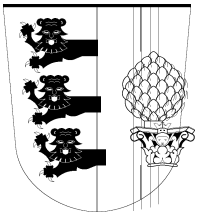
Design and Implementation of COSIMA - A Smart and Speaking E-Sales Assistant

Werner Kießling, Stefan Fischer, Stefan Holland, Thorsten Ehm

Angaben zur Veröffentlichung / Publication details:

Kießling, Werner, Stefan Fischer, Stefan Holland, and Thorsten Ehm. 2001. “Design and Implementation of COSIMA - A Smart and Speaking E-Sales Assistant.” Augsburg: Universität Augsburg.

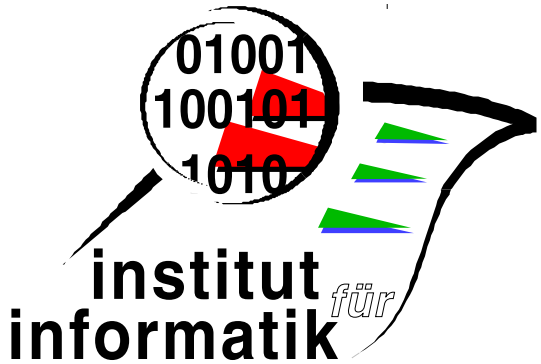
UNIVERSITÄT AUGSBURG



Design and Implementation of COSIMA - A Smart and Speaking E-Sales Assistant

Werner Kießling, Stefan Fischer, Stefan Holland, Thorsten Ehm

Report 2001-1 Januar 2001



INSTITUT FÜR INFORMATIK

D-86135 AUGSBURG

Design and Implementation of COSIMA - A Smart and Speaking E-Sales Assistant

Werner Kießling, Stefan Fischer, Stefan Holland, Thorsten Ehm

Chair of Databases and Information Systems
Institute of Computer Science, University of Augsburg
D-86135 Augsburg, Germany

{kiessling, fischer, holland, ehm}@informatik.uni-augsburg.de

Contact: Stefan Fischer, Phone +49-821-598-2166, Fax +49-821-598-2274

Keywords: E-Sales Assistant, Speaking Avatar, Preference SQL, E-Shopping, Dynamic E-Sales Advice

Abstract

We present a new cooperative user interface for e-shopping in B2C e-commerce. COSIMA is a smart and animated Internet avatar with synchronized dynamic voice output who assists customers through their e-shopping tours and advises them in the spirit of a real salesperson in the old economy. COSIMA benefits from a Preference SQL-based search engine that relies on the Pareto-principle, computing best matching results to the customer's wishes. This enables COSIMA to offer ideal hits or suitable alternatives with one single query, to value these search results in terms of the stated preferences and to generate proper voice output following basic principles of sales psychology. As a sample application we present our meta-search engine for comparison shopping. Since COSIMA constitutes an entire technological framework, it can easily be adjusted to other e-commerce environments, including mobile commerce, by customizing the search engine, the avatar animations and the speech contexts.

1 Introduction

When visiting a web site for e-shopping, the customer wants a quick survey what he can do and how he can do it. Textual hints are often ignored. A further deficit is that the current generation of e-shopping systems often suffers from inadequate search engines, but definitely from the lack of smart sales advice. When shopping for items an e-customer typically has to fill out one or several query forms in a very uncooperative environment. After pressing the "GO"-button he or she far too often encounters the infamous "empty result - sorry, please try again" situations as one extreme, or a flooding with irrelevant results as the other. This is aggravated by the fact that search results are not explained in the context of the customer's preferences. Such a behavior would be unheard of in a traditional department store of the old economy with a friendly shop assistant. Thus, to make e-shops more customer-friendly than they appear today, many things have to improve substantially. It will become essential to listen to customers' wishes more carefully, to help him or her conveniently throughout the

sales process, to present a good selection of products that match his or her preferences and to induce the e-customer to carry out a purchase right now. The popular comparison shop MySimon[1] made a first step by using a static avatar. Even the Deutsche Bank equipped some web pages with predefined voice output and an avatar named Cora[2] . But a much more personal and smart contact is needed for a really good sales advice for e-shopping.
The COSIMA project aims to go much further. In this paper we describe the design and implementation of COSIMA 1.0 as a major step towards better e-shopping. In *Section 2* we start out with our overall design principles, followed by the architecture of COSIMA 1.0 which offers comparison shopping over various real-life e-shops. In fact, COSIMA is the name of our charming avatar or can be interpreted as an acronym for "COmparison Shopping with Interactive Meta search Agents". Thereafter we lead through a sample e-shopping session with COSIMA. *Section 3* deals with the impact of search engine technology introducing Preference SQL as one suitable choice. In *Section 4* the automatic generation of sales contexts for smart product explanations is the main focus. *Section 5* investigates implementation details of our COSIMA 1.0 prototype, including aspects of parallel search agents and web-engineering issues for dynamic speech output. We also report first user feedback we have gained from our comparison shop, featuring the categories books, audio CD's and PC-hardware, and we discuss performance data gathered from our real-life application. In *section 7* we summarize what COSIMA 1.0 has achieved so far and give an outlook on further work within the COSIMA project.

2 Better E-Shopping with COSIMA

Now we present our guiding design principles followed by the architecture of our COSIMA 1.0 comparison application. Then we take the reader on a shopping tour with the avatar COSIMA itself.

2.1 COSIMA's Design Principles

The creation of a smart e-shopping assistant poses a major challenge. COSIMA's road map towards this ambitious end incorporates the following general design principles:
P 1) *Knowing and respecting the customer's preferences:*
Customer wishes entered into the Internet search mask must seriously be interpreted as preferences. Neither should COSIMA interpret the user's input as hard conditions like many search engines do, nor should it be required to enter complex boolean queries what so-called advanced query modes often ask for. Instead, COSIMA has to accept customer wishes in a purely declarative manner. In any case best matching results have to be returned:

- If no perfect match can be found, then best possible alternatives are offered automatically, avoiding the embarrassing *empty result effect*.
- No results are returned which are subsumed by better offers, avoiding the annoying *flooding effect* with irrelevant results.

Solutions that are conform with *P1* are discussed in section 3.
P 2) *Generating smart sales advice dynamically:*
E-sales advice must be dynamically adapted to the degree of match between the customer's pref-

[1] http://www.MySimon.com
[2] http://www.deutsche-bank.de

erences and presented items, augmented by a portion of sales psychology like in the old economy. Section 4 will focus in more details on this topic.

P 3) *Communicating by a charming animation character and by dynamic voice output:*
COSIMA is presumably the first avatar acting like a smart, speaking e-salesperson. If COSIMA wants to present the results or wants to say anything else, she does it via dynamically generated voice output. To provide such a rather complex but intuitive interaction of very different features poses a major architectural design challenge by itself. Its careful treatment will be discussed in section 5.

P 4) *Extensibility:*
The overall system design of COSIMA is based on Internet component technology, rendering it flexible enough to allow for further improvements of the human-avatar interface.

2.2 Architecture of COSIMA 1.0

To prove our visions in a real-life environment, with COSIMA 1.0 we implemented a meta-search engine doing comparison shopping over a variety of existing commercial e-shops. The system architecture is partitioned in three tiers as shown in fig. 1. The animated and speaking female avatar

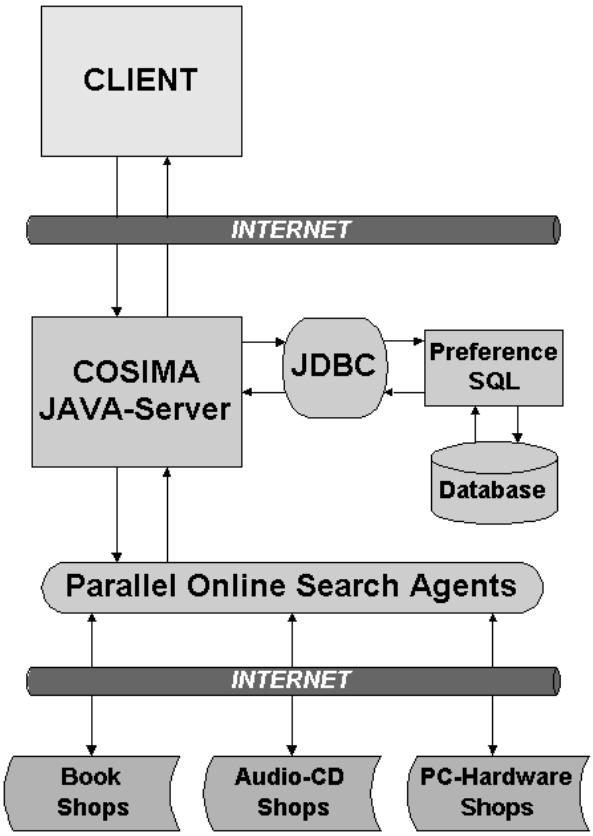


Figure 1: Architecture of COSIMA 1.0

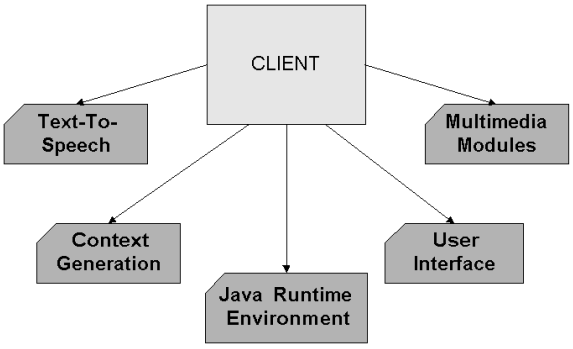


Figure 2: COSIMA-Client

COSIMA advises the customer at the client side to declaratively enter his wishes into a search mask[3].

[3]We decided not to equip COSIMA 1.0 with speech input or chatting capabilities. Once desired (and commercially viable products for speech input exist), all of those can be added in a latter stage in a modular fashion.

This request is sent to the COSIMA-server via Internet. From there Internet agents search in parallel across several connected e-shops and deliver their partial results back to the COSIMA-server. Via JDBC these results are temporarily stored in an SQL database (Oracle 8i) and then filtered with Preference SQL (see section 3). The set of the best matching items is sent back to the client where COSIMA presents them to the customer. The sales contexts and the corresponding speech-audio-files are generated dynamically using information about the customer's preferences and the search results. In fig. 2 we show what software components are located in a COSIMA-client[4]. This separation of client- and server-components has significant advantages:

- The *COSIMA-server* sends only basic product information. No audio-files are transmitted because the dynamic generation of the animations and speech-files takes place in the client only. Thus even a slow 14.4 KBit modem on client side is sufficient for the data transmission.
- The synchronization of COSIMA's lips to the spoken text is computed on the client-side as well and therefore cannot be compromised by low bandwidth or overloaded servers.

Presently the Internet is the major bottle neck w.r.t. e-shopping performance, making it senseless to burden the server-side further. But since many people have access to powerful PC-clients, the client-side can cope with some extra tasks like speech generation and the coordination of the multimedia tasks.

2.3 A Shopping-Tour with COSIMA

Now we present several examples of COSIMA in action.

2.3.1 Welcome, Shopping Guidance and Infotainment

COSIMA is charming and polite like a real salesperson. Starting COSIMA she welcomes the user and guides him to the preferred category. He can choose from books, audio CD's or PC-hardware. In this example he takes the CD-rubric and searches for "Tales of Mystery and Imagination" of "Don Williams", with a preferred price range between 10 and 20 German Marks (Fig. 3). Note, he mixed up the the performer "Don Williams" with "Alan Parsons". The text shown in the status area records all spoken sentences of COSIMA so far. After having pressed the "GO"-button COSIMA shortens the waiting time by entertaining the user. The maximum search duration is user-adjustable in the options-menu. An e-shop not delivering the results in the specified time is ignored for the current search.

2.3.2 Presenting the first result

Sending out parallel shopping agents and filtering the partial results with Preference SQL yields a result as shown in fig. 4. COSIMA tells the customer that she has found nine CDs for him. Since these are all best matches to the stated customer query, COSIMA has a choice which one to start with. Thus with her own *vendor-oriented preference*[5], she selects a first item which is then made palatable by explaining its properties and advantages:

≪*The best hit is a CD of Alan Parsons. The price is within your specified price range. The title contains your specified keywords* **Tales**, **of**, **Mystery**, **and**, *and* **Imagination**. *If You want me to explain another product just click on it.*≫

[4]The COSIMA 1.0 client software can be downloaded from http://www.myCOSIMA.com. Pretty soon it will become available as a browser plugin, too. Note that currently COSIMA only speaks German.

[5]The benefit of owning a search engine is an important issue by itself, see e.g. [5].

Figure 3: Request for a CD

COSIMA suggests that this CD is the best possible hit, by enumerating the matching keywords of the title and confirming that the price is within the preferred price range.

2.3.3 Explaining an alternative

If the user clicks to another product COSIMA explains the differences to the previous one as in fig. 5:

≪*This CD is not from Alan Parsons but from Don Williams. The title is Borrowed Tales. It is two German Marks more expensive. The performer contains two more keywords.*≫

The performer contains the highlighted keywords **Don** and **Williams**. COSIMA names the title of the alternative as well since it differs from the previous one, but as a good sales practice she does not touch that the title does not match as well as before. Of course, COSIMA also computes and emphasizes the price differences.

2.3.4 Explaining a bargain

Our next example is from the book category. This time the customer searches for the book “The Dilbert Principle” from Scott Adams, wishing a paperback version and willing to pay a price around

Figure 4: Presenting the first best matching item

20 German Marks. The result is shown in fig. 6. COSIMA emphasizes that the result is a perfect hit and moreover that the book is a bargain:

≪At the shop Buchwelt I found the perfect hit, because it contains all keywords. The price is even cheaper than you specified. This is a real bargain! According to your wishes the book is a paperback version.≫

Like a real salesperson COSIMA uses a little sales psychology: She denotes the result as a perfect hit since author and title contain all specified keywords. But this book is not available in the desired price range. Instead of stupidly delivering an empty result, COSIMA names the book a bargain because it's even cheaper than the user was willing to pay for it. COSIMA's sales psychology is based on meta knowledge (e.g. cheaper prices are better) as well as domain specific product knowledge (e.g. hardcover is better than paperback).

Satisfying items can be put into a shopping basket as usual. The actual purchase is then carried out through the e-shops in question. As a friendly salesperson COSIMA also says goodbye to the user when he or she quits the shopping session.

Figure 5: Explaining an alternative

3 Product Search with COSIMA

Today there exists a variety of different search engines for the Internet and for product databases of e-shops. As everybody knows from own experience their search behavior varies enormously. Thus the proper choice of a good search engine for e-shopping is an important issue, which becomes even more crucial if better e-sales advice has to be offered on top.

3.1 Requirements for an E-Shopping Search Engine

A fundamental decision has to be made when reacting to customers' wishes or preferences, namely what should be the right technical way of modeling such preferences. To date there exist several different approaches, in particular:

1. Translate customer preferences into *hard* selection conditions.

2. Translate customer preferences into *soft* selection conditions and apply a ranked query model, assigning *numerical scores* between 0 and 1, the latter denoting a perfect match.

Figure 6: Offering a bargain

3. Translate customer preferences into *soft* selection conditions, modeling preferences as *partial orders*.

In practice mixtures between hard and soft selection conditions occur and must be supported, too.
Approach 1 often translates a user query straightforwardly into a SQL query, hence being exposed to the empty-result effect, which is particularly annoying when e-shopping. Ad hoc remedies like parametric search are far from optimal either. In terms of shopping in the old economy this would amount to ask the customer to search by himself iteratively through the warehouse - which is very customer unfriendly and prevents the salesperson from appropriately learning about the customer's wishes. *Approach 2* often has problems with determining the score values and how to interpret them meaningful to the customer. It may also be vulnerable to the flooding effect with irrelevant results.
We argue that search engines implementing *approach 3* are a reasonable choice as a basis for good e-sales advice. A full discussion of this non-trivial problem is beyond the scope of this paper, instead we want to emphasize our point by describing a representative for this approach, namely Preference SQL, in some detail next.

3.2 Preference SQL

Preference SQL aims at putting an end to search hassles for e-shopping. As an extension of standard SQL it does not only support standard hard selection conditions in the `where`-clause of a query, but also soft conditions identified by the new keyword `preferring`. Technically, all preferences can declaratively be modeled by partial orders, which have been shown to be compatible with database technology ([7, 8]). The query evaluation semantics is as follows: First apply all hard filter conditions from the where-clause, then determine the *maximal elements* according to the partial order specified

in the `preferring`-clause. This semantics obviously avoids both the empty result effect (unless the `where`-clause evaluates to the empty set) and the flooding effect with worse elements. In the sequel we use the Preference SQL syntax implemented by ([11])[6].
Preferences come in two flavors, namely *basic* preferences and *compound* preferences. Examples for basic preferences are, e.g., `price around 10`, `price between 10 and 20` or `minimize price`. Compound preferences construct a new partial order from some given preferences. Mathematically there are several well-known ways to do this, like the cartesian product, the lexicographic order, etc. ([4]). The proper choice of course depends on the requirements of the application area. For e-shopping purposes we argue that offering compound preferences that rely on the *Pareto-optimality principle*[7] are an effective choice. In particular this supports *multi-attribute optimizations* and the search for *best-matching alternative* offers, if the perfect match is not available. Preference SQL currently supports two combining operators to construct compound preferences:

- The `AND`-operator models preferences of equal importance and implements the Pareto-optimality principle.
- The ‘,’-operator expresses priorities among different preferences.

As an example take the customer query from our COSIMA session in fig. 3. Applying a stop-word list for the text part, which excludes common words like ‘of’ or ‘and’, the Preference SQL query generated by COSIMA might look as follows:

```
SELECT *,top(price),top(type) FROM Cosima_Server_tmp
PREFERRING
 (title CONTAINS ‘Tales‘ AND title CONTAINS ‘mystery‘ AND
  title CONTAINS ‘imagination‘ AND interpret CONTAINS ‘Don‘ AND
  interpret CONTAINS ‘Williams‘),
  price BETWEEN 10 and 20;
```

Figure 7: Customer wishes translated into one Preference SQL query

All keyword conditions have been chosen to appear on the same priority level, but they are assumed to be more important than the price. The answers computed by the COSIMA server in behalf of this single Preference SQL query are precisely all best-matching items according to the stated customer preferences.
Moreover, Preference SQL can return *quality information*, too. Syntactically this can be expressed in the `select`-clause by using the keywords `level`, `distance` or `top`. For numerical attributes `distance` calculates the (relative or absolute) numerical deviation from the desired perfect value, whereas `level` relates to the level within the underlying partial order (a level of 1 denoting a maximal element). The keyword `top` is shortcut for `level`=1. As an important consequence, the quality of search results can be interpreted meaningfully in terms of the customer’s input preferences. Clearly, this is an added value for an e-sales agent who has to sell a product to the customer.

[6] So far, there are no scientific publications about implementation details of Preference SQL due to pending patents. Sample applications of Preference SQL technology have been reported in [9].

[7] The Pareto principle is folklore in the social and business disciplines since about 50 years. In particular for multi-attribute decision problems it has been applied and studied extensively by these communities (see e.g. [6]).

4 Product Presentation with COSIMA

4.1 Correlation with the Choice of the Search Engine

We just argued already that the proper choice of a search engine is crucial for e-shopping. It does not only influence the quality of the search result, but as a consequence the *quality* of the dynamic e-sales advice. First of all, the e-sales assistant should know the customer's preferences - the more, the better. Second, the completeness of the search results w.r.t. best matches is an essential basis for generating convincing and believable explanations. And third, unless there are perfect matches in stock, suitable alternatives must be available for offer instead.
Given all of that, the dynamically generated sales explanations of the e-sales assistant can be similar to what knowledgeable sales persons in the old economy would tell their customers, including methods from *sales psychology*. As argued before, Preference SQL as the basic for our meta search engine for COSIMA satisfies above requirements pretty well. Now let's give a first insight into the sales explanations of COSIMA 1.0 (which of course can and will be refined considerably in forthcoming versions).

4.2 Sales Contexts of COSIMA 1.0

Let us loosely define a *sales context* as any spoken interaction of COSIMA with her customers. The COSIMA sales contexts are based on templates as described in [12]. These contexts are needed for customer-friendly advice. They are divided in *static* strings for welcome or infotainment and *dynamic* strings for the sales-related explanations.

4.2.1 Static Sales Contexts

There are five different static sales categories:

- **Welcome**, e.g. *≪Hello, my name is Cosima. I'm your smart e-shop assistant.≫*
- **Shopping guidance** may give meta information about an e-shopping portal or mall, including directory or yellow pages service. In our case it suffices to help the user to find his preferred category, e.g. *≪Do you like books, CD's or PC-hardware. Please visit the referring category above.≫*
- **Infotainment** fills the gap while the search is in progress. Cross-selling, entertainment, informing about the latest news or stock quotations, etc., can be offered here. Right now COSIMA talks to the idle customer, e.g. *≪By the way, do you know, the new album of Madonna is available now? It's only 25 German Marks.≫*
- **End of the search phase** is an event where COSIMA seizes the customer's attention: *≪The agents are back. I found six CD's for you.≫*
- **Goodbye**, e.g. *≪Hope you enjoyed your visit. See you soon.≫*

During runtime a suitable sales context of each category is computed randomly, aiming to prevent the waiting customer from getting bored. This modular architecture has several advantages:

▹ Sales categories can easily be extended by adding a new one.

▹ Sales contexts can be scaled just by adding a new phrase to the particular sales category.

4.2.2 Dynamic Sales Contexts

There are two types of dynamic sales contexts, namely explaining the first search result and comparing two search results.

◇ **Explaining the first search result**

To start the sales presentation COSIMA has to select one of the search results. Since our Preference SQL-based search engine retrieves all best matches to a customer's query preferences, there is an additional degree of freedom which of them to offer first. In a traditional shop, the sales person probably would apply some vendor preferences like minimizing large inventories of some products or maximize her own profit margin. COSIMA can likewise apply this strategy, too[8]. Fig. 8 shows the simplified procedure of generating the sales context for explaining an item in the audio CD category. First the selected result and the customer's request are analyzed with regard to matching keywords, price and CD type (maxi-single or album). Depending on the degree of match COSIMA decides between "perfect hit", "bargain", "good hit" and "bad hit". The case of no results is only possible if all connected e-shops returned no results at all. In this situation COSIMA proposes a longer search time adjustable by the user. Each situation causes COSIMA to generate an appropriate sales context.

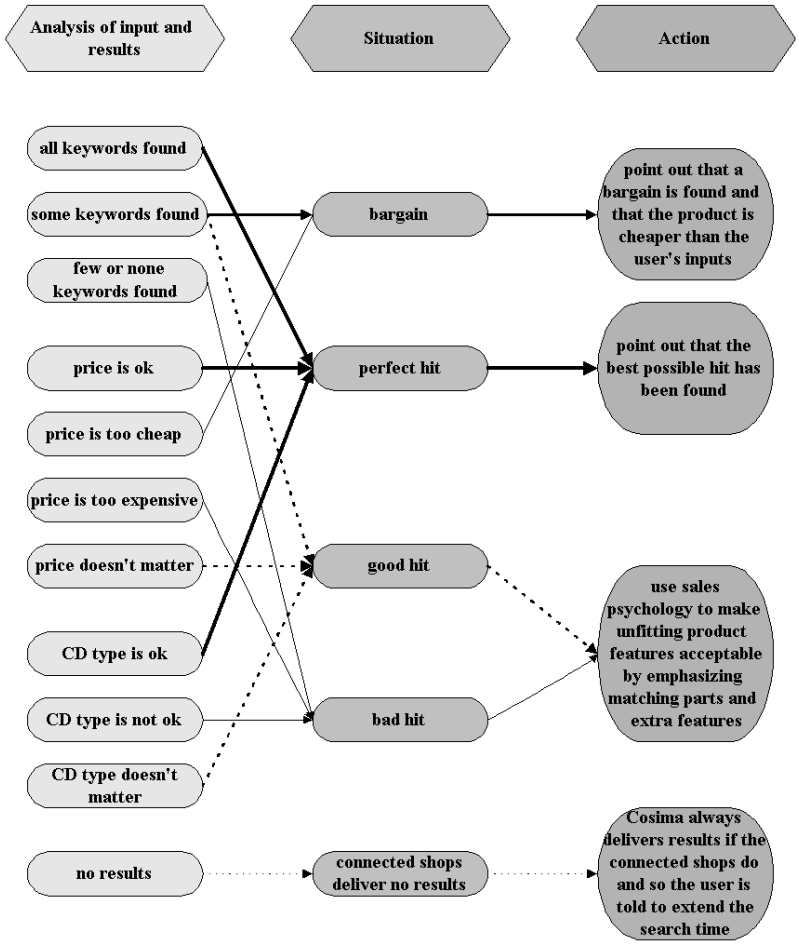


Figure 8: Computing the dynamic sales contexts

[8]Technically, vendor preferences can even be mixed already with the original Preference SQL query.

The sales explanation consists of three parts:

Keywords:	Which of the keywords are found in the result?
Type:	Is the CD of the denoted type?
Price:	Is the price minimal or in the specified range? If not, how large is the mismatch?

The generated sales contexts are merged together to the actual explanation within an overall evaluation of the result by COSIMA, e.g. ≪*I found a bargain for you.*≫ or ≪*I have a perfect hit.*≫. She also uses sales psychology to advertise products not exactly matching the customer wishes:

≪*The CD "Alive" from Pearl Jam is not deliverable right now, but I have another live-CD of the same band. "Live on Two Legs" is even ten German Marks cheaper. Do you want to buy it?*≫

◇ **Comparing two search results**

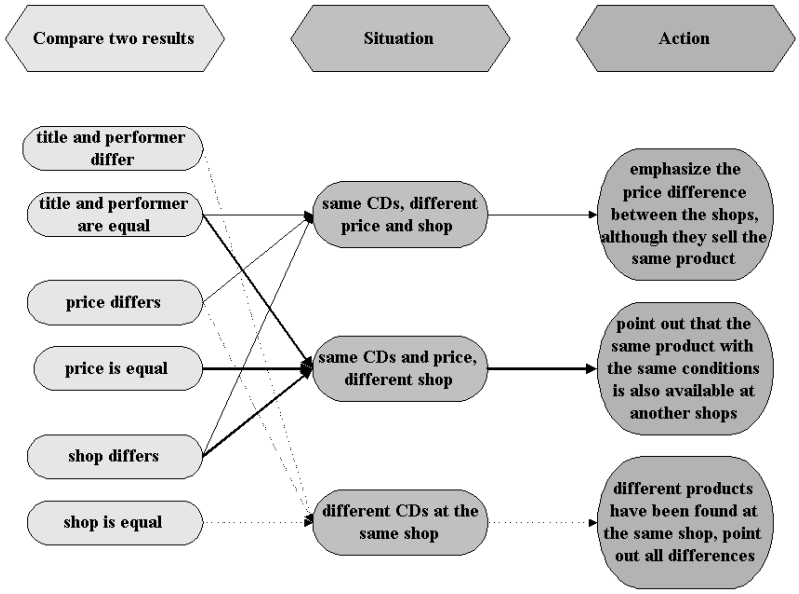


Figure 9: Computing the dynamic comparison contexts

If a perfect match is not available, the Pareto semantics of Preference SQL performs a multi-attribute optimization presenting all best-possible alternatives that can be compared to each other by COSIMA. Fig. 9 illustrates some of the possible situations. An important case are e-shops offering the same item at varying prices. COSIMA computes the price difference and emphasizes that the CD is cheaper or more expensive, respectively. COSIMA pays also attention on special situations not appearing in fig. 9, e.g. alternative CDs from the same performer, CDs from a different performer but of the same title, same CDs varying in the type (maxi-single or album), etc. This sales explanation technology here analogously works in the categories books and computer hardware.

5 Implementation of COSIMA 1.0

5.1 Comparison Shopping with Parallel Internet Agents

If a customer query has been issued at a COSIMA client, it gets translated into a Preference SQL query. However, since each of the connected e-shops[9] has its own search engine, this query must be adapted before our Internet agents can be sent out. Fig. 10 shows the work of our parallel Internet agents.

To circumvent the weaknesses of the search engines of the connected e-shops, we implemented suitable countermeasures in the form of *query relaxation*: To guard against the empty-result effect, all attribute conditions of a COSIMA client query are sent as a disjunctive query (i.e. "or"-ing all conditions) to each connected e-shop. For e-shops that don't even support disjunctions, several parallel subqueries with one attribute condition each are generated. Thus one COSIMA client query may lead to many online-requests to the connected e-shops. The results returned by our agents are temporarily stored in a SQL database at the COSIMA-server. Now Preference SQL can filter these pre-selected items with the Pareto semantics, avoiding the flooding effect with worse results. Only the best-possible results are sent to the COSIMA-client and explained by COSIMA.

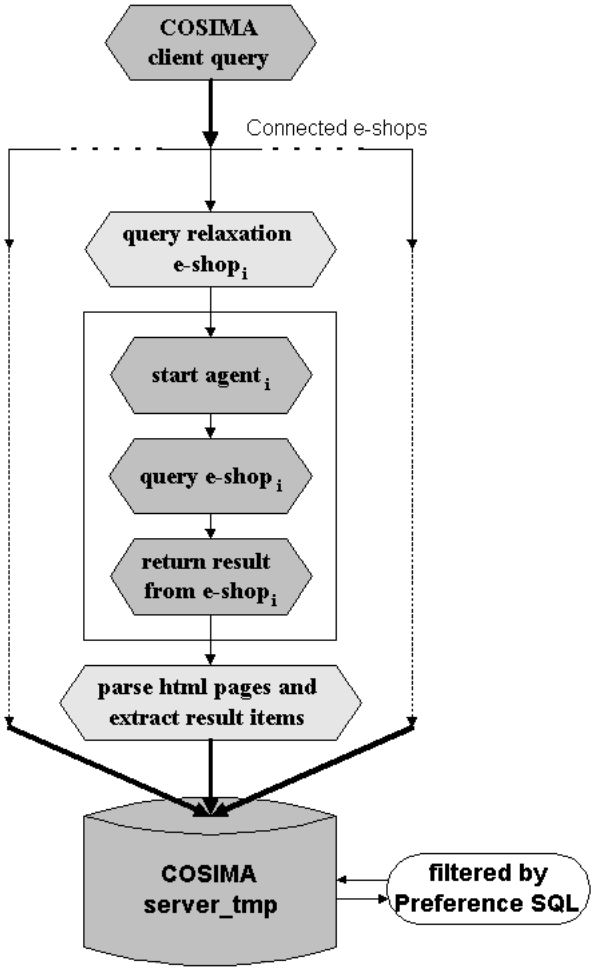


Figure 10: Internet agent delivering results from connected e-shops

We use parallel Java threads to implement these agents. With a standard PC as COSIMA-server (850 Mhz CPU, 512 MB main memory, 10 MegaBit dedicated line) tests showed that we can handle up to 1000 of these threads simultaneously. The online-requests to connected e-shops are done with an interface that uses the Java network capabilities to manage the Internet connections. It provides an implemented method to get html pages via the Internet and two abstract methods to create the uniform resource locator dynamically and to parse the resulting html pages. A new e-shop can easily be included by implementing these two methods.

[9] Currently connected e-shops are:
Category books: BOL, Amazon, buecher.de, Buchwelt, Buchkatalog
Category audio CDs: Alphamusic, Lost and Found, CD and More, CD Station, Zweitausendeins, Musikschule online, CD Wonder
Category PC hardware: Alternate, 24h Comtec online, Mint Data, HD-Computer, Acom

5.2 The Multimedia Avatar

Now we want to provide an insight into some implementation details of our multimedia avatar concerning the avatar COSIMA itself, dynamic speech generation and synchronization issues.

5.2.1 COSIMA - the avatar

Whenever advisory service has to be provided it is much more convenient to have somebody face to face (see [1]). A *persona* looking at the customer embraces several valuable aspects for e-commerce. It is much more personal to the customer. The feeling that someone cares about the customer can be transmitted by an avatar. Moreover, a persona like COSIMA (see fig. 11) entails a recognition effect. Positive effects for an e-shop are obvious, if the avatar represents a likeable character.
Creating such a likeable avatar[10] is a difficult, interdisciplinary task by itself, involving cultural and marketing aspects. Since the present TTS-tools do not perfectly sound like humans, we have chosen a *2.5D-avatar*. The 2.5D-lady COSIMA is an animated cartoon. To bring life to the e-shop it is necessary to have several sequences of the animated persona to present different situations and different emotions. Examples of scenes we need for COSIMA are saying hello or goodbye with a smile, explaining some products and pointing on them, entertaining the user while searching the Internet, being very happy finding a bargain, and confirming changed options in a friendly way. Negative examples that should be avoided are, e.g. a male avatar with a female sounding voice and vice versa. Also a non-animated cartoon does not fit to a speech output via voice. COSIMA compares books, compact discs and pc-hardware. So in our case a young fashioned lady is that what people normally expect.

Figure 11: The avatar COSIMA

5.2.2 Dynamic Speech Generation

Though the usual way of providing sound output is to support predefined sound files in standard formats like *.au*, *.wav* or *.ra*, this implies significant disadvantages:

- It is not flexible. The sentences or parts of them are limited to the prepared files.
- For enlarging the voice database the same person in the same condition is necessary to record the new files.
- The voice database is very storage-intensive.
- There is no way to react in realtime to a new situation you want to talk about, e.g. you find a new product with a name you never heard.
- A transfer over the Internet is often impossible in tolerable time.

[10]The character design for COSIMA is under copyright of the German Research Center for Artificial Intelligence (DFKI), Saarbrücken.

All these problems can be whipped off by using a **Text-To-Speech** (TTS) tool at client side. In our setting sound generation by a TTS-tool takes only very little time (under one second typically). Hence our architecture is superior to common server-side solutions, which often aim at application with lengthy audio text ([14]). Nowadays the speech quality of such TTS-tools is at an acceptable level. The intonation is rather good and to some degree even emotions can be expressed. Dynamically converting short text messages to speech can be done in less than one second even on a PC with a Pentium I 200Mhz processor. Applying such a TTS-tool[11] COSIMA appears to the e-customer to speak in realtime.

5.2.3 The Synchronization Observer

Since all spoken text is generated at runtime, all animation sequences must be flexible in length. The challenge of combining voice and animation is to provide a reasonable good synchronization. Though we only transfer basic product information via Internet we cannot trust in any time to start

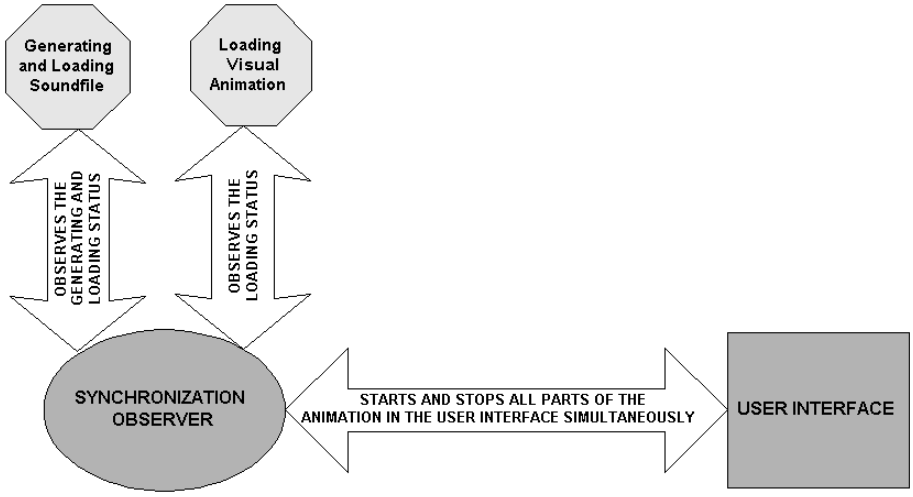


Figure 12: Functionality of the synchronization observer

the animation. Even small differences like half a second between starting sound and animation make a very poor appearance. Therefore we chose to implement our own observer which is in charge of speech and animation synchronization.
Our synchronization observer (fig.12) controls the arriving product data, the conversion from ASCII-text to speech with the TTS-tool and the loading of the temporary sound file into the random access memory (RAM). Concurrently the observer supervises the loading of all needed animation sequences to RAM. When all parts are available in the RAM the synchronization observer starts voice and avatar simultaneously. At the end of the voice output the observer draws a final picture or an idle task animation so that the action of the lips stops right on the point.

[11]The TTS-tool used by COSIMA 1.0 is a combination of the speech synthesizer of Mbrola (http://tcts.fpms.ac.be/synthesis/mbrola.html) and the Text-to-Phone converter of Hadifix (http://www.ikp.uni-bonn.de/˜tpo/Hadifix.en.html).

6 Evaluation of COSIMA 1.0

COSIMA has successfully given her debut at the international computer exhibition fair SYSTEMS 2000 in Munich and was presented several times to a large German public audience in newspapers, radio- and television-broadcast. A freeware client-version of COSIMA 1.0 is in use by several hundred users.

6.1 User feedback

COSIMA users are not selected test user, they are people who heard about COSIMA in radio, TV or print media and downloaded the software from our homepage[12]. Daily we receive about 30 requests to our comparison shop with increasing tendency. To use COSIMA a PC with at least a 400 Mhz CPU, 64 MB RAM and a 14.4 KBit Internet connection is recommended. The users gave a lot of encouragement for our work. We invited them to send any feedback. Major aspects were, e.g. positive comments about our very nice female avatar, that voice output is a very "cool" and helpful idea, that the comparison shop saves money, that query results are very well related, or the desire for more search categories, especially DVD-movies.
In summary, positive feedback and acceptance of this novel cooperative interface for e-shopping was very good. It's worth mentioning that so far we did not have a simple complaint about a bug or crash of the very complex COSIMA 1.0 system.

6.2 Efficiency

Now let's have a look at some performance evaluations under three perspectives.

◇ **Cardinality of Preference SQL result**
The Preference SQL search engine picks up only the best-matching results. Predominantly this num-

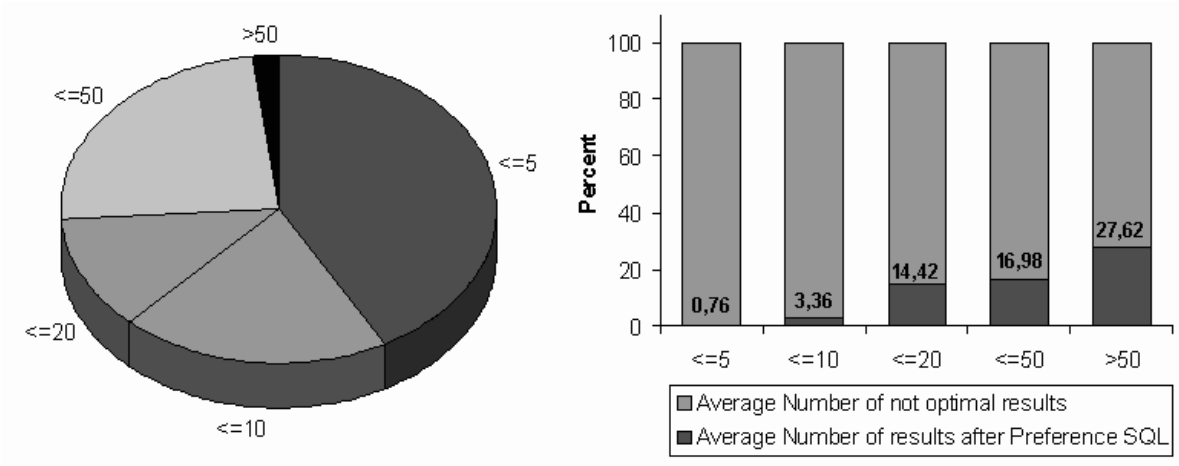


Figure 13: Best matching items per query and filter effectiveness of Preference SQL

ber of the presented results is within the interval of 1 and 20 (see left part of fig.13). Note that this fits very well to the requirements of a sales situation, where the customer wants to have an easy to survey choice.

[12] http://www.myCOSIMA.com

◇ **Filter effectiveness of Preference SQL**

In section 5.1 we described the implementation of our search agents. When a request with very popular words, e.g. “love”, is sent, quite a lot of results came back to the database. We measured up to 300.000 data sets for one request. The median is about 1.400. The filter effect of Preference SQL is very remarkable. The right part of fig.13 displays the median of how many percent of the temporary COSIMA server database are selected by the match-making process of Preference SQL.

◇ **Overall response time of a customer request**

Finally we want to study the *overall response time* for one customer request to COSIMA. In fig. 14 the mean response time for the book agents is depicted.

These values are similar to the two other categories audio CDs and PC-hardware. These figures suggest a value of 10-15 seconds as limit to wait for agent results. In some cases (e.g. overloaded Internet connection) it is useful to increment this maximum search time.

The overall response time for one COSIMA client request is decomposed as shown in fig. 15.

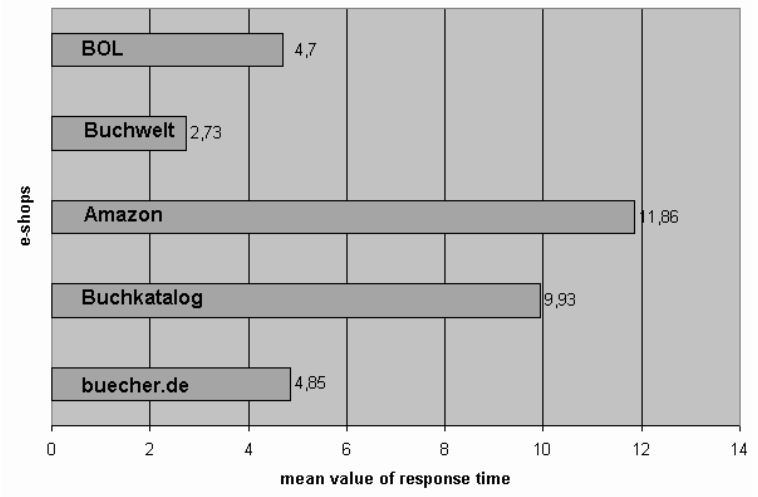


Figure 14: Response time for the book agents

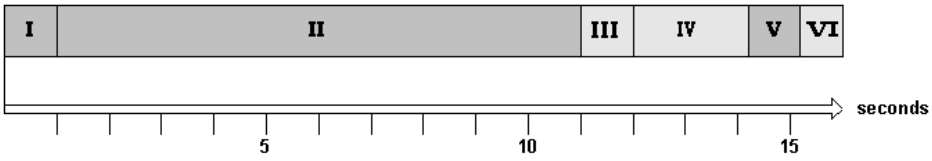


Figure 15: Time expense for one customer query

There are six steps to do in one request:

I Sending the request from a COSIMA client to the COSIMA server depends on the connection (LAN, modem etc.), so we estimate about one second.

II This depends on the option setting for maximal agent search time. Let’s assume a default of 10 seconds.

III Our median of inserting the agent results to the COSIMA server database is 0.9 seconds.

IV The Preference SQL search engine on the average needs 2.2 seconds.

V For sending the results back to the user we again estimate 1 second.

VI The median of speech generation and loading of speech and animation is 0.8 seconds.

In summary, providing the extra benefit of a good sales advice with an advanced multimedia interface to the e-customers only causes an overhead of about 4 seconds. Most of the time is consumed by the search engines of the connected e-shops. Since COSIMA can shorten this waiting time by infotainment, the whole response time of around 13 seconds was never found frustrating by our users.

7 Summary and Future Work

We have presented our ongoing COSIMA project, envisioning a smart and speaking e-sales assistant for better e-shopping. With COSIMA 1.0 we have implemented a meta-search engine for comparison shopping over various existing commercial e-shops. The very successful debut of COSIMA 1.0 at the SYSTEMS 2000 computer fair in Munich, the positive feedback and a steadily growing user community have confirmed our belief in this novel type of cooperative user interface. The main challenges and contributions of COSIMA 1.0 so far can be attributed to sophisticated architectural design issues and careful component selection. In particular, we analyzed the interrelationship of a search engine and the quality of dynamically generated sales advice, pointing out that Preference SQL-based search engines are a very suitable choice. Both the embarrassing empty-result effect and the annoying flooding effect with worse information experienced from many other search engines are avoided by Preference SQL due to the principle of Pareto-optimality. We also demonstrated how dynamic sales contexts, including basic ingredients of sales psychology, can be realized. As another highlight we introduced an efficient component architecture, where the synchronized interplay and speech generation between the animated avatar COSIMA and the spoken sales advice take place on client side. In a nutshell, our charming e-sales assistant COSIMA brings already a bit of the shopping feeling to the new economy that we appreciated so much in the old economy.
Since one of our overall design principles has been that of flexibility, there are numerous orthogonal evolution paths for forthcoming versions of COSIMA, in particular:

◇ *Negotiations:* This is a trendy but important topic, where the overall COSIMA approach fits already very well as a starting point. According to [3], important features of a powerful negotiation architecture would include a search engine that is capable of multi-attribute optimizations and a dialog environment based on the Language-Action Perspective of Searle and Habermas (see [15]). With Preference SQL we are prepared already for multi-attribute optimization, whereas our human-avatar interaction needs to consider a theory of speech acts.

◇ *Sales psychology:* We can add new pieces thereof in an evolutionary way, considering the traditional 4-stage model for performing a successful sales dialog. The impact will even be greater, if we achieve to express emotions for our animated sales avatar ([2]).

◇ *Smartness and chatting:* Our e-sales agent can be made smarter, if she has online access to domain-specific knowledge or meta knowledge, e.g. about audio CDs. Then she can immediately react to query formulations of the customer that don't make much sense and require some expert know-how. In combination with a chatting functionality this would be another step forward towards competent customer advice. Preliminary work on this topic is already in progress in cooperation with a commercial chat-tool manufacturer.

◇ *Speech input:* For mobile commerce speech output as already realized by COSIMA 1.0 would be a great feature. Also offering speech input instead of typing the query into a search mask, would be another great progress for mobile commerce. Since mobile devices like cellular phones will be very powerful clients once UMTS is available, we claim that COSIMA technology is applicable there either.

◇ *Preference mining:* Knowing customers' preferences is at the heart of any successful sales advice. Preference SQL learns such preferences directly from customer input into the search mask. But it is very worthwhile to perform preference mining on server log files as well in order to discover more preferences. This knowledge can be used in future e-shopping sessions for cross-selling, collaborative filtering or personalized infotainment. Preliminary results for extracting preferences (under the partial order model) from server log files using the commercial data-mining tools Cognos and SPSS are already available (see [13]).

◇ *Autonomous e-shopping agent:* Our parallel agent technology can be used as a backbone to implement an autonomous agent, who given a list of shopping items, a price and time constraint can do the errands on behalf of an e-customer. This can even incorporate autonomous negotiating with e-shops and e-auctions or electronic bargain systems as described in [10].

The long-term vision of COSIMA is to create a real interactive shopping feeling with speech recognition, natural language processing and user-adaptive behavior. Certainly COSIMA will never replace a smart real salesperson, but we believe she will make e-shopping much more exciting than it is today. The benefits will be very rewarding, including a higher customer loyalty or more cost-effective call centers, which can delegate from simple advisory tasks to the e-sales assistants and only must connect to high-paid human experts for high-quality advise.

Acknowledgements: We gratefully acknowledge Elisabeth André for the permission to use the avatar "Cyberella" (Character design Avatar ©2000 by DFKI Saarbrücken) and Wolf-Tilo Balke for carefully reading a draft of this paper.

References

[1] Elisabeth André, Thomas Rist, and Jochen Müller. Webpersona: A life-like presentation agent for the world-wide web. *Knowledge-Based Systems*, Vol. 11:25–36, 1998.

[2] Joseph Bates. The role of emotion in believable agents. *Communications of ACM*, 37(7):122–125, 1994.

[3] Sungwon Cho. A framework for the evaluation of an electronic marketplace design with evolutionary negotiation support. In *Thirty-Fourth Annual Hawaii International Conference on System Science (HICSS)*, Maui, Hawaii, January 2001.

[4] B. A. Davey and H. A. Priestly. *Introduction to Lattices and Order*. Cambridge University Press, April 1990.

[5] Rajiv Dewan, Marshall Freimer, and Paul Nelson. Impact of search engine ownership on underlying market for goods and services. In *Thirty-Fourth Annual Hawaii International Conference on System Science (HICSS)*, Maui, Hawaii, January 2001.

[6] Ralph L. Keeney and Howard Raiffa. *Decisions with Multiple Objectives : Preferences and Value Tradeoffs*. Wiley, April 1976.

[7] Werner Kießling and Ulrich Güntzer. Database reasoning – a deductive framework for solving large and complex problems by means of subsumption. In *3rd Workshop on Information Systems and Artifical Intelligence*, volume 777 of *LNCS*, pages 118–138, Hamburg, Germany, Febr. 1994.

[8] Gerhard Köstler, Werner Kießling, Helmut Thöne, and Ulrich Güntzer. Fixpoint iteration with subsumption in deductive databases. *Journal of Intelligent Information Systems*, Vol. 4:123–148, 1995.

[9] Achim Leubner and Werner Kießling. Personalized nonlinear ranking using full-text preferences. ACM SIGIR '99, Workshop on Customised Information Delivery, Berkeley, August, 1999. `http://www.ted.cmis.csiro.au/sigir99/leubner`.

[10] Ting-Peng Liang and Her-Sen Doong. Effect of bargaining in electronic commerce. In *International Workshop on Advanced Issues of E-Commerce and Web-Based Information Systems*, pages 174–181, Santa Clara - California, USA, April 1999.

[11] User Manual Preference SQL 1.3, 1999. Database Preference Software GmbH.

[12] Ehud Reiter. NLG vs. templates. In *Fifth European Workshop on Natural-Language Generation (EWNLG '95)*, pages 95–106, Leiden, Netherlands, 1995.

[13] Stefanie Rinderle. The examination of OLAP-tools for determination of customer preferences in e-commerce (in German). Master's thesis, University of Augsburg, Nov. 2000.

[14] Franck Rousseau and Andrzej Duda. An advanced multimedia infrastructure for www-based information systems. In *International Workshop on Advanced Issues of E-Commerce and Web-Based Information Systems*, pages 108–115, Santa Clara - California, USA, April 1999.

[15] Mareike Schoop and Christoph Quix. DOC.COM: combining document and communication management for negotiation support in business-to-bussiness electronic commerce. In *Thirty-Fourth Annual Hawaii International Conference on System Science (HICSS)*, Maui, Hawaii, January 2001.