An Extension Collaborative Innovation Model in the Context of Big Data

Xingsen Li[a], Yingjie Tian[b], Haolan Zhang[a], Florentin Smarandache[c]

[a]School of Management, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, P.R.China
[b]Research Center on Fictitious Economy and Data Science, CAS, Beijing 100190, P.R.China
[c]Mathematics and Science Department, University of New Mexico, Gallup, NM 87301, USA
Email: lixs@nit.zju.edu.cn, smarand@unm.edu

Abstract The process of generating innovative solutions mostly rely on skilled experts which are usually unavailable and with uncertainty. Computer science and information technology is changing the innovation environment and accumulating big data from which a lot of knowledge is discovered. However, it is a rather nebulous area and still remains several challenge problems to integrate multi-information and lots of rough knowledge effectively to support the process of innovation. Based on the new cross discipline Extenics, we present a collaborative innovation model in the context of big data. The model transforms collected data into a knowledge base in a uniform basic-element format, and then we explore the innovation paths and its solutions by a formularized model based on Extenics. Finally we score and select all possible solutions by 2D dependent function. The model can collaborate different departments to put forward the innovation solutions with support of big data. The model is proved useful by a practical innovation case in management.

Keywords: Extension Innovation Model; Big data; Data Mining; Extenics; Knowledge Management

1. Introduction

In the last few decades, a large number of scholarly efforts and theories on collaborative innovation have been developed, and many approaches have contributed to reveal the nature of innovation process in various degrees (Hippel & Katz, 2002; Cantisani, 2006). Nevertheless, the innovation process still largely remains a black box (Birkinshaw, Hamel, 2008). Although each theory can explain *some* of the mechanisms behind innovations, the general mysteries behind innovation process are still far from resolved. Most of the innovation models makes use of a group of experts, along with amounts of time and depend on personal intelligence which would subject to limitations of individuals themselves, therefore it keeps out of step under the rapidly changed information and knowledge era.

We live in an era when a remarkable number of new information and data is accumulated everywhere. Because of the massive amount of data that is generated almost everywhere, new tools need to be developed in order to manage and analyze the data, especially in the field of management (Chae and Olson, 2013). The information environment is rapidly and continuously changing and uncertain due to global competition, information explosion and advances in new technologies. The Web and other information technology (IT) combined with business and lives have accumulated huge data and information. Information and other technologies are successfully bypassing the main obstacle to technological advance, and when technology support net is fully established and fixed, innovation is not free and autonomous process of applied creativity, but is technically, economically and politically subservient to the "holders and owners" of the support net (Zeleny,2012).

The data is accumulating in almost all aspect of our lives that it becomes huge, muliti-structures and much of hidden valued information. This is the era of big data.

Big data (Bughin, et al. 2010; Chisholm, 2012) is a collection of data sets. It is so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications.

The challenges also include capture, curation, storage, search, sharing, transfer, analysis and visualization (http://en.wikipedia.org/wiki/Big_data). Big data provides materials for mining hidden patterns to support innovation (Manyika, 2011) mostly by data mining (Han, 2006; Olson and Shi, 2007). The interaction research with big data techniques' supports methods for innovation is rare at present (Ji, et al. 2012). Knowledge discovered by data mining is novel and quantitative (Li, Shi and Zhang, 2010; Shi and Li, 2007). However, it still lacks a uniform knowledge management model to support the innovation process effectively (Zhang et al,2009).

Extenics (formerly referred as Extension Theory) focuses to solve contradictory problems by formalization methods based on the concepts of matter-element and extension set (Cai, 1983; Yang and Cai, 2013). Extenics use a uniform 3-dimension-matrix to express information and knowledge, and utilizes extension transformation to represent the properties' change of things, and indicates things with certain attribute can be changed into things that without such attribute. It provides a new view for understand the process of innovation. But it needs more support of information technology, especially for big data.

To improve the quality of collaborative innovation by objective hidden knowledge from big data, we propose a new innovation model combining information technology and Extenics. The rest of this paper is organized as follows. Section 2 discusses the main existing innovation methods and problems. Section 3 provides a systemic model for preparing the input for innovation from big data. Section 4 presents a framework and process about utilization of Extenics to generate innovative ideas or solutions with a new method to score the solutions, followed by a case study in Section 5 and a brief summary and some future research scopes in Section 6.

2. Literature Review

There are a number of factors that affect the quality of collaborative innovation including internal factors such as the will to change, the attitude to new things, the thinking model to overcome habitual domains (Yu, 1995) and external factors such as the information, data, team work and the policy. Among them creative thinking model (Saaty, 2010) and information technology are the most important (Li, et al 2013).

Creative thinking mostly relies on individuals. It cannot be understood using a single, simple, model, it involves multiple, complex, processing operations. The operation of multiple processes, multiple strategies, and multiple knowledge structures makes it difficult to understand creative thinking process (Hennessey, Amabile, 2010). However, effective creativity execution depends on the knowledge available and the strategies people employ in executing these processes (Mumford, Medeiros and Partlow, 2012).

The operation of multiple processes, multiple strategies, and multiple knowledge structures makes it difficult to formulate an understanding of innovation (Hennessey and Amabile, 2010). Declarative knowledge, factual, information and cognitive schema are commonly held to be involved in most forms of creative thinking. Information & Communication Technology (ICT) tools are likely to provide new innovation approaches and effective means to support such new innovation processes. The new approaches for innovation will find their wide application in industry (Sorli and Stokic, 2011).

A lot of innovation methods makes use of various approaches to stimulate innovation of individuals or a group of experts/team members, along with huge amounts of spend of time, fund and human resource and severe dependence upon personal intelligence, so it would subject to limitations of individuals themselves (Vandeven, 1986), keep out of step under the rapidly changed information and knowledge environments.

To support Innovation process, there are some tools as following.

There can be eight core processes for the effective execution of innovation (Mumford, Medeiros and Partlow, 2012): (a) problem definition, (b) information gathering, (c) information organization, (d) conceptual combination, (e) idea generation, (f) idea evaluation, (g) implementation planning, and (h) solution monitoring. Effective execution of these processes, in turn, depends on people applying requisite strategies during process execution and having available requisite knowledge. And it needs to be refined in detail for practical use.

TRIZ was developed to resolve contradictions in technological inventions, with a set of 40 inventive principles and later a Matrix of Contradictions which indicates 39 system factors (Hua, Yang, et al. 2006). It's useful in several specific fields, such as mechanics and electronics, but limited in other fields.

One main problem presented in many existing approaches is they have applied a deductive approach by attempting to abstract common features from historical instances and to obtain general rules for invention. Although this "expert system"-like approach is not without its merit, in reality it is not realistic, because there are so many conditions and variables to match, and for each condition or variable, there are so many possible values to compare, so it may be computationally infeasible. On the other hand, although generative approach has been proposed by some authors, there has been a lack of effective way of generating all possible innovative solutions.

Extenics focus on solving incompatible problems by formularized methods both in management and engineering. Zhou and Li (2010) put forward an Extenics-based enterprise independent innovation model and its implement platform. Declarative knowledge, factual, information and cognitive schema are commonly involved in most forms of complex performance including innovation (Ericsson & Moxley, 2012). By integrate methodology knowledge and information, we need a theory to guide the generation of innovative solutions.

Big data is the next frontier for innovation, competition, and productivity (Manyika, Chui, et al, 2011), it can help to better capture, understand, and meet customer needs (Van

Horn, Olewnik & Lewis, 2012), it's a very important source to knowledge (Snijders, 2012; Hilbert, 2013) and other new discoveries (Kim, Lund, Dombrowski, 2013), but how to use big data to support the collaborative innovation effectively? The data preparation process for innovation remains a black box. Data is big enough but the methods to handle information and knowledge are very limited. Moreover, models listed above pay little attention to data analysis during innovation processes, the methods to score innovation solutions are mostly qualitative and we need more quantitative methods.

The big data and information is so huge that they are beyond human mind's processing capability. So it's time to implement collaborative innovation effectively in big data era. It's necessary to explore the high-efficiency models that would fill up the gaps with innovation process and new big data methodologies. We try to use Extenics to bridge the innovation process with big data technology in this paper.

3. Preparation for Innovation from Big Data

To overcome above problems, we design Extension collaborative innovation model mainly on big data technology and Extenics.

3.1 Data set Collection based on Extenics

Innovation process needs data and knowledge, both explicit and tacit. There are two main sources for collecting data: Internal source, such as MIS, local data base, working tables or other files, External source, such as the Web, public data base, data of other companies with relationships et al. There is huge quantity of data and information still growing dramatically. How to choose the proper data set and process it is a challenge problem.

Basic-element theory describe the matter (Physical existence), event and relations as the basic elements for analysis on incompatible problems —"matter-element", "event-element" and "relation-element". Basic element is an ordered triad composed of the element name, the characteristics and its value, denoted by R= (N, c, v) as matter-element, $I = (d, b, u)$ *as* event-element and relation-element $Q = (s, a, w)$ (Cai, 1983). As the matter-element R= (N, c, v) is an ordered triad composed of the matter, its characteristics and measures, we can develop new concepts from the extensibilities of one of the three sub-elements in the triad.

Multiple characteristics are accompanied with multi-dimensional parametric matter-element, and can be expressed as:

$$M(t) = \begin{bmatrix} O_m(t), & c_{m1}, & v_{m1}(t) \\ & c_{m2}, & v_{m2}(t) \\ & \vdots & \vdots \\ & c_{mn}, & v_{mn}(t) \end{bmatrix} = (O_m(t),\ C_m,\ V_m(t))$$

As to a given matter, it has corresponding measure value about any characteristic, which is unique at non-simultaneous moment.

Furthermore, characteristics of matters can be divided into materiality, systematicness, dynamism and antagonism, which are generally called matter's conjugation. According to matters' conjugation, a matter is consisted of the imaginary and real, the soft and hard, the latent and apparent, the negative and positive parts (Yang and Cai, 2013).

3.1.1. Non-physical Part and physical Part

In terms of physical attribute of matter, all matters are composed of a physical part and a non-physical part, the former is referred to the real part of matter and the latter is referred to the *non-physical* or virtual part of matter. For example, a product's entity is its real part, while its brand and reputation are its *non-physical* part. The empty space is a cup's *non-physical* part, while the glass consists of the cup is its *physical* part.

3.1.2. Soft Part and Hard Part

Considering a matter's structure in terms of the matter's systematic attribute, we define the matter's components as the hard part of matter, the relations between the matters and its components as the soft part of the matter. "Three heads are better than one", three persons are hard parts. Their cooperation relationship is the soft part. Good soft part completely lead to good results.

The matter's soft part has three types: 1) relations between the matter's components; 2) relations between the matter and its subordinate matters, and 3) relations between the matter and other matters.

3.1.3. Latent Part and Apparent Part

Considering of a matter's dynamic property, we trust that any matter is changing. Disease has its latent period, seed has its incubation period of germination, and an egg can hatch into chicken at a certain temperature and after a certain time. The matter's latent parts and apparent parts exist synchronously.

The latent part of some matters may become apparent under certain conditions, for example, students in class currently will become teachers after ten years. There must be a criticality in the process of reciprocal transformation between latent parts and apparent parts.

3.1.4. Negative Part and Positive Part

In terms of antithetic property of matters, all matters have two antithetic parts. The part producing the positive value to certain characteristic is defined as the positive part, and the part producing the negative value is defined as the negative part.

For example, in terms of profits, employees' welfare department, kindergarten, and publicity department, etc. have negative value of measure about profits, being the negative parts of the company, but these parts will improve employees' job enthusiasm and promote company's reputation, so they are the "advantageous" parts of the company.

Conjugate analysis and basic-element theory is a guide for us to collect data and information in a systematical way. Accordingly, we form a detailed data collecting list as showing in table 1(Li et al, 2013):

Innovation activities have their goals and the conditions. The purpose of data collection and processing is to solve problems about how to from the conditions to the goals. Therefore, the data we collected should also relative to the goals.

TABLE I. DATA COLLECTION LIST BASED ON BASIC-ELEMENT THEORY AND ITS CONJUGATE ANALYSIS

	Materiality		Systematicness		Dynamic		Antithetical	
	physical part	*non-physical part*	*soft part*	*hard part*	*apparent part*	*latent part*	*positive part*	*negative part*
Matter-Element	The physical part of matter	non-physical part of matter	Soft part of matter	Hard part of matter	apparent part of matter	Latent part of matter	Positive part of matter	Negative part of matter
Event-Element	The physical part of event	non-physical part of event	Soft part of event	Hard part of event	apparent part of event	Latent part of event	Positive part of event	Negative part of event
Relation-Element	The physical part of relation	non-physical part of relations	Soft part of relation	Hard part of relation	apparent part of relation	Latent part of relation	Positive part of relation	Negative part of relation

It can be seen from the above definition that there are three namely paths for the process of transformation between the "positive" and "negative", field, criteria and element. Accordingly, the data and information we collect will include such 3 paths.

TABLE II. INFORMATION COLLECTION LIST FROM THE VIEW OF THE GOAL

	Goals	Conditions	Pathways
Field	Field of goal	Field of conditions	Field of pathway
Criteria	Criteria of goal	Criteria of conditions	Criteria of pathway
Element	Element of goal	Element of conditions	Element of pathway

Goals and conditions can be matter, event, or relations between matters and events which can be represented with basic elements. So each cell in the table 1 can be a basic element in next level of the information tree.

From table 1 and table 2, we can get a systematic cube for collecting data and information for innovation.

3.2 Data processing paths

The big data preprocess chart is shown in fig. 1 as following.

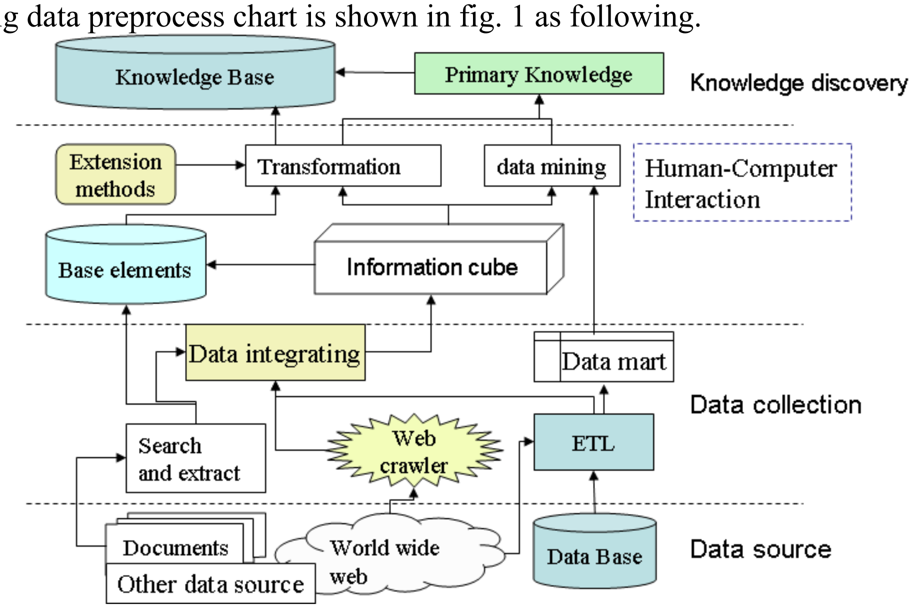

Figure 1. big data collection and processing

There are two main paths located in four levels. One path is to extract data from data base, or use web crawler to collect information from the Web, then transform and clean it into data mart, finally use data mining to discover primary knowledge. Another way is to collect documents and build an information cube by human-computer interaction, then save in base-element data base. Using extension transformation methods, we transform the basic-elements into knowledge base. Base elements and knowledge base will be the input of collaborative innovation model.

3.3 Data Transformation Methods

There are five basic transformations methods in Extenics, which can be used for information transformation by change matter's object, attribute or the value.

a) Substitution transformation

As to basic-element $B_0(t)$=($O(t)$, c, v (t)), if there is certain transformation T that transforms $B_0(t)$ to $B(t)$=(O (t), c, v (t)), i.e. T $B_0(t)$= $B(t)$, then the transformation T is referred to as substitution transformation of basic-element $B_0(t)$.

b) Increasing/decreasing transformation

Increasing transformation refers to increase certain attributes of the element. For example, as to matter-elements M_0= (table A_1, height, 0.8m), M= (chair A_2, height, 0.5m), M is increasable matter-element of M_0, we make TM_0=$M_0 \oplus M$= (table $A_1 \oplus$chair A_2, height, 1.3m), then T is increasing transformation of M_0.

Decreasing transformation refers to decrease certain attributes of the element. In the production process, the reduction of redundant action or work procedures belongs to the decreasing transformation of event-element, which can significantly improve production efficiency.

c) Expansion/contraction transformation

Expansion transformation: Quantitative expansion transformation is multiple quantitative expansion of basic-element. As for matter-element, its quantitative expansion transformation will inevitably lead to expansion transformation of the matter. For example, the volume expansion of a balloon will inevitably lead to expansion of the balloon itself.

Contraction transformation: As for matter-element, its quantitative contraction transformation will inevitably lead to contraction transformation of the matter.

d) Decomposition/Combination transformation

Decomposition transformation refers to divide one object or attributes into several pieces. On the contrary, Combination transformation refers to combine several objects or attributes into a whole one. For example, one action can be executed in several steps.

e) Duplication transformation

Duplication transformation refers to duplicate the basic-element to multiple basic-elements, such as photo-processing, copying, scanning, printing, disc carving, sound recording, video recording, the method of reuse, and reproduction of products, etc. This kind of transformation is extensively applied in the field of information, such as file copying and pasting.

Based on theory of Extension Set, knowledge from data mining can be mined in second level by transformation methods, such as substitution transformation, decomposition or combination transformation et al. For example, decision tree mined the explainable rules, but it is only static know-what knowledge, we still don't know how to transfer class *bad* to class *good.* In order to improve such kind of situation, we focus on a new methodology for discovering actionable know-how knowledge based on decision tree rules and Extension set theory. It is useful to re-mine rules from data mining so as to obtain actionable knowledge for wise decision making. The transformation knowledge acquiring solution on decision tree rules are practically used to reduce customer churn (Li et al 2013).

4. Framework and Process of Extension Collaborative Innovation

4.1 Framework of Extenics based Innovation

The innovation method based on Big data and Extenics would take advantage of specific extension methods to generating new innovative ideas or solutions. The framework and its relevant steps are listed as following:

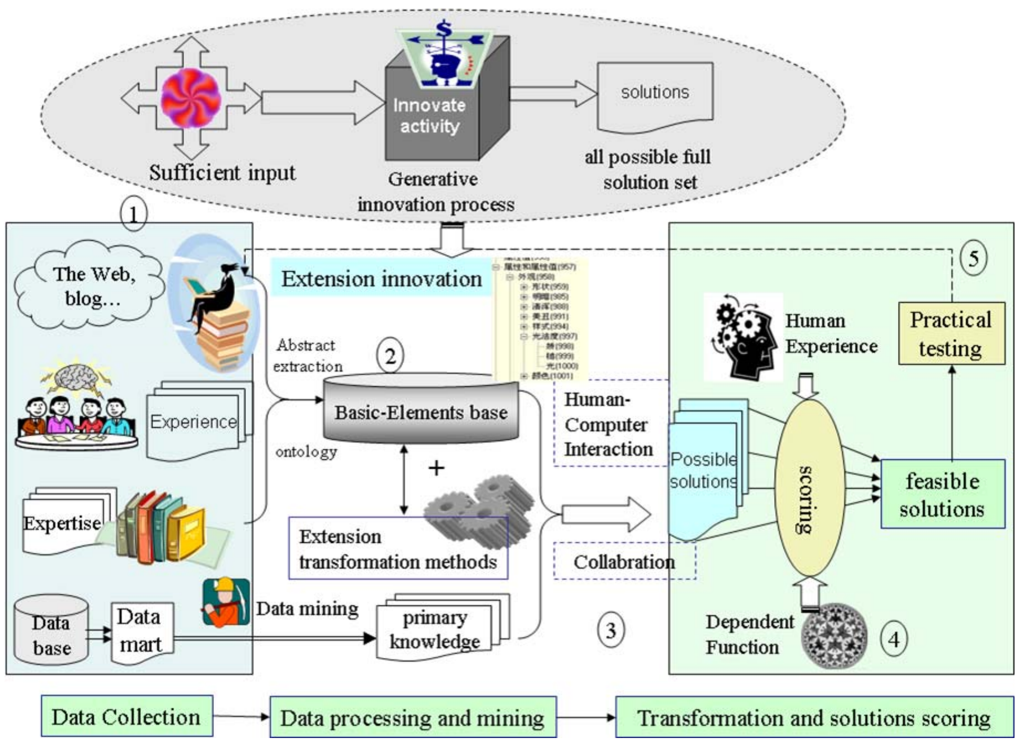

Figure 2. Framework of Extension collabrative innovation model

Step 1. Multi-structure data Collection

Collect data related to the innovation goal G and practical condition L from data base, expertise, tacit knowledge such as experience and the Web, blogs, et al. according to the method presented in section 3.1.

Step 2. Build basic-element base

Describe and transform the information into matter-elements, event-elements or relation-element. Meanwhile, we discover primary knowledge from data mart by data mining. Then, we save them into data base as a basic-element tree supported by ontology. After this step, we could get a systematic cube of integrated information (Li et al, 2009), according to the method presented in section 3.2.

Step 3. Obtain all possible solutions by Extension transformation

Taking basic-elements and primary knowledge as input, Extension transformation methods as methodology (as shown in 3.3), we transform the field, the elements or the criteria of the goals and conditions on basic-elements that are already explored by Step 2, the detailed description will be presented in section 4.2. and get all possible solutions by human-computer interactions based on extension set theory.

Step 4. Scoring and Evaluation by Dependent Function

We score all the possible solutions by superiority evaluation method based on dependent function quantitatively and expert's experience qualitatively. Then result in feasible innovation proposals.

Suppose measuring indicators set $MI=\{MI_1,\ \ MI_2,\ \ \cdots,\ \ MI_n\}$, $MI_i=(c_i,\ \ V_i)$, $(i=1,\ 2,\ \cdots,\ n)$, and weight coefficient distribution is

$$\alpha=(\alpha_1,\ \ \alpha_2,\ \ \cdots,\ \ \alpha_n)$$

According to the requirements of every measuring indicator, establish dependent functions $K_1(x_1), K_2(x_2), \cdots, K_n(x_n)$.

The dependent function value of object Z_j about each measuring indicator MI_i is denoted by $K_i(Z_j)$ for easy writing, and then the dependent degree of every object $Z_1, Z_2, \cdots, Z_m$, about MI_i is

$$K_i=(K_i(Z_1),\ \ K_i(Z_2),\ \ \cdots,\ \ K_i(Z_m)),\quad (i=1,\ 2,\ \cdots,\ n)$$

The above dependent degree is standardized as:

$$k_{ij}=\frac{K_i(Z_j)}{\max\limits_{q\in\{1,2,\cdots,m\}}\left|K_i(Z_q)\right|},\quad (i=1,\ 2,\ \cdots,\ n;\ j=1,\ 2,\ \cdots,\ m)$$

And then the standard dependent degree of every object $Z_1, Z_2, \cdots, Z_m$ about MI_i is

$$k_i=(k_{i1},\ \ k_{i2},\ \ \cdots,\ \ k_{im}),\quad (i=1,\ 2,\ \cdots,\ n)$$

Step 5. Practical testing and feedback

We apply the feasible solutions to practices, collect new data and contrast the results, feedback to the model and update basic-elements, knowledge base or methods.

A framework of Extension innovation model has been used in a famous company in China, a case study is presented as following in Section 5.

4.2 Directions of collaborative innovation

All information and knowledge can be described as basic-element, take matter as example, it has four characteristics and eight aspects as mentioned in Section 3.1. There are four main directions for innovate based on matter analysis as shown in Fig. 3:

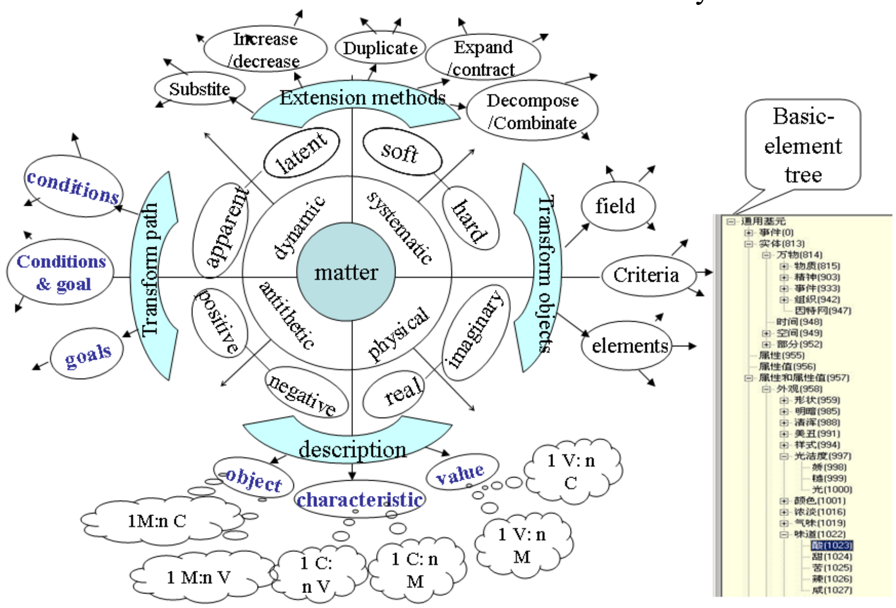

Figure 3. four main directions for collabrative innovattion

[1].From the view of descriptions, we can extend our ideas from object, character and its value, each object has many characters, marked as 1*M: n C,* similarly, each character can has many values, such as color can be red, green, yellow or blue. We mark it as 1 *C: n V*.

[2].From the view of transformation path, we can extend our thinking from goals, conditions and both goal and conditions, which can be described as base-elements. Each path has many characters, values and objects.

[3].Based on basic extension methods, there are five methods to use on the matter, detailed description is in Section 3.3.

[4].From the view of transformation objects, we can transform the elements, such as matter, event or relationships. Or transform the criteria and field. For example, a sales man *F* is regarded as good in company *A*, but scored as bad after he changed to company *B*. *F* is element, the rules of *good* or *bad* is criteria, and *A* to *B* is field. Each change will lead to different results.

After four main transformations, we can get an information tree both for goals and conditions. Similarly, we obtain a innovation solutions tree, which is in coding. The solution tree is shown in Fig.4 as following:

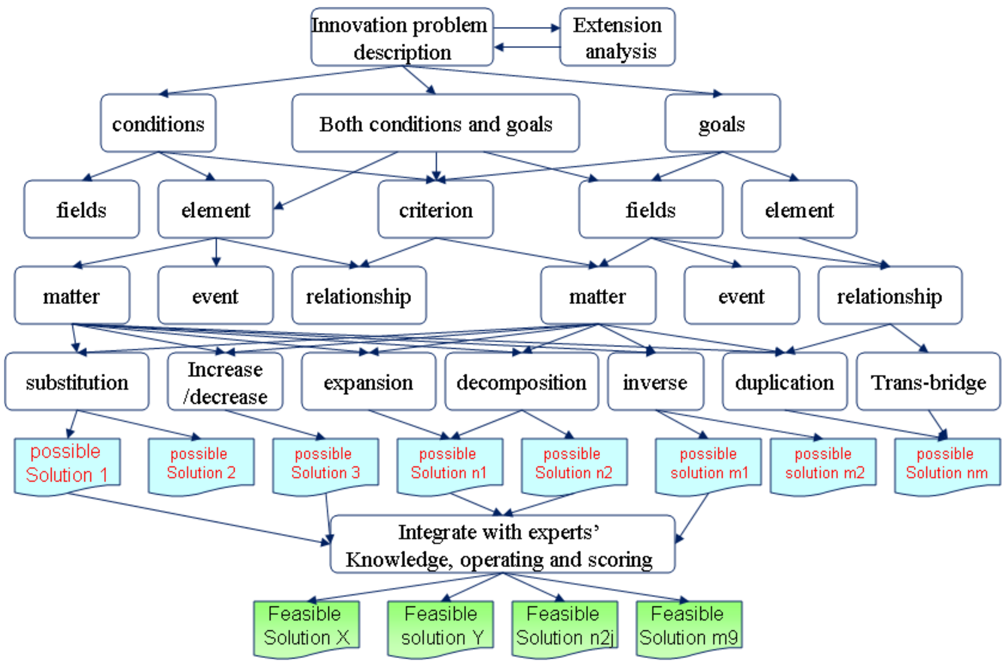

Figure 4. Map of innovation solutions

4.3 Scoring method based on big data

Big data give us a full view to score our innovation solutions. Take 1D and 2D as examples according to Extenics (Smarandache, 2012).

Extension set theory is a set theory that describes the changing recognition and classification accordingly. Extension set describes the variability of things, using the number in $(-\infty, +\infty)$ to describe the degree of how the thing owns certain property, and using an extensible field to describe the reciprocal transformation between the "positive" and "negative" of things. It can describe not only the reciprocal transformation between the positive and negative of things, but also the degree of how the thing owns a property.

4.3.1 Definition of Extension Set

Suppose U is universe of discourse, u is any one element in U, k is a mapping of U to the real field I, $T=(T_U, T_k, T_u)$ is given transformation, we define

$$\tilde{E}(T)=\{(u,y,y') \mid u\in U,\ y=k(u)\in I;\ T_u u\in T_U U,\ y'=T_k k(T_u u)\in I\}$$

an extension set on the universe of discourse U, $y=k(u)$ the dependent function of $\tilde{E}(T)$, and $y'=T_k k(T_u u)$ the extension function of $\tilde{E}(T)$, wherein, T_U, T_k *and* T_u are transformations of respective universe of discourse U, dependent function k and element u (Yang and Cai, 2007).

if $T\neq e$, we define

$$\underset{\bullet}{E}_{+}(T)=\{(u,y,y') \mid u\in U,\ y=k(u)\leq 0;\ T_u u\in T_U U,\ y'=T_k k(T_u u)>0\}$$

As *positive extensible field of* $\tilde{E}(T)$; we define

• $\overset{E}{}_{-}(T) = \{ (u,y,y') \mid u \in U, y=k(u) \geqslant 0;\ T_u u \in T_U U, y'= T_k k(T_u u)<0 \}$

As *negative extensible field* of $\tilde{E}(T)$; we define

$E_+(T) = \{ (u,y,y') \mid u \in U, y=k(u)>0;\ T_u u \in T_U U, y'= T_k k(T_u u)>0\}$

As *positive stable field* of $\tilde{E}(T)$; we define

$E_-(T) = \{ (u,y,y') \mid u \in U, y=k(u)<0;\ T_u u \in T_U U, y'= T_k k(T_u u)<0\}$

As *negative stable field* of $\tilde{E}(T)$; we define

$J_0(T) = \{ (u,y,y') \mid u \in U, T_u u \in T_U U, y'= T_k k(T_u u)=0\}$

As *extension boundary* of $\tilde{E}(T)$.

4.3.2 Dependent Function in *1D*.

Prof. Cai Wen defined in 1983 in *1D* the Dependent Function *K(y)*.
If one considers two intervals X_0 and X, that have no common end point, and $X_0 \subset X$, then:

$$K(y) = \frac{\rho(y,X)}{\rho(y,X) - \rho(y,X_0)}.$$

Since *K(y)* was constructed in 1D in terms of the extension distance $\rho(.,.)$, we simply generalize it to higher dimensions by replacing $\rho(.,.)$ with the generalized $\rho(.,.)$ in a higher dimension.

4.3.3 Extension Distance in 2D.

Instead of considering a segment of line *AB* representing the interval *[a, b]* in *1R*, we consider a rectangle *AMBN* representing all points of its surface in *2D*.
Let's consider two arbitrary points $A(a_1, a_2)$ and $B(b_1, b_2)$.
Let's note by *O* the midpoint of the diagonal *AB*, but *O* is also the center of symmetry (intersection of the diagonals) of the rectangle *AMBN*.
Then one computes the distance between a point $P(x_0, y_0)$ and the rectangle *AMBN*.
compute the distance in *2D* (two dimensions) between the point *P* and the center *O* of the rectangle (intersection of rectangle's diagonals);

considering *P'* as the intersection point between the line *PO* and the frontier of the rectangle, and taken among the intersection points that point *P'* which is the closest to *P*;
The Extension *2D*-Distance will be:

$$\rho((x_0, y_0), AMBM) = d(\text{point } P, \text{rectangle } AMBN) = |PO| - |P'O| = \pm\ |PP'|$$

where $|PO|$ means the classical *2D*-distance between the point *P* and *O*, and similarly for $|P'O|$ and $|PP'|$.

The midpoint (or center of symmetry) *O* has the coordinates $O(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2})$.

Since the *x*-coordinate of point *P'* is a_1 because *P'* lies on the rectangle's edge *AM*, one gets the *y*-coordinate of point *P'* by a simple substitution of $x_{P'} = a_1$ into the above equality:

$$y_{P'} = y_0 + \frac{a_2 + b_2 - 2y_0}{a_1 + b_1 - 2x_0}(a_1 - x_0).$$

Therefore *P'* has the coordinates $P'(\ x_{P'} = a_1,\ y_{P'} = y_0 + \frac{a_2 + b_2 - 2y_0}{a_1 + b_1 - 2x_0}(a_1 - x_0)\ )$.

The distance $d(P,O) = |PO| = \sqrt{(x_0 - \frac{a_1 + b_1}{2})^2 + (y_0 - \frac{a_2 + b_2}{2})^2}$

while the distance

$$d(P',O) = |P'O| = \sqrt{(a_1 - \frac{a_1 + b_1}{2})^2 + (y_{P'} - \frac{a_2 + b_2}{2})^2} = \sqrt{(\frac{a_1 - b_1}{2})^2 + (y_{P'} - \frac{a_2 + b_2}{2})^2}$$

Also, the distance $d(P,P') = |PP'| = \sqrt{(a_1 - x_0)^2 + (y_{P'} - y_0)^2}$.

Whence the Extension *2D*-Distance formula:

$$\rho((x_0, y_0), AMBM) = d(\ P(x_0,y_0),\ A(a_1,a_2)MB(b_1,b_2)N\ ) = |PO| - |P'O|$$

$$= \sqrt{(x_0 - \frac{a_1 + b_1}{2})^2 + (y_0 - \frac{a_2 + b_2}{2})^2} - \sqrt{(\frac{a_1 - b_1}{2})^2 + (y_{P'} - \frac{a_2 + b_2}{2})^2}$$

$$= \pm\ |PP'|$$

$$= \pm \sqrt{(a_1 - x_0)^2 + (y_{P'} - y_0)^2}$$

where $y_{P'} = y_0 + \frac{a_2 + b_2 - 2y_0}{a_1 + b_1 - 2x_0}(a_1 - x_0)$.

4.3.4 Extension Distance in *n-D*.

We generalized in the track of Cai Wen's idea the extension *1D*-set to an extension *n-D*-set, and defined the **Extension *n-D*-Distance** between a point $P(x_1, x_2, \ldots, x_n)$ and the *n-D*-set ***S*** as $\rho(\ (x_1, x_2, \ldots, x_n),\ \boldsymbol{S}\ )$ on the linear direction determined by the point *P* and the optimal point *O* (the line *PO)* in the following way:

a) $\rho(\ (x_1, x_2, \ldots, x_n),\ S\ )$ = the *negative distance* between *P* and the set frontier, if *P* is inside the set *S*;
b) $\rho(\ (x_1, x_2, \ldots, x_n),\ S\ ) = 0$, if *P* lies on the frontier of the set *S*;
c) $\rho(\ (x_1, x_2, \ldots, x_n),\ S\ )$ = the *positive distance* between *P* and the set frontier, if *P* is outside the set.

We got the following **properties**:

a) It is obvious from the above definition of the extension *n-D*-distance between a point *P* in the universe of discourse and the extension *n-D*-set *S* that:
    i) Point $P(x_1, x_2, \ldots, x_n) \in \mathrm{Int}(S)$ iff $\rho(\ (x_1, x_2, \ldots, x_n),\ S\ ) < 0$;
    ii) Point $P(x_1, x_2, \ldots, x_n) \in \mathrm{Fr}(S)$ iff $\rho(\ (x_1, x_2, \ldots, x_n),\ S\ ) = 0$;
    iii) Point $P(x_1, x_2, \ldots, x_n) \notin S$ iff $\rho(\ (x_1, x_2, \ldots, x_n),\ S\ ) > 0$.
b) Let S_1 and S_2 be two extension sets, in the universe of discourse *U,* such that they have no common end points, and $S_1 \subset S_2$. We assume they have the same optimal

points $O_1 \equiv O_2 \equiv O$ located in their center of symmetry. Then for any point $P(x_1, x_2, \ldots, x_n) \in U$ one has:

$$\rho((x_1, x_2, \ldots, x_n), S_1) \geq \rho((x_1, x_2, \ldots, x_n), S_2).$$

Then we proceed to the generalization of the dependent function from *1D*-space to *n-D*-space Dependent Function, using the previous notations.

The **Dependent *n-D*-Function** formula is:

$$K_{nD}(x_1, x_2, \ldots, x_n) = \frac{\rho((x_1, x_2, \ldots, x_n), S_2)}{\rho((x_1, x_2, \ldots, x_n), S_2) - \rho((x_1, x_2, \ldots, x_n), S_1)}$$

5. Case Study

Y Group is one of the world's largest menswear manufacturer group, with a production capacity of 80 million clothing items per year, includes shirts, suits, trousers, jackets, leisure coats, knitted items, and ties. Y has implemented world-class modern production lines and high-end equipment from nations including Germany, the United States, and Japan. Y group's production and supply lines are using state-of-the-art comprehensive computer-operated technologies to enhance the clothing production process.

After thirty years of development, The Group has many subsidiaries, includes factories, sales companies, foreign trade corporation and logistics companies et al. Y has forged a strong vertically integrated clothing chain, integrating the upstream components of textile and fabric production, midstream component of garment creation, and the downstream components of marketing and sales.

Back to 5 years ago, the average annual sales increase of 10%, costs remained nearly unchanged at the previous year's level. However, the average profit of the Group increased only 1.21%, not significantly improved. To find this reason, our data analysis team worked together with the Group's financial sector.

The profit of Y group is the sum of its subsidiaries, denote as

$$P_g = \sum_{j=1}^{n} P_j = P_f + P_s + P_l + P_t + \ldots\ldots, \quad P_g = I_{income} - C_{\cos t}$$

Accordingly, we made a basic-element analysis on profit related attributes, as shown in fig. 5:

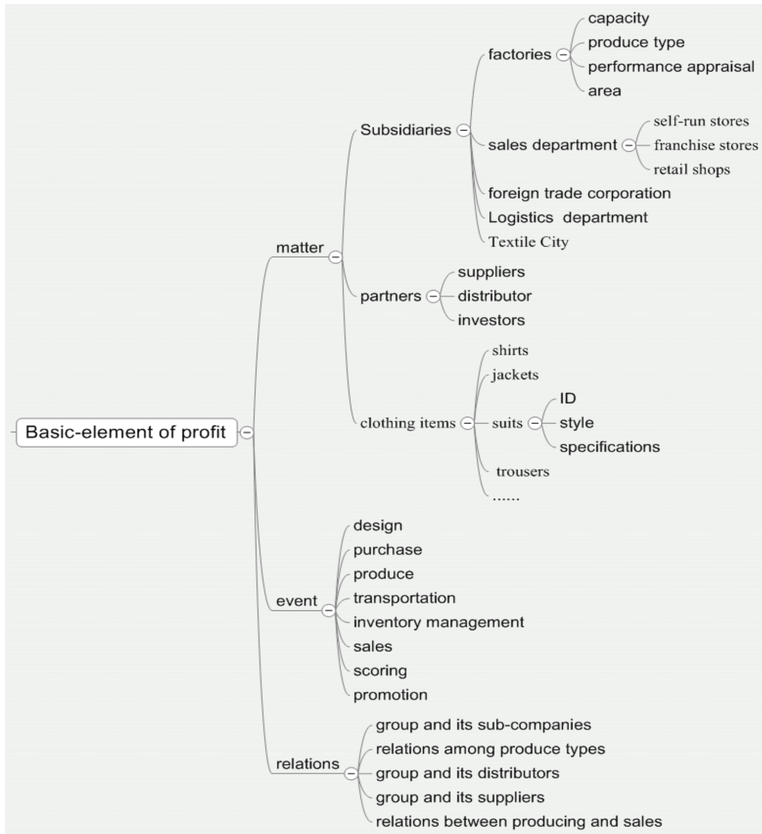

Figure 5. The basic-element of profit analysis

According to the basic-element analysis, we determine the scope of the collection of big data as following:

[1].Total sales, inventory, profit and cost data of the Group for the past five years;

[2].Historical production records of the production plant, raw material records, employee payroll (removal of sensitive personal information), production schedules, storage records et al.

[3].Historical distribution data in logistics department, sales returns data, and loss of productions in self-run stores;

[4].Retail sales records in the stores and shops, records of group purchase, discounts records and inventory balances;

We then integrate data in data warehouse, and calculate a lot of information such as sort of best-selling products, the highest inventory of suits, inventory turnover days, inventory days of supply and other indicators available.

Through analysis of various types of data in the Group's departments, we found that:

[1].The main source of profit of the garment sector is self-store sales and group purchase business confirmed by Group Financial center;

[2].The number of inventory days of supply (similar with Inventory Turnover) is worse. In some areas it is up to 428 days, which means that according to the current average sales, inventory of goods available for sale will be 428 days.

[3].In Production Scheduling small orders are temporarily postponed is as high as 31%, and those into priority processing sequences are mostly OEM orders;

[4].In production plant, the timely delivery rates of group purchase orders have continued to decline, far below that of the OEM production. OEM orders are usually in large quantities and the proportion in the profits of garment plants is also growing.

Based on Extension set theory, we further analyzed the domain and the associated rules. After the analysis of the data, we found that:

[1].Profit is a main KPI (key performance indicator) for the appraisal of subordinate units of the group, in some garment plants annual profit targets are completed more than 95%, the average annual profit growth is of 3.87%, one of the best performance units.

[2].in some sales companies, sales increase at an average rate of 9.01%, while the inventory grows at a rate of 14.2%; and the profit growth rate is only 0.53%.

Based on basic-element theory of Extenics, we collect attributes information related to garment plants:

[1].The produce types are divided into three categories:

- Make-to-stock production, for the production of inventory for all kinds of self-run stores sales;
- Make-to-Order Production, especially for group purchase;
- OEM (Original Equipment Manufacturer). The plants manufacture clothing that are purchased by another international company and retailed under that purchasing company's brand name.

[2].The OEM processing fee is averagely 2-4 Yuan higher than the make-to-order processing fee.

[3].By deep inquiry to managers, we found that in the production plant, the processing priority rules are: high-profit, large quantities of orders will be priority processing, and OEM orders are most meet these two conditions.

Integrating the knowledge mined from the big data and information basic-elements, we draw a knowledge logic chain as follows:

Profit is the primary key performance indicators of production plants, therefore

-> Managers give the top priority of processing the high-profit orders, therefore

-> OEM orders are latently with high profits, therefore

-> processing OEM orders by priority, therefore

-> OEM orders are large and production capacity is limited, so Make-to-Order productions are postponed.

It's reasonable to process high profits OEM orders from the perspective of the garment plants, but it is not understandable from the group's perspective, OEM processing fees is so small that it is almost negligible compared to the profit of a group buy (<= 5%), to earn $4, we lost $80. one reason to cause this "penny wise, pound foolish" phenomenon is the improper set up of profit as KPI for garment plants.

By big data analysis and extension model, we selected a new KPI called ratio of order delivery timely to substitute the profit. the production capacity has been fully utilized in garment plants, timely delivery rate of group buy orders is greatly improved, and OEM orders delivery rate is almost keep the same, while the Group's overall profit improved a lot.

Last year, Y group's total assets are valued at 30 billion RMB. its annual textile and garment sales alone amount to 10-billion RMB. Y was ranked number 113 in China's list of its top 500 manufacturers, and also is the only clothing company that China has recognized as one of its "Advanced Manufacturers."

6. Summary

The paper presents a framework of extension collaborative innovation model with concrete processes and a case study. The model integrates Extenics, data mining and knowledge management, and develops a framework for managing innovation and support team work. By collecting knowledge or information from multi-resources among all departments in enterprises, this platform can build knowledge tree from every employee in varies forms. Knowledge or information related to the problem can find relations by human-computer interaction method. It also considers the new technology and the new innovation method based on Extenics Theory. This particular method combines qualitative analysis which would take advantage of personal intelligence after formalized expression of target innovation problems and quantitative analysis which follows a systematic flow based on accumulated knowledge or innovation patterns. It helps to solve contradictory problems according to the extensibility of basic-element and will be applied in the innovation of management beyond data technology such as data mining and intelligent knowledge management.

Extenics (or Extension theory) can serve as the starting point of a generative approach for collaborative innovation, because it focuses on solving non-compatible problems by formularized methods includes basic-element theory, extension set theory and extension logic (Yang and Cai, 2007). The main features of using Extenics for collaborative innovation can be outlined as follows:

- Extenics provides a system structure for information collection and processing, which can be used as the basis for generation of effective solutions. Due to the completeness of systematic thinking it endorses, Extenics offers an opportunity of capturing important aspects needed for creative and innovative problem solving, which are likely to be ignored by conventional problem solving, or problem solving "by chance."
- In addition to generating possible solutions, Extenics also offers an effective way of evaluate solutions, so that any solutions failed the evaluation will be eliminated from further consideration. This can prevent computational explosion.
- The Extenics-based approach for innovation is a human-computer interactive process: Computers can conduct scalable data storage and mining by making use of various algorithms, far superior than human labors can handle. Nevertheless, the real world is extremely complex, and only humans can capture the dynamics of the real world beyond any algorithms can handle.

From these general features, here we provide an overview on what Extenics can offer for collaborative innovation. Firstly, any innovation must have a goal. In order to achieve this goal, we must acquire sufficient input to generate ideas, from which innovation process can take place in a systematic way, instead of by chance. We put forward a data collection theory in Section 3. Secondly, we present a combined model to process big data, one way is build basic-element base by human, another way is build knowledge base by computer, extension transformation connect them support to generate possible solutions in all directions by a formalized method. Lastly, we extend the dependent function to score multi-attribute innovation solutions in the context of big data.

In the future, we will further test the *n-D* dependent function and compared with other method. Moreover, the basic-element base and knowledge collaborative method need to be integrated with agent based system and enhance the extension innovation model. Since the significant importance of big data, deep research about combination of above methods with web information technology, extension data mining and intelligent knowledge management would be further explored. How to update basic-element base automatically and simulate the knowledge innovation process by intelligent agent is a challenge problem. We define it as intelligent extension innovation in the era of big data.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (#71271191, # 71110107026, #71071151 and #70871111), Zhejiang Philosophy and Social Science Project (#11JCSH03YB), Zhejiang Provincial Natural Science Foundation of China (#Y1110960), Scientific project of Educational Commission of Zhejiang Province (#Y201122111) and Ningbo Soft Science Research Program (#2012A10049, #2011F10008), Zhejiang Federation of Humanities and Social Sciences Circles Program (#2012N122, #2012N123), China.

References

[1] Birkinshaw, J., Hamel, G., & Mol, M. J. (2008), Management Innovation. *Academy of Management Review*, 33(4) 825-845.

[2] Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. McKinsey Quarterly, 56, 1-14

[3] Cai, W. (1983). Extension Set and Non-Compatible Problems. *Advances in Applied Mathematics and Mechanics in China*. Peking: International Academic Publishers,1-21.

[4] Cantisani, A. (2006).Technological innovation processes revisited. *Technovation*, 26, 1294-1301.

[5] Chae, B., Olson, D. L. (2013), Business Analytics for Supply Chain: A Dynamic-Capabilities Framework, *International Journal of Information Technology & Decision Making*, Vol. 12, No. 01, pp. 9-26

[6] Chisholm, M. (2012). Big data and the coming conceptual model revolution. Pdf file in immagic.com

[7] Ericsson, K. A., & Moxley, J. H. (2012). The expert performance approach and deliberate practice: Some potential implications for studying creative performance in organizations. *Handbook of organizational creativity*, 141-167.

[8] Han, J., Kamber, M. (2006), *Data Mining: Concepts and Techniques*, Second Edition, Morgan Kaufman Publishers, San Francisco, CA

[9] Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology*, 61, 569–598

[10] Hippel, E. and R. Katz (2002), Shifting Innovation to Users via Toolkits, *Management Science*, 48(7) 821-833.

[11] Hilbert, Martin (2013), Big Data for Development: From Information to Knowledge Societies (January 24, 2013). Available at SSRN: http://ssrn.com/abstract=2205145

[12] Hua, Z.; Yang, J., Coulibaly, S. and Zhang, B. (2006), Integration TRIZ with problem-solving tools: a literature review from 1995 to 2006. *International Journal of Business Innovation and Research,* vol. 1 (1-2), 111–128.

[13] Ji, Changqing, Yu Li, Wenming Qiu, Yingwei Jin, Yujie Xu, Uchechukwu Awada, Keqiu Li, and Wenyu Qu (2012), Big Data Processing: Big Challenges and Opportunities, *Journal of Interconnection Networks*, Vol. 13, No. 03n04

[14] Kim, J., Arnie Lund, Caroline Dombrowski (2013). Telling the story in big data, *Interactions*, Vol. 20 Iss. 3, 48-51

[15] Kou,G.

[16] Li, X. H. Qu, Z. Zhu et al (2009), A Systematic Information Collection Method for Business Intelligence, *International Conference on Electronic Commerce and Business Intelligence*, 116-119.

[17] Li, X., Y. Shi and L. Zhang (2010), From the information explosion to intelligent knowledge management, Beijing: Science Press.

[18] Li, X., Haolan Zhang, Zhengxiang Zhu, Zhongbiao Xiang, Zhengxin Chen, Yong Shi (2013), An Intelligent Transformation Knowledge Mining Method based on Extenics, *Journal of Internet Technology*, 14 (2): 315-325

[19] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, 1-137.

[20] Mumford, M. D., K. E. Medeiros, P. J. Partlow (2012), Innovation: Processes, Strategies, and Knowledge, *Journal of Creative Behavior*, 46(1), pp. 30–47

[21] Olson, D. and Y. Shi (2007), *Introduction to business data mining*, McGraw-Hill, International edition.

[22] Saaty, T.L. (2010), *Creative Thinking, Problem Solving & Decision Making*, RWS Publications, Pittsburgh, PA, 4th

[23] Shi, Y. and Li, X. (2007), Knowledge Management Platforms and Intelligent Knowledge beyond Data Mining, *Advances in Multiple Criteria Decision Making and Human Systems Management: Knowledge and Wisdom*, Y. Shi et al. (Eds.), IOS Press,2007, 272-288

[24] Smarandache, F. (2012), Generalizations of the Distance and Dependent Function in Extenics to 2D, 3D, and n-D, *Global Journal of Science Frontier Research*, Vol. 12, Issue 8, pp. 47-60

[25] Snijders, C., Matzat, U., & Reips, U.-D. (2012). 'Big Data': Big gaps of knowledge in the field of Internet science. *International Journal of Internet Science*, 7, 1-5. http://www.ijis.net/ijis7_1/ijis7_1_editorial.html

[26] Sorli, M. and D. Stokic (2011), Future Trends In Product/Process Innovation, *International Journal of Innovation and Technology Managemen*t, Vol. 8, No. 4, 577-599

[27] Van Horn, D., Olewnik, A., & Lewis, K. (2012). Design Analytics: Capturing, Understanding, and Meeting Customer Needs Using Big Data. In *ASME International Design Engineering Technical Conference & Computers and Information in Engineering Conference*.

[28] Vandeven, A. H. (1986), Central Problems in the Management of Innovation, *Management Science*, 32(5), 590-607.

[29] Yang, C., W. Cai (2013), *Extenics: Theory, Method and Application*, Science Press, Beijing.

[30] Yu, Po-Lung (1995), Habitual domains, Springer-Verlag, Kansas, USA

[31] Zeleny, M. (2012), High Technology and Barriers to Innovation: From Globalization to Relocalization, *International Journal of Information Technology & Decision Making*, 11 (02), 441-456

[32] Zhang, L., J. Li, Y. Shi and X. Liu (2009), Foundations of Intelligent Knowledge Management, *Human System Management*. 28 (4):145-161

[33] Zhou, Z., X. Li (2010), Research on Extenics-based innovation model construction and application of enterprise independent innovation, *Studies In Science of Science*, 28 (5): 769-776