

Heading towards virtual-commerce portals

Autor(en): **Van Krommer, Robert / Rajman, Martin / Bourlard, Hervé**

Objekttyp: **Article**

Zeitschrift: **Comtec : Informations- und Telekommunikationstechnologie =
information and telecommunication technology**

Band (Jahr): **78 (2000)**

Heft 9

PDF erstellt am: **05.06.2024**

Persistenter Link: <https://doi.org/10.5169/seals-876467>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.



Exploration Programmes:

Corporate Technology Explores Future Telecommunications

Heading Towards Virtual-Commerce Portals

The goal of virtual-commerce is to let customers enjoy e-shopping anytime, anywhere, and anyhow. Considering the fact that most customers do have a voice access, we promote, in this article, the advent of natural speech communication as a business channel additional to e-commerce. Within the InfoVox project, a demonstrator has been built as an intermediate milestone; the target application is a dual Web and voice access to restaurant information and it represents a first step towards virtual-commerce.

Convergence between Internet and telecommunication technologies represents a fruitful area for innovation. The Exploration Programme "Advanced and Value-Added Communication Services" addresses these emerging technologies, and, through the underlying projects, sets a special focus on the following hot topics: voice over IP, multimedia communication & services, interaction & mobility services, innovation in voice services, and virtual-commerce portals.

With its Exploration Programmes, Corporate Technology is exploring telecommunication technologies and new service possibilities with a long-term view of 2–5 years. Further, the expertise built up in the course of this activity enables active support of business innovation projects.

Virtual-commerce, as defined in this article, is a unified service provision of "x"-commerce applications (i.e. e-commerce, m-commerce, v-commerce). Virtual-commerce has emerged from the convergence of

ROBERT VAN KROMMER, MARTIN
RAJMAN AND HERVÉ BOURLARD

telecommunications and the Internet. Several other instances of the convergence phenomenon have already been observed all along the communication path, such as Unified Messaging, or voice over IP. Virtual-commerce is the next step on that convergence path; it takes place at the server side. From a technical perspective, the virtual-commerce portal is an e-market platform that is able to handle a plurality of services accessible through various business channels, including Web, WAP, and voice access. The business framework is defined as an Application Service Provider (ASP) [1] solution applied to e-commerce storefronts.

It is common practice in the telecommunication field to implement platform-specific services. Conversely, virtual-commerce favours a top-down approach of unified service provision, where many services and several access channels share common software components. The purpose is to offer transaction consistency across business channels.

In this article, the InfoVox project (CTI/KTI -4247.1) is described as a virtual-commerce portal with a focus on its

most challenging part, the voice access and its natural speech interface. Furthermore, special attention has been devoted to the related innovation process. In the conclusion section, perspectives and market trends are given to highlight the future potential of virtual-commerce.

InfoVox Project

CT is actively involved in this project. Besides Swisscom AG, the InfoVox team includes the following partners: EPFL (LIA), IDIAP, VOXCom AG, and Omédia AG.

EPFL, the project leader, co-ordinates the three working groups:

- (1) telephone interface and speech recognition,
- (2) human-machine interaction, and
- (3) the database module.

So far, the focus was on designing a

multi-channel access within a real application. For voice access, an innovative approach has been taken with the implementation of a natural speech interface; this lets customers use their own wording in the spoken dialogue. For the combination of Web and voice access, the outcome is that each interface modality needs a specific content presentation. Consequently, parts of the database have been pre-processed for each specific access channel. In our application design, the customer-centric thinking of figure 1 has been favoured in each phase of the innovation process. Furthermore, in the second part of the InfoVox project, an evaluation of the service prototype will be performed by a representative customer subset. The dialogue completion rate will deliver an additional figure of merit to predict the feasibility of virtual-commerce.

InfoVox Project Description and Highlights

Application Description

The target application is a restaurant information service with an elementary transaction feature: the user may perform a reservation. The information content is represented with real data of the restaurants of Martigny. The work of collection and processing the relevant information has been accomplished in the third working group. The schematic view

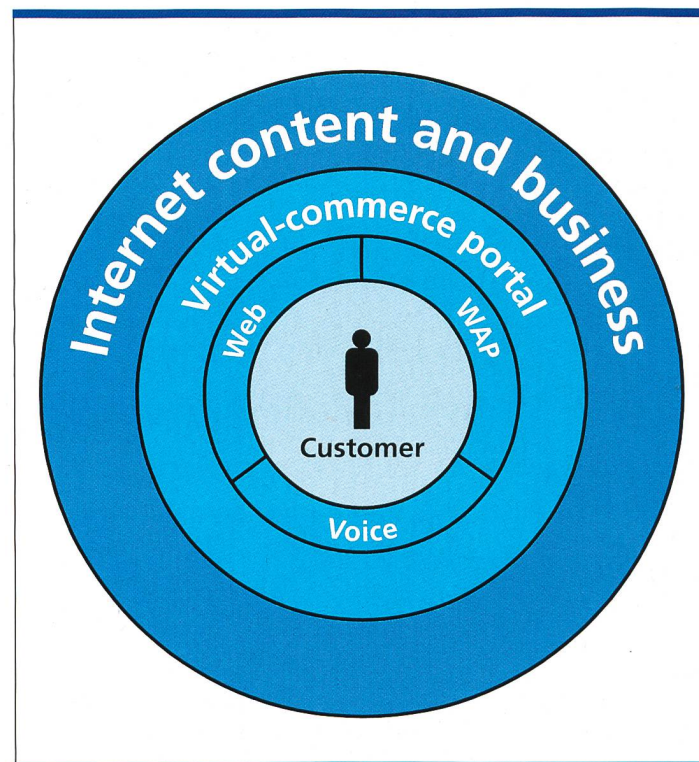


Fig. 1. The customer-centric approach of the virtual-commerce portal.

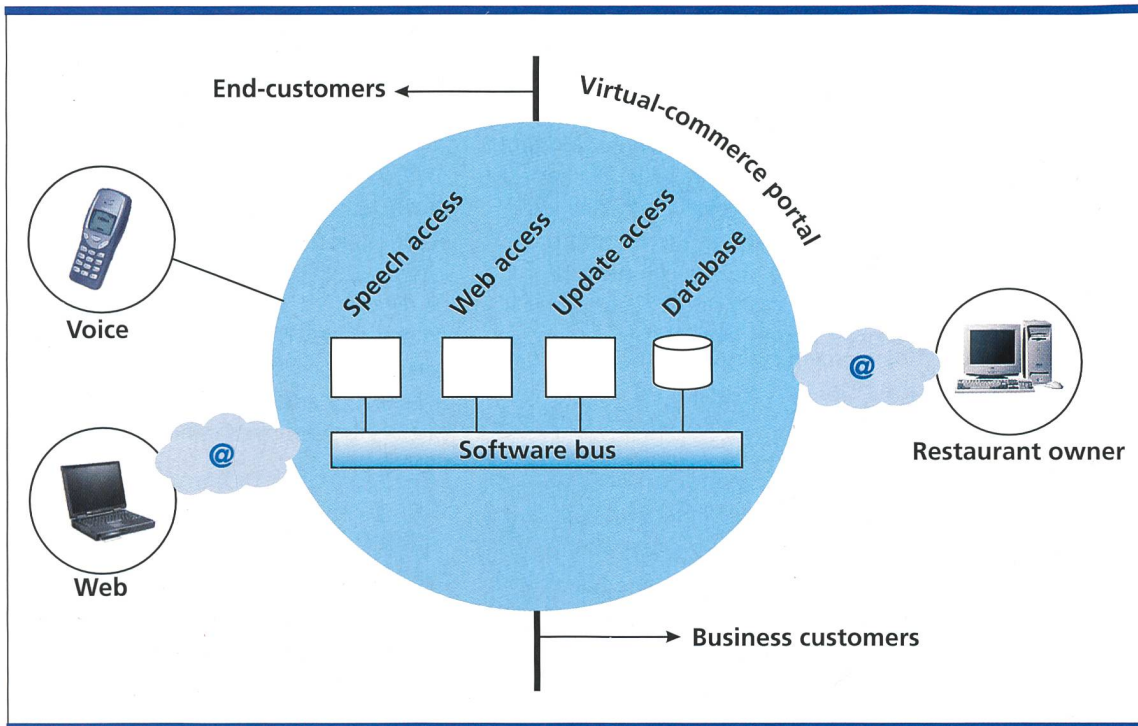


Fig. 2. InfoVox: the restaurant information service with end-customers' multi-channel access and the information maintenance channel for business customers.

of figure 2 depicts the links between the various channels for information access and update. As in an ASP model for e-commerce hosting, there are two types of customers:

- (1) the end-customers, who will access the portal for a reservation or information retrieval, and
- (2), the business customers, who will upload their e-shop to benefit from the portal platform as a unified service provision tool.

Innovation Process

The basic idea of the innovation process is to gradually increase the development/testing costs and therefore to minimise the risk of expensive failure. The current process is split into three phases: the alpha, beta and operational phase. Today's status of the project still belongs to the alpha phase; that is, the system goes through internal cycles of testing, evaluation and improvement. The move to the beta phase is conditioned on the approval of internal experts. The beta phase is characterised by the involvement of a representative sample of external customers within the design cycles. That first acceptance test is envisioned at the end of the project. Another motivation supporting this iterative process is the necessity to increase the amount of collected data that allows the learning algorithms to build accurate language, dialogue and speech models.

Wizard-of-Oz Experiment

In order to address the natural speech interface, our current approach of dialogue design uses a Wizard-of-Oz set-up. This experiment is defined as follows: within the system, a human operator simulates the automatic speech recognition. The user does not perceive that it is a fake recognition system. The goal of the experiment is to collect the necessary bootstrap data to learn the language model as well as to identify the commonly used keywords. For that purpose, all collected data have been manually transcribed.

Natural Speech Interface

Natural speech has the advantage to release constraints on the dialogue module, making it more flexible to the user to introduce complex and natural queries. Indeed, no specific vocabulary or dialogue sequences are strongly imposed by the system to the user, although the system may "suggest" a loose dialogue structure. However, this enhanced interaction has many implications on the design and complexity of the speech recognition module which should be able to accommodate out-of-vocabulary words, as well as sentence structures not matching the predefined syntax. In our project, a state-of-the-art hybrid HMM/ANN recogniser [2] (using Hidden Markov Models together with Artificial Neural Networks) has been trained on a large

telephone database (Swiss-French Polyphone). Such a system is known to be particularly robust and flexible to the lexicon. A first evaluation of an early version of the resulting system shows recognition performance comparable to other recognisers used in similar contexts. Furthermore, additional evaluations of the telephone-channel effects (through a telephone line simulator) have already been performed [3] to measure the expected robustness in real-life situations. As far as dialogue management is concerned, the main goal is to produce an abstract dialogue model that can easily be specialised and tested within the target application. The resulting dialogue model relies on a mix initiative frame-based dialogue strategy, implemented in a form of a "reconfigurable" finite state machine. A first version of the dialogue model has been designed and tested during the Wizard-of-Oz experiment. A refined version of the dialogue model was then produced and implemented in a software demonstrator. The demonstrator is currently used to carry out intensive testing to improve the quality of the dialogue strategy and to increase the fluency and naturalness of the dialogue.

Improved Recognition for Word Sequences

In voice-driven e-commerce applications, a common practice is to identify goods, clients, or bank accounts with a se-

quence of digits or alpha digits. Since the recognition performance on these particular items is very critical, a special recognition module has been developed to address the point. So far, the performance level that has been achieved is a 99,5% recognition rate for single words and 97% for strings with a length of 6 words. Within this specific module, as in the case of the natural speech interface, in-depth work [4] has been pursued to study the robustness in adverse conditions.

Conclusions

Inspired from Web pages and their service description concept, the aim of virtual-commerce is to avoid proprietary multi-channel formalisms, and therefore to rely on available standards, such as HTML, WML, and VoiceXML. The expectation is to benefit from a world-wide effort in designing tools, and consequently, to take advantage of a drastic reduction in time-to-market and development cost. In the emerging market of virtual-commerce, the InfoVox project delivers insights in the work and research needed to bridge the gap between the current pre-competitive status and an ASP solution. Further engineering resources, additional funding as well as industrial support will be required to extend the InfoVox application to a generic virtual-commerce portal.

It is our belief that we are heading towards virtual-commerce portals based on the following trends:

- *Migration of content and related business to the Web*: Any software content's value is directly enhanced by the unparalleled distribution capability of the Internet; therefore, today's appropriate location of content is on the Web.
- *Continuous shopping*: Automated systems are obviously not bound by any time constraints and therefore they offer additional convenience to customers.
- *Self-service*: Many customers become used to service automation, as long as the service itself is clearly understood by them. Moreover, through bypassing human operators, waiting on the phone will become unlikely.
- *One-to-one marketing*: This marketing concept has the aim to target services at customers as precisely as possible. The one-to-one marketing Nirvana is about segmenting audiences down to

individuals and to finally maintain a virtual relationship with each single customer.

- *Personalisation*: In virtual-commerce services, customer behaviour and preferences may be learned by the system; this further improves the quality of service by adapting the settings to the user (e.g. the speech models). Ultimately, every single customer will acquire her/his own personal service.

References

- [1] Richard Comerford, "The new software paladins", IEEE Spectrum, June 2000, volume 37, number 6.
- [2] René Boite, Hervé Bourlard, Thierry Dutoit, Joël Hancq et Henri Leich, "Traitement de la parole", collection électricité des Presses Polytechniques et universitaires Romandes, 1999.
- [3] Sebastian Möller, Hervé Bourlard, "Real-Time Telephone Transmission Simulation for Speech Recognizer and Dialogue System Evaluation and Improvement", 6th International Conference on Spoken Language Processing, Beijing, China, October 16–20, 2000.
- [4] Robert van Kommer et Beat Hirsbrunner, "Apports d'une modélisation par réseaux de neurones multicadences à la reconnaissance de la parole", JEP'2000 Aussois – France, juin 19–23, 2000.

- *The advent of natural speech communication*: Voice is the most natural way for humans to interact; year after year, substantial progress is made in improving the machine capabilities.

Outlook

Having reached the second half of the InfoVox project, the main effort will now be devoted to increase the perceived quality of service. In the short-term and given the successful completion of the beta-test, the restaurant information service will be deployed in Martigny and on the Web. 4

Abbreviations

ASP	Application Service Provider
CTI	Commission pour la Technologie et l'Innovation
HTML	Hyper Link Markup Language
KTI	Kommission für Technologie und Innovation
m-commerce	mobile e-commerce
v-commerce	voice-driven e-commerce
VoiceXML	Voice eXtensible Markup Language
WAP	Wireless Application Protocol
WML	Wireless Markup Language

Zusammenfassung

Die Konvergenz zwischen der Internet- und Telekommunikationstechnologie bringt ein weiteres, höchst aktuelles Thema hervor: Das Virtual-Commerce Portal. Das Virtual-Commerce-Konzept hinterfragt die historische Beschränkung auf dienstspezifische Plattformen. An deren Stelle wird ein Top-down-Ansatz, basierend auf einer Plattform mit Mehrfachzugriff, verfolgt. Dadurch können E-Commerce-Transaktionen auf unterschiedlichen Kanälen (Web, WAP und Sprachkanal) konsistent abgewickelt werden. Im Rahmen des InfoVox CTI/KTI-Projektes wird das Hauptgewicht auf den anspruchsvollsten Kanal gelegt, den Sprachzugriff. Die verfügbaren Zwischenresultate geben wertvolle Einblicke in das zukünftige Potenzial eines Virtual-Commerce Portals.

EPFL (LIA) <http://liawww.epfl.ch>
 IDIAP <http://www.idiap.ch>
 VOXCom SA <http://www.voxcom.ch>
 Omédia SA <http://www.omedias.ch>
 CTI/KTI <http://www.admin.ch/bbt/KTI>

The authors gratefully acknowledge the support of CTI/KTI for carrying out this work.

Robert van Kommer is working at Swisscom AG, Corporate Technology. Since 1992, he has initiated together with research partners several projects in the field of voice-activated services. Specifically, he led the Swiss Polyphone project that consisted in collecting the speech corpus necessary for deploying speech recognition in Switzerland. He is currently pursuing a PhD thesis in the PAI group of Prof. Beat Hirsbrunner at the University of Fribourg.

Martin Rajman is senior lecturer at the Artificial Intelligence Laboratory of EPFL (LIA). He is in charge of the research and teaching activities in the field of Natural Language Processing (NLP). His research work focuses on knowledge extraction out of textual data (Text Mining), probabilistic parsing, NLP interfaces and methods for NLP tools evaluation. Member of ACL (Association for Computational Linguistics), he is author and co-author of several publications and responsible for various academic and industrial research projects in the field of language engineering.

Hervé Bourlard is Director of IDIAP (Dalle Molle Institute for Perceptual Artificial Intelligence, Martigny) and Professor at the Swiss Federal Institute of Technology at Lausanne (EPFL). Dr. Bourlard has been active in speech recognition research and application development for more than 15 years. He has published over 130 papers in international journals and conferences in the field, as well as two books. He is Fellow of IEEE (for contributions in the fields of statistical speech recognition and neural networks).

ETSI and Global VSAT Forum

Next-Generation Satellite Communications Programs

The European Telecommunications Standard Institute (ETSI) and the Global VSAT Forum have signed a landmark co-operation agreement to collaborate on the development of technical arrangements related to advanced satellite-based Very Small Aperture Terminal (VSAT) systems.

The objective of the agreement, signed in Singapore on 6 June 2000, is to promote international co-ordination, with the aim of contributing to the establishment of the global information infrastructure. ETSI and the Global VSAT Forum will adopt a complementary approach in the VSAT domain by co-ordinating the technical work being undertaken by each organization. Karl Heinz Rosenbrock, ETSI Director-General, states:

"A first joint effort between the Global VSAT Forum and ETSI Technical Committee Satellite Earth Station & Systems (SES) was held at the end of 1999 and attracted a good number of participants including satellite operators, manufacturers and representatives of the European Commission. This meeting was a real success and both organizations have decided to strengthen their relationship." He adds: "The Global VSAT Forum is very interested in the ETSI approach to standardization of VSAT terminals. The two organizations will now liaise to extend the use of ETSI TC SES standards worldwide."

The three-year agreement involves close co-operation between the two organizations, with the exchange of satellite-related technical information and participation in programmes that are already well underway at both ETSI and the Global VSAT Forum.

ETSI

The European Telecommunications Standards Institute (ETSI) is a non-profit making organization whose mission is to produce the telecommunications standards that will be used for decades to come throughout Europe and beyond. Based in Sophia Antipolis (France), ETSI unites 773 members from 52 countries

inside and outside Europe, and represents administrations, network operators, manufacturers, service providers, research bodies and users. The Institute's work programme is determined by its members, who are also responsible for approving its deliverables. As a result, ETSI's activities are maintained in close alignment with the market needs expressed by its members. Homepage: www.etsi.org

Global VSAT Forum

The Global VSAT Forum (GVF) is the UK-based, independent, non-partisan and non-profit organization representing every major world region and every sector of the Very Small Aperture Terminal (VSAT) industry, including satellite operators, VSAT network operators, VSAT manufacturers, system integrators, value added and enhanced service providers, telecom carriers and users. The GVF has more than 80 members and serves as the unified voice of the industry in regulatory, policy and trade fora. The association co-ordinates regulatory and policy solutions at the national, regional and global level, and supports educational and promotional programmes in every nation of the world. Homepage: www.gvf.org

ETSI

Florence Pontieux, ETSI Press Officer
 France
 Tel. +33 4 92 94 43 12
 Fax +33 4 92 38 49 32
 E-Mail: press@etsi.fr