



This paper appears in *Managing Modern Organizations Through Information Technology*, Proceedings of the 2005 Information Resources Management Association International Conference, edited by Mehdi Khosrow-Pour. Copyright 2005, Idea Group Inc.

# Discovering a Solution to SME's Communication Problems: VoiceXML Technology

Zhi Yong Xu and Lindsay Smith

School of Multimedia Systems, Faculty of Info. Technology, Monash University, Clyde Rd., Berwick, Australia,  
{zhi.x, lindsay.smith@infotech.monash.edu.au}

## ABSTRACT

Efficient, cost effective communication is an essential business function for every enterprise including small or medium enterprises (SMEs). With the development of Information Communication Technology (ICT), an important concern should be how to use this technology to efficiently address the communication issues of SMEs. Currently there are two major concerns for SMEs. The first is the cost of call centres which make SMEs eager to seek alternative low-cost adaptations of Information Technology with similar or better functionality. The second is to successfully interlink telephone/mobile-phones and web services. From recent developments in Information Technology, we find that there are opportunities for utilizing VoiceXML technology in addressing these issues. This paper describes some of the current communication problems for SMEs, clarifies the functions of VoiceXML, and indicates a solution to some of these communication problems through the use of VoiceXML.

## INTRODUCTION

The question of how to use Information Technology to address existing SME's communication problems is an important task for ICT experts. Call centres, which can either process telephone calls via an operator or answer them automatically, have been established in the communication infrastructure of many large businesses for some time. Although it has been proven that communication systems efficiency is improved by setting up a call centre (IVR, 2004), SMEs cannot afford to establish a Call Centre of their own. In addition, the Web is increasingly popular as a medium for commerce, with an increase in the number of people shopping online. If the SME's Web facilities are interlinked with telephone/mobile phones, it will make them more competitive because it focuses on what the marketplace needs and wants. Therefore, it is important to find a proper solution to these issues by using Information Technology. This paper explores some of the current communication issues and how Voice-XML could provide better solutions.

## CURRENT SME'S COMMUNICATION PROBLEMS

### Definition of Current SME's Communication Problems

The phrase *enterprise communication* refers to the communication that is carried out between enterprise employees, between employees and customers, and between employees and suppliers (Dictionary, 2000, www site). From the history of business communication, *enterprise communication problems* refer to those communication problems that have occurred in each stage of the evolution of enterprise communication system. *Enterprise communication systems (ECS)*, in the view of Satzinger (2002), are what allow employees to communicate with each other and with customers and suppliers.

### Clarification of Current SME's Communication Problems

Although existing SME's communication systems have facilitated the communication between employees, customers and suppliers, there are

still many aspects that need to be improved. The most significant issues are how to make telephone/mobile phones accessible to the SME Web services, and how to make their communication systems operate similarly to, or even better than, a Call Center with more functionality at a more affordable cost.

### *Interlink Between Phone and Web*

Customers, irrespective of age, can more easily use a telephone to contact suppliers or enterprises than any other form of communication, due to the fact that the phone is much easier to operate and more widely available. With the development of mobile technology, mobile phones are extremely popular today (Australian Bureau of Statistics, 2004, www site). According to the Australian Cellular Market Forecast and Analysis 2003-2007 recently released, the number of mobile users had reached 14.9 million, or 75% of the total country's population by the end of 2003, with a projected 83% of the population being mobile users by 2007 (Press release, 2004, www site).

The Web is also popular today, with an increasing number of people shopping via the Internet. To meet all of these challenges, new Web contents and services, particularly ones that are interlinked with the phone system, need to be developed. Creating such a Web service can make an enterprise more competitive due to more intuitive marketing.

Unfortunately, in Australia, little work has been done in merging telephone and Web networks. In other words, there are few enterprises that can provide a voice portal that allows customers using mobile phones or landline phones to retrieve information from their websites.

### *High Cost of Call Centres for SMEs*

The popularity of telephone and mobile phone make it essential for enterprises to seek a solution for efficiently processing these incoming calls. Establishing a Call Centre meets these marketing requirements since it can receive a large number of telephone calls simultaneously and can be accessible without the limitation of time and location. There are two types of Call Centres, one of which is operated by a group of people, processing a large number of telephone calls at the same time. People can ring this Call Centre at any time and from any place. Alternatively a Call Centre whose traffic is processed automatically by computer systems is known as an auto-Call Centre. These computer systems are referred to as automatic dialling systems or predictive dialling systems (IVR, 2004). The popularity of the telephone service means that an enterprise without a Call Centre will be restricted in meeting the needs of its clients by not being able to process a significant number of telephone calls at the same time.

However, establishing and maintaining a Call Centre is quite expensive. A Call Centre that can process a significant number of calls needs to employ a large number of staff. The cost of establishment for an auto-Call Centre is also extremely high. According to Jackson (2003), IVR (Interactive Voice Response) software used in an auto-Call Centre will cost about \$250,000, and companies that deploy it also require \$10,000

per month for professional service expenses (Jackson, 2003, p70). Hence neither of Call Centres can be afforded by small businesses. The Australian Bureau of Statistics (ABS) shows that in June 2001, 1,162,000 or 72% of businesses are deemed to be 'small' enterprises. This represents an average annual growth rate of 9% in the number of small enterprise operators since the previous survey conducted in November 1999 (Australian Bureau of Statistics, 2004, www site).

*Characteristics of SMEs*

From the above analysis it is clear that it is imperative and significant to interlink phone and Web networks for SMEs to meet customer expectations. Unfortunately, the characteristics of SMEs, in particular, the limited number of personnel and their restricted capital impede the development of merged telephone and Web networks.

In addition the restricted capital of SMEs also hinders the establishment of a Call Centre, even though it is recognized that a Call Centre can help satisfy market requirements.

**EXPLORATION OF VOICEXML**

With SMEs having these communication problems, there are some strategies that might be used to address them. VoiceXML technology is amongst the most promising one at present.

**Definition of VoiceXML**

VoiceXML is defined as an XML-based mark-up language for distributed voice applications; much like the way that HTML is a language for distributed visual applications (Sharma, 2002, p.7). In fact it is an emerging standard for speech applications, which uses "tags" to define a call flow, the dialog between a computer and a person over the telephone network. Danielsen (2000) indicated that VoiceXML could offer a mixed service of traditional voice response and innovative initiative (Danielsen, 2000, pp104-106).

**How VoiceXML Works**

Before clarifying how VoiceXML works, it is necessary to note that different developers use interchangeable terminology in VoiceXML systems. For instance, a *Voice portal* is also known as a *Voice Gateway* or a *VoiceXML Platform*, or a *VoiceXML Server*; a *Text-To-Speech translator* is also called a *Text-To-Speech engine*, or *Text-To-Speech machine*, or *TTS system*, or *TTS package* or *Translation system TTS* or *Voice Translator*. Given that this is an emerging technology, there are likely to be many synonymous terms.

To simplify the complex process of VoiceXML, a diagram of the VoiceXML process is shown in Figure 1, based on previous work of Larson, Schmelzer, Chugh & Jagannathan in this area (Larson, 2002, pp91-93; Schmelzer, 2002, pp.919-933; Chugh & Jagannathan, 2003, pp188-189).

This diagram defines the following steps for the operation of a VoiceXML system.

**Making a telephone call:**

- A user presses the buttons on a telephone and initiates a call.
- Voice Portal receiving a call
- The generated signal is received by the dialled Voice Portal and handed to the Input Module for validation of the signal. Whether the signals can be validated depends upon whether the dialled telephone number is registered and stored in the Voice Portal-Input Module. This list of telephone numbers are VoiceXML numbers for every Web Site registered and stored in the Module, which means that every Web Site has a special VoiceXML telephone number stored in its Voice Portal-Input Module.
- If a signal is not able to be validated, then the telephone number is not registered or stored in the Module and the signal will not be sent to any Voice Portal.
- If a signal is validated, it will be received by the Voice Portal Input Module and passed on to the Voice Portal Interpreter & Controller. The Interpreter & Controller then requests a VoiceXML document from the website.

**Processing a call:**

- When the Voice Portal Interpreter & Controller receives the signal, it requests the VoiceXML document from the related website. This website then responds to the request and sends back the required VoiceXML document to the Voice Portal Interpreter. When the Voice Portal Interpreter & Controller receives the VoiceXML document, it will use its own Text-To-Speech engine to synthesise audio, based on the received document. As soon as the synthesised audio is completed, the Interpreter will send it to the Output Module, which then converts the digital audio signal into an analogue format. The analogue audio signal will then be sent to the telephone via a standard telephone line. The telephone receives the signal and plays it. When the user receives the signal from the handset, the dialogue commences.

**Interacting between user and website:**

- When a user receives a response from the website, he or she can speak to the transmitter. When the transmitter receives a human spoken voice message, it sends an analogue signal to the Voice Portal Input Module for conversion and validation. It first converts analogue signals to digital signals via an ADC (Analogue-to-Digital Converter), and then the digital signals are validated by Speech recognition. Whether the digital signals are right or wrong depends upon the spoken words within the grammar registered and stored in the VoiceXML document. The grammar in the VoiceXML document is regulated for every Web Site registered and stored in the Module, which means every Web Site has its own defined VoiceXML grammar temporarily stored in the Voice Portal Input Module. If a signal cannot be understood, it means the grammar is not defined or stored in the Module, and an error message will be passed on to the Voice Portal Output Module which will notify the user. If a signal is correct, it will be delivered to the Voice Portal Interpreter & Controller, which will communicate with the website. This process continues until the dialogue finishes.

**Sample Implementation of VoiceXML**

A sample application of VoiceXML can be a practical demonstration of how VoiceXML works and of its advantages for SMEs. The sample scenario in this paper is the application of VoiceXML to a lunch-order for a take-away shop.

Figure 1. How VoiceXML Works

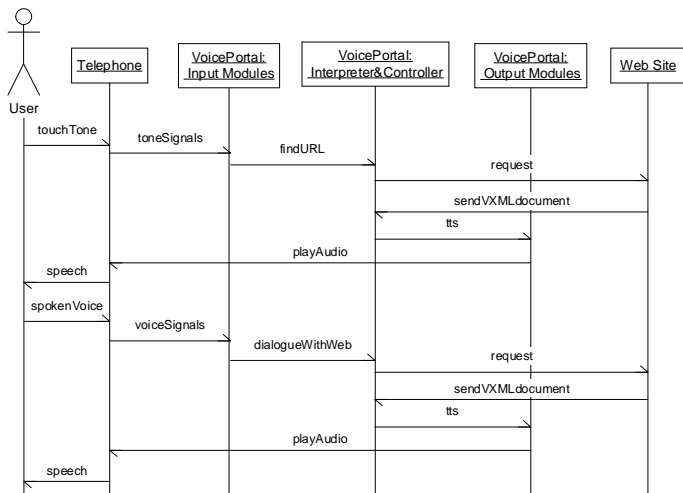


Figure 2. Welcome Sections

```

1 <?xml version="1.0"?>
2 <!DOCTYPE vxml
3 PUBLIC "-//BeVocal Inc/VoiceXML 2.0//EN"
4 "http://cafe.bevocal.com/libraries/dtd/vxml2-0-bevocal.dtd">
5
6 <vxml version="2.0" xmlns="http://www.w3.org/2001/vxml">
7
8 <!-- WELCOME -->
9 <form id="greeting">
10 <block>
11 Welcome to Belgrave Original Charcoal Chicken
12 <goto next="#lunchMenu"/>
13 </block>
14 </form>

```

A small business – Belgrave *Original Charcoal Chicken Shop*, located at Victoria, Australia, is quite busy at lunchtime. They are, however, short of staff who are readily available to take lunch orders by phone at such a busy time, and so they would benefit from an automated system to perform this task. VoiceXML technology could be easily applied as the following scenario illustrates.

When a customer rings, the system will automatically respond to the customer by saying “Welcome to Belgrave Original Charcoal Chicken”. This is illustrated in the Welcome code section shown in figure 2. This VoiceXML application begins with a VoiceXML declaration from line 1 to line 4. the VoiceXML root element `<vxml>` at line 6 is adapted to encompass all other elements including the `<form>` element - a dialogue for presenting information and collecting data. The `<form>` element in turn contains `<block>` and `<goto>` elements, the `<block>` element is a container of executable code, and the `<goto>` element transfers control to another dialogue in the same or different VoiceXML document.

The *Lunch menu code section* follows the *Welcome code section*. The menu code is as shown in figure 3. The `<menu>` element is a dialogue for choosing alternative destinations. The `<prompt>` element is to queue speech synthesis and audio output to the user such as the response by saying “What would you like to order today? ...”. The `<choice>` elements define the available menu items. The `<enumerate>` element is a shorthand for enumerating the choices in a menu. The `<nomatch>` element is to catch a nomatch event. If a users speech cannot be matched to “Lunch pack” nor “Snack Pack” nor “Hawaiian pack”, the system will reprompt the user by saying “You must choose from: Lunch pack, Snack pack, Hawaiian pack”.

The system accepts an order from the menu announced. After a customer makes an order, for example, by saying “Snack pack”, the system responds and executes the snack code section shown in figure 4. The

Figure 3. Lunch Menu Code

```

16 <!-- MAIN MENU -->
17 <menu id="lunchMenu" accept="approximate">
18 <prompt>
19 What would you like to order today?
20 We have lunch pack, lemon chicken and fried rice,
21 Snack pack, chicken and chips,
22 Hawaiian pack, chicken, Pineapple fritter and chips.
23 <enumerate/>.
24 </prompt>
25 <choice next="#lunch">
26 Lunch pack
27 </choice>
28 <choice next="#snack">
29 Snack pack
30 </choice>
31 <choice next="#hawaiian">
32 Hawaiian pack
33 </choice>
34 <nomatch>
35 You must choose from: <enumerate/>
36 </nomatch>
37 </menu>

```

Figure 4. Snack Pack Code

```

87 <!-- Snack pack -->
88 <form id="snack">
89 <field name="lunchOrder">
90 <grammar>
91 <![CDATA[[ yes no ]]]>
92 </grammar>
93 <prompt>
94 chicken baked potatoes and chips,
95 five dollars, Is this OK?
96 </prompt>
97 <filled>
98 <if cond="lunchOrder='yes'">
99 Your order can be picked up from our shop in ten minutes.
100 Your order is recorded under phone number
101 <value expr="session.telephone.ani"/>
102 <goto next="#goodbye"/>
103 </if>
104 <goto nextitem="orderNo"/>
105 </if>
106 </filled>
107 </field>
108 <field name="orderNo">
109 <grammar>
110 <![CDATA[[ yes no ]]]>
111 </grammar>
112 <prompt>
113 OK. what about the chicken, corn and chips.
114 four dollars, Would you like them?
115 </prompt>
116 <filled>
117 <if cond="orderNo='yes'">
118 you can pick up from our shop.
119 <goto next="#goodbye"/>
120 </if>
121 <goto next="#goodbye"/>
122 </if>
123 </filled>
124 </field>
125 </form>

```

`<field>` element declares an input field in a form. The `<grammar>` element at lines 90 to 92 contains a set of utterances that restricts the input of a user. The `<filled>` element, lines 97 to 106 contains a simple conditional logic `<if>` element. If a customer says “no”, the system will give the customer another option. If a customer says “yes”, the system responds by saying “Your order can be picked up from our shop in ten minutes.” The system automatically obtains a customer’s telephone number by using the code at line 101, saying “Your order is recorded under phone number (a customer’s phone number)”, and executing the last section of goodbye code. Finally the dialogue is completed and a new order is output via a SOAP XML packet to the ‘Belgrave Original Charcoal Chicken Shop’ Web service, which initiates a display of the order on the screen of the shop’s computer.

This sample application is built on a free VoiceXML application development environment provided by BeVocal through the Bevocal Cafe (BeVocal, 2004, www site).

## VOICEXML: A SOLUTION TO SMES COMMUNICATION PROBLEMS

### Opportunities for VoiceXML

There are many opportunities for the application of VoiceXML technology to a SME’s communication systems. Sharma (2002), claims that VoiceXML is a one of a variety of innovative technologies available to craft next generation communication networks. Next generation networks refer to the merging of the telephone, mobile, cable telephony and the Internet.

### Benefits of VoiceXML Technology for SMEs

VoiceXML technology makes optimum use of existing phone/mobile services; it only requires basic oral input from the user, and delivers basic audio information back to the user. This is very convenient for customers. Customers do not need a PC, a keyboard, a mouse, a modem or an Internet connection, information from the Web can be available at any time and from any place, particularly when travelling.

With the restricted capital of SMEs, hosting a service with VoiceXML application is cost effective for SMEs due to the fact that the

application of such technology is comparatively cheaper than establishing and managing a Call Centre. For instance, BeVocal provides enterprises with such a service costing \$500/month for 3000 min/month for up to 5 simultaneous calls, \$1000/month for 7500 min/month for up to 10 simultaneous calls, and \$2500/month for 25000 min/month for up to 15 simultaneous calls (BeVocal, 2004, www site). Thus, VoiceXML enabled services of SMEs will be able to operate much more effectively and strengthen their relationship with their customers. In a broader sense, it will be especially applicable to Australian SMEs, given that the statistics clearly indicate that there are already a high percentage of such enterprises.

## CONCLUSION

The history of enterprise communication systems show that they are a dynamic entity, always evolving. This evolution is produced by the continuous need to resolve existing problems within the communications system. The assistance of Information Technology in solving these problems for SMEs is invaluable. If an affordable service, similar in functionality to a Call Centre is available for SMEs, they can provide a competitive service, which at present is only offered by large enterprises. Customers, by using a landline or mobile phone, can contact SMEs anywhere and at any time. Rather than being responded to by a traditional non interactive answering machine, the response is elevated to and automated answering service. Moreover, if a landline or mobile phone can be interlinked to an SME's web site, it means that they can provide their online services for any customer who has a telephone. As the Web service is offered in any place and at any time, the interlinking of Web and telephone systems is a cost effective way of managing this component of an SME's market.

In Australia there is little application of VoiceXML to date, although it has been recognized that there is a need for users to be familiarised with the concept of VoiceXML technology/services before its use will become widespread. VoiceXML developers need to be aware of both the application of this technology as well as how users will relate to VoiceXML in practice. The next stage of research/implementation involves the creation of demonstration VoiceXML based systems to illustrate their role in improving the enterprise communication system.

## REFERENCES

- Ana, O. 2001, 'The power of voice', *InfoWorld*, Vol.23, no. 21, pp. 73-75.
- Anonymous, 2000, 'Business: The power of speech', *The Economist*, Vol. 355, no. 8170, pp. 60-62.
- Australian Bureau of Statistics. (2004). *Australian Bureau of Statistics*, [online], Available: <http://www.abs.gov.au/>, [Accessed: April 5<sup>th</sup> 2004].
- BeVocal. (2004). BeVocal, [online], Available: [http://www.bevocal.com/corporateweb/products\\_sdk.htm](http://www.bevocal.com/corporateweb/products_sdk.htm), [Accessed: September 2<sup>nd</sup> 2004].
- Chugh, J. & Jagannathan, V. 2002, 'Voice-enabling enterprise applications', Eleventh IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 10-12 June 2002, WET ICE, pp.188 – 189.
- Danielsen, P.J. 2000, 'The promise of a voice-enabled Web', *Computer*, IEEE, Vol. 33, no. 8, pp. 104 – 106.
- Dictionary. (2000), *InvestorWords*, [online], Available: <http://dictionary.reference.com/> [Accessed March 3<sup>rd</sup>, 2004].
- Hartman, J.D. & Vila, J.A. 2001, 'VoiceXML builder: a workbench for investigating voiced-based applications', *Frontiers in Education Conference*, 10-13 Oct. 2001, 31st Annual, Volume: 3, pp. S2C - 6-9.
- History- Facts and Myths. (2002). *The Great Idea Finder*, [online], Available: <http://www.cheaplongdistancepro.com/>, [Accessed: April 23<sup>rd</sup> 2004].
- Hocek, A., Cuddihy, D. 2002, *Definitive VoiceXML*, Prentice Hall PTP, USA.
- IVR. (2004). *Our Services and Products*, [online], Available: <http://www.cf-net.com/services/ivr.htm>, [Accessed April 5<sup>th</sup>, 2004].
- Jackson, E. 2003, 'Speaking up for cost savings in the call center: VXML takes on the dinosaur of legacy IVR', *Customer Interaction Solutions*, Vol. 22, no. 2, p. 70.
- Kathleen, W. 2001, 'Speech technology grows up', *Network World*, Vol. 18, no. 34, pp. 27-29.
- Larson, J. A. 2003, 'VoiceXML and the W3C speech interface framework', *Multimedia*, IEEE, Vol. 10, no. 4, pp. 91 – 93.
- Leavitt, N. 2003, 'Two technologies vie for recognition in speech market', *Computer*, Vol. 36, no. 6, pp.13 – 16.
- Miller, M. 2002, *VoiceXML: 10 Projects to Voice Enable Your Web Site*, John Wiley & Sons, Inc., USA.
- Mobile content. (2004). ZD Net Australia, [online], Available: <http://www.zdnet.com.au/news/communications/0,2000061791,20274785,00.htm>, [Accessed April 23<sup>rd</sup> 2004].
- Mohan, T. & McGregor, H. & Saunders, S. & Archee, R. 1997, *Communicating! Theory and Practice*, NSW, Australia, p121.
- Press release. (2004). *IDC analysis the future*, [online], Available: <http://www.idc.com.au/press/detail.asp?releaseid=10>, [Accessed April 23<sup>rd</sup> 2004].
- Quiane Ruiz, J.A. & Manjarrez Sanchez, J.R. 2003, 'Design of a VoiceXML gateway', *Proceedings of the Fourth Mexican International Conference on Computer Science*, Mexico City, Mexico, pp. 49 – 53.
- Satzinger, J. & Jackson, R. & Burd, S. 2002, *Systems Analysis and Design in a Changing World*, Canada, p6.
- Schmelzer, R. et al. 2002, *XML AND Web Services*, Sams Publishing, USA, pp. 919-933.
- Sharma, C. & Kunins, J. 2002, *VoiceXML Strategies and Techniques for Effective Voice Application Development with VoiceXML 2.0*, New York, p7.
- Sharma, C. & Kunins, J. 2002, *VoiceXML Strategies and Techniques for Effective Voice Application Development with VoiceXML 2.0*, New York, p400.
- Srinivasan, S. & Brown, E. 2002, 'Is speech recognition becoming mainstream?', *Computer*, Vol. 35, no. 4, pp. 38 – 41.
- The Staff of DreamTech Inc. 2002, *VoiceXML 2.0 Developer's Guide*, McGraw-Hall Osborne Media, USA.
- Web-enabled call center. (2003). *Kluwer online wireless*, [online], Available: [http://wireless.kluweronline.com/c/journal\\_articles/318479](http://wireless.kluweronline.com/c/journal_articles/318479), [Accessed April 29<sup>th</sup> 2004].

0 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/proceeding-paper/discovering-solution-sme-communication-problems/32660](http://www.igi-global.com/proceeding-paper/discovering-solution-sme-communication-problems/32660)

## Related Content

---

### Challenges in Collecting Qualitative Data for Information Systems Studies

Tiko Iyamuand Irja Naambo Shaanika (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 4480-4489).

[www.irma-international.org/chapter/challenges-in-collecting-qualitative-data-for-information-systems-studies/184156](http://www.irma-international.org/chapter/challenges-in-collecting-qualitative-data-for-information-systems-studies/184156)

### Challenges to the Development of Information Infrastructures: A Synthesis of the Findings

(2012). *Perspectives and Implications for the Development of Information Infrastructures* (pp. 115-135).

[www.irma-international.org/chapter/challenges-development-information-infrastructures/66259](http://www.irma-international.org/chapter/challenges-development-information-infrastructures/66259)

### Analysis of Welding Process Path Planning and Process Simulation Using Digital Twins and Reinforcement Learning

Renping Wangand Huizhen Long (2026). *International Journal of Information Technologies and Systems Approach* (pp. 1-17).

[www.irma-international.org/article/analysis-of-welding-process-path-planning-and-process-simulation-using-digital-twins-and-reinforcement-learning/403591](http://www.irma-international.org/article/analysis-of-welding-process-path-planning-and-process-simulation-using-digital-twins-and-reinforcement-learning/403591)

### EEG Analysis of Imagined Speech

Sadaf Iqbal, Muhammed Shanir P.P., Yusuf Uzzaman Khanand Omar Farooq (2016). *International Journal of Rough Sets and Data Analysis* (pp. 32-44).

[www.irma-international.org/article/eeg-analysis-of-imagined-speech/150463](http://www.irma-international.org/article/eeg-analysis-of-imagined-speech/150463)

### A QoS-Enhanced Model for Inter-Site Backup Operations in Cloud SDN

Ammar AISousand Jorge Marx Gómez (2019). *International Journal of Information Technologies and Systems Approach* (pp. 20-36).

[www.irma-international.org/article/a-qos-enhanced-model-for-inter-site-backup-operations-in-cloud-sdn/218856](http://www.irma-international.org/article/a-qos-enhanced-model-for-inter-site-backup-operations-in-cloud-sdn/218856)