A Collective Intelligence Based Approach to Business-to-Business E-Marketplaces

Youssef Iguider and Hiroyoshi Morita

**The University of Electro-Communications. Graduate School of Information Systems.
Tokyo 182-8585, Japan**

Abstract

This paper describes a novel approach for addressing the issue of business-matching and recommending in business-to-business e-marketplaces. The presented matching system makes use of Collective Intelligence (CI) means for identifying and recommending business opportunities that are best correlated with the user's need. The objective from this approach is to shift the focus from search-oriented matching, toward assistance-providing business matching systems for the next-generation e-marketplaces. At first, the content of business proposals which are submitted to the e-marketplace, along with their companies' profiles are parsed and indexed. The indexed data may be collected from electronically submitted entries, as well as from scanned and handwritten documents. The search engine starts by expanding the queried keywords, to enable an intuitive-like search. The look-up results are then filtered based on compatibility scoring mechanisms, based on CI techniques. The personalized business-matching results and recommendations are later served to the user via a novel visual interactive graphical interface. An experimental system-prototype applying the proposed and described approach is developed and now being experimentally tested, to fully demonstrate the capabilities of the proposed system on real-world data.

Keywords: *Collective Intelligence, e-marketplace, small business, handwriting recognition.*

1. Introduction

1.1 Collective Intelligence (CI)

This research effort is inspired by the vision of Dr. Douglass Engelbart [1] about the use of technology to improve our collective intelligence for the betterment of humanity, by integrating social-cultural strategies with new technology to create a new way to portray information [1]. Dr. Engelbart is often considered the main founder of the field of Collective Intelligence (CI) [2][3]. CI is defined as the capacity of human collectives to engage in intellectual cooperation, in order to build new conclusions from independent contributors [4]. This study is an attempt to apply CI approach to e-marketplaces for small businesses.

1.2 Why Small Business?

The sector of small business has a big impact on nations' economies. Because it usually accounts for over half of all industrial activities; and it is the major source of employment in most countries. In the United States, for instance, small businesses provide 55% of all jobs, and contribute with 54% of all USA sales [5]. In Japan also, the small business sector is the economy's engine [6]. Small and mid-size enterprises (SME) represent 99.7% of the 4,973 thousands companies registered in Japan. These SMEs employ 70.2% of Japan's workforce, and their contribution share to the shipment in the manufacturing market is estimated at 51.1% [6].

1.3 Challenges Faced by Small Business

Although the sector of small business has such an essential role to play in nations' economies, small businesses are too often severely treated by the market's difficult realities. It is a fact that only about 50% of small businesses remain in business after their first 3 years [7]. Small businesses are exposed to bigger threats than larger companies, because they do not have the back-up of extra finance and resources that larger companies possess. Difficulties to commercialize their products; mismatched trading, and the lack of funding partners are often listed among the top challenges that are commonly faced by small businesses, which often lead to their failure [5][7][8].

On the other hand, On a global level, because of the global digital divide and the troubling gaps created by unequal socio-economic levels, small business, especially in rural areas of developing countries, often have less or no access to regional and international trading opportunities, which are offered by Internet and its means of information flow, such as e-marketplaces, electronic commerce, and online social networks, etc. [9][10]. It is feared that the rapid developments in Information and Communication Technology (ICT) which opens up further new global business opportunities in the form of e-commerce, may widen this digital divide and lead small businesses in the

rural areas of developing countries to lag even further behind and lose in the race [11][12].

1.4 Existing e-Marketplaces

Business consulting and business matching for small and large businesses is mostly provided offline. A service that usually comes with more costly expenses than what small business can usually afford. Moreover, the recommended business matches are often geographically limited to local or regional partners. Few online e-marketplaces provide international business matching services. Most of these platforms simply list the same static information from their databases similarly to all users. Without truly taking into account the specific needs and background of the user [13]. In the present paper, we attempt to introduce and discuss a new business-matching and recommending system, which would enable e-marketplaces to provide each user with personalized results that are customized to the specific needs of her/his business. The recommended business matches are served via a novel visual interactive graphical interface. The following section discusses the evolution of information systems, and anticipates the coming trend of information systems, and therefore the motivation behind this research. The section III introduces and describes the proposed system. The section 4 provides an overview about the initial experimental studies. The section V outlines our conclusions and discusses further strategies for addressing the challenges faced by small businesses on a global level.

2. Targeted Information System

Information systems (IS) are combinations of information technology and people's activities using that technology to support operations, management, and decision-making.

2.1 Hardware-Oriented Information Systems

Early IS solutions were operating based on primitive systems that used machine codes and data, to have the central processing unit of computing devises execute specific instructions. These primitive systems often executed one program at a time, and operated mainly as specific hardware-dependent systems. We represent this era of information systems as the “Hardware Management” stage, in Figure 1.

2.2 Software-Oriented Information Systems

Later, in the 60s through the 80s, after hardware capabilities evolved to allow similar software to run on more than one platform, advanced operating systems were born. Which enabled multi-tasking information systems to operate a large amount of software applications. Which help the users perform common tasks and activities in the real world. We represent this era of information systems as the “Software Management” stage, in Figure 1.

2.3 Knowledge-Oriented Information Systems

The expansion of the World Wide Web and the explosion of Internet interactions, led to the constantly increasing production of a huge amount of online data, which is doubling approximately every six months [14]. Therefore, we believe that there is an emerging need for new data management systems, able to take advantage of these large amounts of data, by uncovering new, implicit and potentially useful knowledge from them. And also creating new knowledge out of their interlinked characteristics. This knowledge operating information systems would play a vital role in the information industry. We represent this era of information systems as the “Knowledge Management” stage in Figure 1.

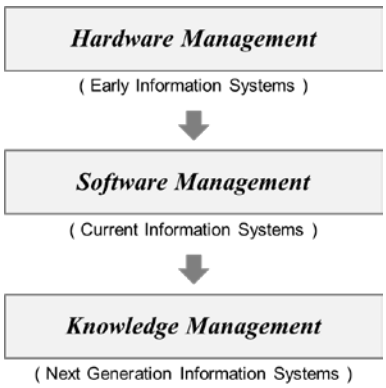


Fig. 1 Evolution of information systems

2.3 CI Based Knowledge-Oriented IS

This work attempts to propose a knowledge managing IS model, based on collective intelligence, which we later apply to develop a new business-matching system for next generation e-marketplaces.
“CI is a multidisciplinary philosophical framework, which integrates social-cultural strategies with new technology to create a way to portray information, with the goal to include, view, and aggregate as much information as possible in order to enable people to act strategically to solve complex problems” [1][15][16].
The basic idea behind our proposed knowledge-managing system is illustrated below in Figure 2. And the process is described below:

Collective-Intelligence Collection: Acquiring data and information from independent users, then provisioning that

data in a way which ensures a later optimal processing.

Intelligence Processing: Converting the collected data into a form suitable for producing intelligence. By conducting various detailed analysis, comparisons and information correlation among the collected data.

Personalized Services: Reducing information overload, by focusing on the consumer's specific need, to interpret the processed information into a finished intelligence product that may help the user draw analytical conclusions.

New Knowledge Creation: Aggregating the collected data and processed information, to systematically and dynamically create new knowledge that may convert lacking information into expanded intelligence.

Customized Expertise Servicing: Conveying expanded intelligence in a usable form, to support user's decision-making with personalized and relevant insights.

Smart & Intuitive User Interface: Creating new ways of structuring facts, and new ways of interacting with the system, is key to extending people's capability to create, manipulate and share knowledge [1].

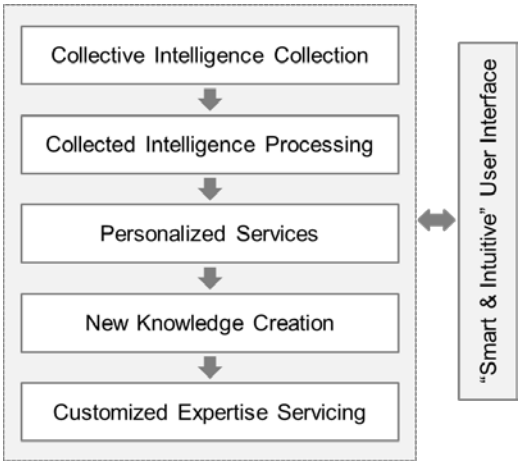


Fig. 2 CI based knowledge management

Actually, we recently discovered that our proposed approach to knowledge-managing systems has an analogical similarity with the five-step approach, called Intelligence Cycle (Figure 3), which is used by the Central Intelligence Agency (CIA) in the United States and many other Intelligence communities. Their approach is apparently well proven for producing and reporting highly accurate Intelligence services [17].

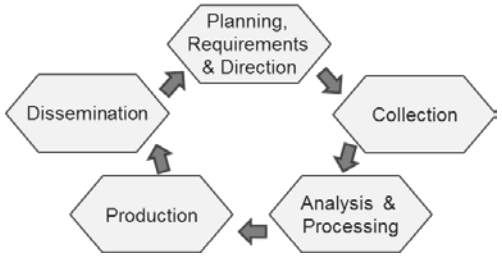


Fig. 3 CIA's Intelligence Cycle.

3. A CI Based System For E-Marketplaces

This section attempts to apply the approach discussed in the previous section, toward designing a CI based new business-matching and recommending system model for e-marketplaces.

Unlike most existing e-marketplaces, where the system usually simply matches companies and lists to the user straightforward search results from available static databases [13], the proposed system model (Figure 4) uses a CI approach toward the process of business-matching and business opportunity recommending. As shown in Figure 4, at first, various data are collected from the collectively submitted users' business (selling or buying) proposals, as well as from their company's profile and background. With the goal to allow traders to have access to the untapped business opportunities which are not available electronically, but on paper. And also with the goal to help bridging the digital divide with traders in many developing countries, especially in rural areas. The system may accept also the information that is captured from handwritten document and transformed to digital data. Moreover the recent advances in on-line data capturing technologies and its widespread deployment in devices like PDAs and notebook PCs is creating large amounts of handwritten data that need to be archived and retrieved efficiently, especially that recognition algorithms and engines are already available for all major language scripts [18][19]. The collected data are then processed and carefully indexed. Then, based on the user's query, and also based on his/her recorded business background and company's profile, the system conducts various analysis and correlation operations, on the user's data vs. the data of the potential candidate partners, and their business proposals and needs. With the goal to reduce the search result overload, and instead convey to the user personalized business recommendations specific to his/her needs, interest and background. The goal is also to enable the user uncover new relevant business opportunities that might not be easily reachable though straightforward search of a static database.

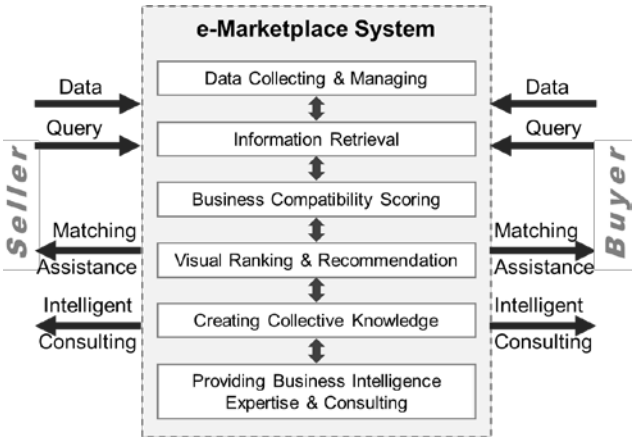


Fig. 4 CI based system for e-marketplaces

Table 1 illustrates an example, where a part of the user's information is scored and matched against the data of other identified business-partner candidates. With the goal to evaluate the user's compatibility with the identified business-partner candidates, and qualify the potential of their proposed business opportunities, the user's attributes are mapped against the attributes of the identified partner-candidates, via several correlation means, including the matching via Euclidean Distance Scoring (1).

Table 1: Compatibility scoring

Attribute #	Attributes Description	User	Candidate Partners				
			P-1	P-2	P-3	…	P-n
A_1	Company Country	$S_0(A_1)$	$S_1(A_1)$	$S_2(A_1)$	$S_3(A_1)$	…	$S_n(A_1)$
A_2	Business Category	$S_0(A_2)$	$S_1(A_2)$	$S_2(A_2)$	$S_3(A_2)$	…	$S_n(A_2)$
A_3	Company Age	$S_0(A_3)$	$S_1(A_3)$	$S_2(A_3)$	$S_3(A_3)$	…	$S_n(A_3)$
A_4	Capital	$S_0(A_4)$	$S_1(A_4)$	$S_2(A_4)$	$S_3(A_4)$	…	$S_n(A_4)$
A_5	Business Volume	$S_0(A_5)$	$S_1(A_5)$	$S_2(A_5)$	$S_3(A_5)$	…	$S_n(A_5)$
A_6	Employees	$S_0(A_6)$	$S_1(A_6)$	$S_2(A_6)$	$S_3(A_6)$	…	$S_n(A_6)$
A_7	Offer type (Proposal)	$S_0(A_7)$	$S_1(A_7)$	$S_2(A_7)$	$S_3(A_7)$	…	$S_n(A_7)$
A_8	Product/Service category	$S_0(A_8)$	$S_1(A_8)$	$S_2(A_8)$	$S_3(A_8)$	…	$S_n(A_8)$
A_9	Minimum order ($)	$S_0(A_9)$	$S_1(A_9)$	$S_2(A_9)$	$S_3(A_9)$	…	$S_n(A_9)$
A_{10}	Targeted region/s	$S_0(A_{10})$	$S_1(A_{10})$	$S_2(A_{10})$	$S_3(A_{10})$	…	$S_n(A_{10})$
A_{11}	Non acceptable countries	$S_0(A_{11})$	$S_1(A_{11})$	$S_2(A_{11})$	$S_0(A_{11})$	…	$S_n(A_{11})$

$$C_{(i,j)} = \sqrt{\sum_{k=1}^{11} \left(S_i(A_k) - S_j(A_k)\right)^2} \qquad (1)$$

Where

$C_{(i,j)}$ - Compatibility between two partner-candidates i and j

$S_i(A_k)$ - Score of a partner-candidate i with regard to the attribute A_k

To comply with the recommendation of Dr. Engelbart (CI's founder) about the importance of creating new ways and symbols for structuring facts, to extend the user's capability of manipulating the created knowledge [1], the matching results and personalized recommendations are later conveyed to the user via a new visual graphic interface, as illustrated by Figure 5 and Table 2.

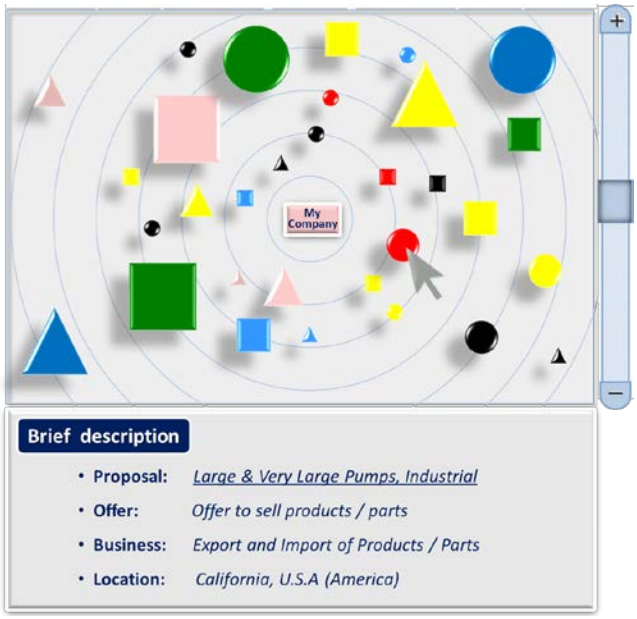


Fig. 5 Visual representations of the business-matching results

Colors, shapes and sizes are used to symbolize characteristics of the candidate partners, to help the user visually explore and interact with the recommended business opportunities via the graphical navigational interface.

Table 2: Graphic symbols interpretation

Shape		"Offer Type" Attribute
Circle	●	SELLING
Triangle	▲	BUYING
Square	■	OTHERS
Color		**"Country Area" Attribute**
Blue	●	EUROPE
Yellow	●	ASIA
Black	●	AFRICA
Green	●	OCEANIA
Red	●	AMERICA
Pink	●	JAPAN
Size		**"Company Size" Attribute**
Small	●	SMALL
Medium	●	MEDIUM
Large	●	LARGE

Automated Business Consulting System: At a later stage, our research aims at creating a CI business-matching system that would enable e-marketplaces to provide automated consulting. The system aggregates the collected data to create new relevant knowledge. By systematically applying best practices used by business experts. A personalized expertize would be generated to support users' decision-making and enhance their market insights and business intelligence. Figure 6 illustrates an example where the targeted automated system would systematically uncover strategic partnering opportunities for the user.

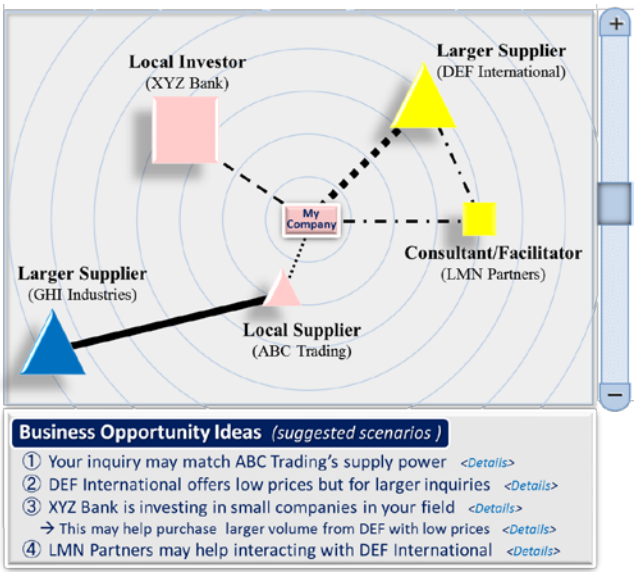


Fig. 6 CI Based Automated Consulting

4. Experimental Studies

To experimentally study the proposed CI based business-matching and recommending system, we are currently developing a prototype that can process advanced business matching and generate recommendations, based on real-word business opportunities.

4.1 JETRO's online business matching database

To study our prototype using real-word data, experimental simulations were conducted based on data collected from JETRO (Japan External Trade Organization)'s online business matching database. JETRO is a Japanese governmental organization, which promotes mutual trade between Japan and the rest of the world. JETRO is running a free online business matching service "Trade Tie-up Promotion Program" <www.jetro.go.jp/ttppoas/> [20]. Which allows business people (especially small and medium companies) to browse through over 20,000 business proposals in various fields.

4.2 Experimental prototype

The components of the experimental prototype are shown in Figure 7. At first an experimental simulation was developed using Python programing language [21], for parsing JETRO's proposals. BeautifulSoup classes were used to screen-scrap and parse the content of JETRO's entries. The result of this parsing led to compiling information from each entry (proposal) in JETRO database.

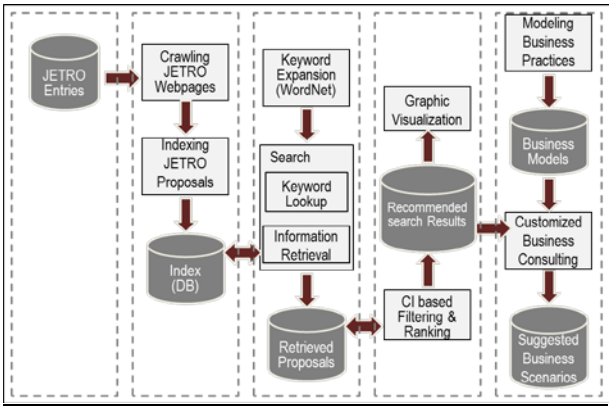


Fig. 7 Experimental business-matching and recommending prototype

4.3 Indexing TTPP business proposals

To process TTPP business proposals in depth, a set of entries (Figure 8) was indexed into a SQLite database. To proceed, the contents were parsed by randomly looping through TTPP URLs to crawl their entry contents. The experimental crawler was designed to ignore a set of words which carry no important meaning. The system was also designed to identify and ignore URL of the proposals that were either expired or deleted by the users. As results, the crawled and parsed TTPP business proposals were automatically indexed into a SQLite database, which was automatically created and saved for further use.

#	Proposal Description	Format
1	Proposal title	Text
2	Proposal URL	URL
3	Proposal No	Number
4	Business Type	Text
5	Offer Type	Text
6	Proposal Category	Text
7	Country / Area	Text

#	Node Attributes	Format
8	Ranking Distance	Number
9	"Offer Type" attribute	Text
10	"Company Size" attribute	Text
11	"Country Area" attribute	Text

Fig. 8 An experimental set of business proposal entries

Figure 9 show samples of indexed data, where the database was populated with, (Figure 9.a) a list of the automatically generated and processed URLs, (Figure 9.b) parsed and processed words, and (Figure 9.c) information about how each word is associated to its corresponding proposal, along with the position of that word within the proposal page.

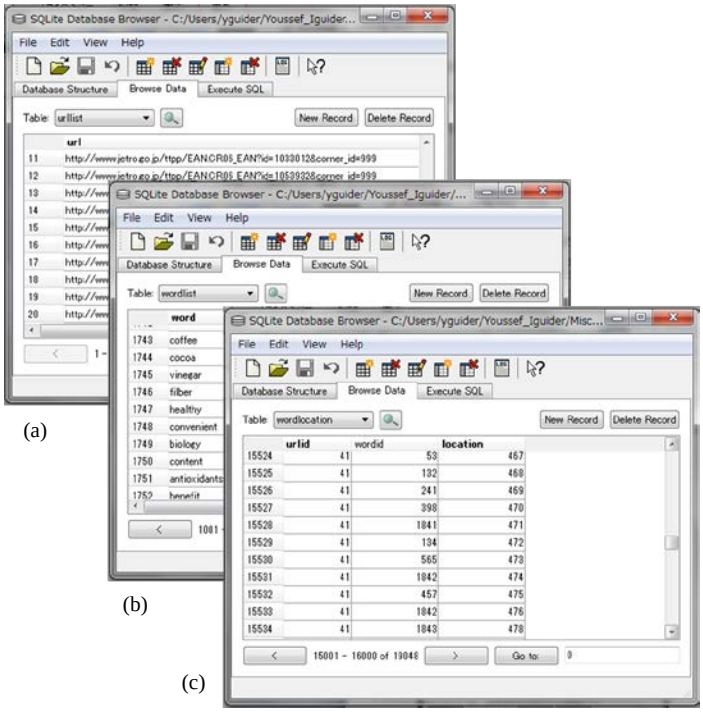


Fig. 9 Sample data indexed into SQLite database: (a) Indexed URLs; (b) Indexed words; (c) Linking words to proposals.

4.4 Keyword expansion based on WordNet

To enable an intuitive-like search, the system uses WordNet to allow identifying additional words that may be relevant to the user's query. This may help uncovering hidden business opportunities that could be of interest to the user. Figures 10,11 illustrate an example where the system is trying to expand a submitted keyword - "carpet".

The expansion of the word "carpet" (Figures 10,11) led to the following:

Similar words: carpet, rug, carpeting
Related Words: Brussels_carpet, Kurdistan, Wilton, Wilton_carpet, broadloom, drugget, flying_carpet, hearthrug, nammad, numdah, numdah_rug, prayer_mat, prayer_rug, red_carpet, runner, scatter_rug, shag_rug, stair-carpet, throw_rug
Related Parts: Edging

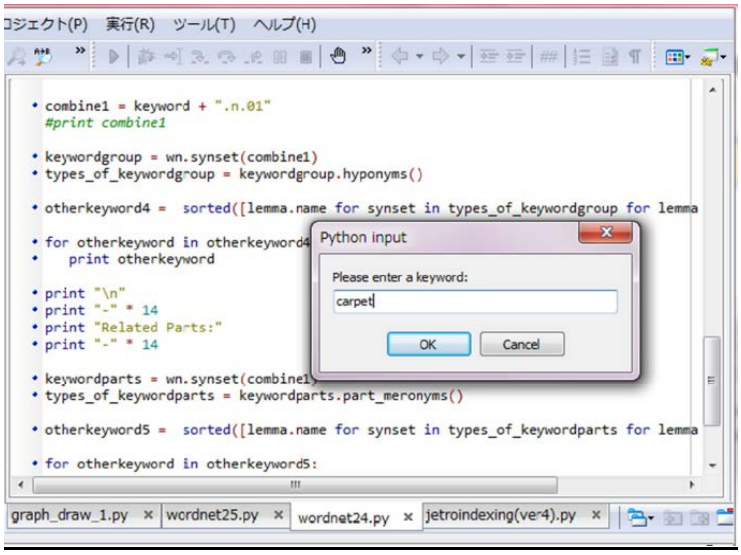


Fig. 10 Submitting a query keyword

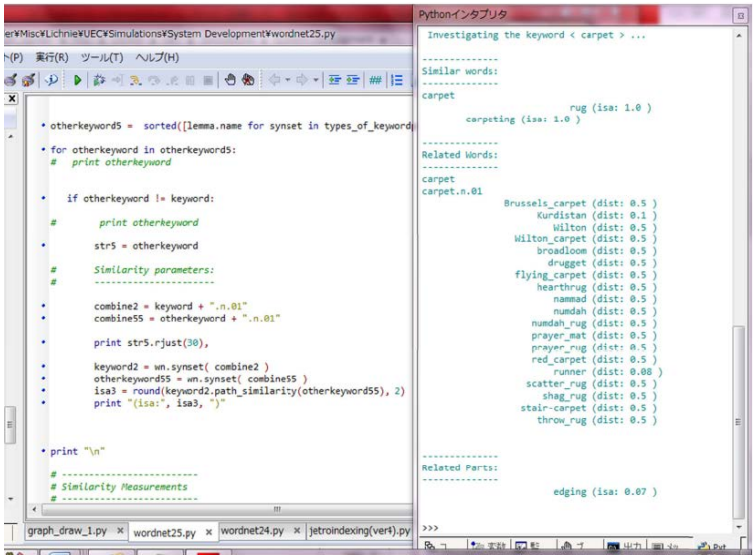
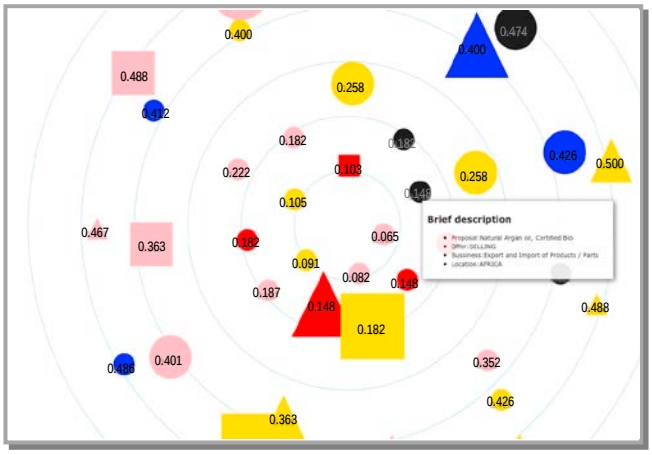
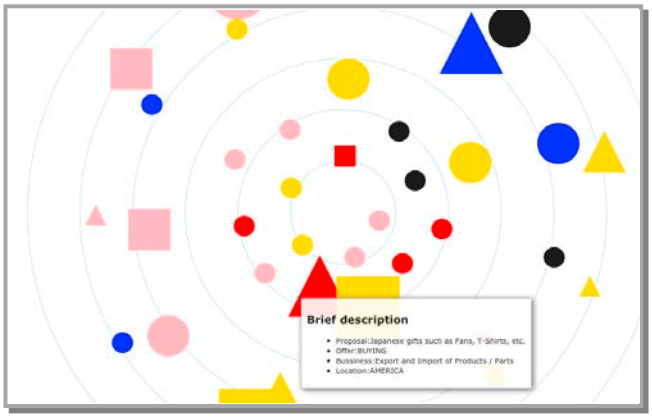


Fig. 11 Expanding the keyword "carpet"

Further, the prototype is simulated to look up the submitted keyword along with the expanded keywords in the indexed SQLite database. The data and content of the identified business proposals are then mapped to the attribute related to the user, by using the CI approach discussed above.
The initial experimental results look quiet promising. Figure 12 presents an experimental visual representation of the business-matching results obtained from processing the experimental set of business proposal entries mentioned above. The results are automatically visualized via an experimental interactive graphical user interface, which was developed using JavaScript InfoVis Toolkit.



(a)



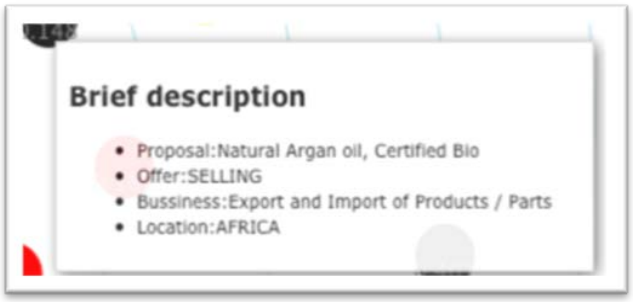
(b)

Fig. 12 Experimental visual representation of business-matching results

The graph is automatically generated based on attributes related to the recommended business proposals. Each proposal is displayed in the form of an interactive node which has a specific shape, color and size. The shape, color and size of nodes are automatically assigned according to the description in Table 2. The nodes are automatically placed away from the center, based on their Distance Scoring according to (1). This enables an intuitive visual navigation of the search results. The nodes are designed to have interactive capabilities with the goal to allow visually exploring and interacting with the recommended business opportunities. For instance, by mouseovering a node, a brief description of its business proposal is displayed, as shown in Figure 13.

Clicking a node, leads to automatically placing in the center of the graph. Meanwhile information about the clicked node is transmitted (as feedback data) back to the system for further processing and compatibility scoring. Several other interactive features are currently being developed to be added to the capabilities of this interactive graphical user interface.

With the goal to have the interactive graphical user interface allow accessing the details of business proposals at any step, our experimental prototype is designed to provide a direct access back to the original business proposals as published on JETRO's TTPP website.



(a) Mouseovering a "selling" node.

(b) Mouseovering a "buying" node.

Fig. 13 Mouseovering an interactive node.

We currently are finalizing the development of the prototype and its visualization system, with the goal to allow us convey the business-matching results and recommendations, via a visual graphic interactive interface, similar to what was described in Figure 5 and Table 2.

5. Conclusion and Discussion

We expect the presented and discussed business-matching and recommending system for e-marketplaces to take the business matching process to a new level. The system makes use of a Collective Intelligence (CI) approach to identify and recommend the best matching and correlated business opportunities for the user. The results are served via a novel visual interactive graphical interface. A system prototype applying the proposed and described approach is being developed and experimentally tested, to fully demonstrate the capabilities of the proposed system on real-world data. Although the prototype is at an early stage, the initial experiments show promising results. We believe therefore that the proposed and simulated approach can addresses many of the marketing challenges, discussed above, that are typically faced by most small businesses.

Handwritten business documents for e-marketplaces: The recent advances in on-line data capturing technologies and its widespread deployment in devices like PDAs and notebook PCs, is creating large amounts of handwritten data that need to be archived and retrieved efficiently. It is important that next generation marketplaces could make use of OCR technologies and handwriting recognition solutions for converting handwritten business documents and other scanned documents (catalogues etc.) into indexable and retrievable data. The increase of such relevant information would enable better correlation, and therefore improve the matching and recommending of business partnerships. Moreover, since handwritten signature remains the most widely accepted biometric means for identity verification in business transaction agreements [22], it makes sense to apply the proven technologies and algorithms which enable author verification and identification via handwriting analysis. To this purpose we are exploring the feasibility to apply our previous works on handwriting recognition [19], and analysis of human handwriting [23][24], with the goal to enable our proposed business-matching and recommending system to make use of the key business information whether it is available electronically or from handwritten documents.

The treats of Digital Divide on small business: Beside the marketing issues discussed earlier, the global digital divide is severely treating many small businesses, especially in the rural areas of developing countries. The Okinawa Charter on Global Information Society (GIS) by the G8, and also the Japan Ministry of International Trade and Industry (MITI) in their Proposal warned that the rapid developments in Information and Communication Technology (ICT) have opened up new global business opportunities in the form of e-commerce, however it is feared that these developments may widen the digital divide, and underdeveloped countries may lag further behind [11][12]. Therefore, to help solving this complicated issue, from our angle, by bridging the digital divide in the e-Marketplaces, we consider allowing the business information available on handwritten documents to be converted into electronic data. Especially that OCR technologies as well as handwriting recognition algorithms and engines are now available for the major language scripts [18][19]. This would not only assist the traders in less developed areas by narrowing down the treat of the digital divide on their small business. But it would also assist the users of e-Marketplaces in industrialized countries to open up and access new markets full of still untapped business opportunities.

References

[1] D. Engelbart , “Boosting Collective IQ,” The Doug Engelbart Institute (www.dougengelbart.org/about/collective-iq.html)

[2] P. Lévy, “From Social Computing to Reflexive Collective intelligence,” IEML Research Program, CRC, FRSC, University of Ottawa. Aug.09.

[3] Stanford University’s archive “Doug Engelbart 1968 Demo,” (http://sloan.stanford.edu/mousesite/1968Demo.html)

[4] T. Segaran, “Programming Collective Intelligence,” O’REILLY ISBN:978-0-596-52932-1, 2007.

[5] The United States Department of The Treasury. Various reports about the Economic Recovery (www.treas.gov)

[6] Japan’s Ministry of Economy, Trade and Industry, “Census of Manufactures,” 2008. (www.meti.go.jp/english/statistics/)

[7] E.J. O’Donoghue, N. Key, and M.J. Roberts, “Does risk matter for farm businesses? The effect of crop insurance on production and diversification,” Report of Economic Research Service, USDA.

[8] B.S. Cromie, “Entrepreneurial Networks,” International Small Business Journal 9, 57-74, 1991.

[9] T.P. Rama Rao, “E-Commerce and Digital Divide: Impact on Consumers, ” Presented at the Regional Meeting for the Asia-Pacific: New Dimensions of Consumer Protection in the Era of Globalisation, Goa, India, September 10-11, 2001.

[10] “Digital Opportunities and the Missing Link for Developing Countries,” Oxford University Press 2001: Oxford Review of Economic Policy, Volume17, Issue2, pp. 282-295, June 1, 2001 .

[11] “Okinawa Charter on Global Information Society”, http://www.g8.utoronto.ca/summit/2000okinawa/gis.htm

[12] MITI’s Proposal for WTO E-Commerce Initiative, “Towards eQuality: Global E-Commerce Presents Digital Opportunity to Close the Divide Between Developed and Developing Countries (2nd Draft)”, http://wwwmeti.go.jp

[13] Y. Iguider and H. Morita, “A Collective Intelligence Based Business-Matching and Recommending System for Next Generation E-Marketplaces,” Proceedings of 2011 IEEE Symp. on Computers & Informatics, pp.489-494, Mar. 2011.

[14] Federal Communications Commission (FCC 98-225), “MEMORANDUM OPINION AND ORDER,” Sep. 14, 1998.

[15] MIT Center for Collective Intelligence (http://cci.mit.edu/)

[16] Wikipedia - The Free Encyclopedia, on “Collective Intelligence” (http://en.wikipedia.org/wiki/Collective_Intelligence)

[17] “INTELLIGENCE COLLECTION ACTIVITIES AND DISCIPLINES,” Operations Security INTELLIGENCE THREAT HANDBOOK, OPSEC, April 1996

[18] S. Jaeger, C.-L. Liu and M. Nakagawa, “The state of the art in Japanese online handwriting recognition compared to techniques in western handwriting recognition,” Internat-l Journal on Document Analysis and Recognition - IJDAR , vol.6, no.2, pp.75-88, 2003.

[19] Y. Iguider and M. Yasuhara, “An Active Recognition of Handwritten Isolated Arabic Characters,” Transaction of the Society of Instrument and Control Engineers. Vol.32, No.8, pp1267/1276, 1996.

[20] Trade Tie-up Promotion Program. The Japan External Trade Organization (JETRO) (http://www.jetro.go.jp/ttppe/)

[21] M. Lutz. “Programming Python,” 3rd Edit. O'Reilly Media, Aug.06.

[22] M.M. Fahmy, “Online handwritten signature verification system based on DWT features extraction and neural network classification,” Ain Shams Engineering Journal, Vol.1, Issue 1, Sep. 2010, pp.59-70.

[23] Y. Iguider and M. Yasuhara, “Extracting Control Pulses of Handwriting movement,” Transaction of the Society of Instrument and Control Engineers. Vol.31, No.8, pp1175/1184, 1995.

[24] Y. Iguider and M. Yasuhara, “A Control-Theoretic Description of Handwriting Process,” Preprint of the 24th Symposium of the Institute of Systems Control and Information Engineers, pp45-48, Nov. 1992.