Benefit from Automated Product Recommendations at kdmz

Michael H. Quade
University of Applied Sciences Northwestern Switzerland
michael.quade@fhnw.ch

ABSTRACT
There are a number of recommender systems with different characteristics (Nageswara Rao and Talwar 2008). For a company it is not easy to select the right system. For the introduction of a recommender system is decisive, what kind of data exists and how suitable this data is for calculating recommendations. The question must be answered which algorithm calculates the best recommendations with these data. Afterwards the recommender system has to be developed and integrated into the existing systems (e. g. e-shop). This paper describes the case where a recommender system was introduced successfully into an existing system. For the description of this case the “eXperience Methodology” was used (Schubert and Wölfle 2007). According to this methodology the company is described first, and then some views onto the final solution, the project and development and to conclude the experiences which were made with the solution.

Keywords
Case Study, Recommender System, Recommendations, Collaborative Filtering, Personalization, Quality Tests, Implementation

INTRODUCTION
Since the first publications about recommender systems in the mid of the 90s (Resnick et al. 1994) the interest in these systems is increasing (Adomavicius and Tuzhilin 2005). A recommendation system shall make it easier for a customer to find products which suit him. Products can be films, music, books, news or as in this case study products for the administration in the public sector. Recommender systems are based on any information about their historical preferences of a customer that can be processed in customer profiles. Nageswara Rao and Talwar (2008) inventory an extensive list of available recommender systems which has been developed by the academia or the industry.

But how shall a small company proceed at the introduction of such a system? Which considerations must be made? What are the steps that shall be taken? This paper gives some answers to these questions by describing the case where a recommender system was introduced successfully into an existing ERP system and e-shop.

To structure the case study the “eXperience Methodology” was used (Schubert and Wölfle 2007). Therefore the paper is structured as following: In the first section the company is described which implemented the recommender system in its environment: What is the company’s history, what it is doing and who its customers are. Why the company started the project and who are the partners. In the second section the final solution of the project is described with four views: the business view, the process view, the application view and the technical view. The third section describes the most important project tasks and the operation of the system. The task is described how the recommender algorithm was selected and how the quality of the recommendations was analyzed. The fourth section describes the experiences which were made during the project and with the recommender system in operation. The final section contains a summary of the conclusions drawn from the case.

The author of this paper was involved in this project. He was one of the concept partners, managed the first phase of the project and was responsible for the knowledge transfer. The project started in Spring 2007 and was completed in Summer 2008.

THE COMPANY KDMZ
The kdmz (Kantonale Drucksachen- und Materialzentrale Zürich) was founded in 1903 and since January 1st, 1997 it has been an office of the finance administration of the Canton of Zurich in Switzerland. kdmz carries out cross-section and mediatory functions and it is the central procurement and service office for the cantonal administration. This includes
organizations and institutions that are sponsored by the public hand or where the canton has an involvement (e.g. the University Hospital Zurich). kdmz can be called a procurement service provider for these organizations (Tanner and Wölfle 2002).

The efficient and effective supply of the public administration with indirect goods at cost-covering prices stands in the centre of the business activity. In the year 2007 the company achieved a turnover of about 45 million CHF with 62 full time jobs.

Business enterprises and private households can obtain printing products like legal texts, official forms and other publications of the canton from kdmz.

**Background, Industrial Sector, Products and Target Groups**

The product range covers four areas: Supplies (office material and cleaning agents), printing products (forms, publications, means of communication), capital goods (IT equipment, photocopier etc.) and services (IT and management consultancy). Approx. 1'500 official forms and pamphlets (publications of the canton) as well as 3'500 office material and cleaning articles are included. kdmz operates production centers of its own as a core competence. This includes a modern digital print center for brochures, books and photocopying, and a center for desktop, printing and confectioning works. Increasingly, additional services are offered to the customers as well like consulting, storage, dispatching services up to and including collection, disposal of special products as well as courier services.

In the year 2007 approx. 98'000 orders with about 5'400 customers from the public administration of the canton and approx. 50'000 customers from business enterprises and the private households were processed. 45'000 of the 98'000 orders were submitted via e-shop.

The fulfillment for the customers in the public administrations is carried out directly by kdmz. The goods are delivered directly with daily tours. Exceptions are only made with very large quantities which kdmz cannot deliver with its own vehicles. As a rule, business customers and private households are supplied by postal service.

**The Trigger of the Project**

With business software the kdmz wants to offer its customers an optimal benefit and to support its staff in their processes efficiently. Four years ago kdmz has, therefore endeavored to integrate new functions in its business software. In 2002 functions like customized product ranges, the mapping of procurement hierarchies and processes as well as various self-service functions were implemented in the online shop (Knechtli 2003).

Already in 1998 kdmz laid the foundation for an efficient platform for the procurement of office materials and office systems with the implementation of an e-shop. The services that kdmz provides its customers via e-business were extended continuously since then.

In order to use its customer data better, kdmz participated as an economy partner in the research project PersECA (Personalization of E-Commerce Applications) of the Competence Center E-Business Basel at the University of Applied Sciences Northwestern Switzerland FHNW. Based on existing customer profiles and data, new functions for personalization should be identified, developed and integrated. With these functions kdmz aimed primarily at generating additional benefit for the customers and further promoting the e-shop as distribution channel.

In the described project kdmz was supported by the following partners:

**IT Partner**

Opacc is a software developer and IT full-service supplier for business automation. Opacc has offered comprehensive computer solutions for sale, enterprise resource planning, purchase, service and after-sales service for more than 20 years. Small and medium-sized companies in the sectors trade and service are its main customers. At kdmz, the business software OpaccOne is used for the handling of all business processes in the trade, in the administrative area of the digital printer center and for the e-shop. OpaccOne and its precursors have already been in use at kdmz since 1998.

**Concept Partners**

The Competence Center E-Business Basel (CCEB) at the Institute for Information Systems at the University of Applied Sciences Northwestern Switzerland FHNW has worked on conception and management for cross-organizational use of information and communication technologies in value-adding areas since 1999. The CCEB brought its knowledge about personalization into the project and led it methodically in the analysis phase.
The Information Systems Research Group of the Department of Informatics of the University Fribourg, Switzerland has been engaged for many years with the use of new technologies in e-commerce, such as recommender systems. The Information Systems Research Group participated in the selection of the procedures used to compute the product similarities.

**ADDITIONAL BENEFIT FROM RECOMMENDATIONS**

Product recommendations can already be found in different e-shops in the B2B and B2C e-commerce. These recommendations are often displayed in the detail view of a product under the heading "Customers Who Bought This Item Also Bought" (Linden et al. 2003, McCarthy and Anagnost 2000, Ling et al. 2005). Such recommendations can be generated in two ways: They were either entered into the e-shop manually or were computed automatically with existing data. A combination is just as conceivable because depending on the company’s interest, recommendations computed automatically could be of limited use. A manual correction makes sense if, for example, a product that appears prominently in the computed recommendations is going to be taken out of the product range.

In addition, personal product recommendations can only be given to customers whose needs are well known. If a company has many customers this is not possible with justifiable effort. To solve these difficulties, kdmz has integrated automatically computed personalized recommendations into their applications.

Within the last few years, different procedures with which recommendations can be computed based on existing data were developed. The procedures at kdmz use transaction data (invoice and credit note positions) for the calculation of personalized recommendations. In the following chapters these procedures are explained in the context of the business view, the process view, the application view and the technical view (Schubert and Wölfle 2007).

**Business View and Objectives**

With the recommendations introduced in its business software kdmz wants to generate an additional benefit for all customers. Furthermore, the customers from the administration of the Canton of Zurich shall be persuaded to adopt a core product range defined by the canton (cf. Figure 1).

The public administration of Zurich optimizes their product range in the context of the project "E-Pro curement". About 2’800 office material articles shall be reduced to a core product range of roughly 1’300 articles. This is made primarily to lower costs by slimmer processes and larger purchasing quantities. The offices which have purchased products outside the newly defined core product range till now shall get appropriate products.

All other public companies and institutions in the canton are not affected by this optimization of the assortment. Furthermore, they can order products from the complete product range. Unlike the public administration they are not obliged to purchase from the kdmz.

![Figure 1. Business Scenario for kdmz](image)
Process View

Before recommendations can be displayed in the e-shop, they must be computed by the system. In the process view the calculation and the use of recommendations are regarded as separate processes.

In principle, the automated calculation of product recommendations is based on existing data. The recommendations are computed in one or several steps, according the mathematical algorithm is used. Depending on the kind of available data different formulae are suitable. The following always applies: The more data are available for calculation the more exact the recommendations become (Sarwar et al. 2000 and Stormer et al. 2006).

The most reliable recommendations are provided by formulae which are based on product ratings as they can be entered by customers on many e-shops: The ratings of the single customers are mathematically compared with each other. The results are product or customer similarities on which the recommendations can be based. The similarities are expressed with numerical values: e.g. with values between 0 (weak) and 1 (strong). A similarity of against 1 would mean that the two compared products are almost identical in their ratings by the customers. If a value is NULL, the value for the similarity is empty (Lemire and Maclachlan 2005).

If ratings are not available, the completed purchases (transactions) can also be considered as rating appraisals. It is assumed that the customers like an article they have purchased if they do not return it after the purchase. Procedures which are based on such ratings are called "Collaborative Filtering" procedures (Peppers and Rogers 1997).

Since no ratings can be given for the products in the e-shop of the kdmz, the transaction data from the sales are used for the calculation of the recommendations. For these calculations kdmz uses the procedure of Deshpande and Karypis (2004) (cf. chapter Development and Roll-Out). The procedure supplies the product similarities between two products in the first calculation step and personalized recommendations in the second step.

Once a month the sales data (invoice and credit note positions) and product data are queried from OpaccOne for the first calculation step (cf. Figure 2). Each time the sales data contain all of the orders of about four years – the current year plus the last three years. From approximately 400'000 data records with an extent of roughly 4'500 products about 20 million similarities are computed.

Figure 2. Calculation process for recommendation

These product similarities are stored in the OpaccOne Business Server for the second calculation step. The similarities would already be suitable to express general, non-personalized recommendations. In the e-shop, for example, an anonymous guest can see similar products as a recommendation in the detail view of a product.

In the second calculation step the product similarities are used to prepare automatically a personal list for defined customers with the 50 best product recommendations (Weng and Liu 2004). These lists are generated by stored queries and routines.
Quade  Benefit from Automated Product Recommendations at kdmz

Per product the similarities to the products already purchased by the customer are added (Weng and Liu 2004). Table 1 shows an example: We assume that a customer has purchased the products i, j and k, but not the products x, y and z. The added similarities yield to the ranking of the recommendations for the customer. The product with the highest value represents the best recommendation. In this example it is product y

<table>
<thead>
<tr>
<th></th>
<th>product x</th>
<th>product y</th>
<th>product z</th>
</tr>
</thead>
<tbody>
<tr>
<td>product i</td>
<td>0.4</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td>product j</td>
<td></td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>product k</td>
<td>0.6</td>
<td>0.4</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Example of the calculation of personal recommendations

When personalized lists of products are built the deposited customer product range is taken into account or the products are replaced with alternatives from the core assortment, depending on the configuration in the customer profile.

Now the recommendations can be used in the sales talk with the customer, e. g. on the telephone and for the advertisement in the e-shop (cf. Figure 3).

Application View

OpaccOne has a layered software architecture and is designed according to the concept of a service oriented architecture (SOA) (Sanders et al. 2008). For all applications the "Business Server" is central. It provides data as business objects and fundamental functions as business services to further integrated applications. The most important further applications are divided up into the parts of “BackOffice” and “FrontOffice”. (cf. Figure 4).

The BackOffice applications support internal users with common ERP functions. At kdmz these are the applications “Sale”, “Enterprise Resource Planning”, and “Management Information System”. The FrontOffice applications are used for the interaction with customers. In use is the application “WebSales” which is based on the OpaccOne “Web Portal”. It offers the customers a variety of self service functions besides the common e-shop functionalities.

For some important customers kdmz supports entire procurement processes. For this purpose the B2B e-shop contains extended functions, for example, for managing the purchasing employees and their procurement budgets by an administrator of the customer or for releasing orders by supervisors.

The self-service functions in the e-shop enable the customer to decide himself whether recommendations are shown to him or not. Per default, recommendations are activated for all customers. With the help of the customer profile in the Back-Office
applications it can be further personalized where in the e-shop the recommendations should be displayed: Everywhere in the product catalog, only in the detail view of an item, only in the shopping basket or in all mentioned places. It is activated as default for all customers that recommendations are shown in all mentioned places.

The calculation of the product similarities is an application which is based on the Business Server. After an initial configuration the application runs fully automatically without any further user interactions.

![Diagram of OpaccOne: Business Server with integrated applications](image)

**Figure 4: OpaccOne: Business Server with integrated applications**

**Technical View**

For the automatic calculation of the recommendations investments in new hard-ware were not necessary. The application "Product Similarity" could be installed on a virtual computer on the existing systems. The technical view, therefore, describes the existing hardware and software environments at kdmz (cf. Figure 5).
The largest part of the systems is operated by kdmz in-house. The systems for the FrontOffice applications are outsourced to the Opacc Hosting Center in Kriens. The FrontOffice applications are connected to the systems of kdmz over a leased line. The access to the e-shop by staff and customers is carried out via the Internet. The customers in the public administration access the e-shop via the network of the Canton of Zurich and the Internet.

Central systems like the Business Server run on redundant servers. Microsoft operating systems are in use on all systems according to the purpose. Windows Server 2003 is installed on servers and Windows XP Professional on clients. Batch jobs, background jobs and smaller applications are primarily executed on 12 virtual computers. Resources like CPU performance or memory can be assigned dynamically to each of the virtual computers as required.

PROJECT AND OPERATIONS

In the context of the FHNW research project PersECA the first workshops for creating ideas took place in spring 2007. The objective of these workshops was to find and to draw up new functions for the personalized customer dialog. Supported by the university, selected staff members of kdmz identified 36 optional functions. Based on the results of a value analysis, ten of them were followed up and worked out in detail. The solution introduced in this case study is based on three of these functions:

- To provide personalized recommendations for the customer.
- To support the streamlining of the core assortment with the help of recommendations.
- To be able to look at the personalized recommendations for the preparation of sales talks.

Development and Roll-out of the Software Solution

Before the software for the calculation of the product similarities was implemented, it was tested with four different collaborative filtering algorithms if the transaction data were suitable. Depending on the procedure, particular factors have more or less influence on the quality of the recommendations. Factors are e.g. the number of purchases or the ratings of a product, the number of customers, who have rated or purchased a product, or the number of the different products and customers.

Quality Tests

The calculations of the quality tests are based on the matrix of customers with products that was generated from the transaction data. Using the number of individual purchases and joint purchases of products a similarity is calculated between the products.

The calculation takes into account the frequency of the purchases with a constant that highly purchased products get weighted less strongly than less purchased.

Before the calculation the issue has to be settled what time period shall be used. At kdmz transaction data is available since 1999. The following facts were taken into the considerations to use not all the data:

- The transaction data are partly archived and therefore in the live systems not available.
- The more data is used, the longer the calculation of the recommendations takes.
- Too old data can distort the result because customers normally change their purchased product range with the time.

For the check of the listed points an analysis of the transaction data was carried out. The calculation of the recommendations was extended step by step with one year of transaction data (cf. Figure 6 and Figure 7).

The quality was measured with 500 to 1’000 by randomly select purchases that were removed from the transaction data before the calculation. The idea behind removal of purchases is that good algorithm calculates product recommendations which represent the removed purchases. The higher the hit rate from recommendation on removed purchases, the better the procedure is. To calculate an average hit rate the quality measuring was carried out approx. ten times per algorithm. It was
examined, considered besides the suitability of the procedures that time period provides good recommendations to transaction data: For four years back turned out for kdmz to be ideal.

In the first step a test matrix was prepared with 500 by randomly select purchases from all data (2007-1999). According to that the calculation was started only with the transaction data from the year 2007. For the further calculations year by year was added gradually.

The result of this calculation can be seen in Figure 6. One recognizes that for all three used similarity algorithms (Deshpande and Karypis 2004, Jaccard 1901) the quality of the recommendations increases if more data is available. In addition the algorithm planned for kdmz – the one from Deshpande and Karypis – has a higher hit rate than the two others.

![Figure 6. Result of the analysis with a test matrix of 1999 – 2007](image)

A second attempt should analyze particularly the third bullet point in the list above. If the preferences of the customers changes over the years, the quality of the recommendations is decreasing with older data. For the check a test matrix was prepared exclusively with purchases from the year 2007.

![Figure 6. Result of the analysis with a test matrix of 1999 – 2007](image)
Figure 7 shows the result of the analysis. It is recognizable that the hit rate remains relatively constant. A slight decrease of the hit rate can be recognized with the Deshpande and Karypis (2004) similarity after about four years.

To mention that at the analysis of older data new products appear only in the list of recommendations since they were bought at least once. If for example the data of the year 2007 is added to the data 2006 and 2005, the product range increases by approx. 300 products.

The analysis has turned out that the transaction data of the last four years (in the example the data of the years 2004 to 2007) for a calculation is a good initial set up.

The quality of the recommendations calculated with the transaction data from kdmz can, however, be measured differently. There are two positions:

- Measuring the quality of a recommender system from a customer perspective: A good recommender system is one that recommends products what the customer likes.
- Measuring the quality of a recommender system from a company perspective: A good recommender system is one that increases turnover and profit. Moreover, with a web application can be investigated, what recommendations were shown to the customer and which he has selected and bought to conclude (see chapter indicators below).

Roll-out

A bigger effort was required for the roll-out for the handling of customer profiles and product data. The customer profiles had to be adjusted at the beginning, considering whether recommendations may be shown or not. Some customers could only order from a restricted assortment at this time. Custom-designed products and accessories to certain products (e.g. toner for laser printers) should not appear in the recommendations. In particular, it should be avoided that accessories are recommended to equipment which they do not fit. Such products were blocked for the use as recommendations.

Regular Maintenance

Due to the experiences with this roll-out, the processes were extended for the entry of new customers and new products. When entering a new product, for example, it has to be decided whether the product shall be blocked for recommendations or not. Accordingly, it has to be specified for every new customer whether she/he shall receive recommendations or not.

For the mentioned project "E-Procurement" alternative products are entered manually to all existing and new products, which are no longer part of the core product range. This classification was already made for a part of the products.

The calculation of the product similarities themselves takes place automatically. Apart from the usual operating controls by the system administrators, the application does not need any maintenance.

EXPERIENCES

The project was completed successfully and the recommender system in online in the systems of kdmz (www.kdmz.ch, access only for registered users). The first experience with automatically computed recommendations indicates that the objectives mentioned and the expectations were achieved. An obvious benefit emerges in the preparation of sales talks because recommendations provide new aspects of a customer.

Internal User Acceptance

kdmz staff members were included in the discussion of the subjective assessment of the quality of the computed recommendations. It was noticed that custom-designed products like forms and envelopes turn up as recommendations. These products were blocked for the use as a recommendation. After an initially skeptical attitude of the staff members towards recommendations the experiences are positive today. The staff is convinced of the quality and the benefit of the recommendations.

Indicators for Customer Acceptance

For the quality measurement of the recommendations, e.g. it is evaluated how often customers have clicked on shown recommendations. Further it is evaluated how often customers opted out the option "Show recommendations" in the self
service functions of the e-shop. Since in July 2008 the first solution has gone into operation, the following indicators have emerged:

- Click rate to "shown recommendations in the product catalog": 0.59%
- Purchase rate to "shown recommendations in the shopping basket": 1.60%
- Number of customers who want to see no more recommendations: 0

The values are high in the comparison with the click rates of 0.18% on advertising on the Internet in Europe (ADTECH 2007). The quality of the recommendations is therefore good. Since no customer has opted out of recommendations, they are apparently no nuisance to the customers.

CONCLUSION

For kdmz and Opacc it was the first project for the introduction of a recommender system. The IT partner surely has had efforts in this project which will not be charged on this scale in another similar project any more. The developed application can almost invariably be introduced with another customer. For another company which would like to introduce a recommender system in its environment, the effort can be hardly reduced. Since requirements and available data are different at every company.

Therefore a competitive advantage results for kdmz from the fact that automated recommender systems are still rarely to be found in the business software of Swiss SMEs. The competitors of kdmz are not using any recommender systems yet. Replacing computed recommendations by alternatives opens up another competitive advantage. Replacing can be used for customers who are not subject to the procurement rules of the canton. These customers can be acquainted with new products or products that kdmz wants to promote.

REFERENCES