Knowledge Experience Base (KEB) for technological transfer

P. Ardimento\textsuperscript{1,2}, M.T. Baldassarre\textsuperscript{1,2}, N. Boffoli \textsuperscript{1,2}, G. Bruno \textsuperscript{1,2}, V.N. Convertini \textsuperscript{1}, G. Visaggio\textsuperscript{1,2}

Abstract This paper presents a framework able to support a cooperative “innovation chain” from an Open Innovation (OI) perspective. To this end we propose a Knowledge Management System (KMS) which consists of a set of processes that make up an Experience Factory (EF) and a platform that is the Knowledge Base Experience (KEB), which collects Knowledge Experience Packages (KEP). We have proposed an approach, called Prometheus, which has been defined in order to project KEPs to assure that: the experience contained in it, even when it is collected through projects carried out with many years of experience with many person years, can be acquired quickly by the user; contains the tools needed to facilitate the acquisition of knowledge that support the innovation to transfer with particular emphasis on how the KMS proposed can help the enterprises in particular Small and Medium ones to survive the current turbulence of the markets. Finally a preliminary evaluation of the proposed approach is proposed in the academic context.

Introduction
The necessity to exchange knowledge between different public and private organizations causes the generation of new approaches to innovation transferring. The proposed approach to knowledge generation and exchanging is based on the Open Innovation [3, 8]. The logic of Open Innovation characterizes the end of the XX century and forecasts that enterprises develop their projects in collaboration with other sites. As contrast to the principles that had characterized the Closed Innovation, the divulging idea is that of: internal R&Ds holding part of a knowledge that comes from external R&D; creation of a business model that assumes a primary role in the knowledge development lifecycle; sharing knowledge with other enterprises in order to achieve higher quality levels.

The PROMETHEUS Project [1, 2, 9, 10, 11, 12] (Practices Process and Methods Evolution Through Experience Unfolded Systematically), is a model of Experience Factory (EF) to collect experimental knowledge in a repository Experience Knowledge Base (KEB) in the form of Knowledge Experience Package (KEP). The KEP is the vehicle suggested for the transfer of knowledge while the EF is the set of processes that make the Open Innovation. This paper describes:

- The proposed approach with particular emphasis on the KEP structure;
- The available tools that can be used to evaluate the competences of the users and to fill the eventual gap.

The rest of the paper is structured as follows: the next section discusses related works and research activities; section 3 presents the proposed approach, focusing particularly on the structure of the KEP, in section 4 we discuss the model for evaluating competences. Section 5 introduces a preliminary evaluation of the efficacy of Prometheus on representing and transferring knowledge and experience with respect to the traditional way. Finally, in the conclusions some

\textsuperscript{1} University of Bari Aldo Moro - Dept of Informatics —Bari -Italy;{ardimento, baldassarre, boffoli, bruno, convertini, visaggio}@di.uniba.it

\textsuperscript{2} SER AND PRACTICES - Spin-off of the University of Bari Aldo Moro
observations are made about the preliminary results obtained, and possible future research pathways are identified.

2. Related Works

The problem of knowledge packaging for better usage is being studied by research centres [4, 5, 6, 7], some of these are enterprises [4, 7], showing the interest of the industrial community to the same problem. Knowledge bases sometimes have a semantically limited scope. This is the case of the Daimler-Benz base [6, 7], that collects lessons learned or mathematical prediction models or results of controlled experiments in the automobile domain only. In other cases the scope is wider but the knowledge is too general and therefore not very usable. This applies to the MIT knowledge base [5] that describes business processes but only at one or two levels of abstraction. There are probably other knowledge bases that cover wider fields with greater operative detail [7] but we do not know much about them because they are private knowledge bases. Our approach focuses on a knowledge base whose contents make it easier to achieve knowledge transfer among research centres; between research centres and production processes; among production processes. The knowledge base must be hybrid, public, as we wish, or private, depending on KEP authors preferences. The public KEB allows to one or more interested communities to develop around it and exchange knowledge. In particular, it must be possible for Small and Medium sized Enterprise (SME) to become members of these communities. In fact, we believe that only membership of these special interest communities can allow SME to adopt Open Innovation and reap the benefits.

3. Proposed Approach

The Authors use the term KEP to refer to an organized set of: knowledge content, teaching units on the use of the demonstration prototypes or tools and all other information that may strengthen the package’s ability to achieve the proposed goal. The knowledge package must be usable independently of its author or authors and for this purpose the content must have a particular structure: distance education and training must be available through an e-learning system. In short, the proposed knowledge package contains knowledge content integrated with an e-learning function. In the proposed approach, the KEP must include all the elements shown in figure 1. A user can access one of the package components and then navigate along all the components of the same package according to her/his training or education needs. Search inside the package starting from any of its components is facilitated by the component’s Attributes.

![Figure 1. Diagram of a Knowledge/Experience package](image)

It can be seen in the figure that the Art & Practices Knowledge Content (KC) is the central one. It contains the knowledge package expressed in a hypermedia form in order to include figures, graphs, formulas and whatever else may help to
understand the content. The KC is organized as a tree. Starting from the root (level 0) descent to the lower levels (level 1, level 2, …) is through pointers. The higher the level of a node the lower the abstraction of the content, which focuses more and more on operative elements. The root node is made up of a “Thoughtful Index” and one or more problems.

The nodes are the answers to the problems: the solution or the solutions proposed for each of the announced problems. Each answer consists of the following: research results for reference, analysis of how far the results on which the innovation should be built can be integrated into the system; analysis of the methods for transferring them into the business processes; details on the indicators listed in the metadata of the KC inherent to the specific package, analyzing and generalizing the experimental data evinced from the evidence and associated projects; analysis of the results of any applications of the package in one or more projects, demonstrating the success of the application or any improvements required, made or in course; details on how to acquire the package.

In line with Open Innovation, the research results integrated by a package may be contained within the same knowledge base or derive from other knowledge bases or other laboratories. If the knowledge package being read uses knowledge packages located in the same experience base, the relations will be explicitly highlighted. As shown in Figure 1, each component in the knowledge package has its own attributes structure. For all the components, these allow rapid selection of the relative elements in the knowledge base. The focus in this work is on the attributes in the KC. In fact, these have been defined during research conducted by the authors and by other authors. To facilitate the research, we used a set of selection classifiers and a set of descriptors summarizing the contents. The classifiers include: the keywords and the problems the package is intended to solve. The summary descriptors include: a brief summary of the content and a history of the essential events occurring during the life cycle of the package, giving the reader an idea of how it has been applied, improved, and how mature it is. The history may also include information telling the reader that the content of all or some parts of the package are currently undergoing improvements.

The interested reader can find further details on the contents of the KEP and the management and use of KEB on the technical report [11].

4. Model for Evaluating Competences

For each competence Prometheus provides a set of learning units. Each learning unit aims to train the user of a KEP on one or more items of the KEP of interest. Therefore, you can attach to each teaching unit a test plan to verify that the user already has, or has acquired, the corresponding part of competence. Such a model predicts that each jurisdiction has an associated evaluation questionnaire and a decision model. In the questionnaires the self-assessment tests result in the user evaluation of the KEP and guidelines to improve the training of the user. The test assesses the competences and skills attained by the user and the gap between them and the ones expected. The model of decision interprets the level of acquisition of the skills of users receiving the evaluation questionnaires and suggests actions to be undertaken to fill any gap between skills expected and skills acquired by the learner. The model decision is made by the decision tables.

Operationally, for each competence C (i) a specific evaluation model is planned (figure 2). In this model, the responses gathered by QC (i) the evaluation questionnaires provided, are interpreted by an appropriate set of decision tables DT (1), ..., DT (n). Note that each DT (k) with k = 1 ... n, aims to interpret the answers of the teaching unit TU (k).
5. Case Study

The aim of the experimentation is to compare the transfer of knowledge represented through KEP collected in Prometheus with the corresponding knowledge transferred through verbal language, analysis and design documentation in an academic context. In particular, the experiment aims at comparing the effectiveness of the learning of two different groups of thesis students that work on the same thesis project.

5.1. Research Goal

The investigation aims to verify the efficacy of knowledge, which needs a final year student to start-up his own thesis work, represented through KEP collected in Prometheus compared with equivalent knowledge expressed by documentation. In particular documentation refers to the thesis work produced previously, from one or more students that have already completed their work. The documentation represents the starting point to start a thesis work of the graduate student.

For efficacy we investigate whether the analysis and extraction of knowledge through a KEP requires less effort than through documentation. With respect to his quality characteristic, we also investigate whether given a topic (from here on indicated as problem) it is more difficult to understand the problem and extract knowledge from documentation rather than a KEP. The following Research Goal has been defined:

Analyze knowledge extraction using a Knowledge Experience Package (KEP) with the aim of evaluating it with respect to efficacy (compared to knowledge extracted from documentation) from the view point of the knowledge user in the context of a controlled experiment on a KEB called Prometheus.

In accordance to the above goal, the following research hypothesis has been made:

H$_{EFF0}$: there are no statistically significant differences in terms of effort for solving problems assigned using KEP rather than documentation

H$_{EFF1}$: there are statistically significant differences in terms of effort for solving problems assigned using KEP rather than documentation.

5.2 Experiment Description

Experiment Variables. The dependent variable of the study is Efficacy. Efficacy indicates to what point the Knowledge Representation criteria is effective (in terms of effort spent) for extracting knowledge and answering a specific set of questions. The independent variables are the two treatments: the problems examined with KEP and with thesis documentation. Two different types of problems were investigated: Software Dependability with GQM, Balanced Score Card with GQM.

A set of 4 questions have been defined for each problem. This has been considered an appropriate number that balances the need for a sufficient amount of data without having to count on an excessive amount of effort and risk to bore a tire experimental subjects. Each question has a different complexity level. Two of the
four questions have analogous complexity levels for both treatments: KEP and documentation. While, for the other two complexities is different between treatments: not complex for a treatment and complex for the other and vice versa. For clearness, the answer is classified as easy to search if it can be localized in a part of the document, not larger than a page; rather it is considered complex if the answer to the question refers to information sparse in multiple parts of the document that cover an area which is greater than a page. Since the structure of KEP and documentation are different, the same question may have different complexity level, depending on the knowledge representation method used.

Selection of Experimental Subjects. The experimental subjects involved in the experimentation are final year students that were getting degree in Informatics course. A total of 12 students have been divided in two groups (GROUP A and GROUP B) with random assignment to each one. Each group was asked to answer questions assigned using, alternatively KEP or documentation. All of the students have previous knowledge on the topic concerning Balanced Scorecard because it is part of their course curricula. While, they have no previous knowledge on the “Software Dependability with GQM” topic.

Experiment Design. The design is represented in table 1:

<table>
<thead>
<tr>
<th>FACTOR B: knowledge extraction problems</th>
<th>Level 1 (Q1)</th>
<th>Level 2 (Q2)</th>
<th>Level 3 (Q3)</th>
<th>Level 4 (Q4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FACTOR A: Knowledge Representation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1: KEP</td>
<td>KEP/Q1</td>
<td>KEP/Q2</td>
<td>KEP/Q3</td>
<td>KEP/Q4</td>
</tr>
<tr>
<td>Level 2: Documentation</td>
<td>Doc./Q1</td>
<td>Doc./Q2</td>
<td>Doc./Q3</td>
<td>Doc./Q4</td>
</tr>
</tbody>
</table>

Table 1: experiment design

The experiment was organized in two experimental runs, RUN1 and RUN2, one per day in two consecutive days. Each run applied the design above. During each run we changed the content of the KEP/documents and the content of the questions used to extract information from the source. Moreover, in RUN1, the KEP/documents content, along with the questions for extracting information, related to Balanced Scorecard with GQM; in RUN2 they referred to Software Dependability with GQM.

Within a RUN, each group was assigned to either one of Factor A and to all the levels of Factor B. The assignments were inverted for the successive run. A summary of the assignments is reported in table 2, as it can be seen, within the same run the subjects use the same topic and the questions are the same.

<table>
<thead>
<tr>
<th>+ RUN 1</th>
<th>RUN 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP A</td>
<td>Subjects of group use KEP on SCORECARD and answer to all questions Q1, Q2, Q3, Q4</td>
</tr>
<tr>
<td>GROUP B</td>
<td>Subjects of group use DOCUMENTS on SCORECARD and answer questions Q1, Q2, Q3</td>
</tr>
</tbody>
</table>

Table 2: summary of assignments for each experimental run

Instrumentation. During each experimental run, for each analyzed problem, experimental subjects were provided the following instrumentation:

- general description of the problem;
• the KEP or set of documents concerning the thesis topic. The package is accessible through Prometheus. The documents are provided in digital version.
• Set of questions related to the topic;
• Data form: in which each experimental subject must report their name, last name, start and end time, and answers to the questions.

Operation. At the beginning of each run, each experimental subject received a complete set of instrumentation (described above). It contained the documents or KEP according to the treatment and group. The students examined the material and answered the questions reporting them on the data form. The start and end time were recorded by the researchers when handing in and collecting the forms.

Measurement Model. Given the above research goals and the research hypotheses, the dependent factors and the measures used to calculate such factors have been defined according to the GQM quality model. The introduced metrics are collected as Prometheus and documents metrics. The metric described in table 3 has been collected on both types of knowledge extraction treatments.

Efficacy Factor is measured as the average of points $P_{ij}$ attributed for answering the $i$-th question of the $j$-th experimental subject. All answers were evaluated according to the interval scale reported in table 3.

<table>
<thead>
<tr>
<th>Evaluation of Question</th>
<th>Point $P_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrong Answer: the $j$-the subject gave a wrong answer to the $i$-th question.</td>
<td>0</td>
</tr>
<tr>
<td>Lacking Answer: the question was not answered by the $j$-the subject</td>
<td>2</td>
</tr>
<tr>
<td>Incomplete Answer: the $j$-the subject gave a partially correct answer to the $i$-th question</td>
<td>4</td>
</tr>
<tr>
<td>Complete Answer: the $i$-th question has received a correct answer by the $j$-th subject</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3: details of efficacy quality factor

Note that the researchers, as domain experts involved in the investigation, corrected all the answers to the questions given by the experimental subjects.

Experimental results. The data collected during the experimentation have been synthesized through descriptive statistics. This allows representing them graphically, identifying possible outliers and deciding if they must be eliminated from the sample. Finally, data has been analyzed through hypothesis testing and validated with respect to a significance level of $\alpha = 5\%$.

Instrumentation. Given the experimental design, a 2X4 analysis of variance with a between-factor (Knowledge Representation: PROMETHEUS vs DOCUMENTS) of two levels, and a within-factor (TOPICS: Q$_1$, Q$_2$, Q$_3$, Q$_4$) of four levels was carried out for investigating efficacy, i.e. an ANOVA repeated measures analysis was carried out. The differences between efficacy values are all significant, as it arises from the p-levels reported in table 4. It refers to the differences between treatments with respect to each question $Q_i$ ($i=1, \ldots, 4$).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>$F$</th>
<th>p-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q$_1$</td>
<td>5,87979</td>
<td>5,879791</td>
<td>7,578096</td>
<td>0,007308</td>
</tr>
<tr>
<td>Q$_2$</td>
<td>14,3907</td>
<td>14,39071</td>
<td>6,149250</td>
<td>0,015247</td>
</tr>
<tr>
<td>Q$_3$</td>
<td>18,7693</td>
<td>18,76934</td>
<td>12,64537</td>
<td>0,000636</td>
</tr>
<tr>
<td>Q$_4$</td>
<td>38,6677</td>
<td>38,66771</td>
<td>25,44427</td>
<td>0,000003</td>
</tr>
</tbody>
</table>

Table 4: Univariate test of significance for Efficacy in RUN$_1$
Also RUN2 points out that overall efficacy is better when Prometheus is used. This result is, in fact, confirmed by the univariate test of significance carried out to investigate the differences in means for each problem, between the two methods used. Table 5 points out that differences are statistically significant. The test points out those differences in correspondence to treatments in each question are statistically significant. So, the observations made in the descriptive analysis are confirmed.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>p-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>5,8276</td>
<td>5,82762</td>
<td>6,155960</td>
<td>0,015194</td>
</tr>
<tr>
<td>Q2</td>
<td>75,4513</td>
<td>75,45134</td>
<td>36,29545</td>
<td>0,000000</td>
</tr>
<tr>
<td>Q3</td>
<td>75,4513</td>
<td>75,45134</td>
<td>36,29545</td>
<td>0,000000</td>
</tr>
<tr>
<td>Q4</td>
<td>11,5977</td>
<td>11,59768</td>
<td>5,787814</td>
<td>0,018445</td>
</tr>
</tbody>
</table>

**Table 5:** Univariate test of significance for Efficacy in RUN2

### 7. Conclusions and Future Works

The educational and business sectors need to be linked to the information cycle. For this reason, the proposed approach PROMETHEUS includes an e-learning System teaching knowledge of the packages and training the user in the use of the demonstration prototypes or tools supporting an innovation. We propose PROMETHEUS, a demonstration platform that integrates a Knowledge Management System and a Learning System, allowing navigation among all its components. The experiment described in this paper has produced empirical evidence on knowledge representation obtained through PROMETHEUS. The evidence collected has confirmed the validity of the structure of a KEP in the tool. Moreover, we can summarize that the structure of the KEP represents explicit knowledge in a more efficacy form with respect to traditional documentation. The university context is generally considered of scarce interest for empirical investigations, as students do not have the same maturity of professional developers. In this specific case student subjects are considered to be appropriate in that knowledge transfer is more critical when previous knowledge of users is low. To this end, experiment has pointed out the efficacy of the KEP technique when the subjects lacked of previous knowledge on the topic. Obviously, in order to generalize the validity of the KEP proposed in this work many replications and further studies extended to other contexts are needed, in particular industrial context. For this reason, the authors intend replicating the experiment, and making instruments and material available to other interested researchers.

### References