Published in: L. M. Hilty, X. Edelmann und A. Ruf (Hrsg.), 2007:
Proceedings of the 8th World Congress on Integrated Resources Management,
Davos (Switzerland) 3.-5.9.2007, CD-ROM , ISBN 978-3-905594-49-2

Metadata Driven Acquisition, Search and Procurement of Secondary Materials within an Internet-based Marketplace

Jörg Bremer[1], Barbara Rapp[1], Ute Vogel[2] and Michael Sonnenschein[2]

[1] OFFIS, Institute for Information Technology, Escherweg 2, D-26121 Oldenburg, Germany

[2] Carl von Ossietzky Universität Oldenburg, Dept. of Computing Science D-26111 Oldenburg, Germany

Speaker: Jörg Bremer

Corresponding author: Jörg Bremer, joerg.bremer@informatik.uni-oldenburg.de

Abstract

We describe the design of an internet-based marketplace for secondary materials, which extends the scope of mere blackboard systems by employing a flexible way to integrate meta-descriptions into the marketplace. In this way the descriptions of the materials can be automatically processed and facilitate further use cases, such as automated procurement of supply and demand or customized search functions for specific residue classes. In this paper, we present our proof-of-concept, with focus on metadata driven acquisition, procurement and search, which supports the opportunity of a self organized formation of recycling networks.

Keywords: recycling networks, internet-based markets, meta-descriptions, procurement, secondary raw materials.

1 Introduction

Apart from desired products, most manufacturing processes result in unwanted by-products. A recirculation of these undesired products by using them as input for other processes often works for bigger companies which produce a sufficiently large range of goods, but rarely for small and medium-sized businesses (SMBs) that need an inter-business network for recycling. Thereby, SMBs usually dispose of by-products as waste. The opportunity of converting by-products to secondary materials can relieve dumping grounds and achieve a reduction of limited raw materials by using secondary instead of primary materials.

Waste exchanges are an already existing opportunity for re-use of secondary materials. According to (Rautenstrauch 1999) these waste exchanges are only suitable for short-term intermittent trade relations. Building up regular material flows as a better alternative requires the formation of a (regional) recycling network. Recycling networks are characterized by purposive collaboration of different businesses with complementary input-output-dependencies. Implications of the required size of the region as well as the number and variety of actors can be found in (Sterr and Ott 2004). To develop the full potential of recycling networks comprehensive information on material flows is required. Nevertheless, the underlying problem remains: How to find a business with a complementary range of products?

At this point, an internet-based marketplace for secondary materials seems appropriate if it is capable of interpreting the captured data. We present the design of such an electronic marketplace where by-products can be offered, secondary materials can be requested, and the underlying application is able to procure between supply and demand. In this way, the opportunity emerges for a self organized formation of recycling networks.

2 Problem statement

Most conventional internet based systems for the exchange of secondary materials (e.g. http://www.waste-exchange.org/) act as blackboard systems like their predecessors in print media. They commonly allow for the posting of (mostly free-) text based advertisements without structured annotations of the data entered. Due to the unstructured description of offered goods, the search for suitable advertisements can only be done manually by text-based searching. Hence, complex queries, which involve more qualified conditions, can not be formulated and, due to homonyms, synonyms, and typing mistakes the search returns only a subset of the suitable offers. The same holds true for requests, where the quality of suitable offers depends on the intended purpose, especially on the production process, which requires these materials as input.

In the case of secondary materials, a qualitatively sound description of offered goods is fundamental: on the one hand, due to their production, materials for recycling are usually not "pure" materials but a mixture of diverse substances. If used goods are traded for re-use on the other hand, the qualities of these goods and their parts have to be described in detail. According to (Baumgarten et al. 2003), a standardised data handling leads to more reliability in such cases. The structure of the quality description, i.e. the number and type of attributes, which are necessary, depends on the kind of material, its processing history as well as its possible intended usage.

Besides good conformance in the quality of offer and request, two other aspects of a successful match between supply and demand are important:

- The transportation of goods is a cost factor that can be determining whether or not a match is acceptable. In order to minimize transportation cost, each match should be qualified with the necessary transportation distance, so that near-by matches can be preferred over matches with high transportation costs.

- Secondary materials can be materials with a "Best-Before"-Date. So, time constraints should be part of the description. Time constraints also involve the fact that secondary materials are produced repeatedly.

The objective of the work presented here, is to design an electronic marketplace for secondary materials which supports these needs. It allows for the configuration of arbitrary, residue-specific kinds of descriptions by meta-descriptions, the collection of advertisements, i.e. supply and demand, which obey the respective meta-description, a manual search for matches, but also an automatic procurement based on arbitrary attributes of the description.

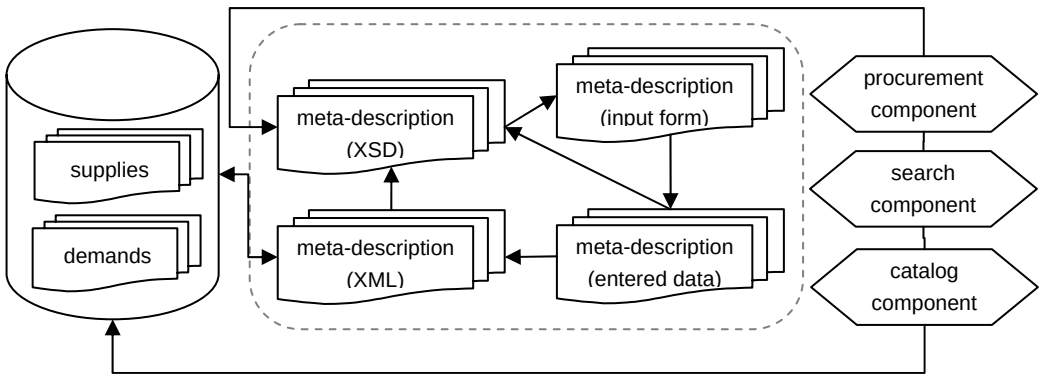

***Fig. 1**: Meta-descriptions as basis for the electronic marketplace*

This electronic market has been prototypically implemented as a proof of concept. To achieve the objectives in a cost effective way and to allow for unrestricted and easy access, a component based web application architecture (Reussner and Hasselbring 2006) has been chosen, allowing for easy and secure design-by-contract based extensibility and thus at the same time serving as an ideal test-bed for the development of future components in the field of support for recycling networks. As a cost-effective open source implementation, our proposed platform is geared to SMBs, but, in the long term, is also aiming at an integration of third party players such as transport or conditioning companies.

Henceforth, the term *residue* is used as a synonym and abbreviated notation for secondary materials and the term *advertisement* is used as a collective term for offers and requests.

3 Meta-description of residues

In our marketplace, metadata is of utmost importance, since it allows for an adequate description of different classes of residues on the one hand as well as automatic post processing on the other. To describe each residue properly we use an XML-based meta-description system. This allows for generation of user-input forms and validation of the user input. Both ensure the homogeneity of related residue descriptions and hence automatic processing.

Our implemented residue management provides assistance for capturing, editing and managing a user's residue related data stock. In addition, repository functionality is offered by the system. This component receives a residue's meta-description from the metadata component by request as input, generates suitable input and editing forms for this residue and ensures a persistent storage in the repository. A taxonomy for residues provided by the market's catalog supports a hierarchal structure of the stored residues.

The market's residue administration has been designed in tight interaction with the meta-description component (Rapp 2005). This means that attributes needed for residue administration are solely provided by the residue component and all attributes required for meta-description are sourced out to the meta-description component. This implementation poses several advantages: On the one hand efficient database queries such as obtaining all offers by advertiser or similar become possible at low cost. On the other hand, it offers the possibility for flexible meta-descriptions that are extensible and exchangeable and allows the integration of such descriptions into the electronic marketplace. The residue administration presented here guarantees adequate and correct data capture by guiding the advertiser through the capturing process and making sure that all registered advertisements are compliant with the corresponding meta-description. This ensures both local and global homogeneity of residue descriptions.

The metadata administration component offers a series of characteristics that are qualified for describing each residue. Contrary to administrative residue attributes, descriptive residue attributes are divided into two classes: those that are exclusively adapted for describing residues from one special residue class, and those that are more comprehensive and qualified for describing residues from each residue class. A suitable data structure that meets the requirements for distinct residue classes, each with attributes in arbitrary number and nuances, is the extensible markup language (XML) in conjunction with XML-Schema (XSD). Both are semi-structured, support data self-labeling and provide the required flexibility.

The implemented component provides different schemata of which exactly one schema covers the universal attributes that can be used for describing all residue classes. Alongside that schema, many residue class specific schemata exist that are appropriate for exactly one

residue class. Every schema allows for a dynamic validation of the entered parameter values towards the expected parameter values defined by the schema itself. Each schema also contains information about how to execute a mapping or matching. By means of this information, it is possible to decide what attributes have to be compared in order to ascertain the identity, similarity or an existing includes relationship between one or more residues. In addition to offering information for residue matching, each schema assists the component in executing matching. Included information for visualisation, e.g. for an HTML representation, allows for generating the desired input forms (e.g. wizard-style) needed for dynamical data capturing at runtime. All input forms generated at runtime are tailored for a specific residue class, because the information needed for displaying the input form is extracted from the specific residue class schema.

The implemented metadata administration allows for the uploading of new residue meta-descriptions at runtime followed by automated translation and compiling. Such a modus operandi provides an opportunity for adding new residue descriptions to the electronic market at runtime and using them directly without having to restart the system. By this process, the new description is added to the taxonomy simultaneously. This allows for extending the system easily.

4 Browsing and searching residues

In a market with different categories of goods and, moreover, distinct descriptions of these goods, it is important to have discriminating means in support of searching. Finding appropriate input for bio-fuel production for example requires other search criteria than querying for metal recycling.

A data independent search which is specific to each class of residues was realised by an XML-based description scheme for defining such search modules. Such an approach allows for an on-the-fly integration of new search modules, whenever a new residue class description is added to the markets taxonomy. Also, a search module for other information is feasible if new components are added.

We denote this component based concept as *search-mandate principle*. This means that a so-called *mandator* defines a search. The input of search parameters, the execution of the search and the presentation of the search results are delegated to the search component. In this way, the search functionality needs to be implemented only once, and the user can always work in a homogeneous and consistent way. To this end, the mandate description provides information about the mandator and its data objects. Information about these objects includes information about which attributes define the search criteria and how to display them. The same applies to the result. An object instance which is used for holding the entered search parameter values indirectly provides information about data format and constraints regarding acceptable values using the underlying metadata. From the information contained in a search-mandate, proper input forms can be generated.

For faster execution, a search index is administered by the search component whereby the consistency between index and data stock is maintained by means of aspect oriented programming (Elrad et al. 2001), i.e. data access calls are delegated ‘through the back door’ to the search component following the cuckoo’s egg design pattern (Miles 2004): calls are made by the search on behalf of the mandator while the index is appropriately updated simultaneously.

In order to support queries with geographical restrictions, i.e. limiting the resulting set to one's own vicinity or to a given region, each address entered into the system is associated with a geographical reference using the zip code.

A geographical representation also serves as an additional orientation for the user while browsing through the advertisements. The catalogue component presents the information from residue meta-descriptions and shows its spatial position in a GIS map. In addition, the catalogues taxonomy forces the existence of an appropriate meta-description in order to assure a consistent data capture and data representation by the catalogue.

5 Automatic procurement of supply and demand

Manually exploring an information system of supply and demand advertisements does not guarantee a distinct identification of dependencies and thus no 'optimal' matching between supply and demand, neither with respect to single advertisements nor for the formation of a recycling network between users of the system. Meta-descriptions can also contain annotations for an automated mapping between corresponding supply and demand. These annotations allow for an autonomous procurement, which associates suitable advertisements with eachother and returns groups of suitable goods. A match is called *perfect*, if supply and demand match with respect to their identifying features. If they differ in some features, e.g. the offered amount, or agree only in attributes which do not allow a reliable match, the match is called *imperfect*. This kind of matching is accomplished by means of calculating a weighted sum of Levenshtein distances (Kruskal 1999) of each assigned value of attributes that have been annotated as suitable for imperfect matching. A threshold value defines if two goods are considered (imperfect) suitable.

A second concern of the procurement component is to pay attention to different and sometimes opposed interests. This means that procurements can be achieved according to multiple economical and ecological criteria. For example, on the one hand, the total transportation effort has to be minimized to serve public interests, i.e. less CO_2 or noise emission. On the other hand, salesmen tend to maximize their profits. Such a Pareto optimal allocation of supply to demand is achieved by a procurement process consisting of three steps:

- First, the set of all advertisements from all residue classes is divided into a set of sets of suitable advertisements. During this clustering, sets which exclusively consist of offers or requests and hence are not placeable are rejected for the following procurement.

- As a result of the previous step, the problem of procuring one set of distinct items from different residue classes has been broken up into a set of smaller, classical transportation problems which can be solved by means of well known algorithms. Here we make use of Vogel's approximation (Reinfeld and Vogel 1958) which does not guarantee an optimal solution, but often obtains one for small to medium sized problems, which can be expected here. For further improvement on the solution's accuracy, the algorithm can easily be replaced by more sophisticated ones or be taken as an initial solution for exact algorithms for refinement. For each transportation problem, the procedure is rerun with a particular objective function to address various environmental and economical interests.

- The second step leads to a specific set of approximate solutions for each transportation problem representing the minimal costs of each optimized objective, each one leading to another distribution according to the respective objective. To achieve a sustainable compromise for all involved parties, a game theory approach has been chosen to find a feasible overall solution. The approach can be found in (Bremer 2005).

This procurement procedure is offered to the markets users as a special service. Every advertisement that is registered with this service takes part in this periodical procedure that periodically results in a submission of proposed co-operations between users by email, personal message board or downloadable spread sheets.

6 Conclusions and Further Work

Applying metadata to the descriptions of residues traded on an internet based market creates new potentials for an integration of service added values. In contrast to currently available blackboard systems, a better support for the input of advertisements, a refined quality assurance, and, above all, an automated procurement of supply and demand respecting various optimization criteria can be achieved. Our prototype demonstrates that these are promising techniques towards enriched internet exchanges for secondary materials and thus may have the potential to support even closed loop networks in the future. Investigations on the stability of and fluctuation within such a network should accompany field test of such systems.

Since it is unlikely to cover the broad range of distinct residue classes in just a single market, it has been our intention, to design and implement a system that allows to be customised to various data standards. Our prototype demonstrates the feasibility of the integration of residue class specific meta-information. In the future, we intend to customise the system for a specific market to analyse additional user's requirements.

References

Baumgarten, H., Butz, C., Fritsch, A., Sommer-Dittrich, T. (2003), Supply chain management and reverse logistics – integration of reverse logistics processes into supply chain management. approaches. IEEE International Symposium on Electronics and the Environment, 19-22 May 2003 Page(s): 79 - 83

Bremer, Jörg (2005), Modellierung und protoypische Realisierung von Komponenten für die Datenexploration und Reststoffvermittlung bei der kooperativen Konzeption eines Sekundärrohstoff-Marktplatzes, BSc Thesis, University of Oldenburg, Dept. of Computing Science

Elrad, T., Filman, R. E., Bader, A. (2001), Aspect-Oriented Programming. Communications of the ACM. Vol. 44, n. 10, October, 2001.

Kruskal, Joseph B. (1999), An overview of sequence comparison. In David Sankoff and Joseph Kruskal, editors, Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, chapter one. CSLI Publications, 1999.

Miles, Russell (2004), AspectJ Cookbook, O'Reilly Media, Inc., Sebastopol, 2004.

Rautenstrauch, Claus (1999), Betriebliche Umweltinformationssysteme, Springer-Verlag, Berlin Heidelberg.

Rapp, Barbara (2005), Modellierung und protoypische Realisierung von Komponenten zur Interaktion zwischen Unternehmen und Marktplatz bei der kooperativen Konzeption eines Sekundärrohstoff-Marktplatzes, BSc Thesis, University of Oldenburg, Dept. of Computing Science

Reinfeld, N.V., Vogel, W.R. (1958), Mathematical Progamming, Prentice Hall, Englewood Cliffs, New Jersey.

Reussner, R., Hasselbring, W. (2006), Handbuch der Software Architektur, Dpunkt Verlag

Sterr, T., Ott, T. (2004), The industrial region as a promising unit for eco-industrial development – reflections, practical experience and establishment of innovative instruments to support industrial ecology. Journal of Cleaner Production 12(2004) 947-965.