How to Acquire Scientific Knowledge for University to Industry Knowledge Transfer

Ioana Suciu, Christophe Fernandez, Amadou Ndiaye
Institut National de la Recherche Agronomique,
INRA UMR 927 (Université Bordeaux 1 – CNRS – INRA)
Bordeaux, France
ioana.suciu@bordeaux.inra.fr,
christophe.fernandez@bordeaux.inra.fr,
amadou.ndiaye@bordeaux.inra.fr

Benoit Le Blanc
Institut Polytechnique de Bordeaux, ENSC
Bordeaux, France
benoit.leblanc@ensc.fr

Catriona Raboutet
Real Cognitic
Bordeaux, France
catriona.raboutet@gmail.com

Abstract - **This paper presents a framework dedicated to support the acquisition of useful and usable scientific knowledge that can be transferred to industrialists. It exposes the seven corpora of knowledge (Initial data, Problem, Hypothesis, Tests, Results, Interpretation and Conclusion) necessary both to acquire and to structure relevant scientific knowledge. The framework is used in the domain of design of tailored biscuits with optimized satiety benefit. The acquired scientific knowledge will be represented and next implemented in an electronic knowledge book that is the channel chosen to transfer it to biscuit-making industrialists.**

Keywords-scientific knowledge acquisition; university-to-industry knowledge transfer; knowledge sheets; electronic k-book.

I. INTRODUCTION

Our work is situated in the field of university to industry knowledge transfer (U2I-KT). The objective of the U2I-KT research is to make the knowledge from scientific results available in a format that is directly accessible and exploitable by industrialists. However, scientific knowledge produced by academic research is intended to scientific communities for an academic use. It is commonly made available through scientific publications, which are not easy to be interpreted and used by industrialists. The difficulty comes from the difference that researchers produce scientific knowledge with the aim of understanding problematic phenomena, whereas industrialists demand this knowledge to manufacture their products. The challenge is to reformulate scientific knowledge created in university with respect to the industrial process of product manufacturing. The U2I knowledge reformulation issue will introduce new ways to structure, to access and to handle useful scientific knowledge, all oriented to its final industrial use.

The U2I reformulation program requests a knowledge base (k-base) of relevant and useful scientific knowledge that should be constituted and made available in a form that eases its future treatment (easy access to knowledge should be envisaged, for example). Scientific knowledge should be captured, modelled, structured and represented within the k-base and then prepared for reformulation (re-structured) with respect to a configuration properly shaped for its industrial use. An appropriate-designed tool should be conceived and implemented to easily supply the scientific knowledge formerly reconfigured. It will be exploited to transmit this knowledge to the industrialists in order to be assimilated and used by them. An electronic knowledge book (e-k-book) is the tool chosen here for the U2I-KT. It is a hypermedia electronic document composed by a set of hierarchical concept maps and knowledge sheets (see Section V). The e-k-book has the principal advantage of providing to industrialists direct access and immediate use of knowledge within the k-base. It can be easily handled as being produced on, published through, and readable on a computational form.

The work presented in this paper deals with the acquisition of scientific knowledge. The framework proposed here to *A*cquire *s*cientific *K*nowledge (AsK) is intended to support the k-base constitution and its structuring from a scientific-oriented perspective. Further work will be carried out on scientific knowledge representation, its industry-oriented reformulation and e-k-book implementation.

The paper is structured in six sections: Section I points out the context and the objective of our work. Section II refers to related works on knowledge acquisition and transfer. Section III depicts a structural view over the scientific knowledge. Section IV exposes the AsK framework based on this view. It proposes an aid to acquire and to structure scientific knowledge that will be reformulated. Section V discusses an application of AsK to acquire knowledge in the domain of design of satiety biscuits. Section VI exposes conclusions and future work.

II. BACKGROUND AND RELATED WORK

As cited by [1], knowledge transfer is a complex process, which includes knowledge transmission, absorption and use. Important types, methods and levels for its exploitation are cited in [2]. However, to transfer knowledge one has to acquire and make it available in a way that is suitable to its absorption and use [3]. The frequently used concepts [4] in the context of U2I-KT [5] are present in the knowledge acquisition, too. The absorptive capacity [6] depends on the ability of knowledge use, but also on the capacity to acquire and assimilate it. Prior knowledge and domain similarity [4] facilitate the absorption, but also the acquisition strategies. While a survey of tools and techniques for knowledge acquisition is presented in [7], the closest related work to our research is found in [8]. Here,

 ISBN: 978-1-61208-181-6

knowledge acquisition is "the activity of capturing, structuring and modelling knowledge from any source for the purpose of storing, sharing or implementing knowledge". Thus, the overall objective of the acquisition process is to develop a k-base operational for the future processing. This k-base contains informal, structured, formal or computational know-ledge related to all the relevant and useful aspects. Accordingly to [9], lack of knowledge pertaining to research utilization can inhibit appropriate and effective use. In this background, the originality of our work consists in proposing a framework to acquire and to structure scientific knowledge for the U2I-KT.

III. WHICH STRUCTURE FOR SCIENTIFIC KNOWLEDGE

Scientific knowledge is the key ingredient in the process of U2I-KT and the key object of our study. Published knowledge produced by scientist researchers as a result of their scientific research is evaluated by a scientific community and released through scientific publications. Scientific papers are thus the vehicle of spreading and supplying scientific knowledge. They are regularly structured into five formal parts: (a) Introduction and Literature Review; (b) Material and Method; (c) Results; (d) Discussion; (e) Conclusion and Future Work. Knowledge contained in these publications concerns existing data used for the research and new data provided by this. Inspired by [10], which cites the main steps commonly used in a hypothetical-deductive scientific method, we assume that relevant scientific knowledge encloses data, information and knowledge related to the following corpora: *Initial data*, *Problem, Hypothesis, Tests, Results, Interpretation* and *Conclusion*. These corpora offer a structural perspective on the scientific knowledge. Commonly, *Initial data, Problem and Hypothesis* are found in the (a) section, *Tests* in the (b) section, *Results* in (c), *Interpretation* in (d) and *Conclusion* in the (e) section of a scientific publication.

Initial data includes theories, models, facts, representations, observations, methods or concepts related to a specific domain. It reveals accepted shared knowledge of a scientific community that outlines the prior knowledge needed for the research. *Problem* suggests the hindrance raised from the contradictions between new facts and old ideas. It is formulated by a research main question. *Hypothesis* conveys assumptions formed on the basis of the research problem. *Tests* include *Material* (*Product* and *Equipment* used for experimentation, *Instrument* required to observe or to measure parameters and *Operator* to perform tasks) and *Method* (*Protocol* and *Model* used for investigation, or production). Tests (experimental, simulation- or modelling-oriented) are set up to: i) test existing theories or probe results from the background observation, ii) answer a question or investigate a problem, or iii) test the hypothesis introduced, to support or to disprove them. *Results* present data or interesting observations made when handling data, and report the accurate results of tests. *Interpretation* renders the results explanations, analysis, reflexions and assimilations. *Conclusion* exposes the impact of the new research findings on the current knowledge and the restatement of the knowledge after their integration.

These 7 corpora of knowledge are necessary to acquire and structure relevant scientific knowledge. Certainly, industrialists are in quest of accurate and well interpreted scientific results. Yet they are also interested in data, information and knowledge about essential methods and materials to use, parameters values to adjust, concepts and theories to be familiar with. These will provide them a better understanding of the scientific results and may ease the reproduction of the tests, if needed. Hence, all the available corpora of knowledge have to be collected and handy at the end of the acquisition. Otherwise, lost or unavailable pieces of knowledge can have a negative impact on the future absorption and exploitation of the scientific knowledge. This can produce an incomplete transfer of knowledge, which can instigate unnecessary rework and delay for industrialists.

IV. ASK FRAMEWORK PRESENTATION

The AsK framework aims to help a knowledge engineer to acquire and structure useful scientific knowledge into a k-base made available for future processing.

Materials used in AsK are related to knowledge sources and tools. Knowledge sources include the available documentary (publications, reports, slides, work sheets...) and human expert resources (interviews). Knowledge tools include techniques and software used for knowledge collection, visualisation or modelling (diagrams, frames, trees, maps, tables, matrices).

The method used in AsK is based on the assumption that to acquire and to structure relevant scientific knowledge, all the 7 corpora of knowledge (*Initial data, Problem, Hypothesis, Test, Results, Interpretation, Conclusion)* have to be examined.

The AsK framework proposes a three-phase protocol: i) acquire and structure knowledge from available resources, ii) identify the lacking necessary knowledge and then acquire it and iii) verify the conditions for acquisition completed.

During the first phase, knowledge is collected and modelled from the initial sources, following the subsequent procedures:

- The research problems for each knowledge domain are identified. Each problem introduces an empty AsK-frame entitled with regard to the given problem. This frame is a skeleton designed to offer a structured view on the scientific knowledge, as described in Section III. Each frame encloses seven layers, which correspond to the seven corpora of knowledge, as shown in Figure 1.
- Knowledge is collected from the initial sources and transcribed into suitable models. Glossaries, process diagrams, dependence relations, concept trees, process maps, attributes' tables, relation matrices, dependence rules are potential knowledge models (k-models) [8] to be put within the k-base. As far as possible, process-oriented models are used, as they ease the absorption by the industrialists. Knowledge modelling provides punctual views of the collected knowledge and delivers localised pictures within the k-base. The k-models reveal knowledge about concepts, hypothesis, methods, products, properties or results. They are placed into the corresponding AsK-frame, within the suitable layer. Knowledge collected and modelled is thus immediately located in its proper context. The procedure leads to a scientific-oriented structuring of the k-base.
- First interviews with experts are set up to confirm the initial AsK-frames titles with respect to their research domain. They can guide to the creation of new frames. This procedure provides the framing of knowledge (allows to assign one ore more experts to each frame, depending on the objective of their research work).

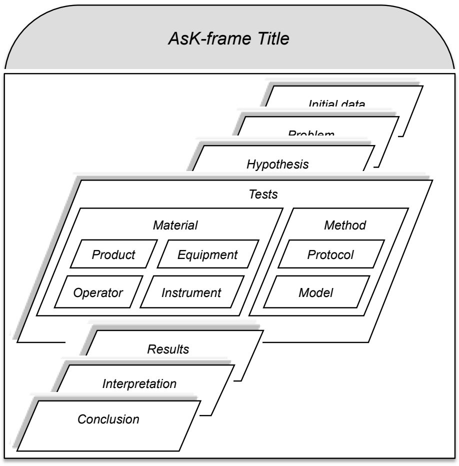

Figure 1. Seven necessary layers in the AsK-frame skeleton

- The initial established k-models are either confirmed or clarified, detailed or infirmed (if knowledge have evolved since then). New knowledge related to these models is collected and submitted to confirmation. The confirmed k-models are checked as relevant in the base
- The proper assignation of the confirmed k-models to a particular layer in the AsK-frame is confirmed as well.

During the second phase, the lacking useful knowledge is first identified and then acquired. The main steps to follow are:

- For a given frame, lacks of useful data, information and knowledge in the corpora of *Initial data, Product, Method, Equipment, Results* and *Interpretation* are identified with respect to the AsK framework.
- New semi-structured interviews are set up in order to collect these corpora of lacking knowledge within each frame. For example, the questions are formulated so that the knowledge obtained from answers directly fit a corpora (AsK-layer), which was lacking in knowledge.
- This new collected knowledge is modelled and the k-models are submitted for confirmation and checked following a similar procedure as cited formerly.
- The triple procedure (collect, model and confirm) is iterated for all the corpora of lacking useful knowledge requested for a given layer of an Ask-frame. Then, the procedure is iterated for all the layers within the frame and for all the frames created for the research domain.

During the third phase, the knowledge engineer verifies all the conditions for "acquisition completed", as drawn here:

- When all the corpora requested for a layer of an AsK-frame are collected, modelled and confirmed, the layer is considered "completed" and checked in the k-base.
- When all the layers of an AsK-frame are "completed", the AsK-frame is declared "completed" and checked.
- When all the AsK-frames within a given domain are "completed", the acquisition of the knowledge related to that domain is considered "completed".
- When all the domain acquisitions are "completed", the global acquisition procedure is declared "completed".

When the final condition is satisfied, the k-base is finished. It contains relevant and useful scientific knowledge, structured in order to ease the access for representation and reformulation.

V. AsK APPLICATION TO A BISCUIT-MAKING DOMAIN

We have applied the AsK framework to acquire knowledge in the domain of design and manufacturing of biscuits with satiety benefit. The work is in progress and it is carried out within a research project with industrial and academic partners.

The materials used as initial sources for collecting knowledge are available documentation (reports, presentations of domain research contributions, notes, and also transcriptions of recorded individual or collective interviews with experts). Three teams of scientific experts (in Formulation, Nutritional characterization and Sensory analysis) and an industrial partner (in the biscuit-making domain) take part in the project.

The method employed to acquire and structure scientific knowledge for each of the expertise domain suits the steps of the protocol mentioned in Section IV. Various k-models are built from initial sources and then submitted to the experts, for individual and group confirmation. The experts are interviewed at different stages during the acquisition process. Figure 2 illustrates one k-model built for dough texture's descriptors.

This k-model is confirmed and placed in the *Product* sub-layer of the AsK-frame "*Dough processability",* created for the Formulation domain. Additional and more specific knowledge is collected regarding to each of the descriptors of the dough texture shown in this k-model. This knowledge is modelled in well-formatted electronic files, which contain pictographic, textual and structure-related information (keywords, links and bibliography used). An example is given for the 'adhesiveness" descriptor, as pointed out in Figure 3. These electronic files are called knowledge sheets (k-sheets) and match the [8] meta-k-models. They are created using a specific computational tool that we have developed to assist the e-k-book implementation. Easy to be accessed, the k-sheets facilitate the knowledge visualization during the acquisition process. They will be properly treated and used as elements in the e-k-book.

K-models and k-sheets are established and then confirmed for all of the texture descriptors. Lack of useful knowledge for "*Dough Processability"* frame is identified with respect to the AsK framework. Then, corpora of knowledge for Hypothesis, Method, Instrument, Equipment and Results are acquired by means of k-sheet models to fill the layers of the given frame.

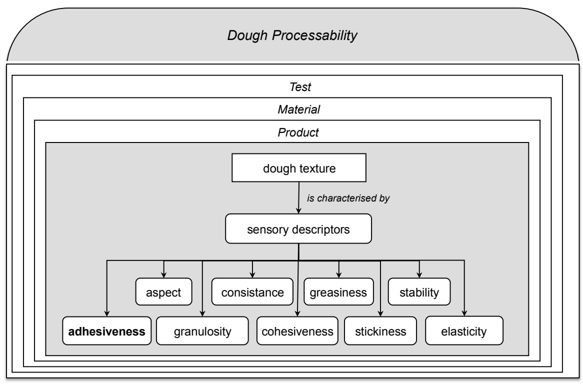

Figure 2. AsK placement for dough texture's characteristics map (k-model)

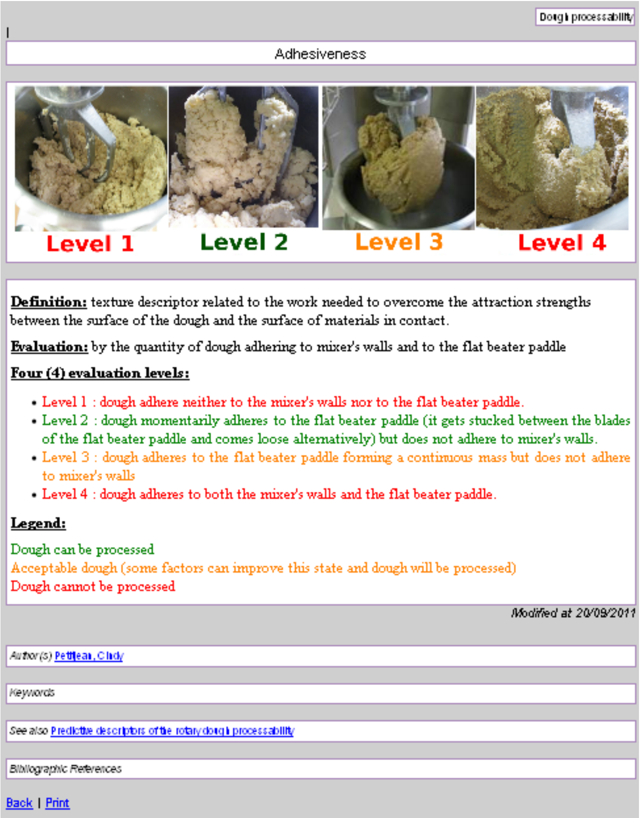

Figure 3. K-sheet collected for the "adhesiveness" descriptor

The example related to the evaluation of dough texture's sensitive descriptors is inspired by [11] and concerns published scientific findings. So far, two cases have been encountered in our project: i) scientific knowledge is already validated by the specific-oriented community and ii) scientific knowledge is in process to be produced and/or published (tests are in process, results are not yet available or published). In the first case, the scientific discourse is clear and it eases the acquisition process. In the second, the discourse is in construction and as a result, the knowledge acquisition can be significantly penalized.

VI. Conclusion and Future Work

We have proposed here a framework (named AsK), which defines the seven corpora of knowledge (*Initial data, Problem, Hypothesis, Tests, Results, Interpretation* and *Conclusion*) that are necessary to be collected and modelled in order to acquire relevant scientific knowledge in the context of a U2I-KT [12]. With regard to these corpora, Ask offers a scientific-oriented [10] way to structure knowledge within the k-base, throughout the collection and the modelling process. Thus, when the acquisition procedure is completed, an organized k-base [8] containing all the necessary [9] corpora of useful and usable scientific knowledge [1] is made available for future treatments (industry-oriented reformulation, representation, and transfer).

The major advantages of using the Ask framework can be summarized as follows: i) it fits for multi-domain knowledge acquisition in parallel; ii) it offers to the knowledge engineer a strategy for planning and preparing the interviews during the acquisition procedure; it helps him to set questions during these interviews, which provide relevant knowledge, structured in a scientific-oriented view, as presented in scientific publications; iii) AsK supplies a global view over the k-base configuration and a granular view on a given k-model within the k-base; iv) it provides a helpful technique to instantly situate an already built k-model into its context within the k-base; v) Ask provides information about the vacant parts of the k-base; it detects lack of or not yet available useful knowledge necessary to acquire; vi) it gives information about the k-base local or global status (in process/completed) at any time of the acquisition process. Moreover, the Ask framework allows listing extra information potentially useful for further research (such as complementary methods to use, new hypothesis emerged, results to correlate). New research programs can be proposed and consequently new related knowledge can be generated, acquired and integrated in the (evolving-designed) e-k-book.

Further work will be dedicated to finish the acquisition process so that all the scientific findings will be available into a relevant k-base. Reformulation strategies and representation formats will be set up to make the acquired knowledge usable by industrialists and an e-k-book will be built for its transfer.

Acknowledgment

This work was carried out with financial support from the ANR France. The authors thank the BISENS project partners for all their contributions to the knowledge acquisition process.

References

[1] T.H. Davenport and L. Prusak, Working knowledge-How Organization Manage What They Know, 1998, MA: Harvard BS Press, pp. 100-104.

[2] J.M. Beyer, "Research utilization: Bridging the gap between communities" in Journal of Management Inquiry, vol. 6, 1997, pp. 17–22.

[3] J. Kang, M. Rhee, and K.H. Kang, "Revisiting knowledge transfer: Effects on knowledge characteristics on organizational effort for knowledge transfer", in Experts Systems with Applications, vol. 37, n° 12, 2010, pp. 8155-8160.

[4] A. Agrawal, "University-to-industry knowledge transfer: literature review and unanswered questions", in International Journal of Management Reviews, vol. 3, issue 4, 2001, pp. 285-302.

[5] A. Hughes, "University-Industry Links and U.K. Science and Innovation Policy" in How Universities Promote Economic Growth (eds. S.Yusuf and K. Nabeshima), ch. 4, 2007, pp. 71-91.

[6] W.M. Cohen and D.A. Levinthal, "Absorptive Capacity: A new perspective on learning and innovation", Administrative Science Quarterly, vol. 35, 1990, pp. 128-152.

[7] J.H. Boose, "A survey of knowledge acquisition techniques and tools" in Knowledge Acquisition, vol. 1, n°1, 1989, pp. 39-58.

[8] N.R. Milton, "Knowledge acquisition in practice. A step-by-step guide", ed. Springer, 2007, pp. 69-110.

[9] C.B. Stetler, "Updating the Stetler model of research utilization to facilitate evidence-based practice" in Nursing Outlook, n° 49, 2001, pp. 272–278.

[10] J.Y. Cariou, "La formation de l'esprit scientifique: trois axes théoriques, un outil pratique: DiPHTeRIC", Bull. APBG, n° 2, 2002, pp. 279-320.

[11] C. Petitjean, P. Roussel, C. Simonnot, S. Berland, P. Aymard, and C. Michon, "Recherche des descripteurs prédictifs de la machinabilité des pâtes moulées", Poster Presentation, 60èmes JTIC, 2009.

[12] E. Rasmussen, O. Moen, and M. Gulbrandsen, "Initiatives to promote commercialization of University Knowledge", in Technovation, n° 26, 2006, pp. 518-533.