



AGINFRA

Europeana as a Resource for Social Scientists In Agriculture and Food

A Europeana Cloud Case Study



White Paper on

“Europeana as a Resource for Social Scientists

In Agriculture and Food”

A Europeana Cloud Case Study

Acknowledgements

Authors: Maritina Stavrakaki (Agroknow/ARIADNE Foundation), Vassilis Protonotarios (Agroknow/ARIADNE Foundation), Nikos Manouselis (Agroknow/ARIADNE Foundation), Giannis Stoitsis (Agroknow/ARIADNE Foundation).

With thanks to: Akis Tsekouras, Veronika Ovcinnikova, Nikos Marianos, Effie Tsiflidou, Katerina Galani, Thodoris Kontogiannis, Andreas Drakos



This White Paper has been compiled as part of the “Europeana Cloud: Unlocking Europe’s Research via The Cloud” project. The Europeana Cloud project is co-funded by EC under the CIP-ICT-PSP programme.

Table of Contents

Executive summary	8
1. Introduction.....	10
1.1. Rationale.....	10
1.2. What you need to know	11
2. Methodology	14
2.1. Profiling user personas	16
2.1.1. Content analysis	18
2.1.2. Information-related challenges and problems.....	18
2.1.3. Description of information services and systems	19
2.1.4. Description of current usage scenarios of information services and systems	19
2.1.5. Design interface wireframes and description of user-envisaged usage workflows..	19
2.1.6. List resulting requirements for outcomes services	20
2.2. Validation of user’s requirements	20
2.3. Designing final wireframes with required features for the information services of the use cases.....	21
2.4. Translating features into requirements for services	22
3. Outcomes	24
3.1. Initial profiling of targeted personas.....	24
3.1.1. Using online questionnaires	24
3.1.2. Interviews	26
3.1.3. Using the Persona Graph.....	26
3.2. Personas and usage scenarios.....	27
3.3. From Personas to Tools	29
3.4. Software and services.....	32
3.5. Evaluation report on integrated tools	38
4. Integration with AGINFRA	42
4.1. Europeana portal.....	42
4.2. The AGINFRA data infrastructure for agricultural research	44
4.2.1. The agAGINFRA FP7 project	44
4.2.2. The AGINFRA global hub of agricultural research	45
4.3. CIARD-RING	46

4.3.1.	Register to CIARD RING	48
4.3.2.	Register a service or dataset	51
4.4.	Recommendations on establishing a liaison between Europeana and AGRINFRA.....	53
5.	Conclusion	56
6.	Annex A: Online questionnaire for the AGRERI case.....	58

Table of Figures

Figure 1. The main workflow followed for the development of the services	15
Figure 2. Empty persona graph with some initial questions to be answered.....	17
Figure 3. Partial view of the online questionnaire	25
Figure 4. Information about a persona to be transferred to a persona graph	27
Figure 5. The first version of the mock-ups for the agricultural researchers scenario	29
Figure 6. The second version of the mock-ups for agricultural researchers scenario with some of the changes highlighted	30
Figure 7. Screenshot from the AGRERI Library.....	31
Figure 8. Screenshot from the External Libraries	32
Figure 9. Screenshot from the AGRERI site	33
Figure 10. Screenshot from results listing after a search query.....	35
Figure 11. Screenshot from view item with related resources	36
Figure 12. AGRERI Discovery Microsite architecture	37
Figure 13. Screenshot from the Europeana portal.....	43
Figure 14. Screenshot from the Europeana portal listing results after a search query	43
Figure 15. Screenshot from the Europeana portal view item with the metadata	44
Figure 16. The AGRINFRA global hub of agri-food research.....	46
Figure 17. Screenshot from the CIARD RING site	48
Figure 18.. Screenshot from the CIARD RING registration process.....	49
Figure 19. Screenshot from the CIARD RING confirmation e-mail.....	49
Figure 20. Screenshot from the CIARD RING e-mail validation.....	50
Figure 21. Screenshot from the CIARD RING service/dataset registration	50
Figure 22. Screenshot from Europeana search query using the keywords “agricultural cooperative”	51
Figure 23. Screenshot from the CIARD RING create information source, dataset.....	52
Figure 24. Screenshot from the CIARD RING created information source, dataset.....	53

Executive summary

This White Paper is presenting how Europeana¹ may be used as a resource for Social Scientists that work in a different scientific domain, the one of agriculture and food. Policy making in this domain is heavily dependent on ethnographic studies and other social science methods and tools. In many cases, historical resources and archives are used as research resource.

In this White Paper, we present how researchers who are working in completely different disciplines, like agriculture and food, could discover and make uses of available interesting data sets from the vast amount of resources available through Europeana. By using these resources and data sets, agricultural (and not only) researchers could be able to investigate various topics using different scientific methods and tools, thus making multi-disciplinary agriculture research more useful and meaningful.

We present the work carried out within the Europeana Cloud project² on a piloted case study: the Agriculture Economic and Policy Research Institute (AGRERI)³ that is the national agricultural economics and policy research institute of Greece. It also discusses how Europeana Research⁴ can establish a connection with the European research data e-infrastructure for agriculture and food called AGINFRA⁵. This work has been carried out with the Europeana Cloud project by Agroknow⁶, a member of the ARIADNE Foundation⁷ that was a partner in the project.

¹ <http://www.europeana.eu/portal/>

² The “Europeana Cloud: Unlocking Europe’s Research via The Cloud” project is a project co-funded by EC under the CIP-ICT-PSP programme (<http://pro.europeana.eu/structure/europeana-cloud>)

³ <http://www.agreri.gr>

⁴ <http://research.europeana.eu/>

⁵ <http://aginfra.eu/>

⁶ <http://www.agroknow.com>

⁷ <http://www.ariadne-eu.org/>

1. Introduction

1.1. Rationale

The Europeana portal (<http://www.europeana.eu/portal>) is the outcome of the efforts of many European countries to make their rich and diverse cultural heritage available online. Working with a very large network of museums, libraries and archives around Europe, more than 40 million digital cultural artefacts are made discoverable and accessible through the portal. The next step was to evolve Europeana from a large aggregator of digital cultural collections into an e-infrastructure which was the aim of the Europeana Cloud project.

One of the challenges is to identify and make available interesting data sets from the vast amount of resources that become discoverable and available through Europeana as well as through other aggregators. Using the above mentioned valuable and rare archives is essential in order to support researchers who work not only with research methods and tools but also with social scientific methods and tools.

Another challenge is to explore how researchers from totally different disciplines can be engaged into using and taking advantage of the collections that become discoverable from an aggregator like Europeana.

To this context, an interesting discipline that could evaluate the importance of such infrastructure is the research community of agricultural researchers at the Agriculture Economic and Policy Research Institute (AGERI) who investigate the economic, cultural and societal factors affecting major agricultural aspects not only in Greece but also in other European countries.

An infrastructure that will be able to integrate and present in a simple and visual way all this information from disperse sources will be of great importance not only to researchers but to anyone who would like to take advantage of all this knowledge. It will also change the way that researchers are currently storing and using the data produced as well as answering and making more sense of the data produced.

Socioeconomics and policy research have become an integral part of agricultural research projects and programs. The concepts, framework and methodological tools and techniques for socioeconomic and policy research are derived from social sciences disciplines which help bring human dimension and issues in the research process and cover a wide range of research issues. Social science can contribute critically in the analysis and conduction of research in crosscutting themes such as poverty, food security and natural resource management. They provide broad framework, research methods and tools for scarce resource allocations, identification of research agenda, research priority setting, adoption and impact assessment studies and can play important

role by identifying policy gaps and constraints to technology development, uptake and scaling-up⁸.

More specifically, Agriculture Policy and Food Policy have always been in the centre of attention, with the aim to tackle many social problems, namely hunger, poverty, major nutrition imbalances and environmental degradation among others. Many major initiatives and institutes have been providing research-based policy solutions in order to sustainably reduce poverty and end hunger and malnutrition worldwide and especially in developing countries, including:

- The Farm Accountancy Data Network (FADN)⁹ which is an instrument of the European Union for evaluating the income of agricultural holdings and the impacts of the Common Agricultural Policy (CAP)¹⁰.
- The International Food Policy Research Institute (IFPRI)¹¹ which is a research centre of the CGIAR Consortium¹² and its research analyses options for policies, institutions, innovations and technologies that can advance a significant transformation of agriculture and food-production systems, to tackle the afore mentioned social problems.

In order to investigate the societal factors that affect major agricultural and food aspects, a number of social science methods and tools have been used, including survey data collection, statistics, analyses at a farm, cooperative and industry level, case studies, focus groups research, guides, interactive maps etc. that produce or contain socioeconomic information and data of great importance for researchers.

1.2. What you need to know

What is the Europeana Cloud?

The Europeana Cloud project¹³ (Europeana Cloud: Unlocking Europe's Research via The Cloud) is a 36-month project co-funded by EC under the CIP-ICT-PSP programme which aims to create a digital infrastructure in order to support researchers from different disciplines discover valuable and rare resources from disperse sources and collections, namely Europeana, a large aggregator of digital cultural collections.

⁸ https://ec.europa.eu/research/social-sciences/pdf/other_pubs/socio-economic-tools-for-sustainability-impact-assessment_en.pdf

⁹ <http://ec.europa.eu/agriculture/rica/>

¹⁰ http://ec.europa.eu/agriculture/cap-overview/2014_en.pdf

¹¹ <http://www.ifpri.org/>

¹² <http://www.cgiar.org/>

¹³ <http://pro.europeana.eu/structure/europeana-cloud>

What is AGRERI?

The Agriculture Economics and Policy Research Institute (AGRERI) is a research institute founded by the Greek Ministry of Rural development and Food that focuses on the economic, cultural and societal factors that affect the agriculture and food sector by conducting high-level research.

What is the AGRERI Discovery Microsite?

The AGRERI Discovery Microsite is one of the tools that has been developed in the context of third year deployment of the Europeana Cloud project. It is a technology that will be embedded in the existing AGRERI website and it will allow users to search for resources from external sources, like the Europeana aggregator, and present results from different sources in the same place with specific facets and filters which will help them narrow down the results according to their specific needs and research interests.

What is the purpose of this White Paper?

The purpose of this White Paper is to provide a generic, high level blueprint to anyone interested in discovering information from heterogeneous data at disperse sources. Through this White Paper, researchers and interested organizations will get a clear picture of the requirements, benefits and implementation process and how they can be powered by and benefited from such a digital infrastructure.

What is the intended audience and what is their benefit?

The White Paper is targeted towards researchers, research teams and institutions from different disciplines that are interested in discover valuable and rare archives from different collections in order to investigate and answer some of the research questions at hand.

What is the relation of this White Paper to the Europeana Cloud project work?

This White Paper is a standalone appendix to the deliverables D3.1 “Document on Personas, scenarios and use cases”, D3.2 “Tools and services”, D3.3 “Evaluation report on integrated tools” and D5.5 “Europeana Cloud Partner Roadmap” of the Europeana Cloud project.

2. Methodology

The methodology followed and described in this document aims to provide the means for extracting information from potential stakeholders of the service that is going to be developed, in this case the users of the AGRERI Discovery Microsite, but it can also be applied to other use cases. The requirements extracted from these stakeholders will be used for shaping the expected outcomes of the service for the end users.

Everyone involved in tasks related to the elicitation of requirements has to organize and implement a number of activities (such as interviews, focus groups or online surveys) aiming at extracting user requirements from selected members of the communities they represent in the project, as described earlier in this document. For example, the initial profiling of the personas may take place in the form of customized online questionnaires or interviews with users featuring the characteristics of interest to the Europeana Cloud project. In both cases, the same set of predefined questions should be used; in the former, the responses are provided by the anonymous user himself/herself while in the latter the responses are recorded by the facilitator (project partner).

The validation of this feedback and elaboration on the responses collected is achieved during dedicated, focused workshops that aim to engage different types of users (personas) and allow them to work closely on the description of solutions that each different persona is facing. During these workshops, participants are asked to describe each persona in details, providing detailed information on its demographics, content-related sources and content-related issues that need to be addressed (per persona). In addition, during the workshop, they are asked to design how an initial design will look like in the form of wireframes, using paper and markers.

The next step of the process has to do with the validation of this initial feedback (referring to both the persona and the wireframes) with feedback acquired through online questionnaires, interviews & workshops. Only after this information has been validated can the project partners involved in this task confirm the suggested personas and design the wireframes using an appropriate software tool.

The proposed methodology, which can be adapted in several instances to meet the requirements of a wide variety of stakeholder, consists of the following steps:

1. Initial profiling of a user persona
 - a. A generic online questionnaire is developed for the collection of initial requirements and the initial profiling of a persona.
 - b. The generic questionnaire is adapted accordingly in order for it to meet the specific needs of each user community, e.g. translation of the survey questionnaire (all the results should be in English, both the report and the interviewees' answers)
 - c. The adapted questionnaires are used for the initial profiling of each persona for the use case described.
 - d. Feedback is also collected from the online questionnaires.

- e. Profiling can also take place through interviews (face to face, through Skype or phone call).
2. Validation of the collected requirements
 - a. Workshops are organized for the validation of the feedback collected regarding the personas.
 - b. New information regarding the personas is acquired.
 - c. Feedback from each workshop is collected through reporting.
 - d. Validation can also take place through questionnaires.
3. Brainstorming on new services, in terms of functionalities, user interface (wireframes) and interaction between the user and the services (usage workflows)
4. Transformation of validated information and user requirements into requirements that will allow the technical team to work on the outcomes to serve the corresponding personas

Reports from all workshops are used for the analysis of requirements that will drive the development of services by the technical team. A schematic representation of this process is shown in the figure below (Figure 1).

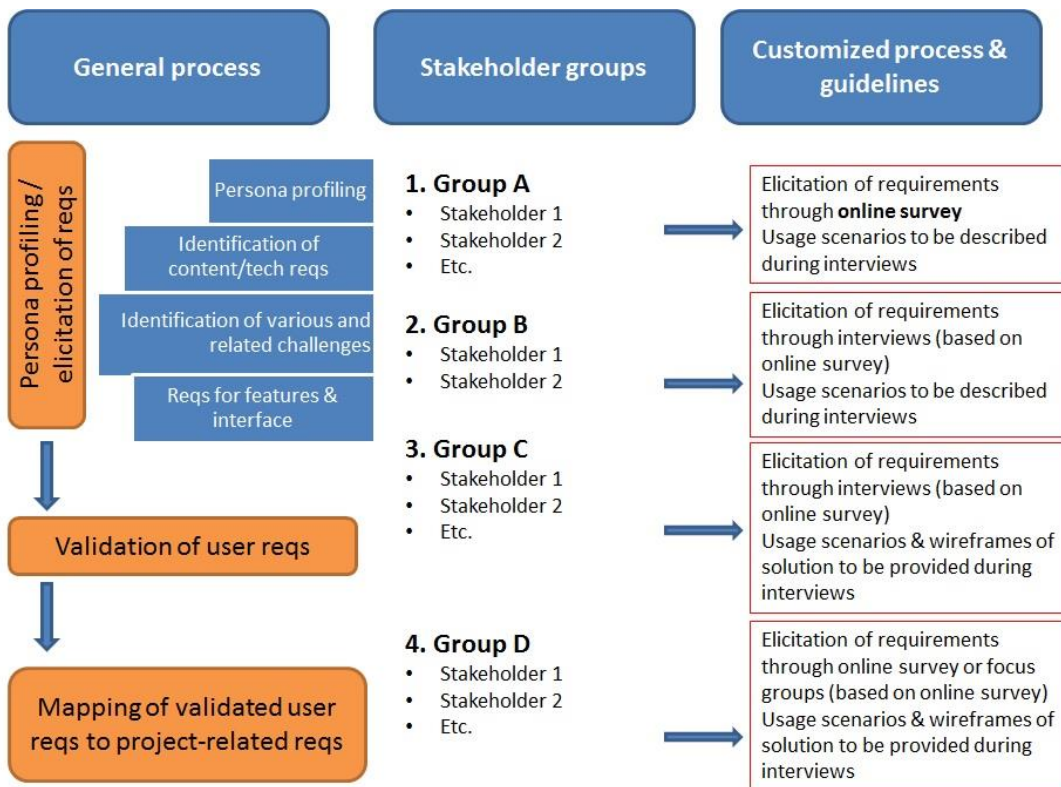


Figure 1. The main workflow followed for the development of the services

2.1. Profiling user personas

The first step in the proposed methodology is the identification of the different user profiles (personas) of interest to the service.

A persona is defined as a fictional character that is created in order to represent the different user types that may use a service. A persona is defined as a representation of the behavior of a hypothesized but validated group of users. The profile of a persona is compiled from information extracted through interviews with real users; this information is collected, evaluated and validated before it is merged with related information from interviews with other users with similar characteristics. This compilation of information is used for developing the profile of a persona that includes behavior patterns, skills, attitudes, and other attributes. The use of personas is common in workflows related to product development and validation, as it allows the design of a product based on the needs assigned to a specific persona.

Usually, the following attributes are used for the development of a persona:

- Persona description and demographics;
- Data types and services of interest to each persona as well as data types and services already being used (own data or data from external source) by the specific persona;
- Features of the envisaged solutions for a specific community or target group.

These different types of information are analyzed in the following sections.

For the recording of these requirements, a template called **Persona Graph** is used. This allows the uniform description of different personas by different people involved in this process. The following figure presents an example of an empty Persona Graph with some initial questions to be answered (Figure 2).

<h3 style="text-align: center;">Persona</h3> <ul style="list-style-type: none"> • Name, title, seniority level • Topic of study • Background, sector/field • Type of affiliated organisation • Works alone or in group (big/small)? • Team setting (Local, national or international team) • Type/expertise of people (s)he works with • Prior experience in the field (also in years) • Willingness to learn new tools 	<h3 style="text-align: center;">Data Requirements</h3> <ul style="list-style-type: none"> • Interesting content type (e.g. publications, datasets, presentations, maps etc.); • Content format (e.g. PDF/DOC, PPTs etc.); • Content volume (no of resources or size in MB/GB); • Content sources (database, repository, web portal etc.) & interoperability options (e.g. OAI-PMH, API, RSS, SPARQL endpoint etc.); • Content Language (content & associated metadata); • IPR and licensing information, referring to use, reuse, adapt and redistribute, etc
<h3 style="text-align: center;">Information Challenges</h3> <ul style="list-style-type: none"> • Problems with data • Problems with existing tools • Problems with people • Other problems/issues 	<h3 style="text-align: center;">Solution Features</h3> <ul style="list-style-type: none"> • Desired feature 1 • Desired feature 2 • Etc.

Figure 2. Empty persona graph with some initial questions to be answered.

The Persona Graph is based on an adapted version of the Lean Canvas, as it was presented by Maurya (2012)¹⁴. The Lean Canvas is a template that aims to include information related to the development of a new product and its business model. It focuses on addressing broad customer problems and solutions and delivering them to customer segments through a unique value proposition. The version used in the requirements' elicitation methodology is adapted in a way that allows the collection of information focused on a project's expected developments, without the need of collecting additional information or going into too many details that would not be substantial for the project.

A Persona Graph contains information that is collected through the use of the online questionnaire, interviews or through the focused workshops, as described in the next sections. The prioritization of the information recorded in each Persona Graph, based on the importance of each statement according to the users, is of major importance. In this direction, in the boxes related to the Data Requirements, Information Challenges and Solution Features the most important (according to the users) statements should be recorded on the top of the list and the least important at the bottom.

¹⁴ Maurya, A. (2012). Running Lean: Iterate from Plan A to a Plan that Works. O'Reilly Media; 2nd edition (March 9, 2012). ISBN-10: 1449305172

2.1.1. Content analysis

Content analysis is the second step of the proposed workflow. It refers to the process that aims to identify, record and analyze the characteristics of the content of interest as well as the content already in use for the personas identified and studied in the task. The following parameters are included in the analysis of the content analysis:

- Content type (e.g. publications, datasets, presentations, maps, etc.);
- Content format (e.g. PDF/DOC files, PPTs, ZIP files, etc.);
- Content volume (number of resources or size in MB/GB);
- Content source (database, repository, website or web portal etc.) and interoperability options (e.g. OAI-PMH, API, RSS, SPARQL endpoint etc.);
- Content Language (referring to the language of the content and its associated metadata);
- Intellectual Property Rights (IPR) and licensing information, referring to use, reuse, adapt and redistribute, among others.

The detailed analysis of the content identified through this process will allow the optimization of the services so that they will meet the specific requirements of specific communities making use of specific content. Additional content sources and types will be identified through the profiling of the user personas using the methodology described in this document.

2.1.2. Information-related challenges and problems

One of the aspects of the definition of each persona and the identification of its attributes is the identification of information- and content-related challenges and problems that a specific persona is facing. These problems may have to do with identifying, retrieving, accessing and managing the content of interest, among others.

The identification of these content-related issues will allow the project partners to work on services that will address these issues and provide meaningful solutions. These solutions may also be proposed by participants of the project's user requirements' workshops or through interviews.

When recording the information in the corresponding box of the Persona Graph, it is really important to indicate the importance of each entry by prioritizing them in the list. In this direction, the most important statement (according to the users) should be listed on the top of the list with the least important one at the bottom of the same list.

2.1.3. Description of information services and systems

A list of available information services and systems will be used as use cases and will be enhanced through the integration of the AGRERI Discovery Microsite services. On top of that, additional ones are identified during the user requirements elicitation. Both existing and new information services and systems need to be described in details by the users during the interviews and related workshop sessions. This information needs to be extracted from the expected end users of the project's outcomes through the interactive sessions of the related workshops to be organized by the project as well as through interviews and other means of extracting requirements. These services and systems are expected to be enhanced through the integration of various functionalities that will enhance the experience of their end users.

2.1.4. Description of current usage scenarios of information services and systems

A set of requirements of high importance to the project will be related to the description of the interaction between the user and the information services and systems. The way that users interact with such services need to be described in details and depicted as a set of steps consisting a workflow. This information will allow a better understanding of the activities of a user when using a service or a system and therefore a better identification of his/her requirements, steps that could be improved or enhanced with the use of the AGRERI Discovery Microsite services, expectations from such services etc.

2.1.5. Design interface wireframes and description of user-envisaged usage workflows

During the interactive sessions of the workshops to be organized by the project for elicitation of user requirements, participants will be asked to share their ideas on the integration of the new, envisaged service or feature, in the existing website. This design refers to the location of the new service box or button in a way that the user will feel comfortable with, e.g. not interfering with the current workflow of the user. These requirements will drive the User Interface Design (UI) that should be considered and implemented by the corresponding team. The wireframes that will result from the workshops will be in the form of a drawing in a flip chart paper.

After describing the content-related issues and proposing potential solutions that would address these issues, usage workflows related to the use of these envisaged services need to be described. More specifically, potential users of the AGRERI Discovery Microsite (e.g. interviewees or participants of the project's workshops) will be asked to describe the interaction between the users and these services in the form of a usage workflow.

A usage workflow refers to the description of all the activities that take place for a user completing a specified task; more precisely, a usage workflow is defined as a use case drawn out into a step-by-step procedure, sometimes accompanied by a flowchart. For example, if a user wants to

extract a figure from a research publication, he/she would have to go through a number of steps, which could be more or less the following:

- the user visits a content repository;
- he/she uses a search term in order to retrieve publications related to his/her work;
- he/she retrieves a number of results;
- he/she filters the more relevant results through the use of filters or facets;
- he/she selects a specific publication;
- he/she goes through the publication and identifies an interesting reference;
- he/she clicks on a button allowing him/her to retrieve only the figures of this publication.

All these steps need to be recorded by the users in details and in a logical order. This will allow the project partners to provide users with a set of functionalities that will meet the expectations of the users, taking them into consideration as User Experience Design (UX) requirements.

2.1.6. List resulting requirements for outcomes services

The feedback and requirements collected during the interviews, workshops and other means organized for this purpose will have to be collected, organized and validated before it is used for the development of the corresponding outcomes, such as e-infrastructure and final online services for end users. All additional related information collected through the aforementioned means will also have to be transformed into the corresponding requirements. The resulting requirements will be classified according to their topic, in the following categories:

- Requirements for the definition and optimal description of the user persona (referring to the demographics of the persona);
- Requirements related to the content-sources already used or of interest to each persona;
- Requirements related to the content-related issues of each persona;
- Requirements related to the expected services of each persona;
- Requirements related to the interaction of a user with the new service(s);
- Requirements related to the integration of each envisaged service in an existing website.

The requirements collected and validated through the process described are used in order to drive the development of the services both at functionality and at user interface level.

2.2. Validation of user's requirements

After the targeted personas have been identified and their initial profiles have been created (including demographics, content-related issues and proposed solutions), they need to be validated in order to ensure their accuracy and eligibility for being used as the basis for the project's envisaged services. This validation may take place in the form of smaller- or larger-scale events.

Small scale events include interviews that can take place either face to face, through phone or online (e.g. Skype calls), as described earlier and focus groups.

Large scale events include meetings with teams of stakeholders and even workshops. They can facilitate a high number of participants and should include sessions for teamwork.

The outcomes of each user requirements' event included the following:

1. input per persona in the Persona Graphs provided by the facilitators. It should be noted that in the case of such events, there should be parallel interactive sessions running per persona.
2. a visualization of the envisaged services to be developed per persona (hand-drawn, using markers and paper)
3. a proposed workflow per persona (hand-drawn, using markers and paper), explaining the use of such as service and the role of different types of users

2.3. Designing final wireframes with required features for the information services of the use cases

Through the aforementioned means of acquiring feedback from the specific user types, a number of expected services are expected to be identified, described and analyzed for each persona within each one of the use cases studied by the project. Such brainstorming sessions should take place in the context of workshops, where different types of potential end users are expected to participate.

As regards the design of wireframes, the following steps were followed:

1. Initial wireframes were drawn by the participants of the events (one wireframe per persona), highlighting the way that these services will be integrated in existing websites, finders and other sources of content.
2. Additional requirements and feedback collected through interviews or online questionnaire to be taken into consideration for improving or revising the initial wireframes
3. Refined wireframes to be designed by the project partners who are involved in this task, using software such as Balsamiq¹⁵. The use of Balsamiq is proposed but partners are allowed to use any related software they feel more comfortable with.
4. Follow-up interviews to take place for the validation of the proposed wireframes, with selected representatives of the targeted users' community (solution validation).

¹⁵ <https://balsamiq.com>

2.4. Translating features into requirements for services

The definition and description of the expected services will provide a number of requirements that are expected to shape and drive the development of the AGRERI Discovery Microsite services. The features of these services will need to be transformed into the corresponding requirements so that the partners of the project will work on and provide the services described by the users.

These features will be collected from the aforementioned means of requirements' elicitation, including the interviews, the online questionnaires and the workshops to be organized by the project partners for this purpose, and will have to be transformed into requirements to be used by the project's technical partners for the development of expected services.

3. Outcomes

3.1. Initial profiling of targeted personas

The initial profiling of personas, including content sources of interest and relevance to them as well as information on content-related challenges and problems that these personas face is probably the most important aspect of this work. The information extracted from these personas will have to be transformed to requirements that drive the development of the project's services.

3.1.1. Using online questionnaires

For the initial profiling of targeted personas, the use of online questionnaires is proposed, as they provide a free, efficient and quick mean for collecting requirements with only limited effort required for setting up the initial questionnaire (Figure 3). While any online survey tool can be used for acquiring requirements, Google Forms¹⁶ is suggested for this purpose; it is a free, widely used tool, allows automatic export of responses in a spreadsheet and easy processing in this form. In addition, a Google Form can be easily and collaboratively revised by team members working on a specific use case, adapted and reused with modifications for serving different needs, only requiring a valid Gmail address. However, the use of alternative tools such as LimeSurvey¹⁷ and SurveyMonkey¹⁸ may also be considered for this purpose.

What is important is that the responses collected through any online questionnaire should be exported and stored as a spreadsheet file that will allow the processing and analysis of the extraction of requirements.

Structure of the online questionnaire

An initial online questionnaire in the form of Google Form is available online¹⁹ and accessible to be used by the project partner responsible involved in the user requirements' elicitation.

¹⁶ <https://www.google.com/forms/about>

¹⁷ <https://www.limesurvey.org>

¹⁸ <https://www.surveymonkey.com>

¹⁹ https://docs.google.com/forms/d/10WQHSwvWsSapNzQDAhaAREXSYOHuNy7aM93ffyy_mW0/viewform

The image shows a partial view of an online questionnaire. At the top, there is a green header with faint icons. Below the header, the main title is "What is your interaction with AGRERI's website?". The text explains the purpose of the survey and asks for feedback. It includes a note that the input is important and asks for 5 minutes of time. A red asterisk indicates a required field. The section "Demographics" asks for the type of organization. A list of radio button options includes: Research Center, Research Institute, University Laboratory, Research & Development Dept. of a private company, Non-Profit Organization, Business Association, Cooperative, Chamber of Commerce, Central Government (e.g. Ministry), Regional/ Municipal Government, Farmer, I work as individual, and Other: [input field].

Figure 3. Partial view of the online questionnaire

Considerations regarding the online questionnaire

A number of additional aspects should be taken into consideration, in order to ensure the maximization of the questionnaire's outputs.

- For each use case, a minimum of fifteen (15) responses is expected. While this number refers to the total number of responses collected by any mean (e.g. including face-to-face or Skype/call interviews), it is important that a number of at least 15 responses be provided for each use case;
- The questionnaire can be used either as an online survey by sharing the public URL of the survey (not the internal one that allows full access to the form) or as a script for direct interviews, i.e. through Skype, phone calls or even face to face.
- The questionnaire can also be used for collecting preliminary feedback and requirements from registered participants of any of the events planned for collecting requirements. Then, participants may elaborate on their initial feedback through the interactive sessions of these events, which will also allow the validation of the personas and the feedback acquired through them.
- The online questionnaire may remain open / active for the collection of as many responses as possible. It can also be used for the collection of feedback on solutions and features at a later stage, when the scenarios for each use case runs.

3.1.2. Interviews

Another mean for the elicitation of requirements is through interviews. Interviews may be conducted with potential stakeholders, using the same set of questions included in the online questionnaire, in any of the following means:

- face to face
- through Skype
- through a phone call

Responses collected through the interviews should be recorded either using the online questionnaire (so that the same spreadsheet will contain all the responses) or in a spreadsheet that contains the same fields as the one provided by the online form, in order to ensure the homogeneity of the templates and facilitate the aggregation of responses from various means. A minimum of fifteen (15) interviews per use case are expected to provide sufficient feedback for the purposes of the task.

3.1.3. Using the Persona Graph

The Persona Graph consists of four (4) different boxes that need to be completed in detail:

1. Persona characteristics (demographics): Includes name, role in team, affiliation, research interests, field of expertise etc.
2. Data/content related requirements: What data/content types are of interest to the specific persona and which is this persona is currently using?
3. Key information challenges: Does this persona face any challenges related to identifying, accessing, retrieving and managing data/content of interest?
4. Features of the solution: What is the expected solution that would help the persona address the aforementioned challenges?

The following figure (Figure 4) provides an example of the information collected by a specific user in order to build the persona profile.


AGRERI use case: Researcher Persona	
 Name: Constantine I.	This profile includes all agricultural researchers (individuals or within organizations) that are looking for agricultural knowledge.
Job Position:	
Story:	His research program focuses on agribusiness economics and collective entrepreneurship—particularly the organizational and capital-acquisition aspects of innovative agribusiness cooperatives, and rural development—with attention to both theoretical and policy concerns.
Needs:	<ul style="list-style-type: none"> • seeks specific, clearly aggregated, indexed information • needs clearly organized and searchable content • looks for information from dispersed sources
Pains:	<ul style="list-style-type: none"> • difficulty to share his research outcomes • needs to combine multiple sources for research data • seeks for relevant information to his core research sector

Figure 4. Information about a persona to be transferred to a persona graph

It should be noted that during the collection of information for each persona, the importance of each need and challenge should also be mentioned and the corresponding input should be prioritized accordingly in the persona profile table. More specifically, the most important statement for each category should be at the top of the list while the least important one at the bottom of the same list.

3.2. Personas and usage scenarios

The following section contains a brief description of user personas and the usage scenarios that have been developed in dialogue with the target user community related to agricultural economics. Real use-cases examples are provided in order to better understand the personas and their needs. Personas and scenarios are listed in no particular order and there is some level of overlap in the scenarios, as in some cases, all personas can have the same needs and face the same problems, although from a different perspective. Something which is actually useful as it helped identify the common issues.

Persona A: Researcher

Costas is a researcher who investigates the social and cultural factors that have a significant impact on the creation, the development, the success or failure of agricultural cooperatives in Greece and worldwide.

Scenario

Constas is specifically interested in investigating how agricultural and farm cooperatives have evolved in Greece over the course of time and he wants to compare the Greek farm cooperatives to the more advanced farm cooperatives in the Netherlands and in the United Kingdom. To this end, he will have to access and study the archives and the sources of the past, such as newspapers, files and archives of local cooperatives and libraries, the proceedings of discussions in the Parliament when the relevant legislation was introduced for the first time and related material. Costas is not aware that all this material is located in a digital format in Europeana as he does not use the portal. Having enhanced the AGRERI website with a search interface that will allow users to discover and navigate relevant digital resources that are coming from large aggregators, as is the Europeana portal, Costas will be able to browse through resources related to farm cooperatives, coming from all over Europe simply by using keywords or the filter facets provided and save a lot of time and effort for his search. He will have access to texts and images that will help him compare the current status of farm cooperatives in Greece to that of other European countries and track their development through the years.

Persona B: Junior Researcher

Maria is a MSc student at the Agricultural University of Athens, in the Department of Agricultural Economics and Rural Development. She is now in the process to develop her MSc thesis which includes both bibliographical and experimental data.

Scenario

In the context of her thesis, Maria needs to find literature and references related to agricultural cooperatives from all over the world, in order to study the resources and reach meaningful conclusions. By comparing the statistical and economic data in these resources, she will be able to perform new analyses and present variable model analyses of the significance of agricultural cooperatives in the rural development of Greece. Having enhanced the AGRERI website with a search interface that will allow users to discover and navigate relevant digital resources that are coming from large aggregators, as is AGRIS and the AgEcon Search, Maria will be able to browse through publications and scholarly literature related agricultural economics and cooperatives, coming from all over Europe simply by using keywords or the filter facets provided in order to better conduct her literature review and proceed with her thesis.

3.3. From Personas to Tools

After analyzing the profile of the persona, the next step is a generic analysis of a scenario for the tool or service to be developed. Based on the persona and the described scenario, the following step is the creation of the first use case scenario, followed by the development of the first mock-ups, shown in the figure below (Figure 5). The mock-ups provide an illustration of what the final tool could look like based on the analysis of the scenario. Since the AGRERI Discovery Microsite is a technology that is embedded in an existing site, the first version of mock-ups included the AGRERI Library page where users can search the available resources. Additional buttons are included to quickly access predefined searches, specific resources etc.

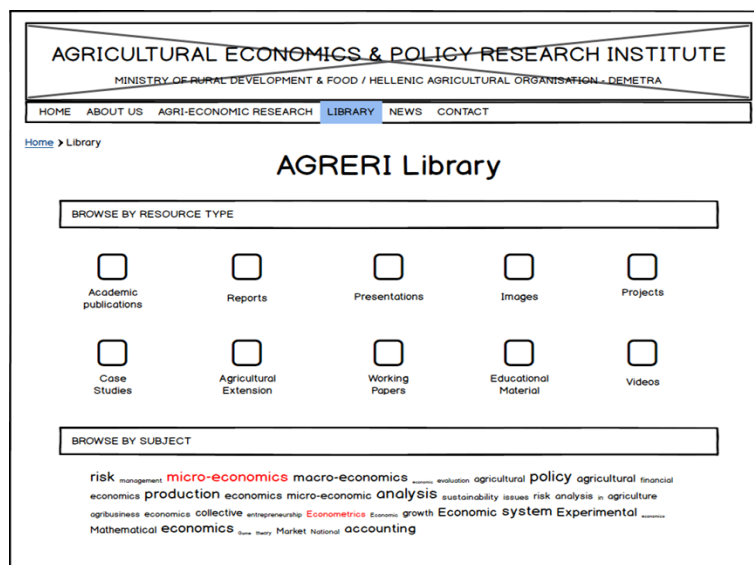


Figure 5. The first version of the mock-ups for the agricultural researchers scenario

This set of first mock-ups was distributed among the identified personas through meetings and workshops, in order to discuss them. The feedback gathered led to a second version of the mock-ups and to respective changes in the use case description. The second version of the mock-ups, shown in Figure 6, especially in the Library section, had the following additional features:

- A main page of the Library where users can browse through thematic areas but more importantly, browse for resources directly through the different repositories;
- A search result page with a faceted search to quickly filter the results per thematic area, per year, per author, per collection;
- When viewing a resource, the option to get related resources to view either from the local repository or from the external collections.

Social & Humanity Resources for Agricultural Researchers

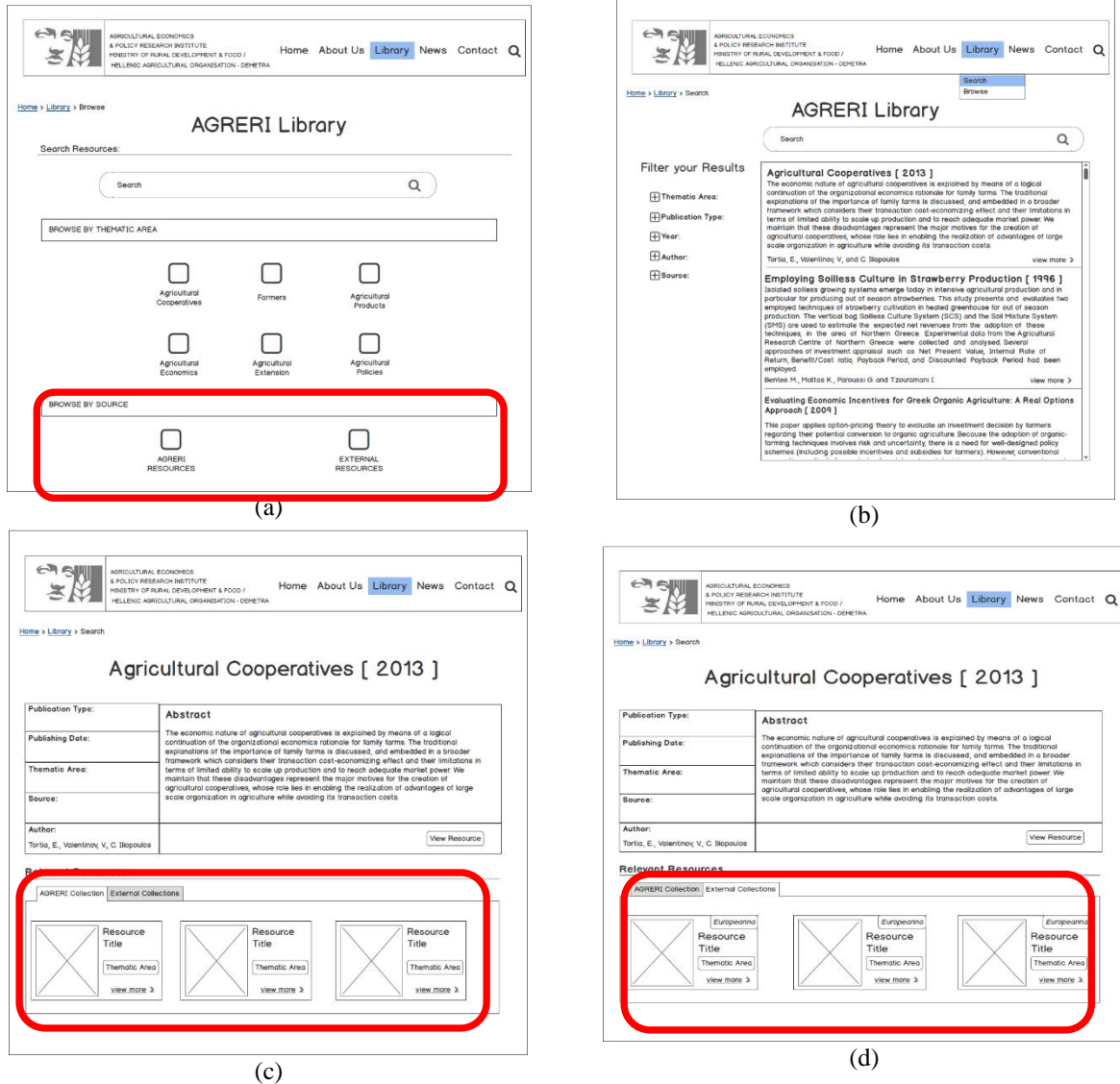


Figure 6. The second version of the mock-ups for agricultural researchers scenario with some of the changes highlighted

As explained earlier, feedback was continuously collected from users in order to update the use case scenarios and mock-ups until a desired outcome was reached and the pilot could be developed. Several meetings and interviews were arranged to discuss the AGRERI Discovery Microsite demo. The mock-ups were presented to the AGRERI team along with a number of questions to guide the users in providing feedback for the technical team through a presentation. When all the feedback was collected, the third and final version of the mock-ups was developed and discussed with members of the AGRERI team, with the change that the Library section was split in the “AGRERI Library” for the collections of the AGRERI team and the “External Libraries” for the collections of the external sources, in the first prototype was presented (Figures 7, 8).

After presenting the first prototype of the AGRERI Discovery Microsite to the AGRERI team, the agricultural researchers provided feedback on the tool. This feedback was considered for the next

iterations. The same procedure was adopted for the testing and evaluation of the service. Once the researchers had access to the tools of the mash-up, they were asked to provide expert feedback on whether and how these tools could support their research and workflow and if the Europeana collections together with the other external sources were useful to the researchers.

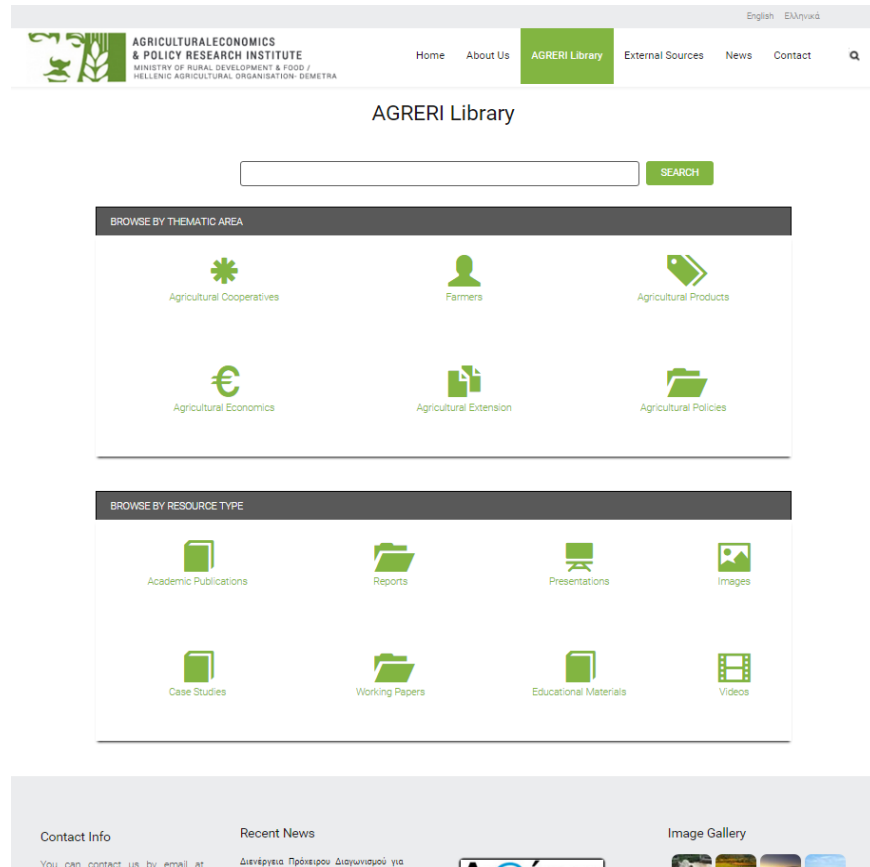


Figure 7. Screenshot from the AGRERI Library

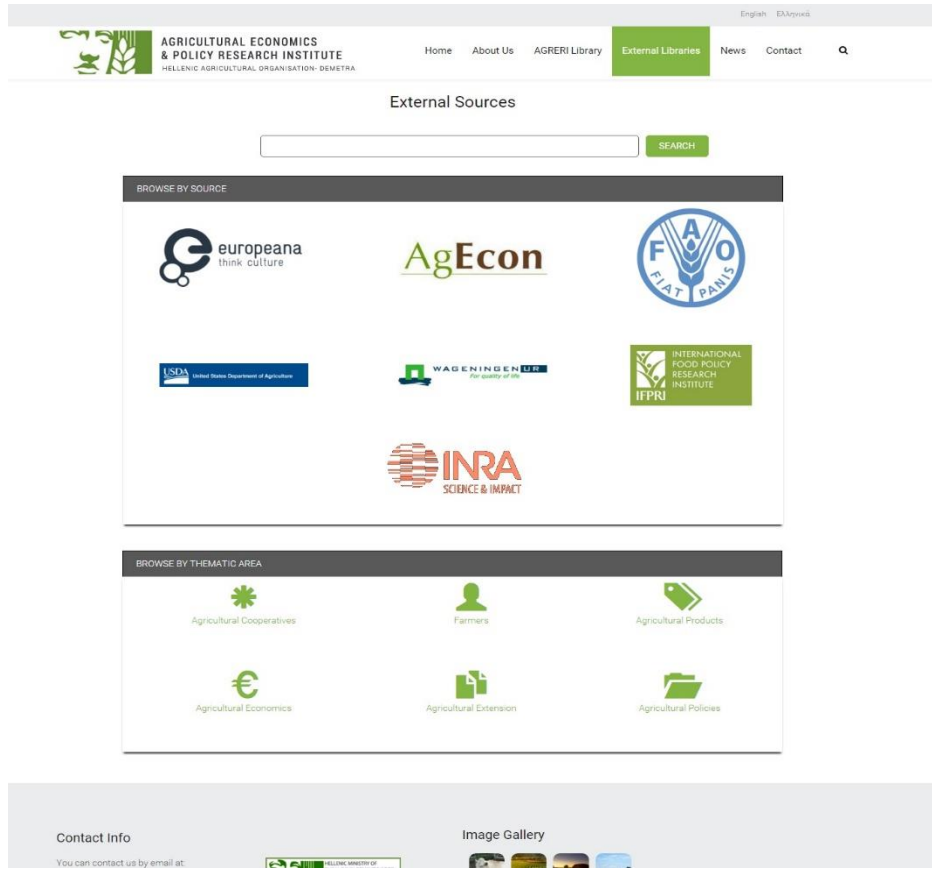


Figure 8. Screenshot from the External Libraries

3.4. Software and services

The AGRERI Discovery Microsite is a Microsite that can be used by researchers in the field of agricultural economics to search and discover relevant resources. The Microsite searches predefined collections of datasets based on user input and presents the results in a uniform way. It comes as a Microsite, built with lightweight web-technologies (HTML, CSS, HTTP, Javascript, AJAX) in order to be easily embedded in sites and web-applications, without the need to make changes for matching the existing technologies of the application. The first prototype of the Microsite is designed based on the needs of the AGRERI agricultural researchers (Figure 9).

The Microsite is a tool in the following two ways: it is graphically designed to be smoothly integrated with the web site of the AGRERI team²⁰ and is built on top of collections that have been requested by stakeholders.

The main usage of the AGRERI Discovery Microsite is a faceted search interface that allows users to search and quickly filter the results (Figure 10) and then, get a more detailed view of the

²⁰ <http://dev.agreri.agroknow.gr/>

resource selected with the related resources of the same thematic (Figure 11). In addition, predefined categories that allow access over specific content (i.e. agricultural economics) are also available.

The prototype has been designed and developed with the constant feedback from the AGRERI team in order to better capture and cover their needs. During the development phase, a number of discussions were organized between the group and technical team in order to gather feedback concerning which collections to search, which categories to use, and which facets the stakeholders would like to use.

In order for the finder to allow faceted search and uniform representation of the metadata from resources coming from different collections, the Microsite uses the existing AGRERI infrastructure to store a repository with all the metadata. In the pilot, the resources and different collections are stored in the repository and in all cases an API is used to filter thematically resources for the dataset. At the same time, and in order to provide this uniform representation and make the resources available throughout the Microsite a transformation process took place, where all metadata records were transformed from their original scheme to an internal format. During this transformation procedure metadata records have also been enriched.

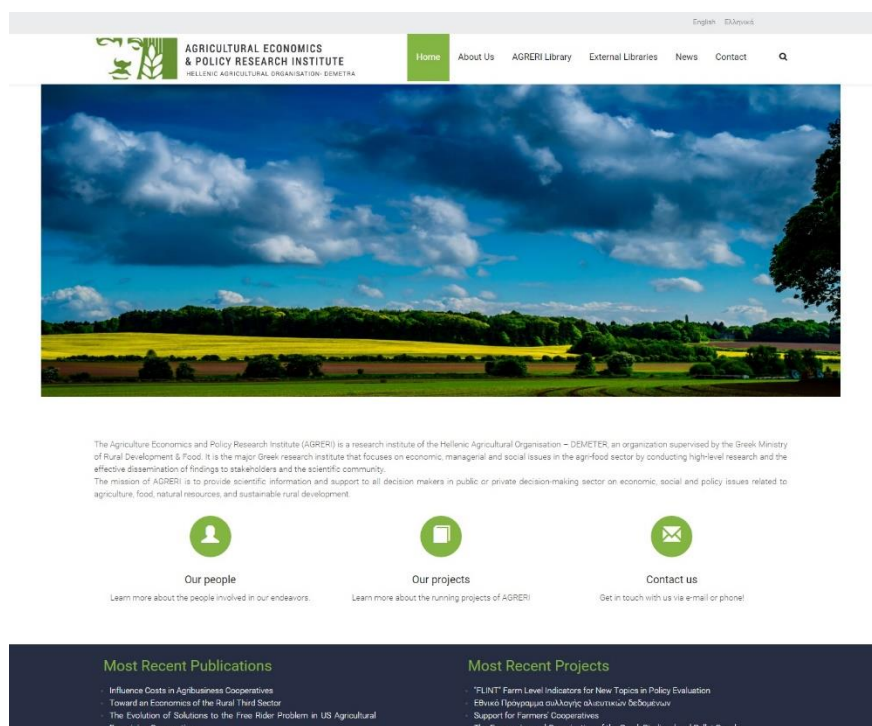


Figure 9. Screenshot from the AGRERI site

Collections in the AGRERI Discovery Microsite

In the AGRERI Discovery Microsite, there are many different collections stored. The AGRERI Library contains resources (publications, reports and other material) of the AGRERI researchers.

The External Libraries contain collections from external sources and this way these collections are made discoverable through the Microsite, as in the case of Europeana. As mentioned in previous sections, the Europeana portal contains and makes discoverable and accessible more than 40 million digital cultural artefacts from across Europe working with a large network of European museums, libraries and archives.

Other content collections of interest to the users available through the Microsite are:

Collections from the AgEcon Search²¹

The AgEcon Search is a free, open access repository of full-text scholarly literature in agricultural and applied economics, including working papers, conference papers and journal articles.

Collections from FAO AGRIS²²

AGRIS (International System for Agricultural Science and Technology) of the Food and Agriculture Organization of the United Nations (FAO) is a global public database providing access to bibliographic information on agricultural science and technology and its content is provided by participating institutions from all around the globe that form the network of AGRIS centers.

Collections from the United States Department of Agriculture (USDA)²³

The United States Department of Agriculture (USDA) is the federal executive department responsible for developing and executing federal government policy on farming, agriculture, forestry and food. It aims to meet the needs of farmers and ranchers, promote agricultural trade and production, work to assure food safety, protect natural resources, foster rural communities and end hunger in the United States and internationally.

Collections from the Wageningen University and Research Center (WUR)²⁴

The Wageningen UR Library contains publications, journals, textbooks, databases on life sciences and natural resources while it develops and maintains a, partly public, knowledge infrastructure contributing to the effectiveness of the research done and education provided by Wageningen UR.

Collections from the International Food Policy Research Institute (IFPRI)²⁵

The International Food Policy Research Institute (IFPRI) provides research-based policy solutions to reduce poverty and end hunger. Its Library captures, organizes and provides access to and exchange of IFPRI's research, through its knowledge repositories and academic networks.

²¹ <http://ageconsearch.umn.edu/>

²² <http://agris.fao.org/>

²³ <http://www.usda.gov/>

²⁴ <http://www.wageningenur.nl/>

²⁵ <http://www.ifpri.org/>

Collections from the National Institute of Agricultural Research (INRA)²⁶


The National Institute of Agricultural Research (INRA) is Europe's top agricultural research institute and the world's number two center for the agricultural sciences generating knowledge and know-how in the fields of agriculture, ecology, nutrition and food sciences.

The screenshot shows the website header for the Agricultural Economics & Policy Research Institute, with navigation links for Home, About Us, AGRERI Library, External Sources, News, and Contact. A search bar is located below the header. The search results are displayed in a list format, with each result showing the title, a snippet of the title, and the publication date. The results are filtered by source, publisher, source type, and thematic area.

Figure 10. Screenshot from results listing after a search query

²⁶ <http://www.inra.fr/>

[English](#) [Ελληνικά](#)



**AGRICULTURALECONOMICS
& POLICY RESEARCH INSTITUTE**
MINISTRY OF RURAL DEVELOPMENT & FOOD /
HELLENIC AGRICULTURAL ORGANISATION - DEMETRA

[Home](#) [About Us](#) [AGRERI Library](#) [External Sources](#) [News](#) [Contact](#) [🔍](#)

Cooperative Farm Unions: The Romanian Model, a Good Beginning

<p>Type: Text</p> <p>Year of publication: Submitted</p> <p>Source: Europeana</p>	<p>Description: An information bulletin produced by Radio Free Europe based on extensive monitoring of communist bloc news outlets and primarily intended for internal circulation. Pieces were generally contributed by the staff of the research and analysis departments and treated current affairs in the region.</p> <p>Authors: Radio Free Europe/Radio Liberty Research Institute (RFE/RL Research Institute)</p> <p>Related Resources</p> <div style="display: flex; justify-content: space-around;"> AGRERI Collections External Collections </div>
---	--

<p>Influence Costs in Agribusiness Cooperatives</p> <p>File: ILIPOULOS-HENDRIKSE-ISMO-2009.pdf</p> <p>Provider: AGRERI</p> <p>agrovoc: cooperatives agricultural cooperatives</p>	<p>Toward an Economics of the Rural Third Sector</p> <p>File: Iliopoulos and Valentinov-IJARGE-2009.pdf</p> <p>Provider: AGRERI</p> <p>agrovoc: socioeconomic development rural development agricultural development economics rural economics</p>	<p>The Evolution of Solutions to the Free Rider Problem in US Agricultural Bargaining Cooperatives</p> <p>File: Iliopoulos (in Rezitis-Editor)-2009.pdf</p> <p>Provider: AGRERI</p> <p>agrovoc: agricultural development economics agricultural economics cooperatives agricultural cooperatives</p>
---	--	--

← Back to search


Contact Info

You can contact us by email at info@agrer.gr

Recent News

Διενέργεια Προχείρου Διαγωνισμού για την προμήθεια κατηγορίας διαπίνης -

Image Gallery






Figure 11. Screenshot from view item with related resources

AGRERI Discovery Microsite architecture

The AGRERI Discovery Microsite is implemented using modules at three layers, namely a) the data Ingestion Layer, b) the repository layer and c) the front end layer (Figure 12).

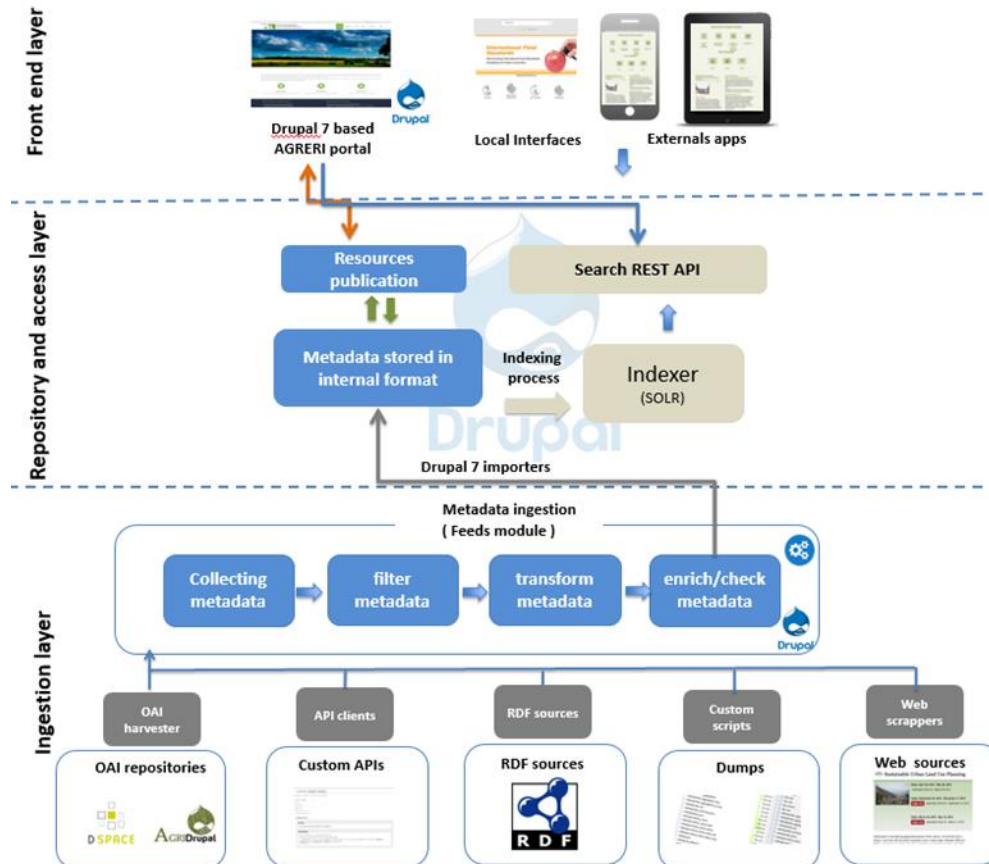


Figure 12. AGRERI Discovery Microsite architecture

More specifically, the following modules and components were deployed in the context of the Europeana Cloud project in order to implement the AGRERI Discovery Microsite.

- **REST API client** that collects relevant content from Europeana. A custom client has been developed to collect the relevant Europeana content in a JSON based format through Europeana's REST API.
- **Data ingestion module** that imports the metadata in AGRIS AP format to the AGRERI site. This module can be used by the researchers to import new relevant content from Europeana. The data ingestion module includes the following components

- **Transformation component** that transforms the data from Europeana format to AGRIS AP that is suitable for agricultural research
- **Enrichment component** that enriches the Europeana metadata with information about the thematic category of the content
- **Data indexing module** that indexes the metadata so they are available in the AGRERI Discovery Microsite.
- **Front end modules** (AGRERI Discovery Microsite) site that allows the discovery of the content from Europeana and connects this content to the existing content for agricultural economics that the AGRERI institution holds.

Ingestion (Data import)

The AGRERI Discovery Microsite offers a data ingestion mechanism. This data ingestion mechanism is responsible for importing both content and metadata records from multiple diverse sources that are publishing the metadata records through custom API (e.g. Europeana API), OAI-PMH protocol, RDF and dump files.

An importer can ingest content from raw files that are located to web apps directory or from a web API through Rest Services and to be imported into the AGRERI site. Also, before the data are stored in the database, the cleaning mechanism, if configured, removes unwanted content or curates it. For example, it can remove unwanted HTML characters. Another important functionality of the importer is the filtering mechanism that could be configured in order to filter out not relative content. For example, if a record is irrelevant to agriculture economics thematics, then this record will not be imported to the AGRERI Discovery Microsite.

In order to provide to the user all the resources of the AGRERI Discovery Microsite, the powerful Apache Solr 1.4 engine is used. Furthermore, the ingested content is indexed with Apache Solr and offered via this engine to the user search interface under a specific schema.

3.5. Evaluation report on integrated tools

The aim of the service developed was to support agricultural researchers investigating societal or cultural topics, who are using social science methods and tools. By working closely with the agricultural researchers of the Agriculture Economics and Policy Research Institute (AGRERI), we were able to understand their data requirements and how they could make use of digitized resources and archives available through Europeana and other sources.

Through interviews, an online survey and a workshop with these researchers, we were able to identify their needs and requirements as well as some core problems that they face when it comes to using search tools:

- problems with navigating and identifying relevant (digital) content from diverse sources
- the need to use specific facets and filters, in order to make it easier to discover resources related to specific themes

In order to help these researchers, the AGRERI Discovery Microsite was developed. It:

- searches predefined collections of datasets based on user input and presents the results in a uniform way
- is a faceted search interface that allows users to search and quickly filter the results
- allows the discovery of the content from Europeana and other sources, and connects this content to the existing content for agricultural economics

Evaluation with Agricultural Researchers

In initial meetings, members of the research community of agricultural researchers of AGRERI discussed with us their workflows, challenges and data requirements.

Several face-to-face meetings and an online survey took place between July 2015 and November 2015, in order to collect user requirements, feedback on the AGRERI Discovery Microsite and also evaluation on the final version of the Microsite, in the form of a hands-on workshop that took place on January 8th, 2016.

General discussion

The members of the AGRERI team had the chance to discuss and reflect on the AGRERI Discovery Microsite. They were very interested in the potential behind the Microsite which could potentially facilitate their work meeting their requirements. The problem they face when searching for resources, as in the case of researchers in almost all disciplines, is that they often end up browsing a lot of irrelevant results or have to use numerous different data sources in order to be able to search different collections. As a result, they often face a duplication of search results. Having a personalized tool such as the AGRERI Discovery Microsite, embedded in their site to search different collections from one point of entrance, would provide them with the ability to reduce the time spent on searching and browsing. They also find it useful to have a uniform way of viewing the metadata of the results, regardless of the initial provider, and value having the metadata in a cleaned format.

Workshops results

The workshop that was organised had the objective to provide a hands-on evaluation of the AGRERI Discovery Microsite not only by the agricultural researchers of the AGRERI team but also with real users interested in the tool. These users belong to different user categories than the five agricultural researchers of the AGRERI team mentioned below, such as agricultural advisors, students of agricultural sciences and farmers (who wished to remain anonymous). Fifteen people in total participated in the workshop that took place on the 8th of January 2016.

All participants were familiar with the AGRERI Discovery Microsite, but from different perspectives. After a briefing on the different personas, scenarios and use cases, participants were

asked to use the discovery Microsite with specific needs in mind, and later a discussion took place in order to evaluate the discovery Microsite in detail.

The users provided feedback on a number of things that could be adjusted to better cover their needs. Not all participants were familiar with the external sources of the Microsite, but all users liked the idea of having a Microsite where external sources can be included with only one query.

The need to filter search results per thematic is very important to them, as well as to filter search results per year. Also, a facet which would enable them to filter results based on the author of the resource could also be of use. Also, something that they find is quite essential in a discovery Microsite is to be able to get relevant resources like the one they were searching for (in the same thematic but also based on keywords), and these relevant resources to be presented beneath the resource selected.

Another comment that was made by the users is related to the type of resources that they are searching for. While most of them are typically searching for publications, they would also like to be able to get results of different content type (such as video, image etc.) which would provide them with a better picture of the available resources that exist in the various external sources.

Feedback was also received on whether the Europeana collections in particular were useful to the users. Not all participants were familiar with the Europeana portal and its collections, but all participants were impressed with the fact that the rich and diverse cultural heritage of European countries is available online through this portal. What is more, they were impressed by the fact that Europeana proved to be quite useful even for researchers in their own discipline, since they had not anticipated to find useful and related resources on agricultural economics and policy thematic areas. Specifically when it comes to accessing rare and useful content, like the one that Europeana portal provides.

4. Integration with AGINFRA

One of the things that we wanted to investigate through the Europeana Cloud project was to explore how Europeana Research²⁷ may be linked to the domain-specific research infrastructures of the agriculture community, through the research data e-infrastructure for agriculture and food (AGINFRA²⁸), in order to make relevant collections discoverable also through the CIARD RING²⁹ registry by both humans and software services, and see if and how the collections in Europeana can have value to agricultural social scientists.

We wanted to investigate whether there can be an automated dissemination mechanism by federating the Europeana and RING registries and how we can use the APIs and support the Europeana portal in order to offer access to Europeana collections through the domain-specific data e-infrastructure of AGINFRA. The process on how a user can register collections from the Europeana portal into the CIARD RING is explained below, together with the issues that may arise through this process and the recommendations that can be made.

4.1. Europeana portal

The Europeana portal is online accessible at <http://europeana.eu/portal/>. The main search box is used to browse through collections by using specific keywords based on the user's needs (Figure 13). The search results appear as a list together with the facets available (Figure 14). The search results have a URL that will be used later to register to the CIARD RING. If the user clicks on one of the listed results, he/she is directed to the page where the metadata of the item are listed (Figure 15).

²⁷ <http://research.europeana.eu/>

²⁸ <http://aginfra.eu/>

²⁹ <http://ring.ciard.info/>

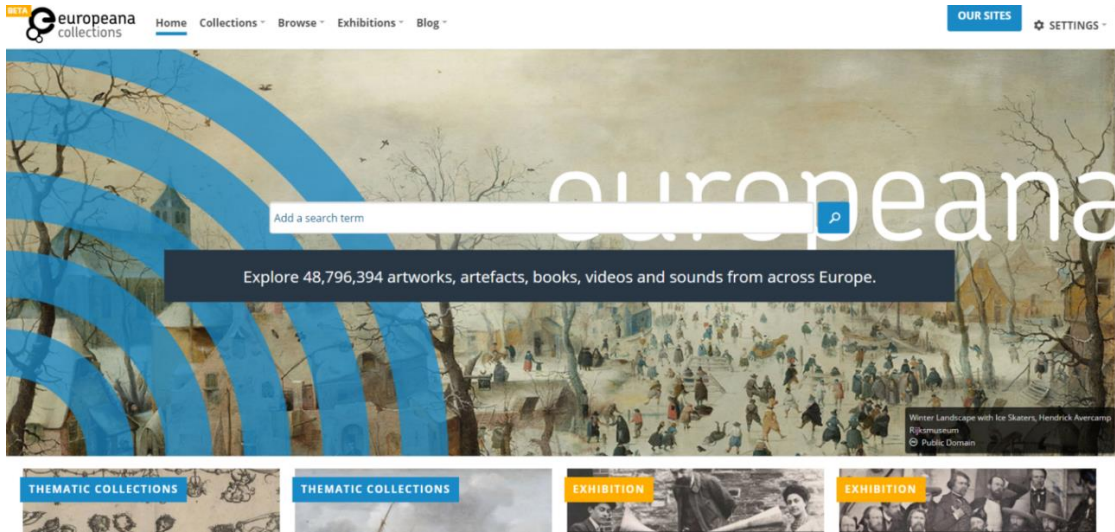


Figure 13. Screenshot from the Europeana portal

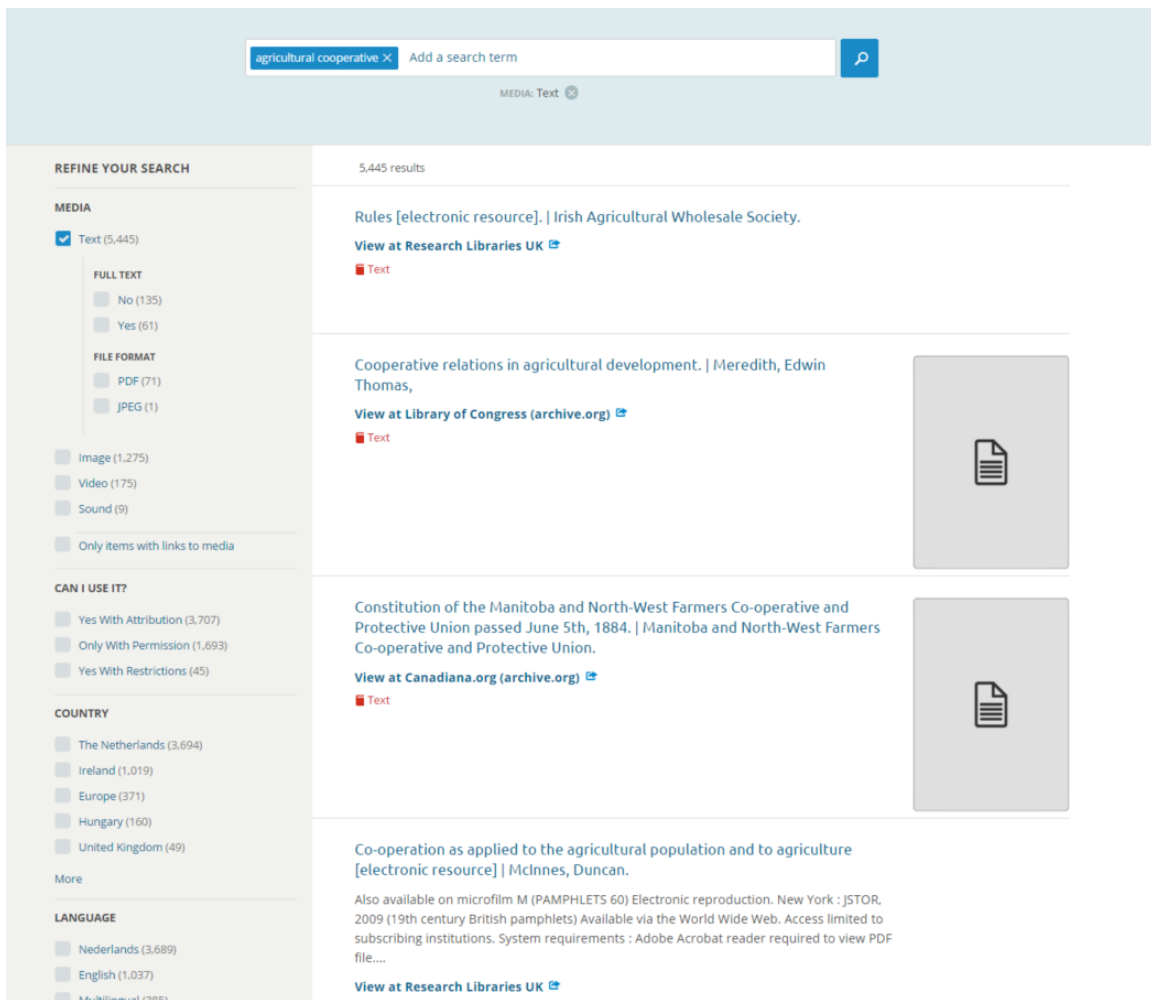


Figure 14. Screenshot from the Europeana portal listing results after a search query

The screenshot shows the Europeana portal interface. At the top, there are navigation links for 'rowse', 'Exhibitions', and 'Blog'. Below this is a large dark grey area with a document icon and a 'DOWNLOAD' button. The main content area is divided into several sections:

- Title:** Cooperative relations in agricultural development. | Meredith, Edwin Thomas.
- People:** Creator: Meredith, Edwin Thomas.
- Classifications:** Type: Text; Subject: Agricultural colleges, [from old catalog], Dept. of Agriculture, United States.
- Properties:** Language: English.
- Time:** Creation Date: 1920, 20th century CE/1920s; First Issued: 1920, 20th century CE/1920s.
- Provenance:** Publisher: Govt. print. off., Washington; Identifier: oai:eu.bhl:BHLUS:ITEMS/000000089491, DLC:agr:20001900; Institution: Library of Congress (archive.org); Provider: BHL Europe; Providing Country: europe; First Published in Europeana: 1970-01-01; Last Updated in Europeana: 1970-01-01.
- Copyright:** Rights: <http://www.europeana.eu/rights/rr-f/>.

Additional information on the right side includes 'FIND OUT MORE' (linking to the Library of Congress Archive), 'CAN I USE IT?' (indicating 'Rights reserved - Free access'), and 'SHARE' options for Facebook, Twitter, and YouTube.

Figure 15. Screenshot from the Europeana portal view item with the metadata

4.2. The AGINFRA data infrastructure for agricultural research

4.2.1. The agAGINFRA FP7 project

The agINFRA FP7 project³⁰ was an Integrated Infrastructure Initiative (I3) project that introduced the agricultural scientific communities into the vision of open and participatory data-intensive science. In particular, agINFRA aimed at designing and developing a scientific data infrastructure for agricultural sciences to facilitate the development of policies and the deployment of services that promote sharing of data among agricultural scientists and develop trust within and among their communities. agINFRA worked towards eliminating existing obstacles concerning the open access to scientific information and data in agriculture and improved the preparedness of agricultural scientific communities to face, manage and exploit the abundance of relevant data that is (or will be) available and can support agricultural research.

³⁰ <http://www.aginfra.eu/project>

During the lifetime of the project, agINFRA worked towards the following aspects:

- shared e-infrastructure, tools and services: agINFRA developed and made available the shared e-infrastructure required for agricultural research resources (content/data) and services;
- higher interoperability of data: agINFRA promoted a higher level of interoperability between agricultural and other data resources (e.g. through deploying a Linked Agricultural Data Layer);
- improved research data services: agINFRA allowed for service improvement so that agricultural researchers can produce and transfer novel scientific and technological results for effective outcomes in the agricultural sector.

4.2.2. The AGINFRA global hub of agricultural research

After the end of the agINFRA project, and with contributions from various key organizations at a global level, agINFRA evolved into a global hub of research and extension information related to agriculture and food security³¹. The AGINFRA global hub consists of the following, among others:

- a global atlas of agricultural research & extension (including institutions, people, publications, data sets, projects, courses, OERs);
- a semantic layer of processing, enriching & interlinking research information from distributed, heterogeneous sources & formats;
- a catalogue of software components (open source software stack & APIs) that anyone may use to process research information;
- a help desk service to support institutions & projects that wish to publish their research information openly;
- a set of data-rich service and application demonstrators for specific case studies, such as food safety, vitis, crop composition etc.

The following figure (Figure 16) provides an overview of AGINFRA as a global agri-food research data hub.

³¹ A. Drakos, V. Protonotarios, N. Manouselis, "agINFRA: a research data hub for agriculture, food and the environment", <http://f1000r.es/593>] F1000 Research, 4:127 (doi: 10.12688/f1000research.6349.1), 2015.

AGINFRA as a global agri-food research & extension atlas

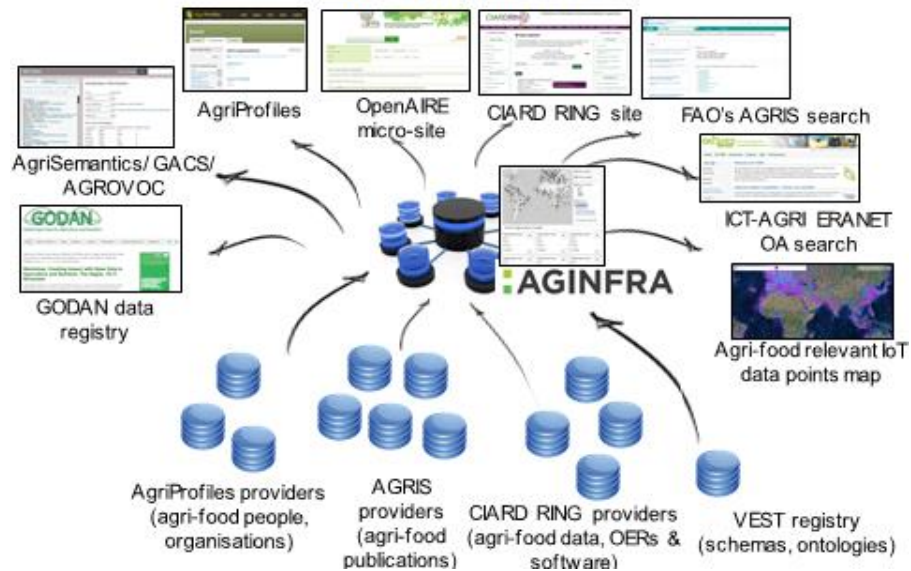


Figure 16. The AGINFRA global hub of agri-food research

AGINFRA aims to facilitate the work of various stakeholders in the agri-food research context and at a global level by identifying, describing and enhancing the accessibility of different components that are substantial for the agri-food research community.

4.3. CIARD-RING

The CIARD Routemap to Information Nodes and Gateways (CIARD RING)³² is a project implemented within the Coherence in Information for Agricultural Research for Development (CIARD)³³; <http://www.ciard.info>) initiative and is led by the Global Forum on Agricultural Research (GFAR)³⁴. The CIARD-RING is a core component of the AGINFRA e-infrastructure.

The RING is a global directory of web-based information services and datasets for agricultural research for development (ARD). It is the principal tool created through the CIARD initiative to

³² <http://ring.ciard.info>

³³ <http://www.ciard.info/>

³⁴ <http://www.gfar.net>

allow information providers to register their services and datasets in various categories and so facilitate the discovery of sources of agriculture-related information across the world.

The RING aims to provide an infrastructure to improve the accessibility of the outputs of agricultural research and of information relevant to ARD management. The functions of the RING consist of:

- providing a map of accessible information sources with instructions on how they can be used effectively;
- providing a dataset sharing platform for agriculture;
- providing examples of services that show good practices on implementing “interoperability”;
- clarifying the level and mode of interoperability of information sources;
- providing instructions for building enhanced integrated services that repackage information in different ways.

In the context of the AGINFRA, the RING provides a machine-readable Linked Data layer to meet the following requirements:

- Datasets registered in the RING have to be found by applications;
- Applications have to be able to read all the metadata about datasets and filter datasets according to their needs
- Applications have to find enough technical metadata in the RING to:
 - Identify datasets with a specific coverage (type of data, thematic coverage, geographic coverage);
 - Identify datasets that comply with certain technical specifications (format, protocol etc.);
 - Access the dataset and get the data;

This machine-readable layer can for instance support the data aggregation workflows of external services.

The CIARD-RING is online accessible at <http://ring.ciard.info/> (Figure 17).

The screenshot shows the CIARD RING website homepage. At the top, there is a navigation bar with language options (English, 简体中文, Français, Русский, Español) and user actions (Login, Register). The main content area is divided into several sections:


- What is the RING:** A directory of information services and datasets in agriculture. It states that all providers of agricultural information are invited to register their services and datasets on this portal.
- New services:** Total: 1103. Examples include Amazonian Biota (Federal University of Amapa, Network: AGRIS, Added on: 28/01/2016 - 4:43pm) and Technical Centre for Agricultural and Rural Cooperation (Technical Centre for Agricultural and Rural Cooperation, Network: AGRIS, Agriculture - General/All Documents (DLIOs), Added on: 27/01/2016 - 12:05pm).
- New datasets:** Total: 309. Examples include ISI-MIP (Wageningen University and Research Centre, Added on: 29/11/2015 - 10:01pm) and AgMIP (Wageningen University and Research Centre, Added on: 29/11/2015 - 10:00pm).
- New data model:** A diagram showing the flow of data from Global Catalog to Global Database, which is linked to Global Catalog, Global Database, and Global Database.
- New providers:** Total: 561. Example: Technical Centre for Agricultural and Rural Cooperation.
- Distribution by domain:** A bar chart showing the distribution of services and datasets across various domains: Engineering, Technology and Research; Education and Extension; Information Management; Plant Production and Protection; and Forestry.

Figure 17. Screenshot from the CIARD RING site

4.3.1. Register to CIARD RING

In order to register services and datasets in the RING, a user must first register to the service. By creating an account (Figure 18), a confirmation e-mail is sent to the user's email address (Figure 19). After the e-mail validation (Figure 20), the user can register a service or dataset (Figure 21).

EN Español Login Register



A directory of information services and datasets in agriculture

Home All info services Datasets Software Participants Open AGRigate Networks Indexing criteria About How to Linked Data agINFRA

You are not logged in.

First name *

Last name *

Username *

Spaces are allowed; punctuation is not allowed except for periods, hyphens, apostrophes, and underscores.

E-mail address *

A valid e-mail address. All e-mails from the system will be sent to this address. The e-mail address is not made public and will only be used if you wish to receive a new password or wish to receive certain news or notifications by e-mail.

Confirm e-mail address *

Please re-type your e-mail address to confirm it is accurate.

Password *

 Password strength:

Confirm password *

Provide a password for the new account in both fields. Password must be at least 5 characters.

Country *

- Select a value -

Organization Show row weights

Figure 18.. Screenshot from the CIARD RING registration process

From ring-helpdesk@ciard.net

Subject Account details for [redacted] at The CIARD R.I.N.G.

To [\[redacted\]](mailto:[redacted])

[redacted],

Thank you for registering at The CIARD R.I.N.G..

Please validate your account by confirming your email address: just click on the following link or paste it to your browser:
http://ring.ciard.info/user/validate/19011/1453477140/R8xuEhHgH1JSNw7BK7A2GT6-T_pBgXcc1S5czYE-4

In the future, you can login at:
<http://ring.ciard.info/user>

username: [redacted]
 password: Your password

-- The CIARD R.I.N.G. team

Figure 19. Screenshot from the CIARD RING confirmation e-mail

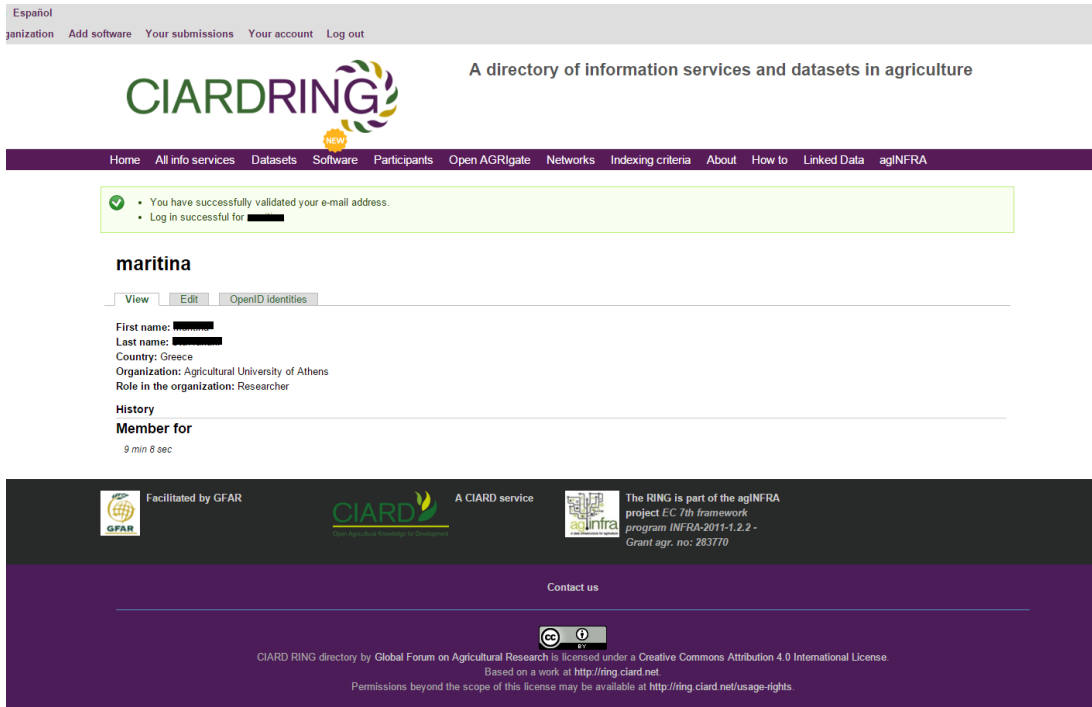


Figure 20. Screenshot from the CIARD RING e-mail validation

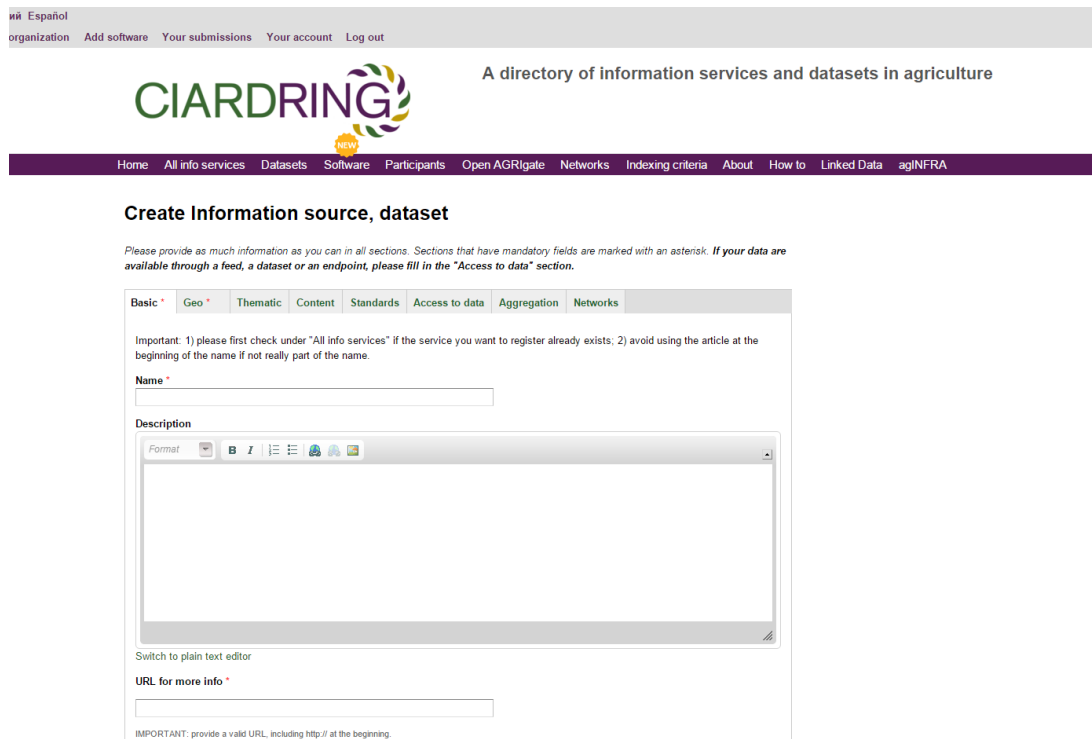


Figure 21. Screenshot from the CIARD RING service/dataset registration

4.3.2. Register a service or dataset

After registering to CIARD RING, the user can create an information source or dataset. In our case, we wanted to register a search query from Europeana using the keywords “agricultural cooperative” with the URL <http://www.europeana.eu/portal/search?q=agricultural+cooperative> (Figure 22).

Figure 22. Screenshot from Europeana search query using the keywords “agricultural cooperative”

The user must complete all fields marked with an asterisk (*) in every tab and section and any other fields suggested (Figure 23). These include:

- Basic: name, description, URL, responsible body, language etc.
- Geo: location country, geographic coverage.
- Thematic: domain, other topics, specific topics (AGROVOC) etc.
- Content: type of service, languages covered, audience, open access mandate etc.
- Standards: metadata set(s) used, KOSs used etc.
- Access to data: distribution(s), special instructions etc.

- Aggregation: harvesting protocol, required format etc.
- Networks: belongs to networks etc.

Figure 23. Screenshot from the CIARD RING create information source, dataset

After completing all the necessary fields, it becomes apparent how a simple initial search query in Europeana using specific keywords (i.e. agricultural cooperative) can become registered in the CIARD RING (Figure 24). This way, the content which exists in Europeana on agricultural cooperatives becomes available as a research dataset through AGINFRA and therefore, through AGINFRA-powered services like AGRIS where domain-specific and not only researchers can find it. Something which is extremely important taking into consideration the fact that AGRIS has more than 200.000 visitors per month, and thus enhancing the visibility and re-use of aggregators like the Europeana portal.

The screenshot shows the CIARD RING website interface. At the top, there is a navigation bar with links for 'our submissions', 'Your account', and 'Log out'. Below this is the CIARD RING logo and the tagline 'A directory of information services and datasets in agriculture'. A secondary navigation bar contains links for 'Home', 'All info services', 'Datasets', 'Software', 'Participants', 'Open AGRgate', 'Networks', 'Indexing criteria', 'About', 'How to', 'Linked Data', and 'agINFRA'.

The main content area is titled 'Agriculture Cooperative Text collection'. It includes a search bar with 'View' and 'Edit' options. Below the search bar, there is a section for 'Owner' (Agricultural University of Athens) and a description of the collection. To the right, there is a 'Dataset distributions' section with a note that no distributions are available for this service.

The 'General' section provides details such as 'Type: Local search engine', 'URL for more info: Link', 'Started in year: 2016', and 'Contacts: Email: Contact person by email'. The 'Content' section lists 'Domain: Agriculture - General/All', 'Specific topic (AGROVOC): agricultural cooperatives', and 'Currency: Current / updated'.

The 'Geographic' section indicates 'Geographic coverage: Global' and 'Country where the service is managed: Greece'. It includes a map of Europe with Greece highlighted in red. Below the map, it specifies 'Region: Europe' and 'GFAR region: Europe'. At the bottom of the page, there is a footer with logos for 'Facilitated by GFAR', 'A CIARD service', and 'The RING is part of the agINFRA project EC 7th framework program INFRA-2011-1.2.2 - Grant agreement: 303719'.

Figure 24. Screenshot from the CIARD RING created information source, dataset

4.4. Recommendations on establishing a liaison between Europeana and AGINFRA

The information provided in the previous sections highlighted the potential of the linking between Europeana and domain-specific data registries like the CIARD RING of AGINFRA by providing a specific example with the steps that a user needs to follow in order to register a possible dataset. Through this process, a number of limitations occurred which led to some recommendations in order to establish a more stable liaison and content exchange and discovery mechanism to help AGINFRA communities find collections of relevance and of interest in Europeana.

Indeed, Europeana provides access to content of interest to agricultural researchers so this content needs to be more easily accessible to them; however, Europeana does not include domain specific collections (e.g. ones on agricultural topics). In this context, new, alternative and innovative ways that will allow sharing of content of interest among researchers and other types of stakeholders will not only facilitate the sharing of high-quality content but at the same time it will significantly enhance the reuse of the Europeana content and its use in different domain specific applications. In the specific case described in the previous sections, a new collection is created based on the search results for a specific query.

This envisaged integration of services like CIARD RING in the specific use case presented in the document, as well as domain specific services from other domains, will require the consideration of a number of issues that were identified during the analysis presented earlier. Such issues are the following:

- Definition of the person responsible for the collection, so that it can be attributed and contacted when needed.
- Identification of the owner of the collection, so that he/she can be contacted for use and licensing issues.
- While the process can be completed by any user, a user with basic knowledge on metadata, licensing etc. may more easily complete the fields.
- Each URL is a dynamic list of results that are compiled after the query. Any revision in the content of the Europeana collections will result in the corresponding changes in the search results.
- A user has to register his/her organization if it doesn't already exist in CIARD RING. This might be an issue in the case of users that do not represent an organization but act as individual users.

In this context, addressing the aforementioned issues can provide a set of recommendations that will facilitate the linking of Europeana with domain specific services, like CIARD RING, as similar issues may have to be addressed in other contexts as well. This linking may be further enhanced by direct communication between the technical team of the Europeana and technical teams of other services that want to be linked to Europeana. This will ensure the optimal technical interoperability, in cases where automatic metadata and content sharing takes place. Last but not least, the definition of ownership over collections, like the one mentioned earlier in the case of CIARD, as well as the allowed uses through a well-defined licensing schema will further facilitate the use and reuse of Europeana content through third party services.

5. Conclusion

This White Paper aimed to provide a generic, high level blueprint to anyone interested in discovering information from heterogeneous data at disperse sources. The specific piloted case study, which was carried out within the Europeana Cloud project and was described in this document, presented how Europeana may be used as a resource for Social Scientists and researchers that work in different scientific domains and disciplines, in this case in the domain of agricultural economics, how they can make use of the available resources and data sets and how they can be benefited from such digital infrastructures.

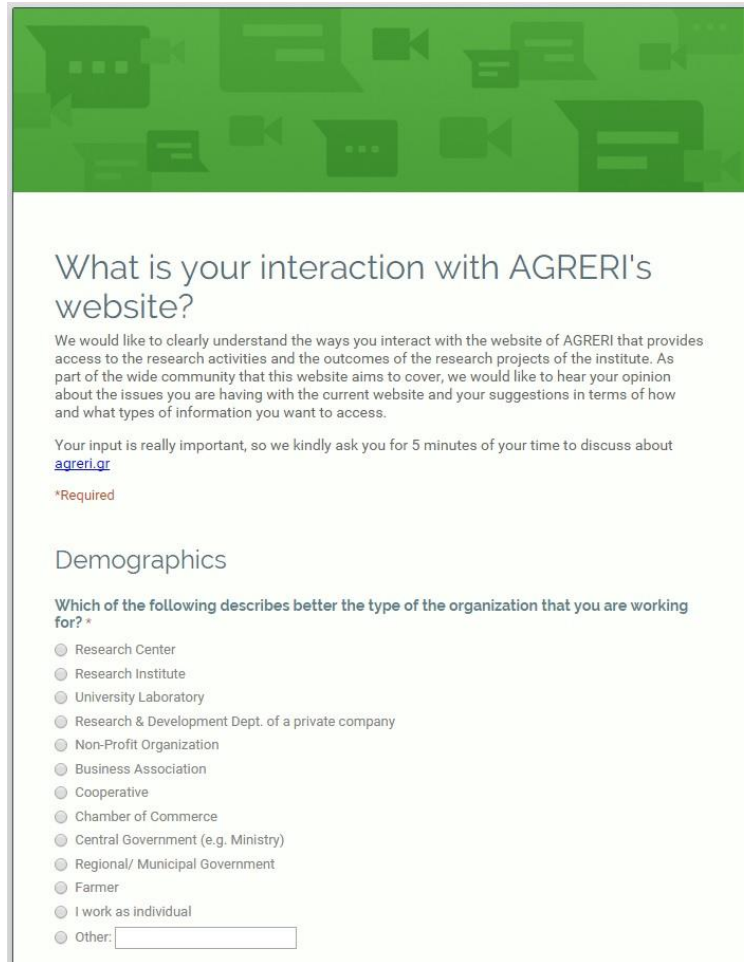
The methodology followed for the elicitation of requirements in this piloted case study, which was described in this document, could also be used for the elicitation of requirements from targeted end-users in other case studies. The methodology consists of a number of well-defined steps that should be completed for the successful acquisition of requirements, as they were presented in this document.

At the same time, a potential connection between Europeana and the domain-specific research data e-infrastructure AGINFRA was investigated, in order to examine how to make the relevant collections discoverable through AGINFRA's specific registries and services. This will allow to further facilitate the sharing of high-quality content and at the same time it will significantly enhance the reuse of the Europeana content and its use in different domain specific applications.

6. Annex A: Online questionnaire for the AGRERI case

Available at:

https://docs.google.com/forms/d/10WQHSwvWsSapNzQDAhaAREXSYOHuNy7aM93ffyy_mW0/viewform



The image shows a screenshot of a Google Forms questionnaire. At the top, there is a green header with various icons representing communication and research. The main title of the form is "What is your interaction with AGRERI's website?". Below the title, there is a paragraph explaining the purpose of the survey: to understand how users interact with the AGRERI website and to gather feedback on the current website and suggestions for improvement. A note states that the input is important and asks for a 5-minute time commitment. A red asterisk indicates that the following question is required. The question is "Which of the following describes better the type of the organization that you are working for? *". Below this question is a list of radio button options: Research Center, Research Institute, University Laboratory, Research & Development Dept. of a private company, Non-Profit Organization, Business Association, Cooperative, Chamber of Commerce, Central Government (e.g. Ministry), Regional/ Municipal Government, Farmer, I work as individual, and Other: [text input field].

What is your interaction with AGRERI's website?

We would like to clearly understand the ways you interact with the website of AGRERI that provides access to the research activities and the outcomes of the research projects of the institute. As part of the wide community that this website aims to cover, we would like to hear your opinion about the issues you are having with the current website and your suggestions in terms of how and what types of information you want to access.

Your input is really important, so we kindly ask you for 5 minutes of your time to discuss about ageri.gr

***Required**

Demographics

Which of the following describes better the type of the organization that you are working for? *

- Research Center
- Research Institute
- University Laboratory
- Research & Development Dept. of a private company
- Non-Profit Organization
- Business Association
- Cooperative
- Chamber of Commerce
- Central Government (e.g. Ministry)
- Regional/ Municipal Government
- Farmer
- I work as individual
- Other:

In which thematic are you activated? *

- Agricultural Economics
- Crop production
- Sustainable agriculture
- Organic agriculture
- Aquaculture & Marine biology
- Food science
- Food processing & packaging
- Biodiversity
- Management of natural resources
- Agricultural engineering
- Animal husbandry
- Natural sciences (e.g. Biology, Chemistry)
- Environmental sciences
- Nutrition
- Social Sciences
- Humanities
- Other:

What is your role in the team? *

- (Full/ Associate/ Assistant) Professor
- Lecturer
- Researcher
- Post-doctoral researcher
- PhD candidate
- Special technical laboratory staff
- Laboratory teaching staff
- Advisor
- Scientific Associate
- Agricultural Extension Officer
- Farmer
- Civil Servant
- Other:

Which is the main funding source of your team? *

- National / State funds
- EU funds
- Private sector / Industry (directly)
- Other:

Your interaction with agreri.gr

What is the main reason for visiting agreri.gr? *

- Find news from the agricultural economics sector
- Search for new publications
- Search for case studies
- Search for events in the agricultural economics sector
- Find new calls / announcements
- Search for statistics
- Search for new research outcomes
- Job announcements
- Farmers' Media Library
- Explore new ideas
- Other:

Please mention here the content types that you are interested to get access through the AGREER website: *

- Publications
- Economical data
- Statistical data
- Announcements/ Calls
- Farmers' media library
- Researchers' profiles
- Organisations profiles
- Forthcoming Events & Seminars
- Reports
- Presentations
- Journals / Articles
- Case studies
- Digitalised newspapers of EU regulation
- Cooperatives' Announcements
- Other:

What is the most preferred filter for the search service of the website? *

- Content type (e.g. publication)
- Author
- Year
- Keywords
- Thematic Area
- Region
- Other:

Do you have a specific website / repository to present the progress of your research activities, researchers' profiles and the research outcomes? *

- None
- Institution website
- Laboratory website
- Project website
- Other:

How do you update the existing institution / laboratory / project website with the research outcomes of your team / project? *

- Manually by myself
- Someone else from my team
- The knowledge management team of my institute
- I don't know
- Other:

Challenges

Which of the following are the most important challenges that should be faced in order to give you access to the most updated and reliable information in agricultural economics research? *

- Difficulty in discovering, tracking and monitoring the research outcomes (publications and reposts) of researchers in the agricultural economics sector
- Difficulty in discovering early enough the scientific publications that researchers in agricultural economics are working on (draft version)
- Difficulty in finding and downloading the raw (primary experimentation) data for a specific agricultural economics-related topic
- Difficulty in finding and downloading the processed data (i.e. graphs) for a specific agricultural economics-related topic
- Difficulty in finding the core research teams in the agricultural economics sector
- Difficulty in discovering the major research projects and research initiatives in the area of agricultural economics
- Difficulty in connecting your laboratory / project website with other repositories and websites (e.g. FAO AGRIS) for the dissemination of the research activities of the team / project
- Lack of time to discover all appropriate dissemination channels for your researchers' publications, datasets, activities or other outcomes
- Time consuming to continuously disseminate the scientific publications and research outcomes of your project / team
- Other:

Suggestions

If the AGRERI website can provide you access to the most updated information about the research activities of the institute and the related content from the research community in agricultural economics, which of the following features should be included: *

- To search and find publications based on the researchers' names for a specific agricultural economics topic
- To search and find all the relevant publications according to agricultural economics topic
- To get recommendations, according to specific terms I am searching
- To be updated when other researchers start working on publication(s) or on experiment(s)
- To discover automatically if other researchers work on a publication or experiment by looking on social media (blog, Slideshare, Twitter, Facebook)
- To discover automatically if other researchers have presented or published something through their Slideshare profile
- To have a visualized map with all the research institutions, researchers and research projects in the area of agricultural economics
- To explore researchers, research institutes and research projects & initiatives that are relevant in agricultural economics research
- To get access to raw (i.e. excel files) data for a specific agricultural economics topics
- To get access to processed data for a specific agricultural economics topic
- To have access to figures from the publications and to be able to find them using a search mechanism
- To have access to more multimedia content (photos, videos)
- To have access to more data regarding the activities of Agricultural Cooperatives
- To explore more relevant publications from external sources
- To have access to more case studies & relevant legislation on European Agricultural Cooperatives
- Other:

Never submit passwords through Google Forms.