Journal of Computer Science 9 (10): 1396-1405, 2013
ISSN: 1549-3636

doi:10.3844/jcssp.2013.1396.1405 Published Online 9 (10) 2013 (http://www.thescipub.com/jcs.toc)

BUSINESS INTELLIGENCE SERVICE BASED ON ADAPTIVE USER MODELING AND GROUPING

Jinhyung Kim, Myunggwon Hwang, Do-Heon Jeong, Sa-Kwang Song and Hanmin Jung

Department of Computer Intelligence Research,
Korea Institute of Science and Technology Information, Republic of Korea

Received 2013-04-19, Revised 2013-07-31; Accepted 2013-09-17

ABSTRACT

In this study, we suggest the mobile business intelligence service based on adaptive recognition of user intention and usage patterns. This service is named as InSciTe adaptive and based on text mining and semantic web technologies. This service supports not only technology-focusing analysis and prediction but also adaptive recognition about user's intention by semi-automatic user modeling process. By the adaptive user modeling, this service can provide more suitable service flow and more proper analysis results based on user's intention.

Keywords: Business Intelligence, Adaptive Recognition, User Intention, User Modeling, Analysis and Prediction

1. INTRODUCTION

Business Intelligence (BI) is the ability of an organization to collect, maintain and organize knowledge. This activity produces large amounts of information that can help in developing new opportunities. Identifying these opportunities and implementing an effective strategy can provide a competitive market advantage and long-term stability.

BI includes diverse technologies such as online analytical processing, analytics, data mining, process mining, complex event processing, business performance management, benchmarking, text mining, predictive analytics and prescriptive analytics. The goal of modern BI deployments is to support better decision-making in businesses; thus, a BI system can be called a Decision Support System (DSS). Although the term "business intelligence" is synonymous to "competitive intelligence," BI uses technologies, processes and applications to analyze mostly internal, structured data and business processes, whereas competitive intelligence gathers, analyzes and disseminates topical information with a focus on competitors (Ranjan, 2009; Azma and Mostafapour, 2012; Li *et al*., 2012; Cheung and Li, 2012). Therefore, the final objective of BI is precise analysis of massive amounts of related information and predictions for the effective establishment of the strategy and blueprint of companies. However, as the amount of information increases exponentially every year, data analysis and predictions based on that information become more difficult.

Until now, several studies regarding BI have focused on technology analysis and predictions, such as Foresight and Understanding from Scientific Exposition (FUSE) DARPA, 2009, Combining and Uniting Business Intelligence with Semantic Technology (CUBIST) (Klai *et al*., 2012), Text Mining Software for Technology Management (Point, 2009). These projects aim to support decision making through analysis and pattern recognition of scientific information. However, existing services provide uniform analysis results without considering usage patterns or intentions; therefore, users cannot acquire user-adaptive analysis results or customized services (Kim *et al*., 2012a; 2012b; 2013; Lee *et al*., 2013).

Corresponding Author: Do-Heon Jeong, Department of Computer Intelligence Research, Korea Institute of Science and Technology Information, Republic of Korea

In this study, we recommend the InSciTe system to provide information analysis and predictions about the field of science and technology from 2010. The InSciTe system supports four key services focusing on technology analysis and prediction: Technology trends analysis, element technology analysis and convergence technology discovery. In addition, the system includes a five-step user-modeling process for adaptive recognition of user patterns and intentions.

This study is organized as follows: In section 2, we illustrate the user modeling for adaptive recognition of user intention and in section 3, we suggests an user grouping process based on user modeling results. Section 4 presents the key technology analysis and prediction services in the InSciTe system and we conclude the research in the section 5 with plans for future work.

1.1. User Modeling

InSciTe Adaptive includes a user modeling and grouping process to recognize user patterns and to customize the system to the user's specific needs and intention. InSciTe Adaptive supports a stereotype-based user modeling method, which is a more adaptive modeling method compared to a static user modeling method.

Stereotype-based user models use demographic statistics. Based on the gathered information, users are classified into common stereotypes and the system adapts to these stereotypes. Therefore, the application can make assumptions about a user, even though data about that specific area is not available, because the prior demographic studies have already established that users who match the same stereotype are likely to have the same characteristics. However, stereotype-based user models rely mainly on statistics and do not take into account that personal attributes might not match the stereotype.

In InSciTe Adaptive, the user modeling process consists of five phases. In the first and second steps, the system presents simple questions to the user and collects basic information about user intention. Based on the user's selection in the first and second step, the system automatically provides suggestions regarding detailed service functions in the third step and the user can refine the suggested results that were previously refined by the system to confirm the user's needs. In the fourth step, the system makes a decision about which service the user wants. In the fifth step, the system decides the group in which the user is included.

The first step is a fully manual step and the system presents the same question to all of users. The second and third steps are semi-automatic steps and the system first refines the question, based on the user's selection in the previous step. The fourth and fifth steps are fully automatic and the system makes a decision regarding service and flow.

1st Phase (Key Category Selection Step-Manu al Process)

In the first step, a user selects a key category such as "Technology" or "Organization." InSciTe adaptive includes sub-services regarding technology-to-technology analysis (such as technology trends analysis and convergence technology discovery), technology-to-organization analysis (such as agent levels analysis) and organization-t o-organization analysis (such as agent partner analysis). For example, if a user selects "Technolog y," InSciTe Adaptive can recognize that the user wants to obtain technology-to-technology analysi s and technology-to-organization analysis with a focus on a specific technology. However, if a user selects "Organization," the system understands that the user's intention is to obtain organization-toorganization analysis and organization-to-technology analysis with a focus on a specific organization. At the end of the first step, the system can roughly understand the user's preference and confirm the target elements for analysis.

2nd Phase (Constitution Element Selection Step-Semiautomatic Process)

The second step aims to understand the user's intention and needs in detail using several technology elements (such as "a ssociated technology" and "convergence technology") and organization elements (such as "collabor ating organization" and "competing organization"). Based on the user's category selection in the first step, the system can refine the question in the second step in order to determine the details of the user's preference with regard to the element.

3rd Phase (Constitution Function Decision Step-Semiautomatic Process)

The third step involves various types of functions. Each function is related to an element in the second step and to a service in the fourth step. Based on results from the first to the third steps, the system can more accurateely determine the user's preference and intention. Further, from the analysis result, the system can decide which service the user wants and which service flow is suitable for him/her.

4th Phase (Service Decision Step-Automatic Process)

From the first to the third steps, the syste m-suggested elements and functions are based on the user's manual selections. However, the fourth step is executed by the system automatically, based on the analysis of user responses obtained in the first three steps. In the fourth step, the system decides which service the user wants to use and need.

5th Phase (User Group Decision Step-Automatic Process)

After deciding on the service that the user wants to use, the system finally decides the service flow that is suitable for the user, based on the user's intention.

1.2. User Grouping

In this system, we defined seven user groups based on the user's intention. To define user groups more precisely, we conducted a survey of researchers and analysts regarding R&D analysis and prediction and collected real-world requests from them. Algorithm for user modeling and grouping is as **Table 1**. Then, we classified and defined the user groups based on the intention of each user, to use in an adaptive R&D analysis and a prediction system. Service scenarios are difference based on user group and those are as **Fig. 1-5**.

Definition 1. Field Trends Analysis Group

The field trends analysis group comprises users who want to analyze R&D trends in general and in abstract. They have no information about any specific technology or organization and they are only interested in general categories, such as information technology, bioinformatics and computer science. Because each category includes numerous core technologies, the system has to first select and suggest a few emerging technologies among various technologies in the desired category. Users can choose one of the suggested technologies for a nalysis and prediction.

Definition 2. Technology Trends Analysis Group

The technology trends analysis group comprises users who want to know the viability of and the trends af fecting a specific technology. They are obviously interested in a specific technology and want to acquire detailed information about that technology and other related technologies. Users can check the viability of their own technology and discover emerging technologies for future use. In addition, they want to analyze past information, as well as future predictions, about the given technology.

Table 1. Algorithm for user modeling and grouping

1. Begin
2. function Sel_Cat(UCi: User Intention for Category)
3. return Constitution_Element
4. Enum UC[ ] = {Technology, Organization}
------to recognize technology-focusing trends-----
5. if UCi equals *Technology*
6. then Show(CE(UC_i))
7. for s=1 to N(CE(UC_i))
8. Sel_CE(UE_j: User Intention for Element)
9. return Constitution_Function
10. Enum UE[ ] = {elementTech, similarTech, conver geTech, competeTech, isadomainTech, substituteTech, succ eedingTech}
11. for j=1 to j_{max}
12. SET SETCF[j] to CF(UEj)
13. for k=1 to k_{max}
14. Sel_CF(UF_k:User Intention for Function)
15. return Service
16. Enum SV[ ]={Technology Navigation, Technology Trends, Element Technology,
17. Convergence Technology}
18. function Final_Decision( )
19. return Final_Service_Flow:
20. SET SET_{SV}[k] to Sel_CF(UF_k)
21. for t = k to 1
22. if (SET_{SV}[k]==SET_{SV}[t])
23. delete SET_{SV}[k]
24. SET k to k-1
25. for m =1 to m_{max}
26. if (SET_{SV}[k] is subset of SF_m)
27. then Final_Service_Flow = SF_m
------to recognize organization-focusing trends-----
28. if UC_i equals *Organization*
29. then Show(CE(UC_i))
30. for s=1 to N(CE(UC_i))
31. Sel_CE(UE_j: User Intention for Element)
32. return Constitution_Function
33. Enum UE[ ] = {competeOrg, collaborateOrg, simil arOrg, supplyOrg}
34. for j = 1 to j_{max}
35. SET SET_{CF}[j] to CF(UE_j)
36. for k = 1 to k_{max}
37. Sel_CF(UF_k:User Intention for Function)
38. return Service
39. Enum SV[ ]={Organization Levels, Organizati on Partners}
40. Final_Decision( )
41. END

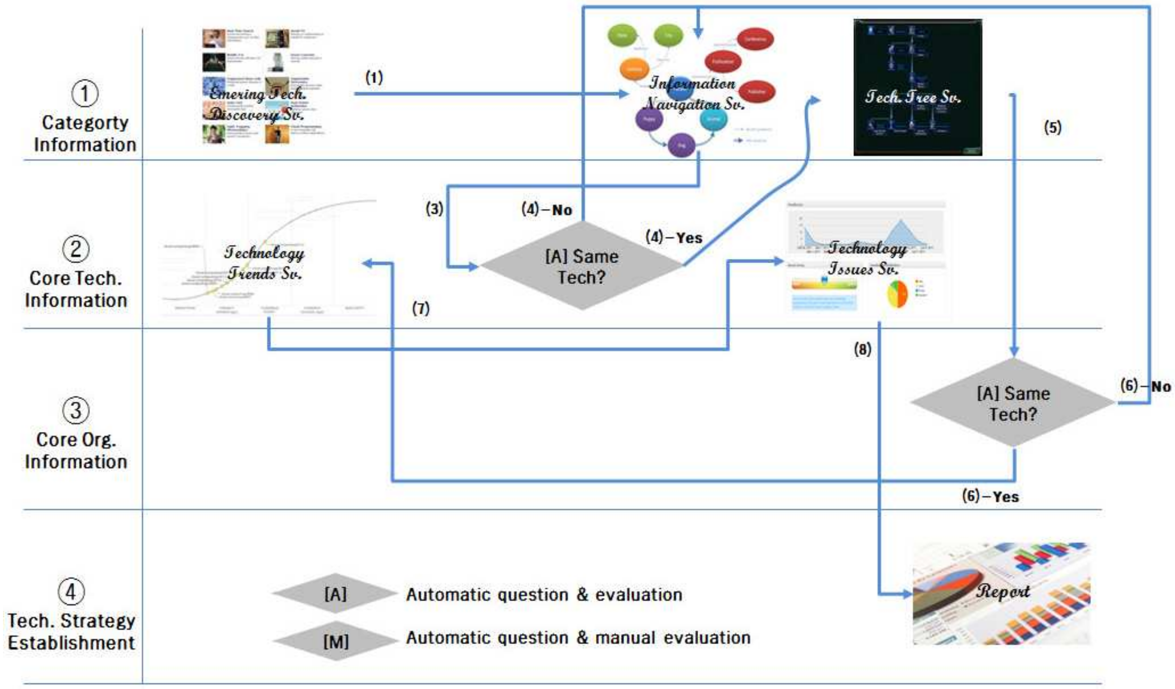

Fig. 1. Service scenario for user group 1

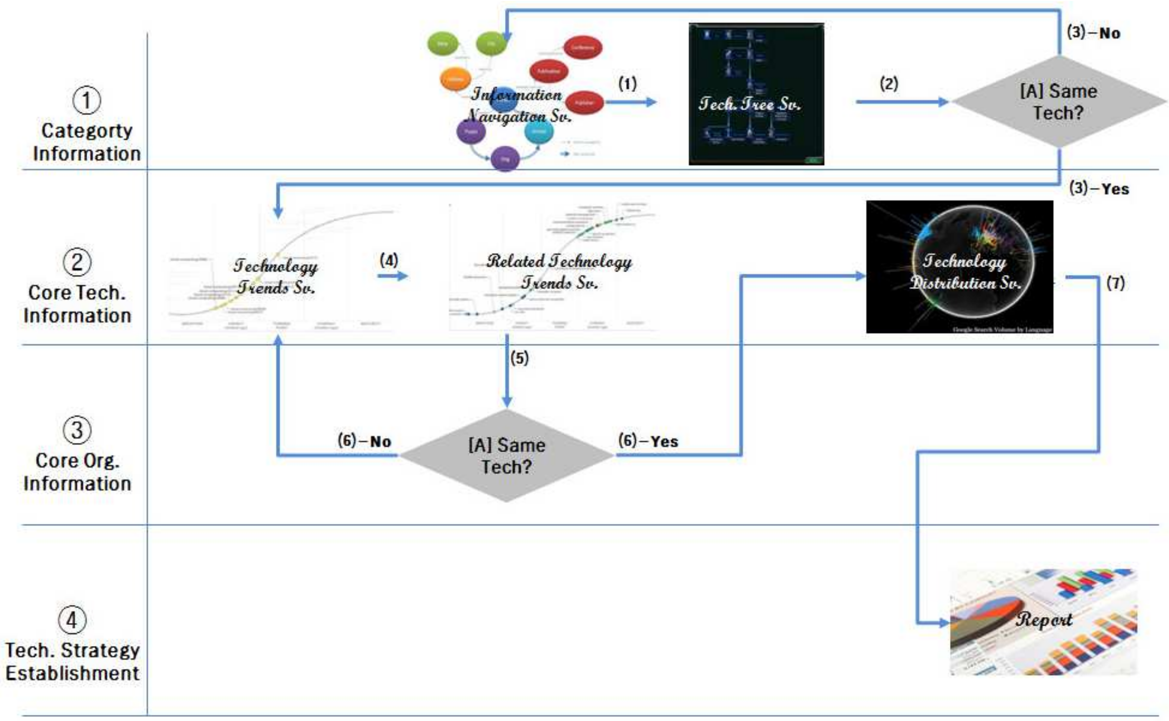

Fig. 2. Service scenario for user group 2

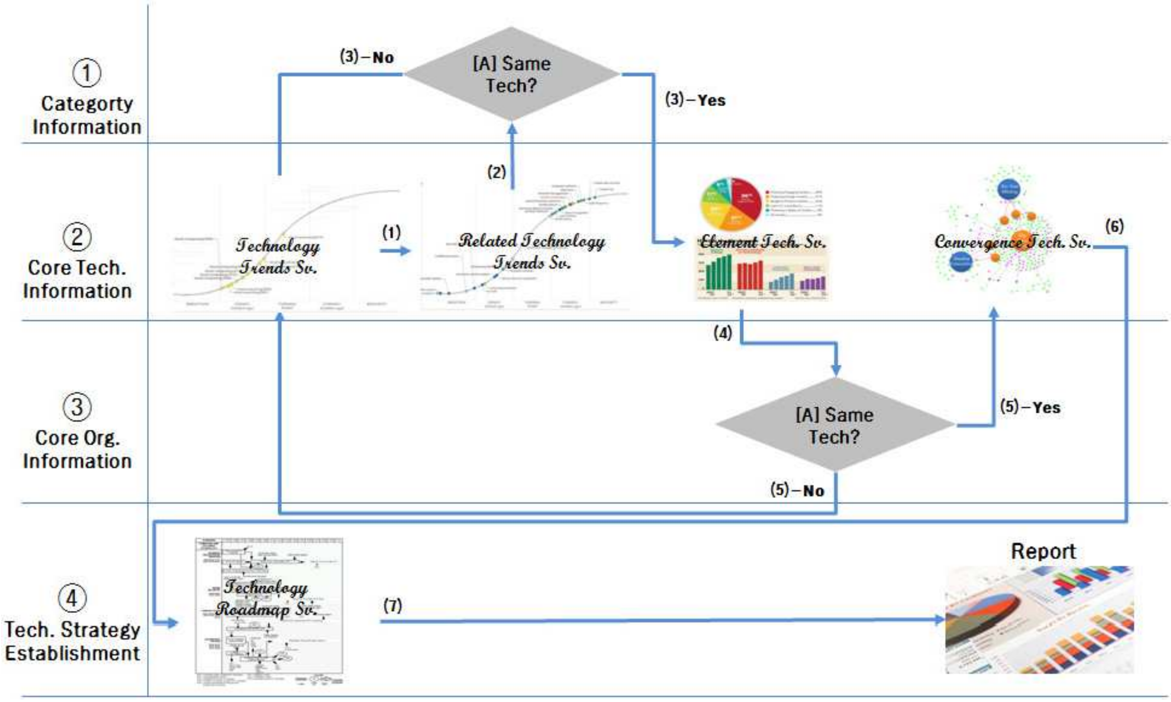

Fig. 3. Service scenario for user group 3

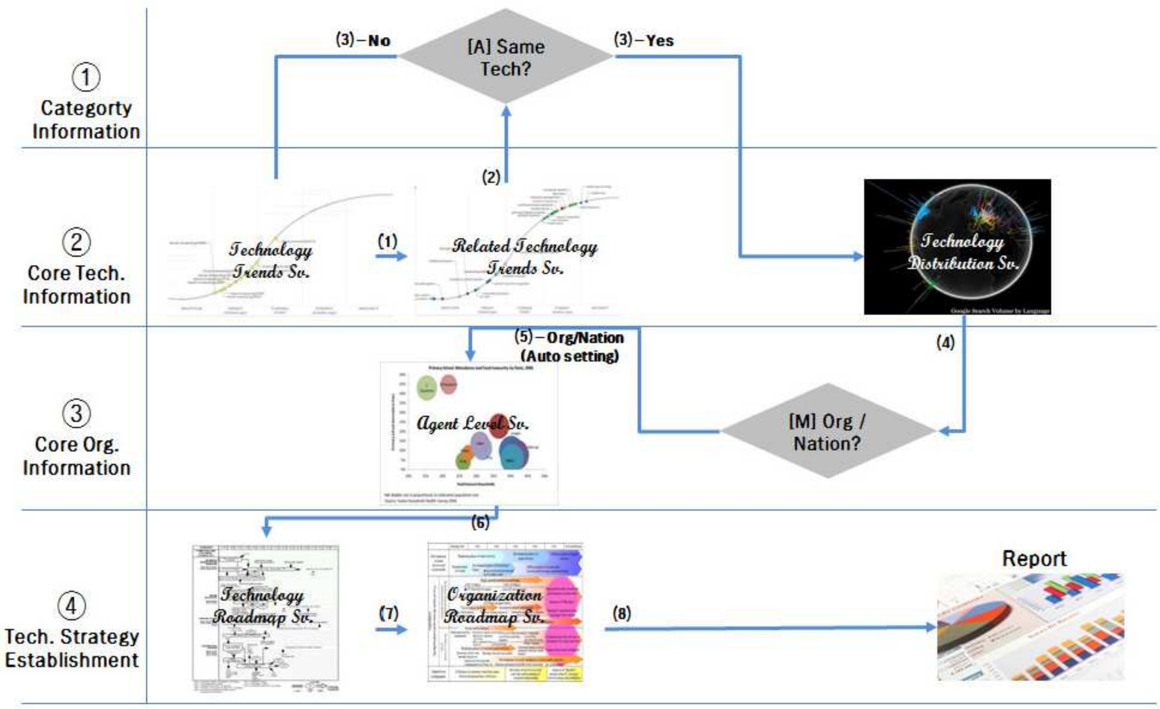

Fig. 4. Service scenario for user group 4

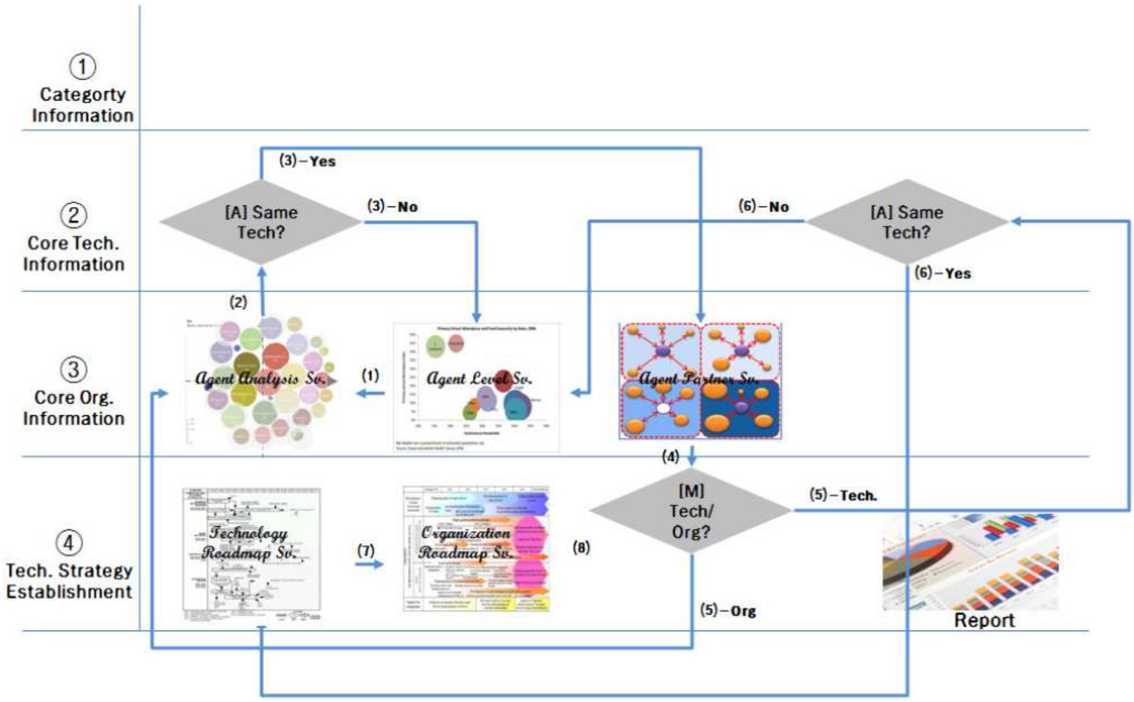

Fig. 5. Service scenario for user group 5

Definition 3. Convergence Technology Discovery Group

The convergence technology discovery group comprises users who want to discover emerging technologies for business expansion. Users have to consider which technologies they already have and which technology is the best fit for them. If the system suggests a technology that is similar to a conventional core technology already owned by the user, the user can combine the two technologies and expand their business with ease. In a dition, the system has to consider the viability of each technology to help users choose a more suitable techn ology for the future.

Definition 4. Technology Validity Confirmation Group

The technology validity confirmation group comprises users who want to verify the validity of a technology that they currently own. The system analyzes the technology and determines where it is in the technology life cycle. In addition, the system determines whether the technology is in the "blue ocean" (non-existent industry or market) or "red ocean" (existing industries and markets). If the technology is viable and in the "blue ocean," users can choose to continue to invest in that technology.

Definition 5. Leading Organization Benchmarking Group

The leading organization benchmarking group comprises users who want to discover the organization that is leading in terms of a specific technology and to benchmark that organization. With the assumption that the organization that leads in terms of a specific technology has something special that makes it succeed, users w ant to know the key secret of the leading organization. Therefore, the system suggests a list of related technoogies used by the leading organization and users can b enchmark them to analyze the success of the business.

Definition 6. Organization Trends Analysis Group

The organization trends analysis group comprises users who want analyze a specific organization in detail. Users want to learn about the organization's investments in technologies, other collaborating organizations and competing organizations. The system presents a list of technologies that an organization has invested in, as well as the ranking and degree of participation for each technology. In addition, the system discovers other organizations that collaborate or compete with the specified organization.

Definition 7. Roadmap Establishment Group

The roadmap establishment group comprises users who want to establish a simple technology/organization roadmap. Users in this group are obviously interested in a specific technology and they want to view a summary of the past, present and future of the specified technology and related technologies.

1.3. Technology Intelligence Services

InSciTe Adaptive includes three types of collection information, such as papers, patents, web resources and semantic ontology information. Papers cover IEEE proceedings, journals and various kinds of international journals in the computer science and the bioinformatics fields. Patents include those from the US, EU and Japan. Web resources cover news and magazine articles from 15 websites. Based on the data collected from the aforementioned sources, we construct a body of semantic ontology data in the quadruple form. The number of datasets collected is as follows.

In InSciTe service, we construct semantic information based on collection information such as papers, patents and web resources described in **Table 2**. By constructing ontology, we can support more diverse and rich analysis and forecast service based on relation information in ontology. We define 5 objects such as technology, organization, product, nation, person and diverse relation among them.

Especially, ontology schema in InSciTe service is different generally used ontology schema. Because time is really important factor in business intelligence regarding proper timing, we construct quadruple-form ontology not triple-form as (subject-predicate-object-time).

The technology trends analysis service represents the emerging phase of a given technology and related technologies. The emerging phase consists of five steps: irruption, frenzy, turning point, synergy and maturity. These steps have been defined as the "Great Surges of Development" by Carlota (2007). Irruptionimplies the emergence of a new technology and frenzy represents the mobilization of financial capital to explore the potential, resulting in the development of a range of business models. The turning point represents a financial crash and recession and synergy represents the emergence of new institutions and industry structures for regrowth and new technologies. Maturity represents the final step towards stability.

For the technology trends analysis service, we use ten feature sets that are based on the metadata informati on of papers and patents and that consist of diverse growth rates. Each feature set is a combination of papers and patent information.

Table 2. Dataset in InSciTe service

Papers	Proceedings	723,821
	Journals	9,041,378
Patents	US	4,963,647
	EU	1,111,853
	Others	1,540,315
Web	IDC	670
Resources	Wikipedia	4,975,178
	Gizmag	17,833
	EtnTws.com	14,679
	Technewsworld	10,099
	New York Times	125,570
	BBC	38,728
	Fox News	11,158
	CNN	20,154
	USA Today	39,502
	Other	10,000
Semantic	Triples (after inference)	498,361,449
Data	Triples (before inference)	375,935,081

In InSciTe Adaptive, we created an answer set for the decision-making process of the emerging phase of technologies based on Gartner's Hype Cycle (Leary, 2008) information. There are approximately 300,000 technologies in InSciTe Adaptive, but we constructed the answer set using only approximately 300 technologies that are included in Hype Cycle from 2007 to 2012. Based on the constructed answer set, we also created a two-level decision tree to achieve greater accuracy with the decision. In particular, the "Irruption" and "Synergy" phases are determined in the early stages of the decision tree, whereas the "Frenzy" and "Turning Point" phases are concluded towards the end of the tree. Decision accuracy for the "Irruption" and "Synergy" phases is high, whereas that for the "Frenzy" and "Turning Point" phases is not. Therefore, we use two separate decision trees to guarantee higher accuracy. The creation of a decision tree is based on the C4.5 algorithm (Du *et al*., 2011; Yi *et al*., 2011), which is used to generate a decision tree based on one that was developed by Ross Quinlan. C4.5 is an extension of Quinlan's ID3 algorithm. The constructed decision tree is an optimized machine learning method. For machine learning of the decision tree, we use the WEKA tool, the C4.5 decision tree algorithm and decision tree induction. The WEKA tool (WEKA, 2010) is a machine-learning and data-mining tool coded in Java and developed as open-source freeware by the University of Waikato in New Zealand. It supports classification, clustering, association and visualization. Because the decision trees generated by C4.5 can be used for classification, C4.5 is often referred to as a statistical classifier. Algorithm for analysis of technology trends is as **Table 3 and 4**.

Table 3. Feature sets for technology trends service

$S(Pp) = \{Pp_1, Pp_2, \ldots, Pp_n\}$, $S(Pt) = \{Pt_1, Pt_2, \ldots, Pt_n\}$

$FS_{abspoluteGraowthrate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}\} = (AN^k_{Pp} - AN^{k-1}_{Pp}) / AN^{k-1}_{Pp}$

$FS_{RelativeGrowthRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}\} = (N^k_{pp} - N^{k-1}_{Pp}) / N^{k-1}_{Pp}$

$FS_{AuthorRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}, author\} = (A^k_{Pp} / AA^k_{Pp}) * 100(\%)$

$FS_{AuthorGrowthRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}, author\} = (AA^k_{Pp} - AA^{k-1}_{Pp}) / AA^{k-1}_{Pp}$

$FS_{DomainRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}, domain\} = (D^k_{Pp} / AD^k_{Pp}) * 100(\%)$

$FS_{DomainGrowthRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}, domain\} = (AD^k_{Pp} / AD^{k-1}_{Pp}) / AD^{k-1}_{Pp}$

$FS_{JournalRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}, Journal\} = (J^k_{Pp} / AJ^k_{Pp}) * 100(\%)$

$FS_{JournalGrowthRate}(S(Pp)^k) = \{number_{Pp}, date_{Pp}, Journal\} = (AJ^k_{Pp} / AJ^{k-1}_{Pp}) / AJ^{k-1}_{Pp}$

$FS_{AbsoluteGrowthRate}(S(Pt)^k) = \{number_{Pt}, date_{Pt}\} = (AN^k_{Pt} - AN^{k-1}_{Pt}) / AN^{k-1}_{Pt}$

$FS_{RelativeGrowthRate}(S(Pt)^k) = \{number_{Pt}, date_{Pp}\} = (N^k_{Pt} - N^{k-1}_{Pt}) / N^{k-1}_{Pt}$

$FS_{InventorRate}(S(Pt)^k) = \{number_{Pt}, date_{Pt}, inventor\} = (I^k_{Pt} / AI^k_{Pt}) * 100(\%)$

$FS_{InventorGrowthRate}(S(Pt)^k) = \{number_{Pt}, date_{Pt}, Inventor\} = (AI^k_{Pt} - AI^{k-1}_{Pt}) / AI^{k-1}_{Pt}$

$FS_{ApplicantRate}(S(Pt)^k) = \{number_{Pp}, date_{Pp}, Applicant\} = (A^k_{Pt} / AA^k_{Pt}) * 100(\%)$

$FS_{ApplicantGrowthRate}(S(Pt)^k) = \{number_{Pp}, date_{Pp}, Applicant\} = (AA^k_{Pt} / AA^{k-1}_{Pt}) / AA^{k-1}_{Pt}$

$FS_{PatentFamikyRate}(S(Pt)^k) = \{number_{Pp}, date_{Pp}, PatentFamiky\} = (P^k_{Pt} / AP^k_{Pt}) * 100(\%)$

$FS_{PatentFamikyGrowthRate}(S(Pt)^k) = \{number_{Pp}, date_{Pp}, PatentFamiky\} = (AP^k_{Pt} / AP^{k-1}_{Pt}) / AP^{k-1}_{Pt}$

Table 4. Decision tree optimization algorithm

```
----------------top-down decision tree induction--------------
function GROW_TREE(T: set of examples)
returns decision tree:
t* := optimal_test(T)
p := partition induced on T by t*
if stop_criterion(p)
then return leaf(into(T))
else
for all Pj in P:
trj := GROW_TREE(Pj)
Return node(t*, ∪j{j,trj})
------------------single node refinement------------------
for all candidate tests t associated with the node:
for all examples e in the training set T:
update_statistics(S[t], t(e), target(e))
Q[t] :=compute_quality(S[t])
t* := argmax(t)Q[t]
partition T according to t*
```

The element technology analysis service illustrates sub-technologies for developing a specific technology. In addition, it suggests the distribution, portion and importance of element technologies so that users can understand which element technology is more important than others and which additional technologies need to be researched further for use with the specified technology. To extract element technologies, we use three properties of the technology, which are highlighted in red in **Fig. 6**. The portion and importance information of element technologies are represented by papers, patents and web resources separately, because each type of data has a different publication date range.

For example, in case of popular technology like 3G and 4G networks, web news and reports are published these days; however, many papers and patents about them were already published approximately 5-10 years ago. **Figure 7** shows a normalization graph that illustrates the time differences among papers, patents and web resources. Generally, it takes approximately two years from application/submission to publication in papers and patents. However, articles arepublished almost immediately in web resources. At the beginning of research, many papers and patents focus on the emerging technology, but as research on that technology stabilizes with many studies, web coverage increases.

The convergence technology discovery service represents two technologies that can be combined to create a new technology. If two technologies share many element technologies, we assume that they can be combined easily. However, if the element technologies of two technologies are almost same, we can suppose that those two technologies were already co-researched and are almost the same technology.

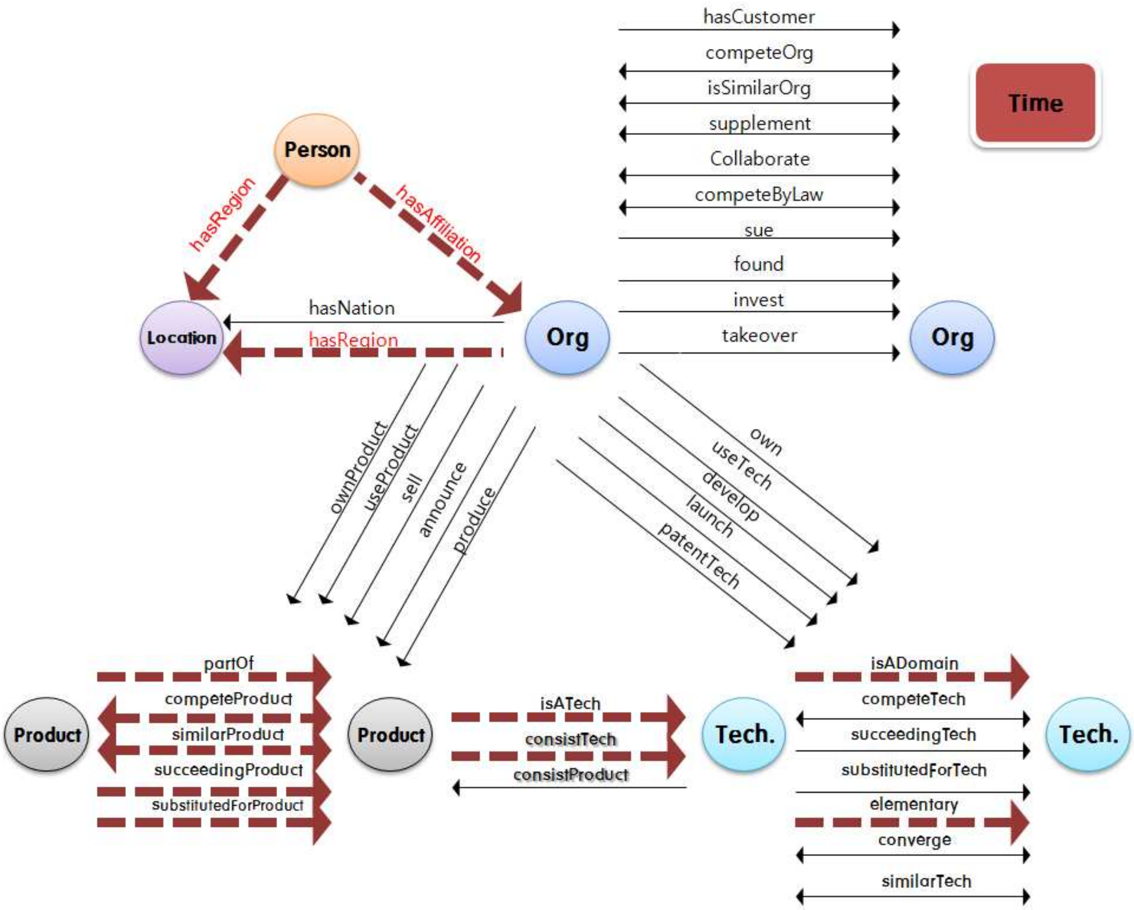

Fig. 6. Ontology schema in InSciTe service

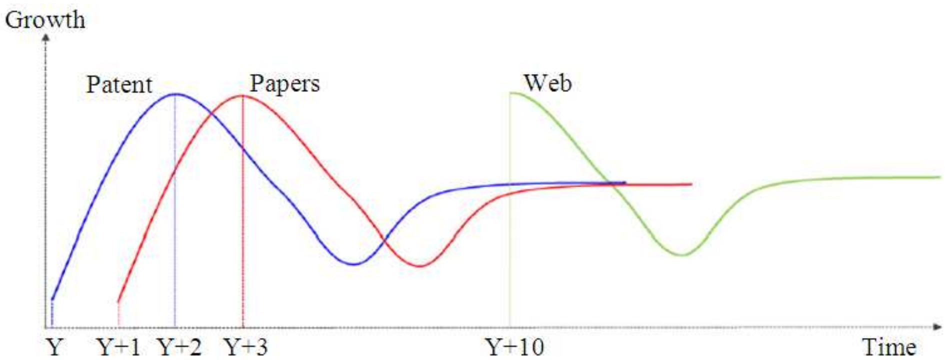

Fig. 7. Time deference among papers, patents, web

In this service, we suggest convergence technologies that share some element technologies but no more than a threshold number. Algorithm for finding convergence technology is as **Table 5**.

Table 5. Algorithm for Discovery of Convergence Tech

----------Convergence Technology Discovery----------
1 function Conver_TECH(T: technology)
2 returns (Convergence_Tech, Convergability):
3 s:=N(Element(T));
4 for s=1 to s_{max}
3 ET[s]:=Element(T);
4 n:=N(SupEle(ET[s]))
5 for n=1 to n_{max}
5 $SET_{ET[s]}$[n]:= SupEle(ET[s])
6 CT[t]:=∩ $SET_{ET[s]}$[n]
---------------Convergability calculation----------------
7 for t=1 to t_{max}
8 Rt=Element(CT[t])/s_{max}
9 Return (CT(t), Rt)

2. CONCLUSION

In this study, we suggest a business intelligence system that focuses on technology analysis and prediction, as well as adaptive recognition of the user preference and intention. InSciTe Adaptive reflects the user's precise needs and intention by applying a five-step user modeling and grouping process. In addition, by using diverse text mining and semantic web technologies, we extracted valuable information from three types of data, paper, patents and web resources and constructed ontology to understand the relationships among the objects. InSciTe Adaptive provides several types of technology analyses and prediction services and helps users make better business decisions.

In future work, we plan to validate service models and analysis results from each technology analysis service. Nowadays, diverse organizations, such as Gartner, MIT and UC Berkley, publish technology analysis and prediction results as emerging technologies and combinable technologies. Therefore, we need to compare the analysis and prediction results from our studies with those of other organizations.

3. REFERENCES

Azma, F. and M. Mostafapour, 2012. Business intelligence as a key strategy for development organizations. Procidia Technol., 1: 102-106. DOI: 10.1016/j.protcy.2012.02.020

Carlota, P., 2007. Great surges of development and alternative forms of globalization. Universities of Cambridge and Sussex, U.K.

Cheung, C.F. and F.L. Li, 2012. A quantitative correlation coefficient mining method for business intelligence in small and medium enterprises of trading business. Expert Syst. Appli., 39: 6279-6291. DOI: 10.1016/j.eswa.2011.10.021

Du, M., S. Wnag and G. Gong, 2011. Research on Decision tree algorithm based on information entropy. Adv. Mater. Res., 267: 732-737. www.scientific.net/AMR.267.732

Ranjan, J.R., 2009. Business intelligence: Concepts, components, techniques and benefits. J. Theoretical Applied Inform. Technol.

Kim, J., Lee, M., Sung, W., Song, S., S. Song and H. Jung, 2012b. Toward discovering and predicting technical opportunities and technology trends. Adv. Inform. Sci. Service Sci., 4: 161-167.

Kim, J., M. Hwang, D.H. Jeong and H. Jung, 2012a. Technology trends analysis and forecasting application based on decision tree and statistical feature analysis. Expert Syst. Applic., 39: 12618-12625. DOI: 10.1016/j.eswa.2012.05.021

Kim, J., S. Lee, J. Lee, M. Lee and H. Jung, 2013. Design of TOD model for analyzing technology trends and predicting future trends. Inform. J., 16: 1371-1378.

Klai, S., Sevinc, E., Fontaine, B., Jacobs, C., Muller, C, 2012. Proceedings of the International Conference on Space Operations, (CSO' 12).

Leary, D.E., 2008. Gartner's hype cycle and information system research issues. Int. J. Account. Inform. Syst., 9: 240-252. DOI: 10.1016/j.accinf.2008.09.001

Lee, M., S. Lee, J. Lee, H. Jung and J. Kim *et al*., 2013. Strategic decision-making support service based on technology opportunity discovery model. Inform. J., 16: 693-698.

Li, Z., L. Li and D. Lian, 2012. Business intelligence in enterprise computing environment. Inform. Technol. Manage., 13: 297-310. DOI: 10.1007/s10799-012-0123-z

Point, V., 2009. Text Mining Software for Technology Management-Search Technology.

WEKA, 2010. Data Mining with Open Source Machine Learning Software in Java.

Yi, W., M. Lu and Z. Liu, 2011. Multi-valued attribute and multi-labeled data decision tree algorithm. Int. J. Mach. Learn. Cybernet., 2: 67-74. DOI: 10.1007/s13042-011-0015-2