IDENTIFYING OUTPUT INTERACTIONS AMONG IS PROJECTS – A TEXT MINING APPROACH

Christian Meier

University of Paderborn
Warburgerstraße 100
33098 Paderborn

christian.meier@wiwi.uni-paderborn.de

Abstract

In the literature, there is anecdotal as well as empirical evidence for the existence and the business impact of output interactions among information systems projects. While a lot of sophisticated optimization models have been suggested which already provide for the consideration of output interactions when selecting information systems project portfolios, the necessary data required for their application in business practice are usually not available to the planner. There is a lack of techniques in the literature on how to identify output interactions already at the time, a portfolio is planned. We attribute this lack to the rather semantical nature of output interactions. We contribute to filling the identified gap by conferring semantic clustering – a technique originating in the text mining literature – to the field of information systems project portfolio selection. A prototypical decision support system is developed that uses latent semantic analysis and hierarchical clustering to identify potential output interactions among information systems project proposals based on semantic similarities within their goal descriptions. This paper focuses on the design of the developed prototype and argues that latent semantic analysis represents a very promising technique for the identification of output interactions among information systems projects.

Keywords: Information Systems, Project Portfolio Selection, Project Interactions, Latent Semantic Analysis, Semantic Clustering.

1 Introduction

The selection of the right information systems (IS) projects to form an adequate project portfolio has become an increasingly "important and recurring activity in many organizations" (Archer and Ghasemzadeh, 1999). An often neglected requirement in this selection process is the consideration of project interactions. Three types of interactions can be distinguished: (1) overlap in project resource utilization, (2) technical interdependencies, and (3) effect interdependencies (referred to as output interactions[1] in the following) (introduced by Aaker and Tyebjee, 1978, adopted by, e.g., Santhanam and Kyparisis, 1995; Lee and Kim, 2001; Eilat et al., 2006). Considering these types of interactions may constitute "valuable cost savings and greater benefits" to an organization (Santhanam and Kyparisis, 1996). It is a challenging and time consuming but important requirement to identify and account for interactions among IS projects in order to avoid making unfavorable project portfolio selection (PPS) decisions (Lee and Kim, 2001).

There is anecdotal as well as empirical evidence for the existence of output interactions. For example, based on a data set of 623 U.S. firms, Aral et al. (2006) name complementarities between the implementation of Enterprise Resource Planning, Customer Relationship Management, and Supply Chain Management Systems as an explanation of performance gains. On a data set of 927 German firms, Engelstätter (2009) finds similar results. He observes positive effects among these three enterprise software systems when they are used together. Engelstätter attributes this observation to possible complementary effects among these software systems, in the following referred to as *complementary output interactions*. Besides complementary output interactions, ESI International (2009) reports from a global survey among 470 project and program management professionals that "71% of respondents report redundancies and conflicts in respect to project priorities". In the following we refer to these redundancies in the project portfolios as *competitive output interactions*. Both,

[1] In the following, we speak of an output interaction, if within the outputs of two or more projects there is an overlap in the provided project goals or services with the result that the business value impact of projects is non-additive.

complementary and competitive output interactions may cause that the business value impacts of projects are non-additive (see, e.g., Fox et al., 1984; Eilat et al., 2006). While the aforementioned studies investigate the existence and impact of output interactions from an *ex post* point of view, to the best of our knowledge no research has been conducted that aims at the *ex ante* identification of output interactions. Considering the reported effects and their expected business value impact, an ex ante consideration of output interactions could substantially affect the portfolio selection decision.

Numerous articles can be found in the literature that already incorporate output interactions into Operations Research (OR) decision models (e.g., Aaker and Tyebjee, 1978; Santhanam and Kyparisis, 1996; Lee and Kim, 2001; Stummer and Heidenberger, 2003; Carazo et al., 2010). However, the time-consuming identification of output interactions is mostly left unsupported with the portfolio planner. This severely hampers the application of these models in business practice. The lack of contributions to the identification of output interactions can be attributed at least partly to the rather semantic nature of output interactions. In contrast to, e.g., resource requirements, a project's planned outputs and goals tend to be formulated in a textual and less structured form. In addition, the effects of output interactions become visible only after the corresponding projects have been conducted, while the effects of overlap in resource utilization or technical interactions may be observed already during conduction. However, indications about possible connections among the projects' goals may be already found within the informal linguistic information of the textual descriptions in the project proposals at the time the portfolio is planned.

These descriptions serve the purpose of communicating the project's goals to co-workers and decision makers. Thus, we expect output interactions to be found within the semantics of these descriptions. To date these interactions have to be identified manually by domain experts. Especially in large project environments where potentially a large number of output interactions may occur, their manual identification by a human expert can become very challenging and time consuming. For example, there are theoretically already over 1m potential interactions among 20 projects. In various application domains, latent semantic analysis (LSA) (Deerwester et al., 1990), an information retrieval technique from the text mining literature, could be successfully applied to identify semantic similarities among a set of text documents. Typically, LSA is applied in the context of search engines (e.g., Berry et al., 1995) with the goal to identify documents best matching a certain search query.

In this paper, we constitute a starting point for a more detailed ex ante identification of output interactions within IS project portfolios by applying LSA to the domain of IS project portfolio selection (IS PPS). Thus, the main contribution of this paper is the development of a prototypical Decision Support System that confers well established concepts from the text mining and information retrieval domain to the field of IS PPS. In a cumulative research tradition, we base our prototype on an approach called *semantic clustering* presented by Kuhn et al. (2005), which uses LSA for the identification of semantic topics in source code, and adapt it the new conditions arising from the application domain of IS PPS. We follow the Design Science research paradigm (Hevner et al., 2004) and contribute to the literature by addressing the following research question: *How can the identification of potential output interactions in IS project portfolios be adequately supported by semantic clustering?*

2 Related Work

Our research is based on two different streams of literature: The literature on interactions in PPS and the literature on text mining techniques for the identification of semantically similar topics in text documents. The former emphasizes the importance of project interactions (e.g., Santhanam, R., Kyparisis, 1996; Lee, J.W., Kim, 2001; Eilat et al., 2006) and defines different interaction types (Aaker and Tyebjee, 1978; Kundisch and Meier, 2011a). Further it shows how to incorporate the different types of interactions into sophisticated optimization models (e.g., Santhanam, R., Kyparisis, 1996; Lee and Kim, 2001; Carazo et al., 2010). While all of these approaches provide very useful techniques for modeling and solving PPS problems under consideration of interactions, they have been built under the (implicit) assumption that the necessary information for identifying and assessing interactions is available to the planner. This may (at least partly) apply to resource interactions, for which few approaches have already been developed to support their identification (e.g., Kundisch and Meier, 2011b). However, especially for output interactions this assumption is hardly met in practice. As discussed above, planned outputs and goals tend to be formulated in a textual and rather unstructured form. Problems of polysemy and synonymy within the textual descriptions additionally hamper the IS supported ex ante identification of output interactions.

In the text mining literature, promising techniques are suggested that may help to overcome some of the problems mentioned above. This stream of literature focuses on how to extract information from textual data automatically. The articles closest related to our work apply LSA (Deerwester et al., 1990), e.g., for mapping readers to documents based on their background knowledge about the documents' topics (Wolfe et al., 1998), and for the identification of related topics in software source code documents (Maletic and Valluri, 1999; Kuhn et al., 2007). Because of similarities in their problem structure, for our research the article of Kuhn et al. (2007) is of particular interest. The authors propose an approach called semantic clustering to identify similarities among variable identifiers in software source code. They employ LSA and clustering to group source code documents with similar vocabulary together. Kuhn et al. apply their technique to two different case studies with mixed results. The comparably small size of the processed documents as well as size and quality of the vocabulary in source code documents lead to difficulties in the application of semantic clustering to their application domain. The authors state that larger documents, the use of natural language instead of artificial identifier names as well as a larger vocabulary are conditions for better results. In IS PPS these conditions are widely met, which constitutes IS PPS as a promising field of application for semantic clustering. Therefore, in a cumulative research tradition, we adapt semantic clustering presented by Kuhn et al. (2007) and apply it to the domain of IS PPS.

3 Prototype Design

A first assessment of the application domain of IS PPS exhibits that the basic conditions for a successful application of semantic clustering to IS PPS seem to be met. Project proposal documents typically serve the purpose of communicating the projects goals and requirements within an organization and are usually formulated in natural language. The most interesting information about the projects goals and outputs often is embedded within the semantics of the proposals. The same project goals may be expressed in many different ways by different individuals so that a simple comparison of the words used to describe these goals often will not be sufficient for an automatic identification of output interactions. Thus, important information with regard to output interactions may not be identified by simply comparing key words in different proposals. In other contexts, LSA has demonstrated its ability to overcome these difficulties and to identify the semantic topics in a set of documents (Landauer et al., 1998). This is achieved by decomposing the large vocabulary from the candidate documents into a considerably smaller set of factors which can be interpreted as linguistic topics. Based on these factors, the proposal documents now are clustered and adequately presented to the planner. Particularly among the documents clustered together this way, we expect output interactions to be found. While this paper is mainly concerned with the design of the prototype, the validation of this hypothesis will be the subject of a full research paper version. Our procedural approach can be decomposed into five conceptual phases (see Fig. 1), namely *text-preprocessing*, *singular value decomposition (SVD)*, *clustering*, *labeling*, and *visualization*. The five phases and the necessary adaptions to Kuhn et al. (2007) are briefly discussed in the following.

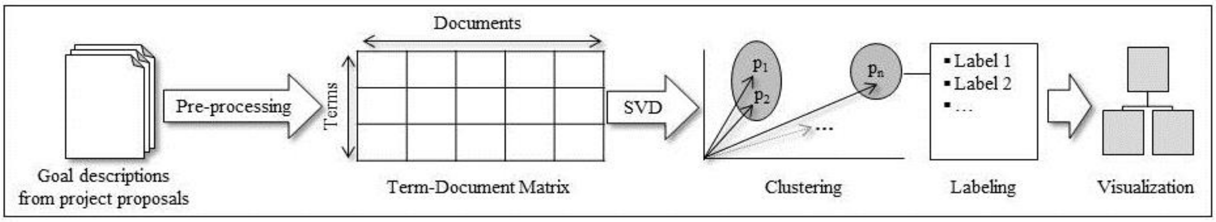

Figure 1. Identification process.

We extract the goal description from each project proposal document as input for our analysis and parse it into a list of words. The vocabulary in the documents originates from natural language, which favors the application of semantic clustering. Still, in the project proposal documents a considerable amount of noise is present due to different linguistic styles of the applicants, words with low semantic relevance, the frequent use of domain specific terms and potentially varying document lengths. As the output quality strongly relates to the quality of the inputs (Kuhn et al. 2007), we implemented an elaborate *pre-processing* to improve input quality. We remove numbers, special characters and single letters and subject the proposal documents to a stemming process (e.g., 'systems' is reduced to 'system') using the 'NHunspell Framework' and the free 'Open Office dictionary'. To remove words with high occurrence frequencies, but rather low semantic relevance, we implemented a comprehensive stop word list as well. This list already contains approximately 1.000 generic words (e.g. 'the', 'and', 'of'). In an organizational context, it can be expected that within the proposal documents domain- or company-specific words, abbreviations and phrases (e.g., company or department names, acronyms for company initiatives) have been assimilated into the corporate language to a certain degree. These words and phrases may not contribute to the identification of semantic similarities and have to be identified and added to the stop words list by the portfolio planner in order to improve input quality. The resulting set of words is then arranged into a *term-document matrix*, where the rows represent the terms and the columns the documents. The cell entries represent the raw occurrence of a specific term in a given document. At last, assuming the goal descriptions are differing in length, potential contortions are handled by a widely used normalization and weighting procedure (Dumais, 1991).[2]

After the pre-processing, *SVD* – a form of principal component analysis – is applied to the processed term-document matrix to reduce the noise in the data by reducing the number of factors based on which the documents will be clustered later on. The result of the application of SVD is an approximation of the original term-document matrix that is reduced by noise in the input data and thus, can be interpreted as a "better model of the text corpus" (Kuhn et al., 2007). The reduced matrix can be represented by a vector space model in which the similarity between two documents can now be acquired by calculating the angle (usually the cosine) between their corresponding vectors.

Clustering represents the key-feature of our prototype. Typically, in IS PPS the number of output interactions (and thus the number of clusters to build) is not known ex ante to the portfolio planner. Thus, we are handling a so called *unsupervised categorization* problem. Popular clustering algorithms as, e.g., k-means clustering, are not applicable without further ado. We therefore implemented an agglomerative hierarchical clustering (see, e.g., Hastie et al., 2011), which generates a tree-shaped dendrogram. It produces a "hierarchical representation in which the clusters at each level of the hierarchy are created by merging clusters at the next lower level" (Hastie et al., 2011). In each step, the documents or clusters exhibiting the highest semantic similarity are merged (see Fig. 2). This form of visualization enables the portfolio planner to facilitate a better understanding of the relationship structure between the project proposals by presenting the underlying hierarchy of the clustering

[2] In case that the PMO restricts goal description in project proposals to a certain number of words or even defines minimum and maximum description length a normalization/weighting may not be required.

decisions, instead of being confronted with a single, intransparent solution. In large project environments, the tree structure may become incomprehensible. Therefore, it may be helpful for the planner to get an idea which hierarchy level represents a good clustering solution and to only present her a relevant excerpt of the tree structure. Even if the optimal number of clusters is unknown, numerous techniques can be found in the clustering literature that can be helpful for this task. Milligan and Cooper (1985) provide an overview of 30 stop criteria to heuristically calculate a good clustering based on the coherency *within* and the separation *between* clusters. Therefore, we implemented the Calinsky and Harabasz (1974) index, which performed best in this study, into our approach as well. While this often may not result in the best possible clustering level for the identification of output interactions from an ex post point of view, we are at least able to suggest a promising hierarchy level based on which the planner can start further analysis. In future research, the visualization of the results as well as a comparison of the performance of different stop criterions with respect to the field of IS PPS have to be thoroughly evaluated. Figure 2 presents an exemplaric clustering result for a set of 17 documents. The optimal clustering levels according to the Calinsky and Harabasz (1974) metric for this example are highlighted.

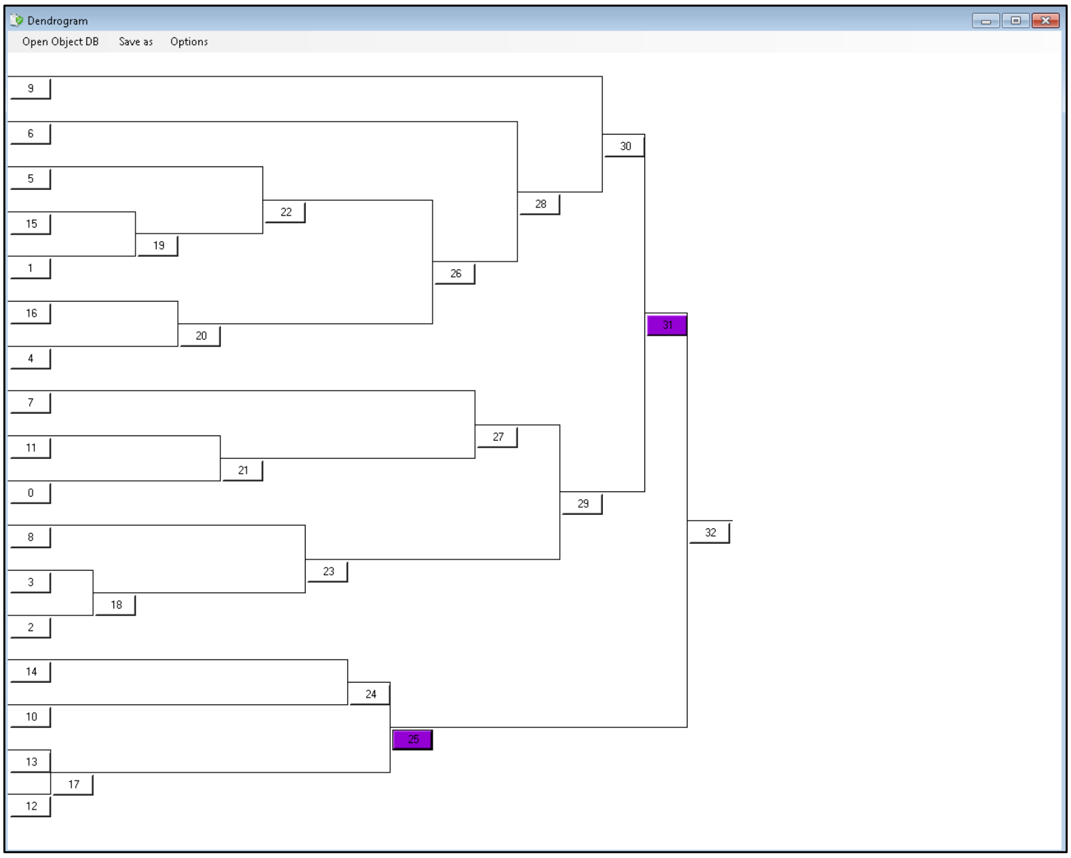

Figure 2. Dendrogram (optimal clusters highlighted)

Within clustering, the proposal documents have been grouped based on semantic topics they share. These topics represent rather abstract linguistic concepts derived from aggregation of the actual vocabulary used in the documents. To be helpful for the planner, we now have to identify the actual vocabulary from our proposal documents which best defines the topic for the corresponding cluster. Therefore, based on the weighting formula presented in Kuhn et al. (2007) each cluster in the clustering hierarchy is labeled with the *n* most relevant terms from the vocabulary which best describe the topic of that cluster. In a small pre-test we have observed that the number of these top words that are necessary to understand the underlying semantic topic varies from cluster to cluster. Therefore, in addition to the weighting formula of Kuhn et al. (2007), we have already implemented two proprietary labeling strategies as well as a parameterized input for the number of top words the clusters are labeled with. The evaluation of how many top words are adequate in our application domain and which of the labeling strategies provides the best results will be subject for future work.

4 Evaluation

In the literature, typically large corpora (sets of text documents which have been annotated by experts to establish a gold standard for comparison) are employed to evaluate the functionality and information retrieval quality of text mining tools. These tools typically are used to classify all or a subset of the documents from the corresponding corpus into clusters. Then, the classification result is compared to the annotation provided by the experts (usually referred to as 'gold standard') and one or more quality measures are calculated for the solution. One crucial aspect in this evaluation is the availability of such a corpus. In some fields much effort has been spent to develop such domain specific annotated corpora like, for example, the GENIA corpus (Ohata et al. 2002) in the molecular biology domain. In the field of project portfolio selection however, there is no such annotated set of documents available. Project proposal data typically constitutes sensible information for organizations and thus, is rarely publicly available. If such a data set is available, annotating the proposals through experts is a costly and time consuming task. Therefore, in a first step we split the evaluation of the proposed approach into two parts, from which part one – the technical evaluation – is presented in this paper. The technical information retrieval quality of the prototype will be evaluated based on the well-known annotated Reuters-21578 standard test set for categorization procedures. We use the Reuters-21578 test set because it is publicly available and well documented[3], and it has been used to evaluate text categorization procedures numerous times in the literature (e.g., in Massey 2005, Pessiot et al. 2010). To obtain the highest possible comparability to other classification techniques used on the Reuters-21578 test set, we follow the evaluation framework suggested by Massey (2005).

In a second step, to evaluate the performance of our approach for the intended domain, in future research we plan to apply the approach to a real world data set of IS project proposals. Therefore, a data set as well as an annotated reference solution from domain experts has to be acquired.

4.1 Data Set

The *Reuters-21578* set has originally been annotated by the Carnegie Group, Inc. and Reuters Ltd. and contains numerous news articles from Reuters newswire 1987. Today, there are at least five different variants of this dataset. Evaluation on different datasets aggravates the comparability of concurring approaches. To improve comparability with other approaches from the literature, we use the 'Modified Apte' (ModApte) split from the Distribution 1.0 – as suggested by Massey (2005) – which consists of a training set including 9.603 documents and a test set comprising 3.299 documents. This split has been used to evaluate supervised learning approaches (for example, by Dumais et al., 1998; Li und Yamanishi, 1999) in the past. Massey (2005) suggest using this split for unsupervised text clustering as well and provide results for three different clustering approaches, e.g., Adaptive Resonance Theory (ART) neural network, k-means und spherical k-means. The dataset contains 93 different Topics and the documents have been assigned to one or more of these topics by a professional indexer beforehand. The dataset comprises documents that are associated with multiple topics and the set exhibits a skewed topic distribution (some topics are overrepresented. Some of the documents in the set have not been assigned to a topic at all. While these documents could have been excluded from the test set, for better comparability and in line with Massey (2005), we have retained them within the set. They have been labeled with "no topic" instead.

[3] http://www.daviddlewis.com/resources/testcollections/reuters21578

Furthermore, each documents in average is assigned to 1,2 topics. 42 of the 93 topics are associated with only 10 out of 3299 documents in the test set. All the aforementioned properties make the Reuters-21578 test set a very demanding clustering task.

4.2 Evaluation Criteria

In order to compare our results to the performance of other algorithms, we use the widely known micro-averaged *F1*-measure used, for example, in Cutting et al. (1992) and Massey (2005). The F1-Measure (see Eq. 3) represents a balance between the two Measures *precision* and *recall*. Precision measures the number of correctly retrieved documents in relation to the overall documents retrieved by the approach and thus, the discriminatory ability of the approach. Recall measures the number of documents retrieved correctly in relation to the number of documents that should have been retrieved. Thereby, it is a measure of completeness of the approach. Precision and recall can thus be calculated as depicted in Eq. (1) and (2). The number of correctly assigned documents is denoted by *a* (true positives), and the number of incorrectly assigned documents by *b* (false positives), respectively.

$$Precision = \frac{a}{a+b} \quad (1)$$

$$Recall = \frac{a}{a+c} \quad (2)$$

The *F1*-measure can then be calculated as depicted in Eq. (3). *F1* can assume values from [0;1] with x ϵ $\mathbb{R}$.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (3)$$

To calculate the values for the measures, the set of topics $S = \{S_j \mid j = 1, \ldots, M_s\}$ has to be known. M_s corresponds to the number of topics in the test set ($|S| = 93$ for the ‘ModApte’ Split). During annotation, each document has been assigned by hand to one or more of these topics (including ‘no topic’) by professional indexers.

The output of the prototype presented in this paper consists of a set of clusters $C = \{C_i \mid i = 1, \ldots, M\}$ with M being the number clusters identified by the prototype. In cases where no reference annotation is available, the number of clusters identified by the prototype does not necessarily correspond to the number of real clusters in the dataset. To align our evaluation to Massey (2005), we fixed the number of clusters our prototype will find to M_s for this evaluation experiment.

Moreover, to be able to calculate the global *F1*-measure for the entire solution, we first have to find a matching for the cluster sets C_i and the topic sets S_j. Massey (2005) suggests calculating an individual *F1* value (we refer to as $F1_{ij}$) for each cluster-topic combination and assign a cluster *i* to the topic *j* for which the matching yields the best individual $F1_{i*}$.

Once this sub-problem is solved, we are able to calculate the global, weighted *F1*-measure ($F1_{glob}$) for the entire solution as depicted in Eq. (4)[4].

$$F1_{glob} = \frac{\sum_{j=1}^{M_s} |S_j| \cdot F1_{i*}}{\sum_{j=1}^{M_s} |S_j|} \tag{4}$$

By calculating $F1_{glob}$ this way, the occurrence frequency of the different topics in all documents is considered and the local F1 values are weighted accordingly. Thereby, distortions induced by the occurrence of a high number with only a few documents assigned to it are attenuated.

4.3 Evaluation Results

Table 1 exemplarily shows the top 10 topics from the test set and the number of documents associated with them originally in the data set. Column represents the cluster found by the prototype that is associated with the corresponding topic. Columns four, five, and six show the corresponding precision, recall and the local $F1_*$-value for our solution, respectively.

---

[4] For complexity reasons, we decided to use a greedy approach to find such a matching instead of solving an assignment problem.

Table 1. Top 10 most frequently assigned topics

Topic	Number of Documents	Matching Cluster	Precision	Recall	Local *F1*$_*$
earn	1087	C_{26}	0,93	0,65	*0,76*
acq	719	C_{68}	0,98	0,08	*0,15*
noTopics	280	C_9	0,64	0,28	*0,39*
crude	190	C_2	0,75	0,31	*0,44*
money-fx	179	C_{28}	0,78	0,30	*0,44*
grain	146	C_{16}	0,66	0,43	*0,52*
interest	131	C_{10}	0,63	0,22	*0,33*
trade	117	C_0	0,59	0,60	*0,60*
ship	89	C_{62}	0,73	0,48	*0,58*
wheat	71	C_5	0,25	0,03	*0,05*

The global F1-value for our solution – calculated according to Eq. (4) – is F1glob = 0,41. Figure 3 shows the results of our prototype (denoted as 'Own') in comparison to the three approaches evaluated in Massey (2005). Unfortunately, there is no detailed information available about precision and recall for the different approaches presented in Massey (2005). Therefore, we only are able to compare the global F1-values achieved by the different approaches. For reference, we also included the performance of a supervised reference approach (based on a Support Vector Machine) presented in Massey (2005) to figure 3.

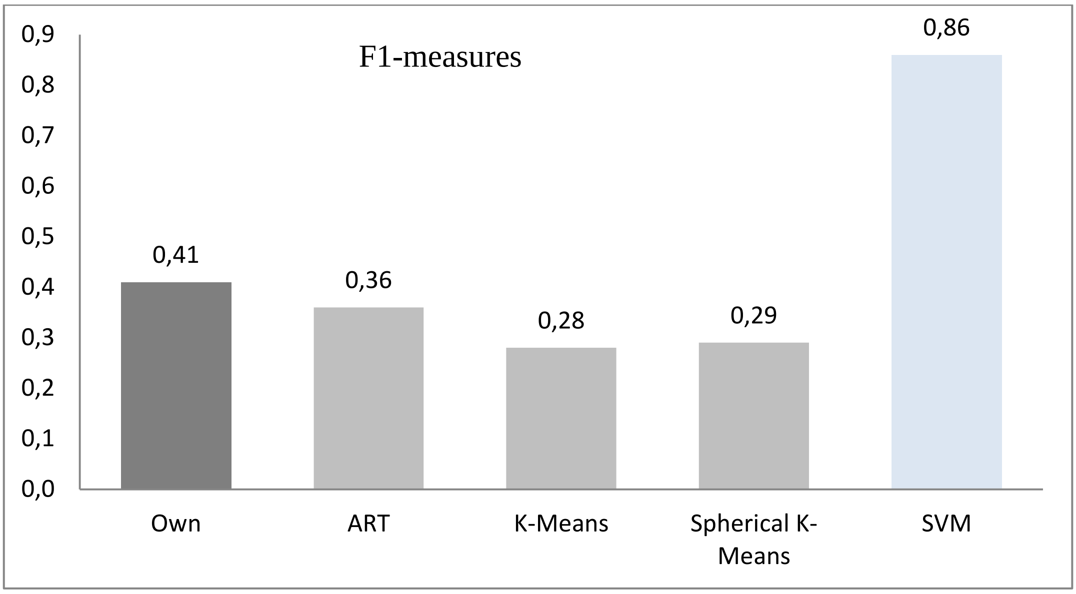

Figure 3. Evaluation results.

4.3 Discussion

The evaluation results suggest that the prototype provides state-of-the-art classification performance for unsupervised text classification procedures on the Reuters-21578 'ModApte' split. It was able to provide better classification solutions than the three approaches evaluated in Massey (2005). The results also suggest a first proof of concept for the functionality of the prototype. However, there are several issues coming along with this first evaluation that need to be discussed. First, provided that a sufficient training set can be acquired for a given application domain, results derived by supervised or semi-supervised approaches (for example, Support Vector Machines, Conditional Random Fields) generally are superior to those produced by unsupervised classification techniques. If such a training set is available, supervised approaches should be used instead of the presented unsupervised approach. Second, using only one test set may be – even if conducted with great care – not sufficient for an elaborate evaluation of the prototype, because there may be a natural compatibility between a particular data set and a certain algorithm for clustering. Thus, in future work, the developed prototype should be evaluated in a similar manner using additional annotated test sets. Third, the Reuters-21578 set is – while widely used and highly valuable – far from perfection. Due to the topic distribution within the set, there is a strong overrepresentation of some topics, e.g., *acquisition* and *earn*. A large part of the classification performance depends on the correct classification of these ubiquitous topics. In addition, some documents lack topics at all and are thus classified by a proxy named ‘no topic’.

Fourth, the presented evaluation is just the first part of the evaluation agenda. In future work, the fit of the approach for the task of identifying interactions in IS project portfolio selection has to be evaluated. Therefore, a domain specific test set could be produced and annotated in a similar to the *Reuters-21578*. Then, the measures and evaluation setup used above could be employed to evaluate the potential of the approach for the mentioned task. If such a test set could not be generated or acquired, the approach could alternatively be used to process and cluster unannotated data from the domain. The results could then be discussed and evaluated together with experts from the field (e.g., experienced portfolio planners).

5 Summary and Future Research

In the literature, there is anecdotal (e.g., Aaker and Tyebjee, 1978) as well as empirical (e.g., Aral et al., 2006; Engelstätter, 2009, ESI International, 2009) evidence for the existence and the business impact of output interactions among IS projects. While a lot of sophisticated optimization models have been suggested which already provide for the consideration of output interactions when selecting IS project portfolios, the necessary data required for their application in business practice usually is not available to the planner. We find a lack of techniques in the literature on how to identify output interactions ex ante to the portfolio selection process and attribute it partly to the rather semantical nature of output interactions. With this paper, we contribute to filling the identified gap by conferring semantic clustering – a technique originating in the text mining literature – to the field of IS PPS. We develop a prototypical DSS that uses LSA and hierarchical clustering to identify potential output interactions among IS project proposals based on semantic similarities within their goal descriptions. This paper focuses on the design of the developed prototype and argues that LSA represents a very promising technique for the identification of output interactions among IS projects. For practitioners, the resulting prototype may serve as a tool to identify output interactions in a structured and potentially more profound way and to include them into their portfolio decisions. We expect our approach to perform particularly well for the identification of competitive output interactions, as this type of interaction seems to be less subtle than complementary output interactions. In addition, the hierarchical representation chosen in this paper may highlight relationships within the organizations project landscape which may have not been recognized explicitly before. For researchers, the presented approach may constitute a starting point to incorporate the identification of output interactions into new or existing approaches. However, to advance this work, several points have to be addressed in future work.

As necessary for design science research, the general applicability of the approach has to be evaluated thoroughly. In addition to the evaluation conducted in this article, we plan to apply the approach to a real world data set of IS project proposals. Therefore, a data set as well as a reference solution from domain experts has to be acquired. As already discussed in section 4.3, additionally the approach could be evaluated by discussing and evaluating it with domain experts in a qualitative manner. This approach could also highlight indication for further incremental improvements.

Further, the development of the prototypical DSS discussed above comes along with several design choices. These choices have to be assessed against numerous alternatives in the future in order to evaluate the applicability of the presented technique to the problem at hand. It has to be determined how the exclusion of domain- and company-specific stop words/phrases influences the solution quality of the approach and how these stop words may be identified automatically by the prototype. In addition, a so called relevance feedback (Dumais, 1991) may be implemented which allows the planner to define, which of the identified interactions are relevant and which can be neglected in a further iteration. The labels of the irrelevant clustering results could be added to the stop word list and be ignored in further iterations. Finally, different stop criteria for the clustering procedure as well as the labeling quality have to be evaluated together with domain experts.

References

Aaker, D.A., and Tyebjee, T.T. (1978). A Model for the Selection of Interdependent R&D Projects. In IEEE Transactions on Engineering Management, 25(2), 30-36.

Aral, S., Brynjolfsson, E., and Wu, D.J. (2006). Which came first, IT or Productivity? The Virtuous Cycle of Investment and Use in Enterprise Systems. In Proceedings of the 27th International Conference on Information Systems, AIS, Milwaukee.

Berry, M.W., Dumais, S.T., and O'Brien, G.W. (1995). Using linear algebra for intelligent information retrieval, SIAM Review, 37(4), 573–597.

Calinsky, R.B., and Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics, 3, 1-27.

Carazo, A.F., Gómez, T., Molina, J., Hernández-Díaz, A., Guerrero, F.M., Caballero, R. (2010). Solving a Comprehensive Model for Multiobjective Project Portfolio Selection. In Computers & Operations Research, 37(4), 630-639.

Cutting, D.; Karger, D.; Pedersen, J.; Tukey, J. (1992). Scatter-gather: A cluster-based approach to browsing large document collections. Proceedings of SIGIR'92.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R. (1990). Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science 41(6), 391-407.

Dumais, S. T.; Platt, J.; Heckerman, D.; Sahami, M. (1998). Inductive learning algorithms and representations for text categorization. Proceedings of ACM-CIKM98, 148-155.

Dumais, S.T. (1991). Improving the retrieval of information from external sources. In Behavior Research Methods, Instruments and Computers, 23, 229–236.

Eilat, H., Golany, B., Shtub, A. (2006). Constructing and Evaluating Balanced Portfolios of R&D Projects with Interactions: A DEA based Methodology. In European Journal of Operational Research, 172, 1018-1039.

Engelstätter, B. (2009). Enterprise Systems and Labor Productivity: Disentangling Combination Effects, In International Journal of Engineering Research and Applications, (forthcoming).

ESI International Inc. (2009). View from the Ground: The Project Manager Perspective on Project Portfolio Management Effectiveness. http://www.esi-intl.co.uk/resource_centre/white_papers.

Fox, G.E., Baker, N.R., Bryant, J.L. (1984). Economic Models for R and D Project Selection in the Presence of Project Interactions. In Management Science, 30(7), 890-902.

Friedman, J.H., Hastie, T., Tibshirani, R., (2001). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edition, Springer-Verlag, Heidelberg.

Hevner, A.R., March, S.T., Park, J., and Ram, S. (2004). Design Science in Information Systems Research. In MIS Quarterly, 28(1), 75-105.

Kuhn, A., Ducasse, S., and Gîrba, T. (2007). Semantic clustering: Identifying topics in source code. Information and Software Technology, 49, 230–243.

Kundisch, D., and Meier, C. (2011a). IT/IS Project Portfolio Selection in the Presence of Project Interactions - Review and Synthesis of the Literature. In Wirtschaftsinformatik Proceedings, 477-486, Zurich.

Kundisch, D., Meier, C. (2011b). A new Perspective on Resource Interactions in IT/IS Project Portfolio Selection. In Proceedings of the European Conference on Information Systems, Paper 174, Helsinki.

Landauer, T.K., Foltz, P., Laham, D. (1998). An introduction to latent semantic analysis, Discourse Processes, 25(2-3), 259-284.

Lee, J.W., Kim, S.H. (2001). An Integrated Approach for Interdependent Information System Project Selection. In International Journal of Project Management, 19, 111-118.

Li, H.; Yamanishi, K. (1999). Text Classification Using ESC-Based Decision Lists. In: Proceedings of International Conference on Information & Knowledge Management, 122-130.

Maletic, J.I., Valluri, N. (1999). Automatic software clustering via Latent Semantic Analysis. In Proceedings of the 14th IEEE International Conference on Automated Software Engineering, FL.

Massey, L. (2005). Evaluating and comparing text clustering results. In Proceedings of International Conference on Computational Intelligence, Calgary.

Milligan, G.W., and Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika, 50(2), 159-179.

Ohata, T., Tateisi, Y., Kim, J., Mima, H., Tsujii, J. (2002). The GENIA Corpus: An Annotated Research Abstract Corpus in Molecular Biology Domain. Proceedings of Human Language Technology Conference.

Santhanam, R., Kyparisis, G.J. (1996). A Decision Model for Interdependent Information System Project Selection. In European Journal of Operational Research, 89, 380-399.

Stummer, C., Heidenberger, K. (2003). Interactive R&D Portfolio Analysis with Project Interdependencies and Time Profiles of Multiple Objectives. In IEEE Transactions on Engineering Management, 50(2), 175-183.

Wolfe, M.B.W, Schreiner, M.E., Rehder, B., Laham, D., Foltz, P.W., Kintsch, W., Landauer, T.K. (2009). Learning from text: Matching readers and texts by latent semantic analysis. Discourse Processes, 25(2-3).