@ntNet-DataStore
impact 1	Reduced cost of access to botanical data	Share of species referenced by PN in relation to the total n° of species in France	Shift from 2% in 2010, what measure in Oct. 2012 a key moment, 76% in Jul. 2015	Pl@ntNet
impact 1	Increase in the popularity of this research theme and the renown of the research team	N° of people able to see plantnet in the media	Shift from 1 million in 2013 to 7 million in 2015	Media
		New contacts who approached the plantnet team from outside the realms of research (per domain)	38 potential partnerships	
		New contacts who approached the plantnet team from the world of research (per discipline)	13	
		Evolution in the N° of research teams involved and registered at assessment forums (lifeCLEF and planttask)	2011: 72 and 8, 2015: 140 and 7	
		N° of scientific communications from the pl@ntnet team by type of communication	. A dozen newspaper articles, . Over 30 communications at information technology conferences . Nearly 30 communications at conferences on biology, ecology	
		N° of quotes from plantnet and life clef in scientific publications	41	Google Scholar
impact 1	Evolution in the credibility and visibility of the association of botanists	Evolution in the N° of members of TB	Constant rate of growth since its creation in 2002, PN's impact is not significant	
		Quality of images selected by TB	40 000 images shared by users via mobile Pl@ntNet, 55% of images were rated. The rates ranged from 1 to 5. 23% of images were rated below 3, 47% equal to 3 and 30% above 3.	
impact 1	Evolution in the skills and involvement of the association of botanists	Share of recent and former members of TB who consider that they have improved their expertise	65% longstanding members, 89% recent members	impress survey
		Monthly evolution in the number of ratings and comments made by TB members	Increasing n° of commentaries, fewer ratings for image quality	impress survey
impact 2	Increase in users' environmental awareness through their improved	% users who use plantnet for leisure who have developed an environmental awareness	64% already were aware, 23% are more so, 12% no change (rate of response 1% n° 627)	impress survey
		% users who use plantnet for leisure who have improved their botanical skills	65% a little or a lot, 34% not at all (Total N° 712)	impress survey

These indicators (which can evolve over time) will be monitored primarily using the analyses of surveys that we will be able to conduct among the actors affected (or potentially affected). This approach is part of the planned evolutionary development of the innovation. The aim is to further our understanding of the user's experience, to channel their involvement more effectively and meet their expectations. With this approach, we hope to provide more satisfactory answers to the research questions that interest us, but also to promote our work and its integration in different fields beyond research (primarily education and natural resource management).