An Empirical Study of the Effect of Coherent and Tailored Document Delivery as an Interface to Organizational Websites

Cécile Paris, Mingfang Wu, Anne-Marie Vercoustre, Stephen Wan, Peter R. Wilkins, Ross Wilkinson

▶ **To cite this version:**

Cécile Paris, Mingfang Wu, Anne-Marie Vercoustre, Stephen Wan, Peter R. Wilkins, et al.. An Empirical Study of the Effect of Coherent and Tailored Document Delivery as an Interface to Organizational Websites. The Adaptive Hypermedia Workshop at the 2003 User Modelling Conference, Jun 2003, Pittsburgh, USA, pp 133 - 144. inria-00090190

HAL Id: inria-00090190
https://inria.hal.science/inria-00090190

Submitted on 31 Aug 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Empirical Study of the Effect of Coherent and Tailored Document Delivery as an Interface to Organizational Websites

Cécile Paris, Mingfang Wu, Anne-Marie Vercoustre, Stephen Wan,
Peter Wilkins, Ross Wilkinson
CSIRO, Australia

Abstract:

Increasingly, people find information about corporations through their web sites, as opposed to reading printed brochures. Typically, the only tools they have to find that information is searching and browsing. But can we do better with a modest user model and coherent delivery? Can we provide, through the web, more relevant and coherent information about a corporation, information specially tailored to specific clients and presented in a useful and understandable manner? Our conjecture was that a tailored hypermedia system offering tailored and coherent information would indeed provide a good alternative to the current search-and-browse facilities, also allowing corporations to be more responsive to their clients' needs.

In this paper, we present a preliminary study aimed at testing this hypothesis. We designed an experiment aimed at investigating the effect of delivering coherent and tailored information in response to user's query, in comparison to delivering the results of conventional search engines. We employed our experimental system, PERCY, which exploits information retrieval, user modeling and discourse-oriented technologies to retrieve and select parts of existing documents from CSIRO website and assemble them back into a coherent and tailored webpage. Our experiment consisted of asking users to evaluate and compare the results produced by PERCY and those produced by a typical search engine, given a query about the some project or research topic at CSIRO. Our experimental results indicate that users preferred the output of Percy, as it attracted more their attention and they understood the results better.

Keywords:

information retrieval, discourse model, evaluation, tailored delivery, coherent delivery

1. Introduction

Increasingly, people find information about corporations through their web sites, as opposed to reading printed brochures. Typically, the only tools they have to find that information is searching and browsing. It is quite often hard for visitors with very different backgrounds and goals to readily find the information they need, especially when they are not familiar with the corporate organization or its website structure. This is because search results are usually neither tailored to the visitors' needs, nor presented in a way that is easily understandable and recognizable to the visitor.

Can we do better with a modest user model and coherent delivery? Can we provide, through the web, more relevant and coherent information about a corporation, information specially tailored to specific clients and presented in a useful and understandable manner? While tailoring of search results has been exploited in the past [5,6,7], the effect of coherent delivery

has not. Yet, the content *and* the presentation of the information to be delivered are equally important. On the one hand, content needs to be relevant and targeted to the advocated community. On the other hand, the organization and appearance of the presentation should help its understanding by the reader and enhance its usefulness for the reader. From the point of view of the author (or the corporation), it should also facilitate the delivery of the intended message to the audience in a clear and succinct way.

Within this context, our conjecture was that there might be a more effective way to provide the information that users needed. In particular, we believed that producing a coherent document about the relevant information would be more easily understood and more useful to the users. Our conjecture was thus that a tailored hypermedia system offering tailored and coherent information would indeed provide a good alternative to the current search-and-browse facilities, allow corporations to be more responsive to their clients' needs, and also provide the right information to the right people.

We observed that most corporations (including ours) typically have paper-based brochures intended to inform the readers of the work performed within the organizations. (CSIRO, for example, has a set of brochures about the different research areas and projects in the organizations.) These brochures form coherent documents that not only provide the relevant information but also introduce the corporation to the readers, provide additional information when appropriate (e.g., contact details). There are two additional important aspects of typical corporate brochures: first, the information is not simply listed as bullet points but rather unfolds in a natural progression of topics. Second, a corporation usually prepares a set of different brochures to send different messages to the different targeted communities. Taking our own organization -- CSIRO (Commonwealth Scientific and Industrial Research Organization) -- as an example: to the general public, we would like to send the message that CSIRO is doing good science and is beneficial to the community; to the university students, we would like to say that CSIRO is an excellent organization with which to study and work; finally, to the business/market developers, we would like to help them to identify which CSIRO research and development outcomes may have potential business interests, and how CSIRO could transfer these outcomes to their business applications. CSIRO thus has different sets of brochures, each presenting the appropriate message while also addressing the needs of the targeted audience.

We hypothesized that such a way of presenting the information would also be useful as an interface to a website instead of the traditional search-and-browse facility. Our system, PERCY, thus aims at delivering such brochure-like documents as a result of a query. Because PERCY generates these documents automatically on the fly, it can actually produce brochures which are much more tailored to the individual needs than paper-based brochures can be. The approach proposed in PERCY is to capture not only a visitor's potential information needs but also the information about the visitor’s personal background, so that the retrieved documents (or parts of documents) can be filtered and tailored. The information can then be organized and presented according to *discourse rules*. With this delivery method, we aim to provide the visitors with more appropriate and readable information than the output of a conventional search engine.

In this paper, we present an experiment we performed to validate this hypothesis. The remainder of this paper is structured as follows. First, we introduce PERCY, our tailored hypermedia system that produces tailored and coherent information about a specific topic in the context of a company web site. We then present our experiment, which aimed at evaluating the value of such delivery system. We conclude with a short discussion of the results.

2. PERCY

Information retrieval is a proven technology for searching large document collections. However, the search result is usually a list of discrete units that are not easy for users to assimilate. In contrast, the discourse-oriented technologies for adaptive hypermedia systems aim at ensuring that the result is coherent and appropriate through explicit text planning or manual authoring. In our work, we have been designing a new approach to tailored information delivery by integrating both technologies [10, 11]. As shown in Figure 1, this approach exploits a user profile, a set of discourse rules (which represent how one presents information in a particular domain) and a set of web data sources. Given a user query, the system uses the discourse rules and the user model to decide which information to present and how to present it in a coherent way.[1] PERCY is a specialized application embodying this approach. As previously mentioned, it produces coherent brochure-like documents about a corporation, personalized to the user's queries and user type. Our current prototype uses CSIRO as the test corporation.

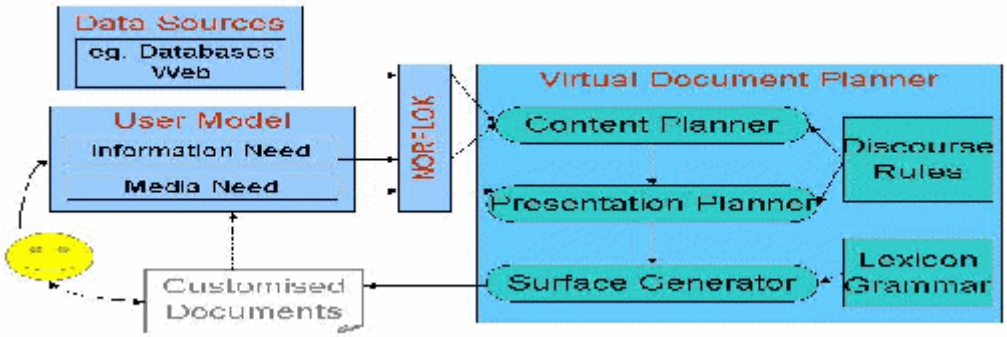

Fig. 1. The System Architecture of PERCY

The first task for PERCY is to obtain a visitor's profile through a web form. The profile is then stored in the user model for information retrieval, filter and presentation planning. Based on a requirements analysis performed in conjunction with professional science communicators, business development personnel and our website managers, we have identified for CSIRO the following characteristics as being important:

Stereotypical information:

- Current job position (e.g., a scientist or a research student or a CEO);
- Industry sector (e.g., science, engineering, finance);

Individual information:

- Interests (e.g., data mining, collaborative technology), which forms the initial query sent to the system; and
- Preferred delivery channel (e.g., web page, hand-held device).

1Our approach also allows for the delivery of the information on a variety of delivery channels, including paper, hand-held devices and the web [2].

The stereotypical information is used at the high level to decide what kinds of information to deliver. For example, a scientist and a CEO may use the same set of keywords to describe their interests when they visit the CSIRO website. However the research scientist may want to know more about the research activities and technical publications, while the CEO may want to know the potential business opportunities. The information about "industry sector" is important to guide PERCY to focus on the most relevant activities in the appropriate sector whenever possible. The individual information is then exploited to decide what information to deliver and how to deliver it. The information about "delivery channel" helps PERCY ensure the document produced is appropriate in its layout to the delivery platform (e.g., a web screen is different from a printout). We will discuss only web-based delivery in this paper.

Given a query about a topic, the virtual document planner (VDP) first creates a discourse tree by calling the content planner (or discourse planner). The discourse tree represents which information is to be included in the text and how it should be organized. It uses metadata to retrieve relevant portions from existing webpages. The discourse planner, modeled on the one described by Moore and Paris [9], uses a library of discourse rules (or plans). These indicate how a discourse goal can be achieved, and how various segments are related to each other to form a coherent whole[2]. The VDP then calls a presentation planner, which decides how best to express the discourse tree. The presentation planner decides on the exact amount of the information to be included in the document and, if necessary, the navigation needed. It also decides upon a presentation suitable for the medium chosen by the user, and augments and annotates the discourse tree appropriately. Finally, the surface generator traverses the final discourse tree to generate the text, delivering information retrieved from web sources, generating new sentences when appropriate, and, finally, producing device specific layout tags.

The discourse rules used by PERCY were produced from a study of corporate brochures. The document which results from PERCY thus typically starts with a brief introduction of the corporation. Then a summary of the search results is presented. This summary explicitly states which query terms were matched or not matched. To improve the search results, PERCY exploits a thesaurus. The final document specifies the synonyms that were added as keywords to the search, to ensure the user is not confused by the results. Finally, summaries of specific projects relevant to the query are given, together with contact information. When the document is a hypertext presentation, links to details are included. Overall, the structure of the document produced reflects that of a paper brochure of the corporation.

Figures 2 and 3 show part of PERCY's output for two queries[3]. Example A is generated for a *scientist* who is interested in *language technology* and in the *science* sector. Example B is generated for a CEO who is also interested in *language technology*, but in the *tourism* sector. Comparing the two examples, we can see that Example A contains more descriptive information about research topics, and that it includes selected publications. On the other hand, Example B gives information about how CSIRO relates to industry. Both examples present the selected information in a coherent manner, reflecting the structure of a conventional brochure, which is a natural (and expected) way to present information about a corporation.

2The Rhetorical Structure Theory (RST) [8] proposed by Mann and Thompson is used here to represent coherency amongst text segments.

3Because of space limitation, Figure 2 and Figure 3 show only the top two matched projects.

3. The Evaluation

3.1 The hypothesis

A visitor gets information from a corporate website either through browsing or searching. PERCY is intended to complement these two mechanisms. With the PERCY system, a visitor can quickly get a bird's eye view of an interesting topic, then can either search the website with a more focused topic, or browse the website through the links suggested by PERCY.

We want to evaluate how useful the PERCY is as a personalized corporate information delivery mechanism, compared with the normal browsing and searching mechanisms. As the effectiveness and speed of obtaining information by browsing largely depends on the usability of that website, comparing PERCY with the website would make the result inclusive. In this experiment, we chose to investigate whether users would *prefer* a tailored and coherent delivery mechanism over the conventional search engine results, and, in that case, we also wanted to try to identify why this may be so.

3.2 Evaluation criteria

There are three important indicators that reflect the effect and effectiveness of an information delivery mechanism for a corporation. The first indicator is a visitor's interactive experience with the delivered information. The second indicator is the visitor's attitudes towards the received message. And the third indicator is the further interactions with the corporation caused or influenced by the visitor's attitude and belief. The third indicator is indirect, and hard to test in a laboratory setting. In our experiment, we concentrated on the first two indicators.

The first two indicators are measured through two sets of questionnaires to assess the visitors' satisfaction with the delivered content (i.e. the quality of the tailored content) and the presentation format [3] (i.e. the quality of the presentation), the visitor's intention (whether subjects have intention to get further information), and preference. Given two systems being evaluated and compared, one questionnaire (the post-system questionnaire) gets the subjects' attitude independently on each test system, while the other questionnaire (the exit questionnaire) gets the subjects' attitude of the second system compared with the first system.

3.3 Experimental setting

Test systems The two test systems are the CMIS (CSIRO, Mathematics and Information Science) website search engine (P@NOPTIC) [1, 4] and PERCY used on the web. To a user, PERCY is very much like a search engine. The only difference is that the user sends only a query to a search engine, while PERCY requires users to provide minimal amount of personal information in addition to the query.

Subjects Only one type of users (university students) was considered in this experiment. So the stereotypical information for all subjects was set as research students from science sector. During the experiment, these students were asked to play the role of research students, interested in the four given research topics. For each topic, subjects were required to find out what CMIS is doing in that area, who are the key players, and other research related issues. They were allowed a maximum of 5 minutes for each topic.

Twenty subjects were recruited. Nineteen of them were university students (from the second year and above) and the other one was a fresh graduate working in a research environment. All subjects were from the computer science department of an Australia university. The average age of the subjects was 21.65. On average, they had 5.33 years online searching experience; their familiarity with CSIRO and CMIS was very low (1.95 and 1.6 on average respectively, on a five-point Likert scale). Fourteen of our subjects had never searched any CSIRO's website before; two of them had searched the CMISO's website for summer studentship position and staff profile, and they claimed that they had obtained what they had searched for.

Experimental design Four search topics (or queries) were selected from a wide range of CMIS research topics, namely: human computer interaction, mathematic modeling, image analysis and language technology; on the assumption that our subjects would have general knowledge of these topics. As we focused on testing the effect of the delivery, the query for each topic was fixed. No mechanism for providing relevance feedback or for supplying a new query was provided; subjects were restricted to explore the two different types of deliveries. While subjects were interacting with a delivery, they were allowed to browse away from the presented interface to any degree.

To reduce the effect of appearing order of systems and topics, the Latin-square experimental design was adopted in this experiment, as shown in Table 1. In this design, four test topics are divided into two blocks (B1 and B2). Topic blocks and systems are then rotated. Each subject uses each system on one block of topics. Subjects are required to fill in an entry questionnaire, a post-system questionnaire (refer to Appendix I) after each system, and an exit questionnaire (refer to Appendix II) at the end.

Table 1. Experimental design

Subjects	**System, Topic**
No. 1-5<	Percy, B1 Search, B2
No. 6-10	Percy, B2 Search, B1
No. 11-15	Search, B1 Percy, B2
No. 16-20	Search, B2 Percy, B1

3.4 Experimental results

We used two questionnaires: a post-system questionnaire and an exit-questionnaire. The post-system questionnaire was to test whether the users obtained the information they desired, in a form that was useful, and whether the output of the system served as a good introduction to what the organization was doing on a specific topic. The post-system questionnaire included 9 questions, which meant to evaluate the systems with respect to the content of the information provided, its format, its coherence in terms of presentation, and, finally, whether based on what they obtained, they would like to have further interactions. The exit questionnaire was to elicit users' preferences between the two systems. We first present the results from the post-system questionnaire, followed by those of the exit questionnaire.

3.4.1 Results from the post-system questionnaire

The results from the subjects' responses to post-system questionnaire are shown in Table 2. Each question in this questionnaire measured certain aspects of the delivered information. (See the Appendix for the questionnaires themselves.)

Table 2. The average of the responses to the post-system questionnaire across all subjects

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Search	3.50	3.60	3.35	3.70	3.65	3.50	2.75	3.95	3.15
Percy	3.65	3.30	3.65	3.55	4.00	3.80	3.50	3.90	2.90
p <	NS	NS	NS	NS	NS	NS	0.06	NS	NS

In general, there were no significant differences[4] between two systems on all questions in the post-system questionnaire. However, PERCY tended to get higher scores for the coherence of the presentation of the information, especially a strong tendency appears for question 7 ("I think the presented information serves well as a useful online brochure"), which meant to measure coherence and readiness of the information delivered. This is encouraging and may point to the fact that users liked obtaining information in a coherent and well presented manner. We examine the results in more detail below.

Content Questions 1-4 measured the content of the delivered information. For questions 1 ("The system provides sufficient information") and 3 ("The information provided by the system meets my need"), PERCY got higher scores than the search system. For questions 2 ("The system provides sufficient information") and 4 ("The system provides me comprehensive information"), the search engine got higher scores. It is hard to interpret these 2 results, especially as there are no significant differences. The higher scores for the search engines for Questions 2 and 4 may be related to the amount of information presented: the search engine presents every piece of information related to a project, while PERCY selects only the most important information given the query and the user.

Format Questions 5 ("The structure of the presented information is clear to me") and 6 ("I think the presented information is organized in a useful format") measured the format of the delivered information. For these two questions, PERCY got higher scores than the search system. While we can notice a tendency,, we cannot derive strong conclusions yet as the differences are not significant.

Question 7 ("I think the presented information serves well as a useful online brochure") provides a measure for both content and format, and was meant to validate whether the delivery was an appropriate introduction or interface to a corporation. Again, the PERCY system has the higher score than the search system. The difference is significant at 0.06. While this may be because subjects were likely to compare the delivery with their own model (or expectation) of a brochure, it provides validation for the use of personalized and coherent delivery as an interface to a corporation's website.

Intention Questions 8 ("I would like to get more information on some specific projects presented") and 9 ("I would like to get more information because I have not got the needed information") were aimed at understanding the subjects" desired further actions. Both systems rated the same on question 8. For question 9, however, more subjects from the search system thought they want to get more information because they had not obtained the needed information. This is curious and interesting since, at the same time, subjects gave higher scores to the search system for questions 2 and 4, as we saw earlier. These seemingly contradictory results may indicate that the *amount* of information presented is not related to the *quality* of the information.

4All significant tests are based on the two tailed t-test.

3.4.2 Results from the exit questionnaire

Besides the post-system questionnaire, subjects were also required to fill in an exit-questionnaire based on their experience of the two systems. This exit-questionnaire was to evaluate their preferences, since it was based on their experiences with both systems. In contrast, the post-system questionnaire was purely based on their experience of that particular system. (Again see the appendix for the questionnaire.)

Preference The subjects' responses to the exit-questionnaire are listed in Table 3. They show that subjects ranked PERCY higher than the search engine for all questions. Notice that for question 6, a lower score indicates more focused delivery. Significant differences are found for questions 4 ("The information delivered by the second system attracts my attention better"), 5 ("The second system provides a better explanation on why a piece of information is presented") and 7 ("Overall, I prefer to use the second system as an online brochure of the searched topic").

Table 3. The average of the responses to the exit-questionnaire across all subjects

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**
Search	2.70	3.00	3.10	2.80	3.30	3.50	2.70
Percy	3.70	3.60	3.50	4.20	4.20	3.40	4.00
p <	NS	NS	NS	0.01	0.03	NS	0.04

4. Discussion

It is interesting to note that the two systems were evaluated as very similar in terms of the quality and quantity of information delivered (as shown for the results of the post-system questionnaires). Yet, subjects did prefer the system that tailored information to their needs *and* presented content in a coherent fashion. Such a kind of information delivery mechanism allowed the subjects to focus their attention better on their interested topics, to understand better why a piece of information is presented, and finally, to have a better overview of the organization's projects in the focused area. The study thus suggests that producing coherent and tailored hypermedia based on a user stereotype is an effective way to deliver information about an organization, probably more effective than the traditional search and browse mechanisms. We believe that, as the amount of information available for users to look at increases, such a means of delivery could in fact become increasingly valuable, as the need for focused attention (and fewer distractions) increases.

The tailored hypermedia we produced necessitated only a small user model and a retrieval engine coupled to a discourse planner. Yet this enables our system delivers tailored information in a coherent manner (in this case, resembling a corporate brochure), which was found useful by the users in comparison to the output of a conventional search engine. This approach thus provides an attractive interface to an organization's website.

The experiment is of course still limited, in terms of the number of subjects, their types, and the scope. For example, while subjects judged that PERCY attracted their attention better, whether this increased focus can generate further interaction with the corporation or the corporate website needs to be explored further.

References

1. CSIRO, http://www.panopticsearch.com
2. Colineau, N. and Wan, S. (2001) Mobile delivery of customized information using Natural Language Generation. In *Monitor* (Special Issue on Wireless Communication Special), 26 (3), September-November 2001. pp 27 - 31.
3. Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly (13)*, 1989, pp. 319-340.
4. Hawking, D., Bailey, P. and Craswell, N. (2000). Efficient and Flexible Search Using Text and Metadata. CSIRO Mathematical and Information Sciences, TR2000-83.
5. Korfhage, R. R. (1984) Query Enhancement by User Profiles. In *Proceedings of the 7th annual international ACM SIGIR conference on Research and development in information retrieval*. Cambridge, England. pp. 111-121.
6. Laine-Cruzel, S., Lafouge, T., Lardy, J.P. and Abdallah, B. N. (1996). Improving Information Retrieval By Combining User Profile and Document Segmentation. *Information Processing & Management* Vol. 32(3), pp. 305-315.
7. MacKinnon, L. & Wilson, M. (1996) User Modelling for Information Retrieval from Multidatabases. In C Stephanidis (Ed.) *Proc. 2nd ERCIM Workshop on "User Interfaces for All"*. Prague, Czech Republic, November 1996.
8. Mann, W. C. and Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. In *Text* 8(3): 243-281.
9. Moore, J. D. and Paris, C. L. (1993). Planning Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information. In *Computational Linguistics*, Cambridge, MA. Vol 19(4), 651-694.
10. Paris, C., Wan S., Wilkinson R. and Wu M. Generating Personal Travel Guides – and who wants them? (2001). In *Proc. of the International Conference on User Modeling (UM2001)*; Sonthofen, Germany, July13-18, 2001.
11. Wilkinson R, Lu S, Paradis F, Paris C, Wan S and Wu M. (2000). Generating Personal Travel Guides from Discourse Plans. In *Proc. of International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*. Trento, Italy, August 2000.

Appendix

Appendix I – Post–system questionnaire

Q1: The system provides me accurate information. 1-5

1	2	3	4	5
Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

Q2: The system provides sufficient information
Q3: The information provided by the system meets my need.
Q4: The system provides me comprehensive information.
Q5. The structure of the presented information is clear to me.
Q6: I think the presented information is organized in a useful format.
Q7: I think the presented information serves well as a useful online brochure.
Q8: I would like to get more information on some specific projects presented.
Q9: I would like to get more information because I have not got the needed information.

Appendix II – Exit questionnaire

You have just used two different systems that deliver the searched information in two different forms. In answering each the following questions, please use the **first system** as your basis of comparison, i.e. how good is the **second system**, as compared to the first one, in the following aspects.

Q1: The second system enables me to focus on the search topic better.

1	2	3	4	5
Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

Q2: The second system makes my search task easier.
Q3: The second system delivers the higher quality information.
Q4: The information delivered by the second system attracts my attention better.
Q5: The second system provides a better explanation on why a piece of information is presented.
Q6: The second system provides more comprehensive information.
Q7: Overall, I prefer to use the second system as an online brochure of the searched topic.

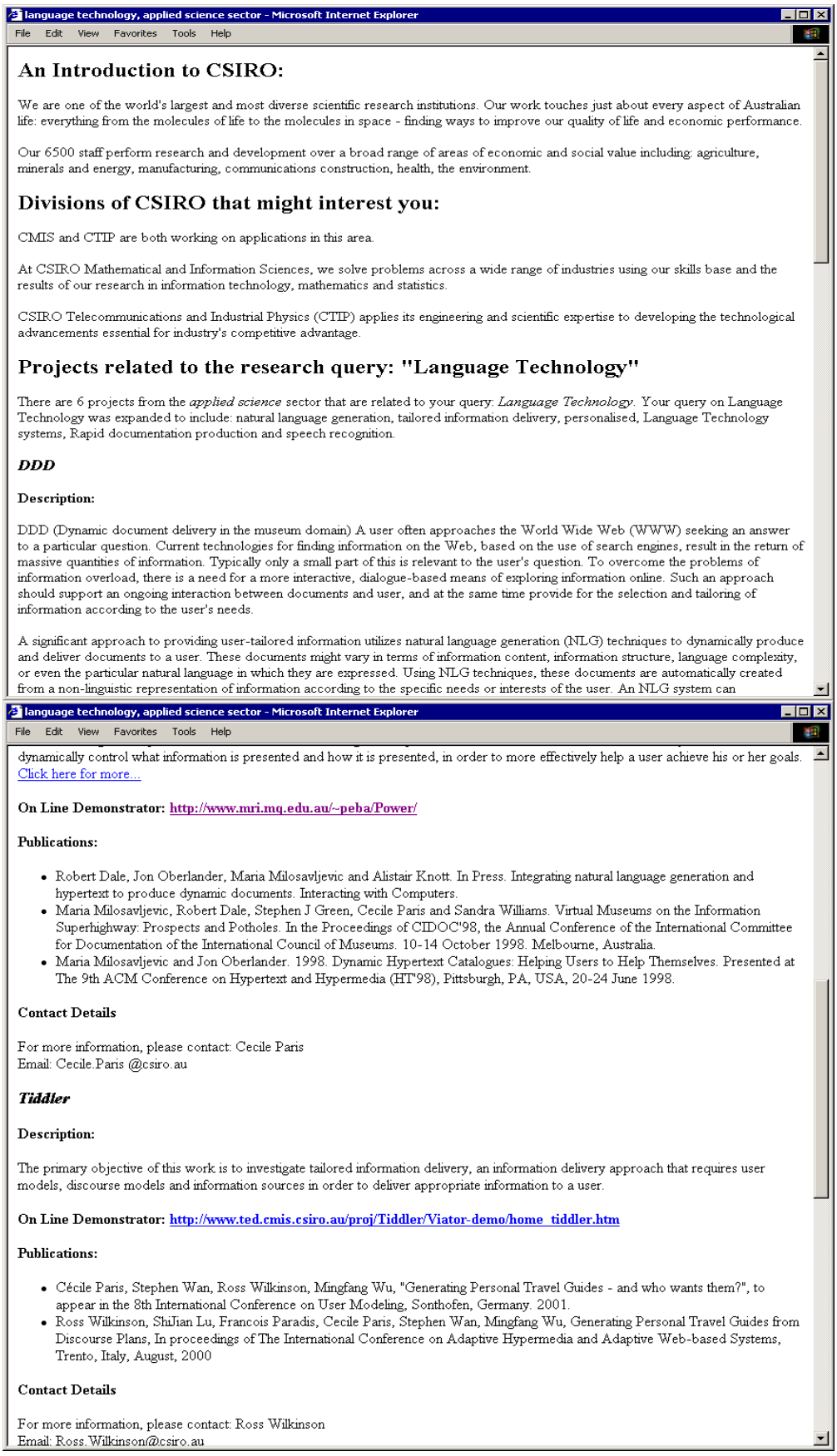

Fig. 2. Example A: the delivery for a scientist, with query in Science sector

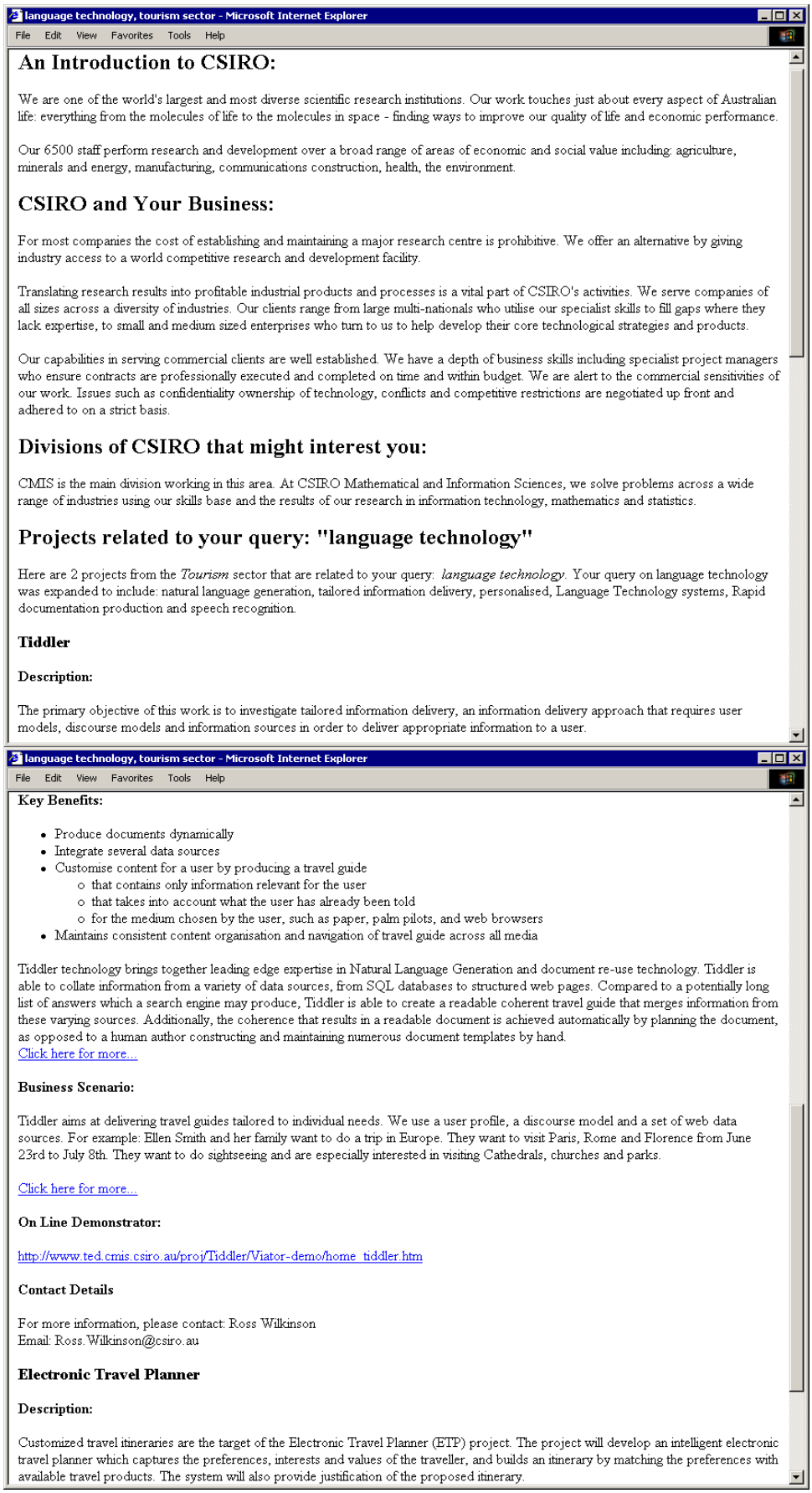

Fig. 3. Example B: the delivery for a CEO, with query in the Tourism sector