Eco-MCPS: a Newly Developed Web-Based Database for Eco-Materials in Japan

Riko Ozao\textsuperscript{1}, Takashi Sawaguchi\textsuperscript{1}, Hideya Ishida\textsuperscript{2}, Masatoshi Iji\textsuperscript{3}, Teruo Furuyama\textsuperscript{4}, Yoshikazu Shinohara\textsuperscript{5}, Katsutoshi Yamada\textsuperscript{5} and Kohmei Halada\textsuperscript{5}

\textsuperscript{1}SONY Institute of Higher Education, Atsugi 243-8501, Japan
\textsuperscript{2}Waseda University, Graduate School of Environment and Energy Engineering, Honjo 367-0035, Japan
\textsuperscript{3}NEC Corporation, Eco-Material, Nano Electronics Research Laboratories, Tsukuba 305-8501, Japan
\textsuperscript{4}Nippon Steel Corporation, Tokyo 100-8071, Japan
\textsuperscript{5}National Institute for Materials Science, Tsukuba 305-0047, Japan

A novel Web-based database on eco-materials was constructed by the database subcommittee of Ecomaterial Forum. Enquiries were made to companies to obtain about 800 product data, and the data were organized into the Web-based database: “Eco-M.C.P.S” (more simply denoted as Eco-MCPs), which was implemented by PostgreSQL (RDBS) and PHP (server side script). Eco-materials and related products can be searched by the product category or the environmental category. Some products are indexed by key categories, and are related with each other by the key, which makes the products comparable or associated with other products. The products or companies can be searched in the search page by typing in case-sensitive free keywords, such as “RoHS”, and the search can be refined by adding product categories or environmental categories. Searches can be made for eco-labeled products. Association rules which may reflect social interest level on ecomaterials were found by data mining the access log data of the database. [doi:10.2320/matertrans.MK200710]

(Received August 23, 2007; Accepted October 23, 2007; Published November 25, 2007)

Keywords: Web database, ecomaterial, environmental categories, data mining, association rule

1. Introduction

The term “ecomaterial” was created in Japan by Prof. Ryoichi Yamamoto in 1991. Ecomaterials were first defined as materials capable of improving the environment throughout their life cycle by six factors:\textsuperscript{1} (1) avoiding and/or reducing the use of non-renewable resources; (2) enhancing the material closed loop by recycling and reusing waste; (3) increasing resource efficiency; (4) using more durable materials with fewer maintenance requirements; (5) promoting the use of renewable resources and energy; and (6) minimizing adverse impacts on biodiversity and eco-systems.

This concept involved not only material scientists and engineers, but also economists and social scientists to work together. In 15 years, not only considerable advances were seen in technology, but also methodologies and tools had been developed.

Ecomaterials Forum (since 2005) has set an ecomaterials guidelines project. More specifically, ecomaterials should themselves be materials with optimal physical and/or chemical properties and good technical performance; moreover, they should provide significant environmental improvement, and no tradeoff of the environmental burden throughout the whole life cycle.\textsuperscript{2} The entire life cycle impacts of ecomaterials must be considered. Furthermore, ecomaterials must be concordant with at least one of the six vectors\textsuperscript{3} which characterize the ecomaterial: (1) high environmental treatment efficiency; (2) high recyclability; (3) minimal hazardous substances, (4) high productivity, (5) minimal environmental impact production process, and (6) green environmental profile.

Scientists and engineers have been working on the advancement of sustainable technologies, and considerable knowledge and technology are now accumulated in companies’ websites\textsuperscript{4} or environmental reports. However, these data are scattered and are not categorized. Eco-Products Directory\textsuperscript{5} compiles product data for ecomaterials, but is available only in printed version or CD-ROM.

The database subcommittee of Ecomaterials Forum decided to develop an web-based database for ecomaterials: “Eco-M.C.P.S” (denoted simply hereinafter as “Eco-MCPS”). This database is easily accessed by public, and is based mostly on published facts provided by the companies. In order to make the materials’ function more easily understood, products and components using the ecomaterials are included in the database and linked by keys. Furthermore, in this site, analysis and data mining can be made on the accumulated access log data.

The aim of this paper is to describe the unique futures of Eco-MCPS, which has been constructed on a concept completely different from other materials databases (see the reference links listed in http://eco-mcpsdb.sntt.or.jp/v3/link.php), and to show the potential usefulness of Eco-MCPS as a tool for analyzing social (public) interest level on ecomaterials.

2. What is Eco-M.C.P.S?

Eco-M.C.P.S Database (Eco-MCPS) includes product data not only for Eco-Materials, but also for environment-conscious Products, Components, and Services using the Materials (which are referred to as Eco-M, Eco-P, Eco-C, and Eco-S, hereafter). The Ecomaterial Forum database subcommittee has discussed the basic concept of the database, and has decided to make it open to public. For the data acquisition, paper-based enquiries were made to companies. The enquiry paper consisted of two parts: description of the product (Eco-M, Eco-P, Eco-C, or Eco-S) and the categories of the product (concerning the environmental impact reduction, the environmental performance, and the distinguished life cycle stage) were asked in the first part, and the vector(s) which characterize the product as an ecomaterial was/were
asked in the second part. As a result, about 800 replies were obtained for the first part. The database now contains 198 facts for Eco-M, 136 facts for Eco-C, and 466 facts for Eco-P. Furthermore, about 200 data were accumulated for the second part. Figures 1 to 3 show the breakdown of each of the categories. Figure 1 shows that about one third of the eco-products are from the electric and electronic field (home appliances and OA/IT equipment), and that about 20% are machines and equipment. The breakdown given in Fig. 2 shows that about one fourth are from electronic parts and 20% are automobiles. It is assumable from those results that about 60–70% of the eco-products (components) are related to electric and electronic industries and automobiles. Similarly, in Fig. 3, polymers and metals, which each account for 37% and 25% of the eco-materials, are closely related to biopolymers and light-weight structural materials, respectively, for use in automobile and electric and electronic equipments.\(^6\)

2.1 Construction of Eco-M-C-P-S
The Web database server was implemented using PostgreSQL (free open source) using PHP as the scripting language. This relational database comprises a product table, a company table, an environmental category table, etc. (Table 1).

Figure 4 shows the top page of the original version of Eco-MCPS (the latest version is now available at http://www.eco-mcpsdb.sntt.or.jp/v3/). The constitution and the transition of pages are as in Fig. 5. The products can be accessed either from the product category or the environmental category (from the viewpoint of environmental impact reduction, the environmental performance required for resolving the problems, or from the life cycle stages of interest). The environmental impacts considered are: (A-1) global warming, (A-2) air pollution, (A-3) hazardous substances, (A-4) wastes, and (A-5) resource consumption. The
environmental performance is categorized by (B-1) recyclability, (B-2) longer life, (B-3) resource saving, (B-4) higher performance, (B-5) energy saving, (B-6) environment cleaning, (B-7) use of reproduced materials. Concerning the life cycle of the product, six stages as follows are set: (C-1) extraction, (C-2) material and parts preparation, (C-3) product design, (C-4) production, (C-5) servicing, and (C-6) waste. Some products are related to each other by providing "keys"; for instance, if cell phone is categorized as OA/IT equipment, and if the ecomaterial used for the casing is a bio-polymer usable for other components such as computers, this product is linked to "housing" of OA/IT equipment.

2.2 Keyword search

Searches can be made by simply typing keywords from the Search page (Fig. 6). Since the keyword search is case sensitive, desired results are obtained only with the exact word, such as "RoHS". This returns 20 hits, and by refining the search field to "components" and "electronic", 9 hits can be obtained in electronic components. Company searches are also made by this page. Furthermore, Eco-labeled products can be searched from this page. From the link page, access can be made to other materials database.

3. Recent Trends in Eco-Materials as Viewed by Eco-MCPS

3.1 Access log analysis

Answers for the enquiries to well-known companies in Japan showed that the development of eco-materials is closely related with automobile industries or with electronic and electrical industries. Furthermore, by analyzing the access log to the database, the public interest on eco-materials can be understood. Access to the database had been growing till December 2006, when the DB made public via a news release and hit a maximum (Fig. 7). Since then, there is a steady access of more than 1000 views per month. For instance, the most viewed page during September 1, 2006 to March 31, 2007 was "Eco-keitai", i.e., a cell phone with its casing made from 100% bioplastics using Kenaf; the second was a CFRP-made car hood, and the third was a chrome-free steel sheet. These are not on the top of the categories, so such a frequent access can be related to the public interest. Furthermore, the most frequently searched keyword was "RoHS", the second one was "mobile", and the third was "ink". The first and the second most searched keywords are frequently mentioned in the presentations and are also of high interest. However, it is assumed that the third keyword was input by some people interested particularly in this field.
3.2 Association rule mining

3.2.1 Theory

The web database is a set \( I \) of \( k \) items, i.e., \( I = \{i_1, i_2, \ldots, i_k\} \) (i.e., \( k \) web pages), and a transaction (or an access) \( T \) to the Web database is a set of items (in other words, viewed pages), \( T \subseteq I \). The Eco-MCPS database access log \( D \) is regarded as a set of transactions. Association rule is defined as \( A \Rightarrow B \) (i.e., if \( A \), then \( B \)), where \( A, B \subseteq I \) and \( A \cap B = \emptyset \) (empty set). Whether the association rule is strong or not must be judged by a support (i.e., the ratio or percentage of the combination of \( A \) and \( B \) in the total transactions), a confidence (i.e., the percentage of the combination of \( A \) and \( B \) occurring under the condition of \( A \)), and a lift (i.e., the ratio of the combination of \( A \) and \( B \) occurring under the condition of \( A \) with respect to \( B \) occurring without the condition of \( A \)). Let us denote the frequency of appearance of an item set, \( A \), as support\(_{\text{count}}(A) \), then, the support, the confidence, and the lift are defined as follows.

\[
\text{support}(A \Rightarrow B) = P(A \cup B) = \frac{\text{support}_{\text{count}}(A \cup B)}{|D|}\\
\text{confidence}(A \Rightarrow B) = P(B | A) = \frac{\text{support}_{\text{count}}(A \cup B) \times \text{support}_{\text{count}}(A)}{\text{support}_{\text{count}}(B)}\\
\text{lift}(A \Rightarrow B) = P(B | A) / P(B) = \frac{\text{support}_{\text{count}}(A \cup B) \times \text{support}_{\text{count}}(A)}{|D| \times \text{support}_{\text{count}}(B)}
\]

where \( P(X) \) represents the probability of \( X \), \( P(B | A) \) the probability of \( B \) occurring under the condition of \( A \), and \( |D| \) is
the frequency of the database. A strong association rule is supported by high support, confidence, and lift.

### 3.2.2 Networking based on association analysis

Figure 8 shows strongly associated pages after clustering the data based on K-means method, which is a common method of clustering by assuming a centroid of a cluster, determining the distance of each object to the centroids, and grouping the objects based on minimum distance by iteration. Desirable and consistent results can be obtained by starting with two centroids and then expanding the number of centroids. The data was analyzed with a minimum support of 1%, a minimum confidence of 10%, and a lift of 5%. The network was drawn with Virtual Mining Studio software (Mathematical Systems, Inc.) for 5 clusters based on the similarity in confidence. The top 50 associations were selected from those having the highest support after sorting in descending order. In the network, the line width is proportional to the confidence of the association; the wider the line width, the stronger is the confidence of the association.

The pages in cluster numbered 1 are strongly associated to each other because of keyword search. These two pages come to the top by searching “fibers (in Japanese)”. It is presumed that some users were interested in fibers for composites during the period of September 2006 to June 2007. The pages in cluster 2 are associated to each other by the LCA stage; it is

---

**Fig. 6** Search and refined search results for “RoHS” related data.

**Fig. 7** Access log statistics for Eco-MCPS DB.
likely that the DB users were interested to LCA stages of (C-5) servicing (in use) and (C-6) waste. It is also to be noted that the LCA stage of (C-1) extraction is associated to the environmental performance of (B-2) longer life. Furthermore, it can be safely said that less product data is available on environmental performances such as (B-1) recyclability and use or (B-7) regenerated (reproduced) materials; this is the reason why only the other performances, i.e., (B-2) longer life, (B-3) resource saving, (B-4) higher performance, (B-5) energy saving, and (B-6) environment cleaning, appear in the network. It is also interesting that a very strong association is found between (B-3) resource saving and (A-4) waste, (B-3) and (B-2) longer life, and (B-3) and (A-5) resource consumption. It is presumed from these results that the users are more interested in waste management and saving resource consumption.

Another distinguished strong association starts from “brass products”, which is listed at the top of the Eco-material category – Metals. The second and the third products in the same list are the “non-Fe solar battery material” and “analytical instrument”, which associations are therefore understandable from the nature of listing. The other three pages, i.e., “home appliances”, “automobile”, and “air conditioners”, are also listed in Metals.

In association with the environmental performance, the most likely clicked category should be global warming (top item in the right hand side of environmental impacts in the Eco-MCPS), but unexpectedly, (A-2) air pollution is associated to the “brass product”. This thereby leads to the top item listed in this category, i.e., “road construction materials”, which is actually “an air pollution reducing pavement blocks using photo-catalytic effect”. This association again leads to the presumption that “global warming” has become such a vague concept or universal concept that one may not expect to obtain an effective solution from these products. Therefore this category was not clicked and the second on the list of environmental performance was clicked.

However, there is no particular relation concerning other associations, such as the “battery” and “magnetic disk (MD)” with “road construction materials”. These are therefore presumed due to the attributes of a Web database that those listed at the top of the ecoproducts (electric appliances) are most likely to be clicked in the early stages of search. Similarly, the pages in the cluster numbered 3 at the bottom are associated to each other because they are the recently added new data, and are therefore listed at the top.

4. Conclusions

The above results show that this database is not a simple compilation of material properties data. The access log data of Eco-MCPS was analyzed using data mining methods such as association analysis, and it has been found that the public interest level is higher for environmental impacts such as (A-4) wastes in association to (A-5) resource consumption, which are also related to (B-2) longer life and (B-3) resource saving. Thus, implementing the database with products capable of providing solutions for these environmental impacts and environmental performances may respond to public needs and provide a network site contributing to the

Fig. 8 Network diagram for clusters based on association rules.
advancement of material technology. Whether a product page is viewed frequently or not greatly depends on the order of listing; thus, further modifications will be made on page design and functions. It is desired in the future to add new products and property data of the eco-materials. It also seems important to expand the eco-services field and to make the database accessible in English.

Acknowledgement

R.O. is grateful for the support by THE WATANABE MEMORIAL FOUNDATION FOR THE ADVANCEMENT OF TECHNOLOGY, and SONY Institute of Higher Education.

REFERENCES

7) J. Han and M. Kamber: Data Mining: Concepts and Techniques, (Morgan Kaufmann, 2001).
8) S. Chakrabarti: Mining the Web: Discovering Knowledge from Hypertext Data, (Morgan Kaufmann, 2003).