

From TCM database to TCM informatics

Wang yuan

Shanghai Innovative Research Center-TCM
Shanghai Institute of Organic Chemistry, AC



**A TCM Database serviced to Medicine
Development and Research**

From TCM database to TCM informatics

**Some inspiration for Development of
Library and Information**

From TCM database to TCM informatics, It belongs to the Category of Traditional Chinese Medicine. But in today's presentation, we are going to discuss how to accommodate the impact followed by the development of Information technology and to explore new development idea and pathway in the field of library and information through the special fields of TCM. Although there are many different fields of science and technology, but I think the idea and method of library and information are common for the different fields.



TCM makes great contribution for promoting the increase gradually in number or quantity of population in China

TCM has rich accumulation of effective clinical case for thousands years. It is a great treasurehouse

Some medicines are invented based on TCM

- **Trichosanthin for AIDS**
- **Artemisia for malarial**
- **Agrimopholum for schistosome and cephalocyst**
- **indirubin for antitumor**

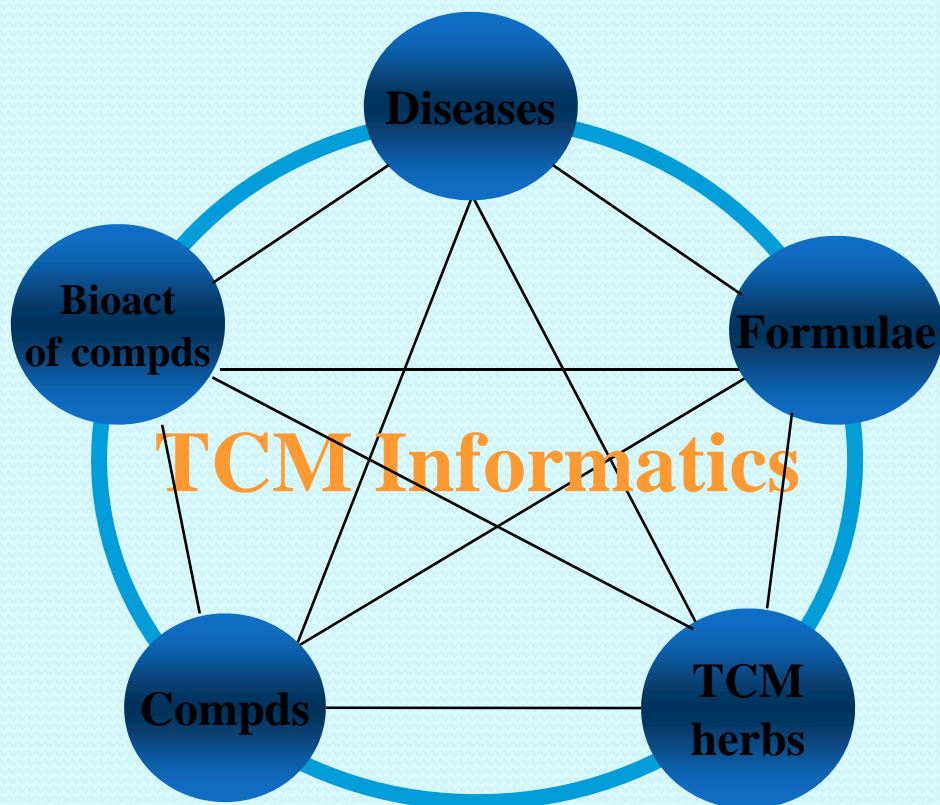
Necessity of TCM Modern Research

- **It is a complex mixture system**
- **We can't elucidate the mechanism of effectivity of TCM**
- **Big dosage, It is not convenience for oral**

Obstacle for TCM modernization

- Far from modern science
- No database
- Less cooperation
- Poor standardization
- Formulation type not accepted
- Complexion of mixture system

Crosslinked data structure



Diseases	1,400
Formulae	180,000
TCM herbs	9,000
Compds	30,000
Bioact of compds	120,000

→ Public services

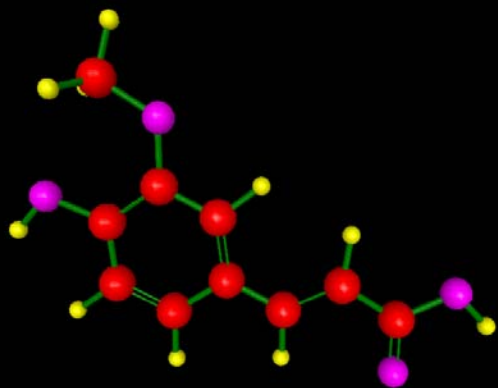
→ Consult services
for data analysis

→ R&D of
Drugs

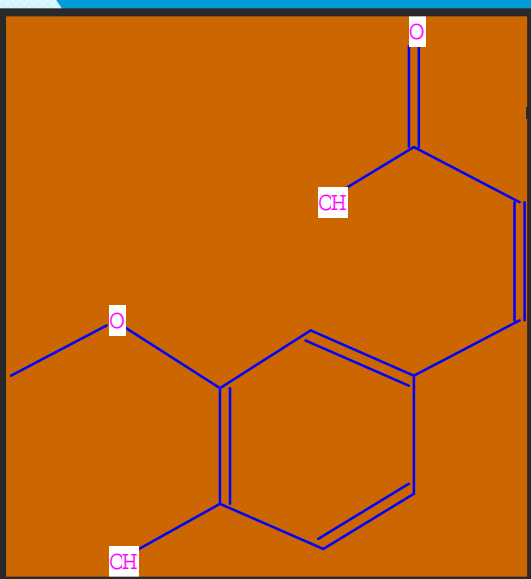
Distinguished Features

- Multidisciplinary system deal with chemistry, biology, medicine, pharmacy and information
- Integrating the ancient TCM experience with modern science and technology
- Rich data type
 - text
 - image of plant
 - topological graph of chemical structures
 - VRML of 3D structures
- Powerful retrieval function
- Crosslinking data structure
- Data mining-oriented data analysis

There are 2D and 3D chemical structures

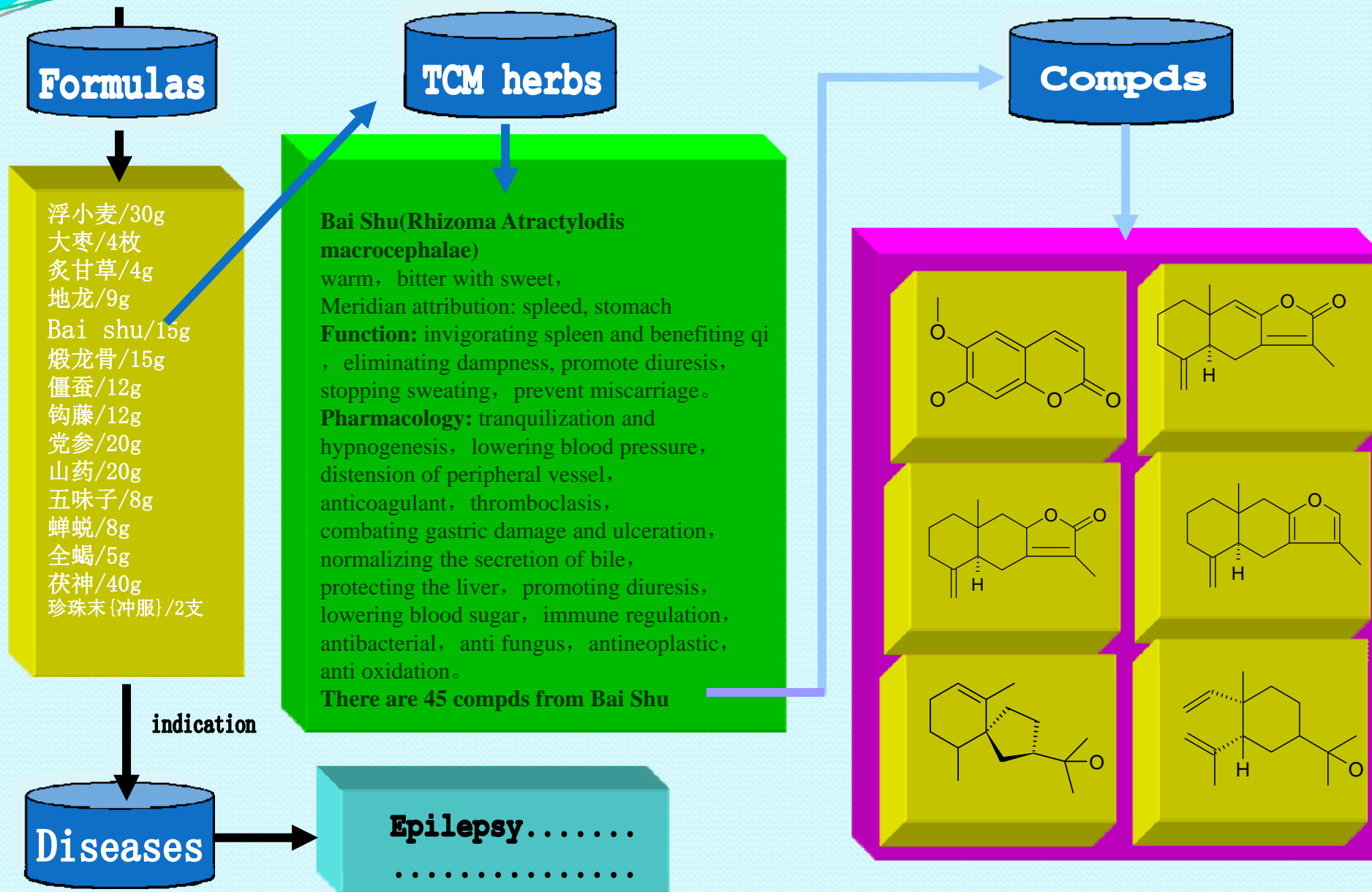


流水号	2
中药代码	5092
数据表属性	pharmacology
药理作用	inhibiting platelet aggregation
实验物质	
结构不确定化合物	
结构确定化合物	ferulic acid
实验动物	Human\$rabbits\$rats
给药和方法	in vitro and in vivo
药理描述	he aq. ext. of Angelica sinensis and its ingredient ferulic acid inhibited blood platelet aggregation and serotonin release by blood platelets of rat in vivo and in vitro. Thus, inhibition of platelet aggregation and 5HT release from platelets may be involved in the therapeutic effect of this ext. on cerebrovascular disorders.
URI	0513487019800321015006a

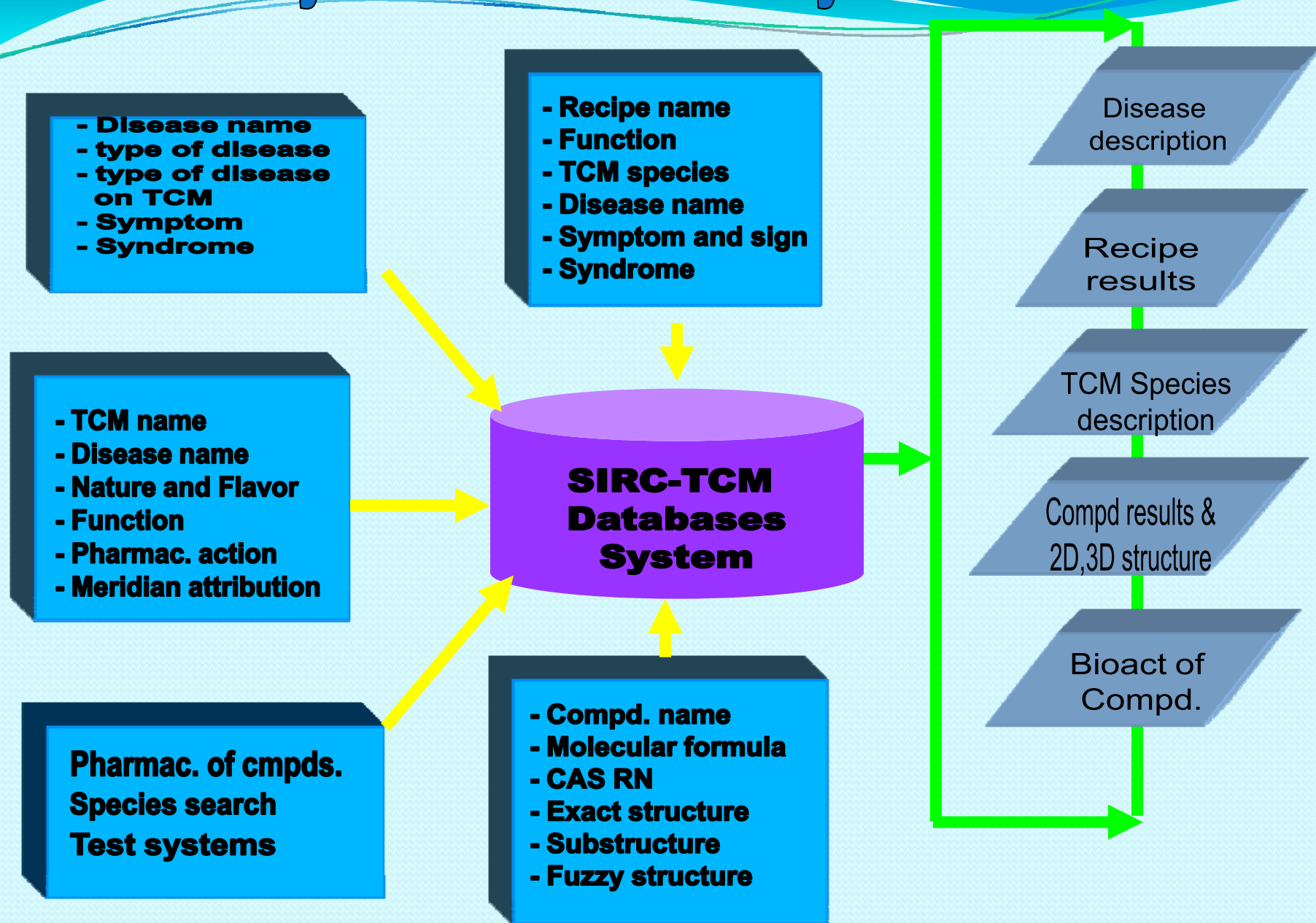


流水号	
化合物名	ferulic acid
CAS 号	1135-24-6
来源中药	Angelica sinensis (Min dictriect of Gansu)
部位	water-sol. fraction
百分含量	
性质	
生物作用	
生物作用描述	
URI	0513487019790529014009a

Integration of ancient TCM experience and theory with modern chemistry and medicine concepts



Search Ways of the Database System



The Strategy of Data Collection

- Collecting data selectively and guidedly
 - To collect recipes with clinical data from TCM journal since 1950
 - To collect compds isolated from TCM species
 - TCM species are different with natural plant
 - variety of plant or animal have to fit TCM species
 - part of plant as TCM species have to be considered
 - 枸杞子 (qou qi zi) fruit of *Lycium barbarum* L.
 - 地骨皮 (di gu pi) root bark of *Lycium barbarum* L.

Query

alzheimer

Recipes

278
Recipes

TCM species

280
TCM species

Cmpds from
TCM species

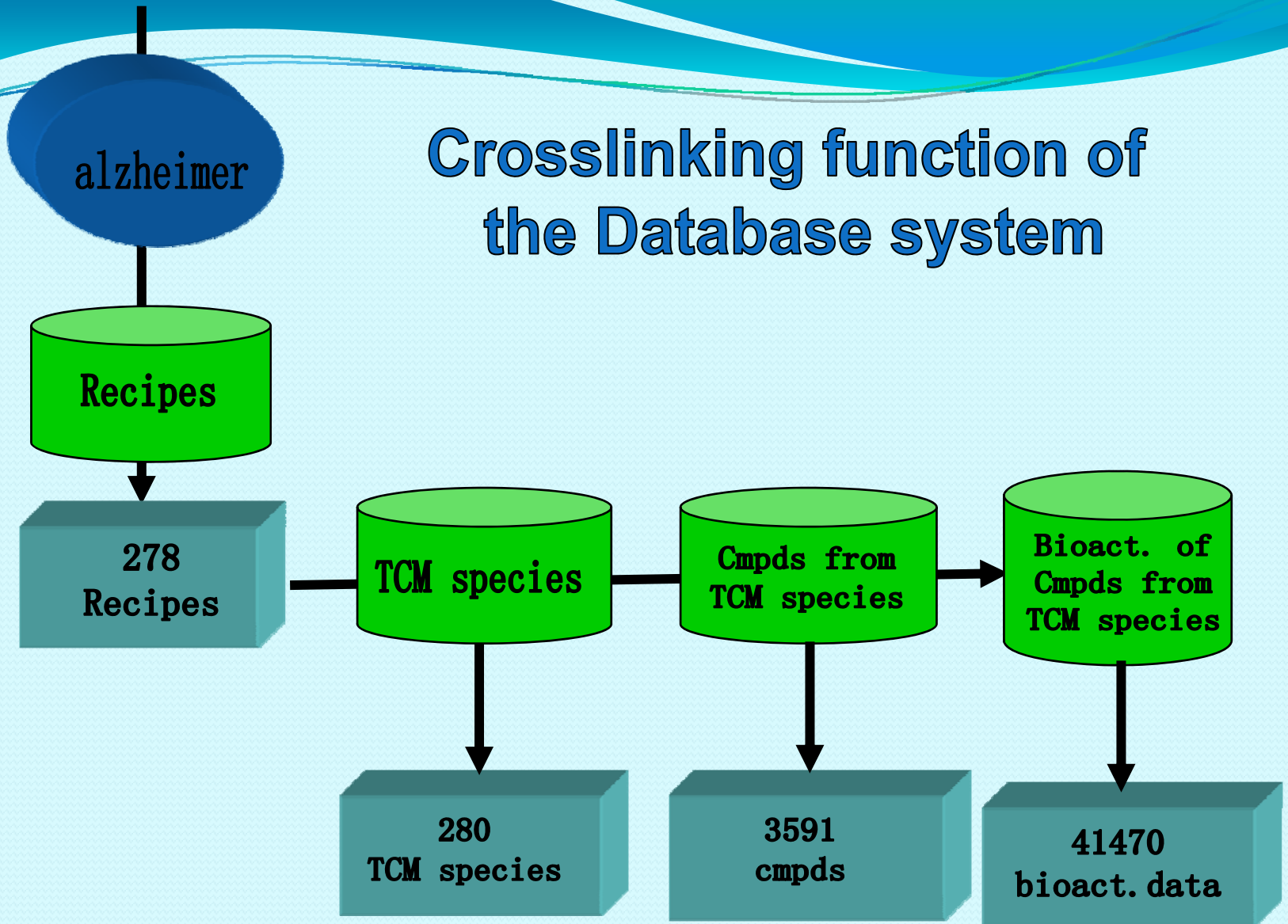
3591
cmpds

Bioact. of
Cmpds from
TCM species

41470
bioact. data

Crosslinking function of the Database system

Integrated results output



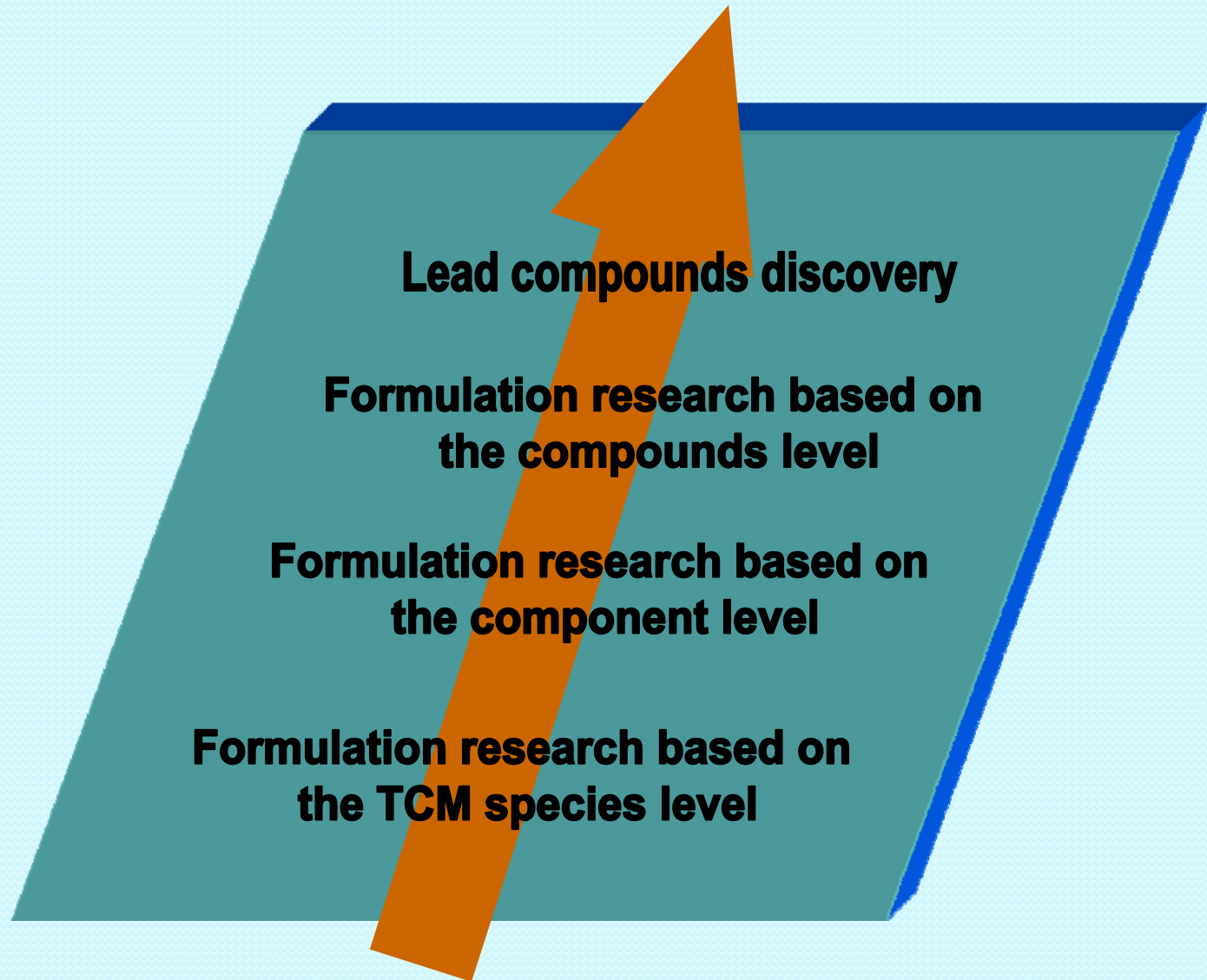
Advantage and disadvantage of the database system served only for retrieval

- Advantage
 - can get comprehensive data deal with various database
- Disadvantage
 - So much results are displayed in a crosslinking multi-database system
 - The relation between the data haven't been explored

TCM Informatics

- **Main purpose is to explore the data relationship between different databases in the TCM field.**
- **To support the R&D of medicines based on TCM**
- **To study data analysis methods based on TCM data**
- **To supplement new databases to the system according to needs of medicine research and development of science**

Four Level Services for R&D of Drugs based on the TCM

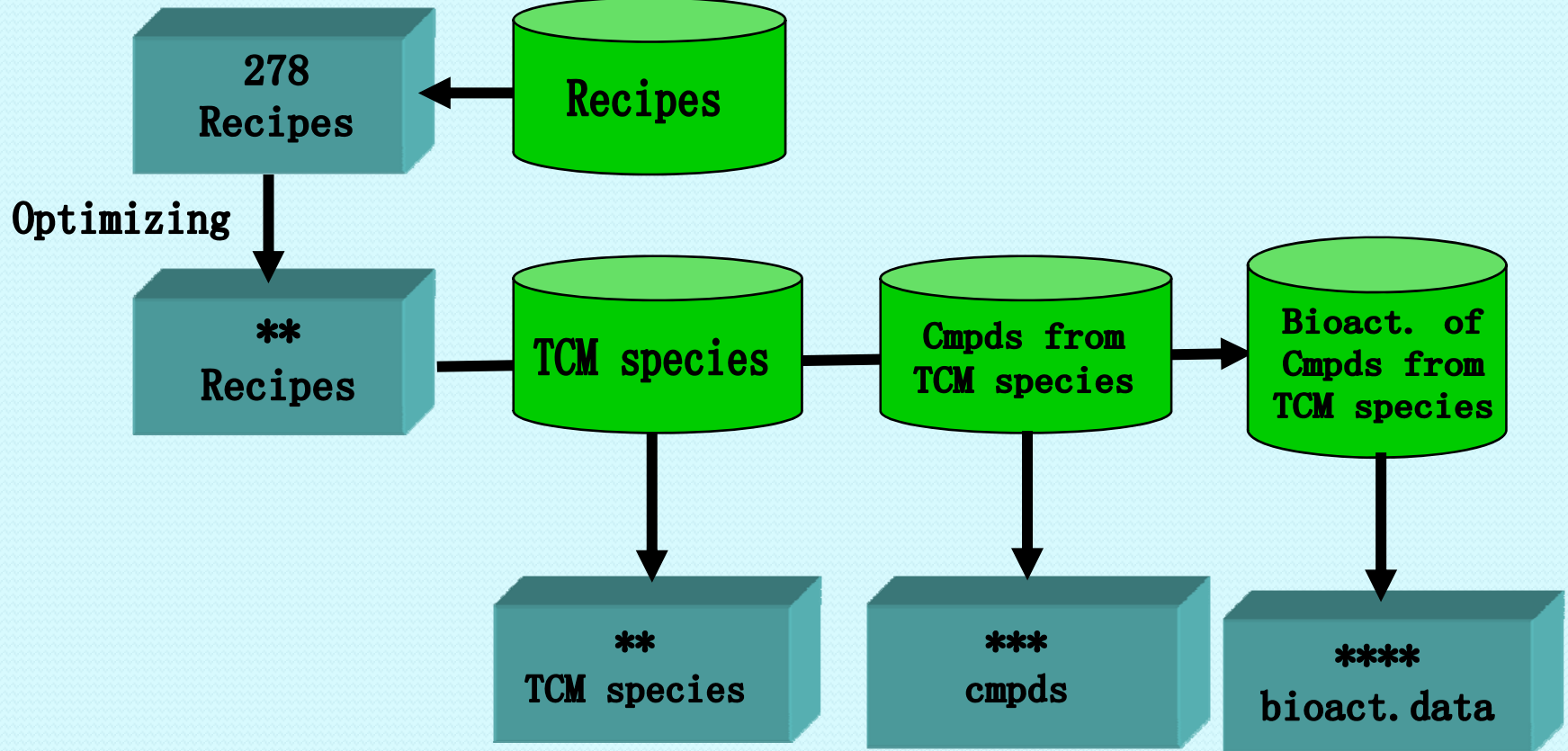


Function of TCM informatics

- Integrating and optimizing the search results
- Frequency statistics of TCM species used in the recipes
- Development research of modern application of classic recipes
- Computer-aided formulation design based on the database system
- Help to find the start and entry key of TCM research projects based on the database system
- Help to research the substance basis of TCM species and recipes efficacy



Optimization based on the coincidence of TCM species in the recipes



Optimized results output

Frequency statistics of TCM species used in the recipes

- **This is a basic function for many data analysis demand**
 - **To determine common used TCM species**
 - **To determine common used TCM species indicated in designated disease**

Frequency statistics of common used TCM species

TCM name	Frequency	TCM name	Frequency	TCM name	Frequency
Gan Cao	35959	Chai Hu	20604	Ze Xie	12438
Dang Gui	35268	Dang Shen	20419	Shu Di huang	12311
Fu Ling	32101	Xian Di huang	19058	Shan Yao	11798
Bai Shao	25514	Chen Pi	17278	Huang Lian	11362
Huang Qi	24463	Huang Qin	17114	Mai Dong	10977
Dan Shen	23283	Tao Ren	14264	Ban Xia	10714
Bai Shu	23070	Mu Dan Pi	14058	Jin Yin Hua	10487
Chuan Xiong	20997	Hong Hua	13135	Da Huang	10387
Chi Shao	20619	Gui Zhi	20419	Yu Jin	9996

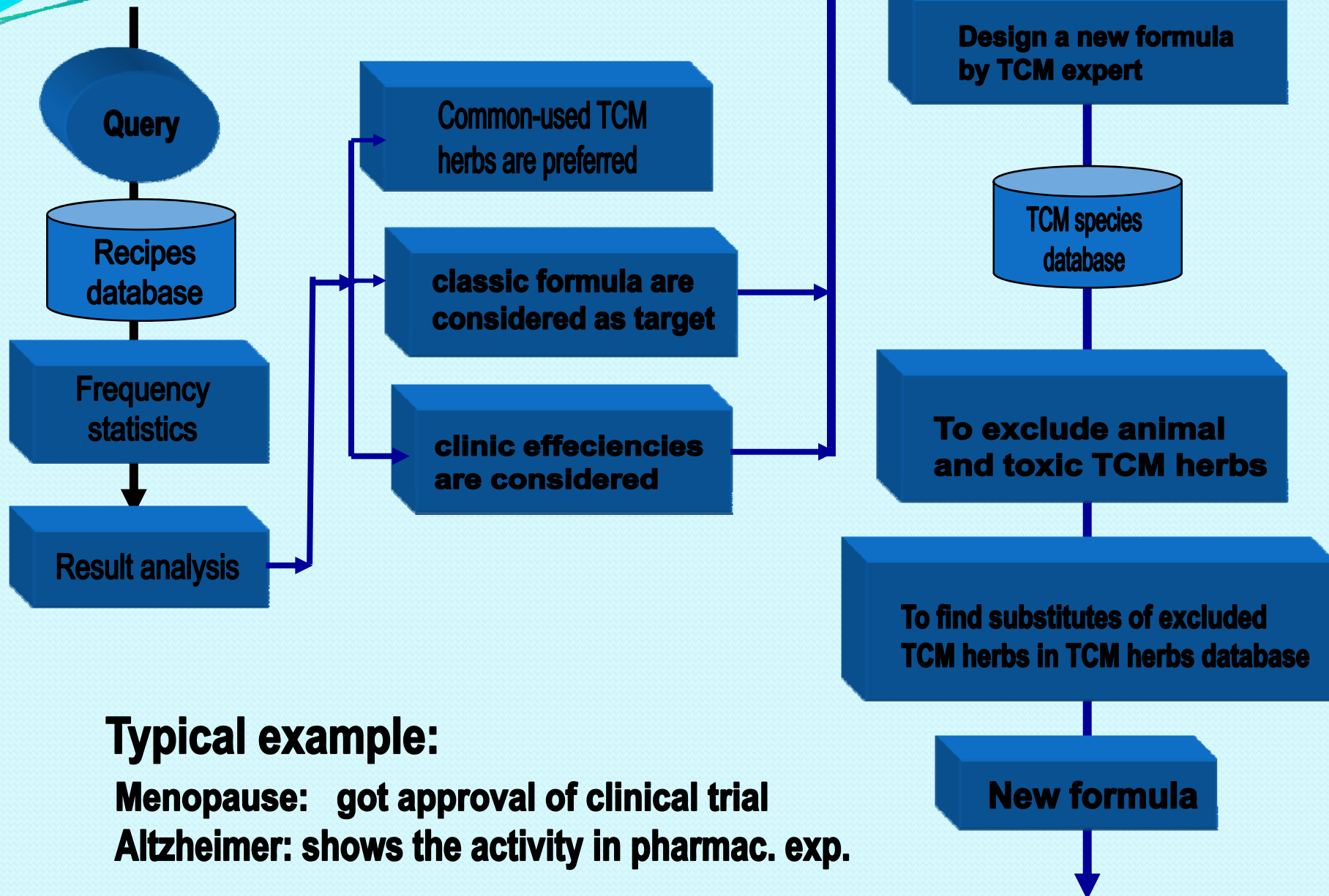
The statistic result in 134500 recipes

Frequency of TCM species used in Alzheimer's recipes

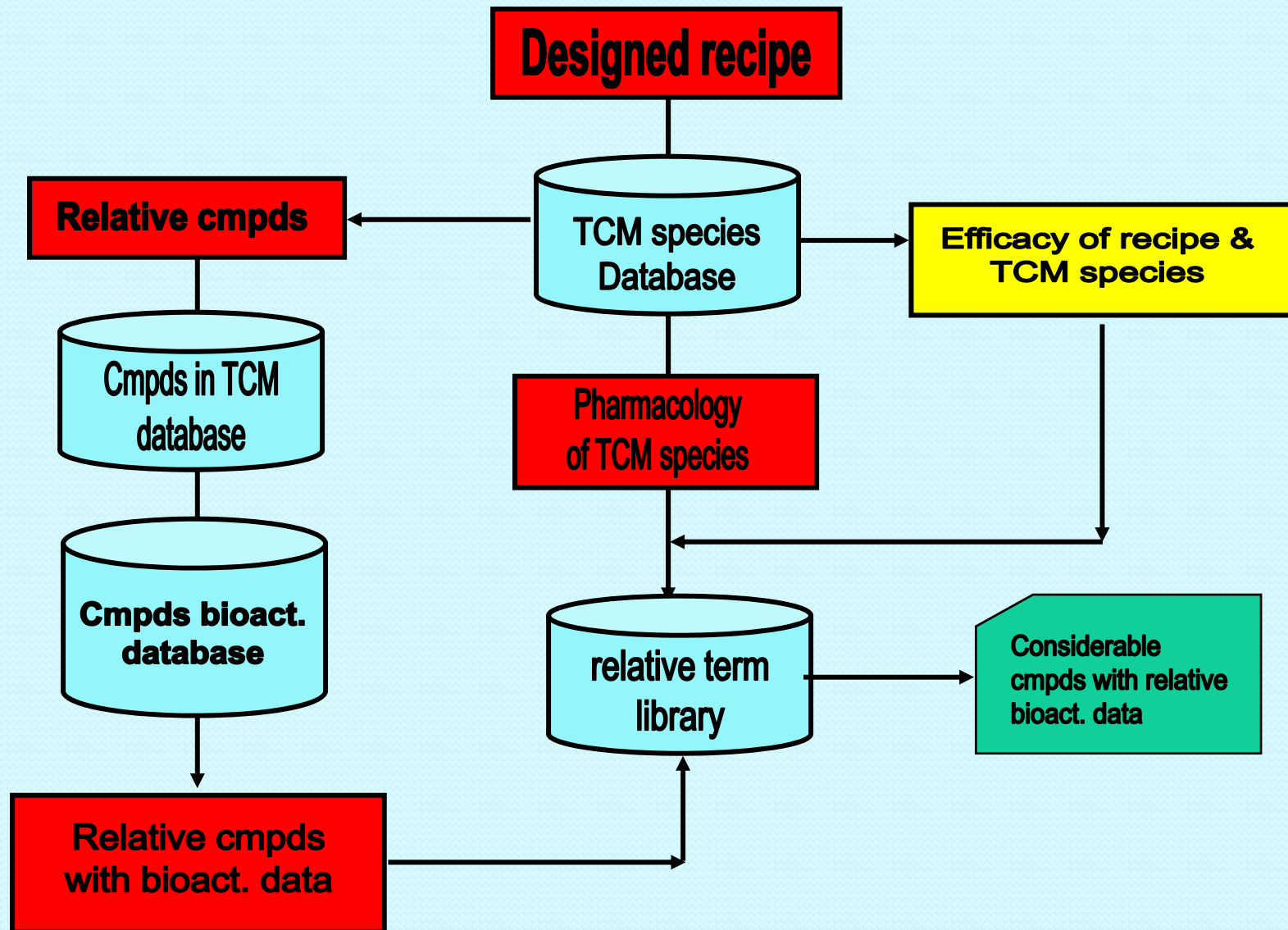
TCM herbs	Freq	TCM herbs	Freq	TCM herbs	Freq
Rhizoma Acori Tatarinowii	18	Tulip Agesneriana	6	Fructus Alpiniae Oxyphyllae	5
Radix Polygalae	13	Rhizoma Atractylodis Macrocephalae	6	Rhizoma Pinelliae	5
Radix Salviae Miltorrhizae	12	Polygonum Multiflorum Thunb	6	Peach kernel	5
Rhizoma Chuanxiong	10	Lumbricus	6	Fructus Ligustri Lucidi	5
Radix Astragali	10	Radix Angelicae Sinensis	5	root of Pilose Asiabell	5
Fructus Lycii	8	Cornus Officinalis Sieb. et Zucc.	5	Dioscorea Japonica Thunb	4
Radix Rehmanniae Preparata	8	Radix Paeoniae Alba	5	Concretio Silicea bambusae	4
fresh rhizome of Adhesive Rehmannia	7	root of Thomson kudzu vine	5	Abalone shell	4
Poria Cocos	7	Arisaema with bile	5	Semen Ziziphi Spinosae	4
Ginseng	6	all-grass of Longspur Epimedium	5	rhizome of Siberian Solomonseal.	4

We selected 5 TCM herbs in the above table to combine a recipe showing pharmacological activity

NewTCM Formulation

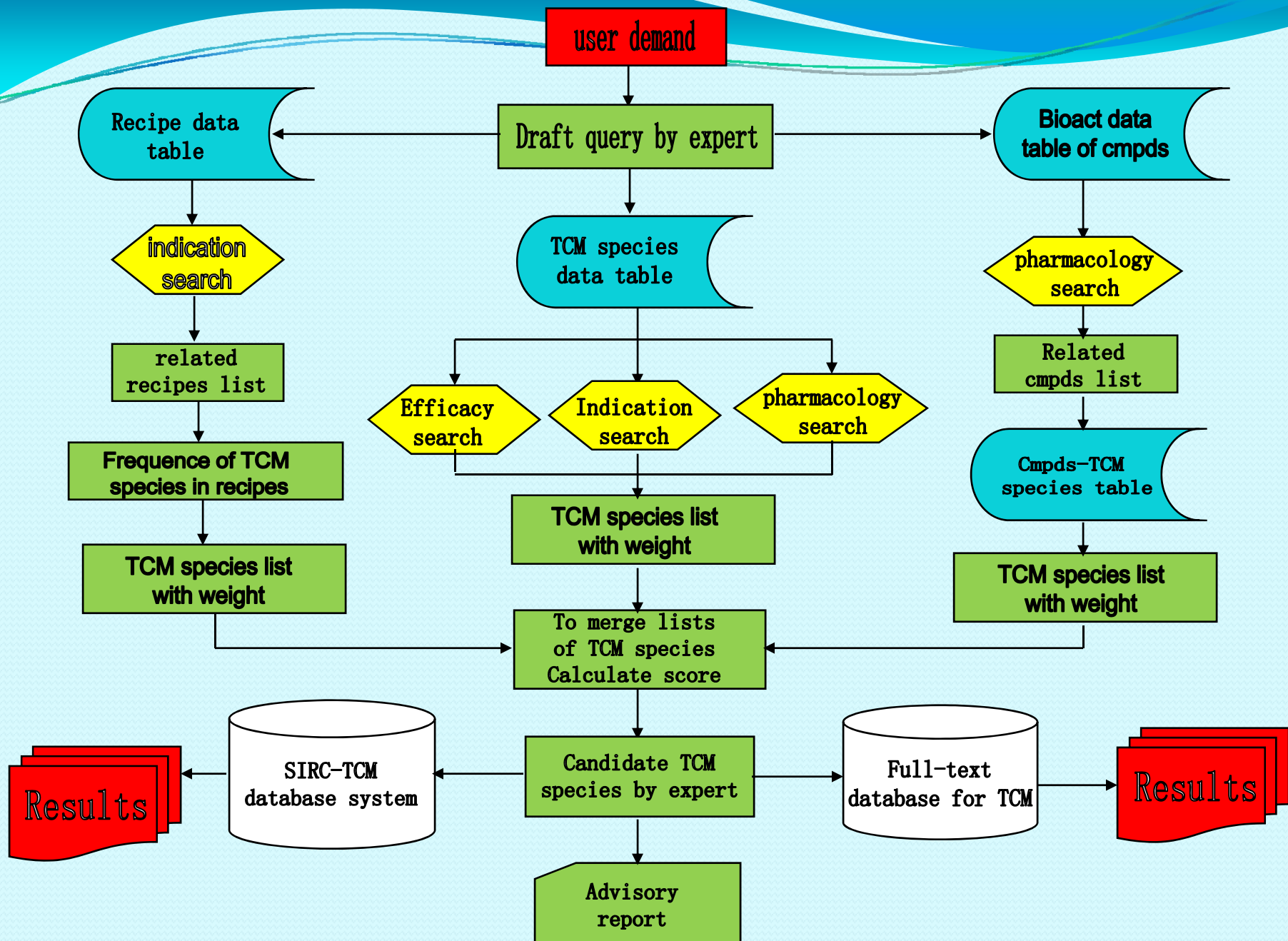


Formulation design in the effective component and compounds



Help to find the start and entry key of TCM research projects

- This method is developed based on demand provided by users
- More and more pharmacy co. pay attention to the TCM field, but they don't know to choice what kind of TCM species as research object
- We have developed a comprehensive analysis method to provide related TCM species according to the data of recipe, TCM species, compounds and bioactivity of compounds for selecting by TCM expert



Study on the substance bases of complex recipes

- Complex recipe is the essence of TCM
- The effectivity of TCM haven't be elucidated difficulty
- Substance basis of complex recipes are key projects of TCM modernization
- The compounds and their bioactivity in the TCM species are the basis of these research projects.
- A very important work is to create proximity relation between TCM terms and modern pharmacological terms

To create relative link between different kind of terms

hepatitis
hepatic cirrhosis
adiposis hepatica
hepatauxe

inhibiting activity of Alanine aminotransferase in serum
inhibiting activity of ALT in serum
increasing activity of microsome cytochrome P450
lipidperoxidation
inhibiting liver fibrosis
reinforcing uptake of amino acid by hepatocyte
reinforcing uptake of nucleotides by hepatocyte
promoting synthesis of ribonucleic acid of hepatocyte
neutrizing lipidperoxidation
increasing blood fluid in liver
preventing fatty degeneration of liver
preventing hepatocyte necrosis
improving protein metabolism

保肝=protection the liver

疏肝化癥=dispersing the depressed liver-qi and dissolving stasis

滋补肝肾=enriching and supplementing liver and kidney

滋阴柔肝=nourishing yin to nourish the liver

化癥疏肝=removing stasis to disperse the depressed liver-qi

疏肝通络=dispersing the depressed liver-qi and dredging the collateral

疏肝养血=dispersing the depressed liver-qi and nourishing blood

滋阴疏肝=nourishing yin and dispersing the depressed liver-qi

清肝解毒=clearing away liver and removing toxic substances

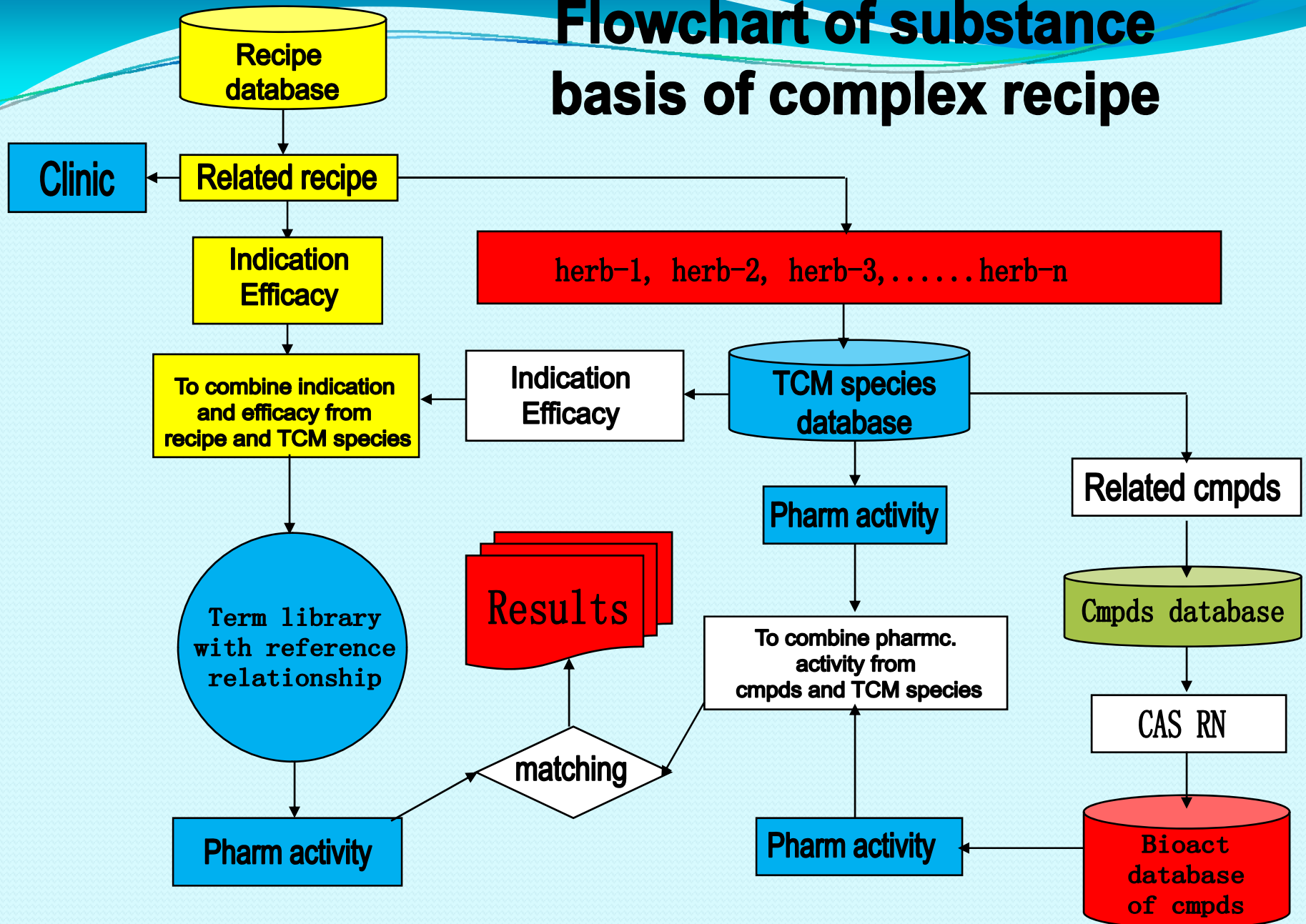
疏肝散结=dispersing the depressed liver-qi and dissipating the mass

养血养肝=enriching the blood to nourish the liver

柔肝软坚=nourishing the liver and softening hard masses

清泻肝胆=clearing and purging the liver and gallbladder

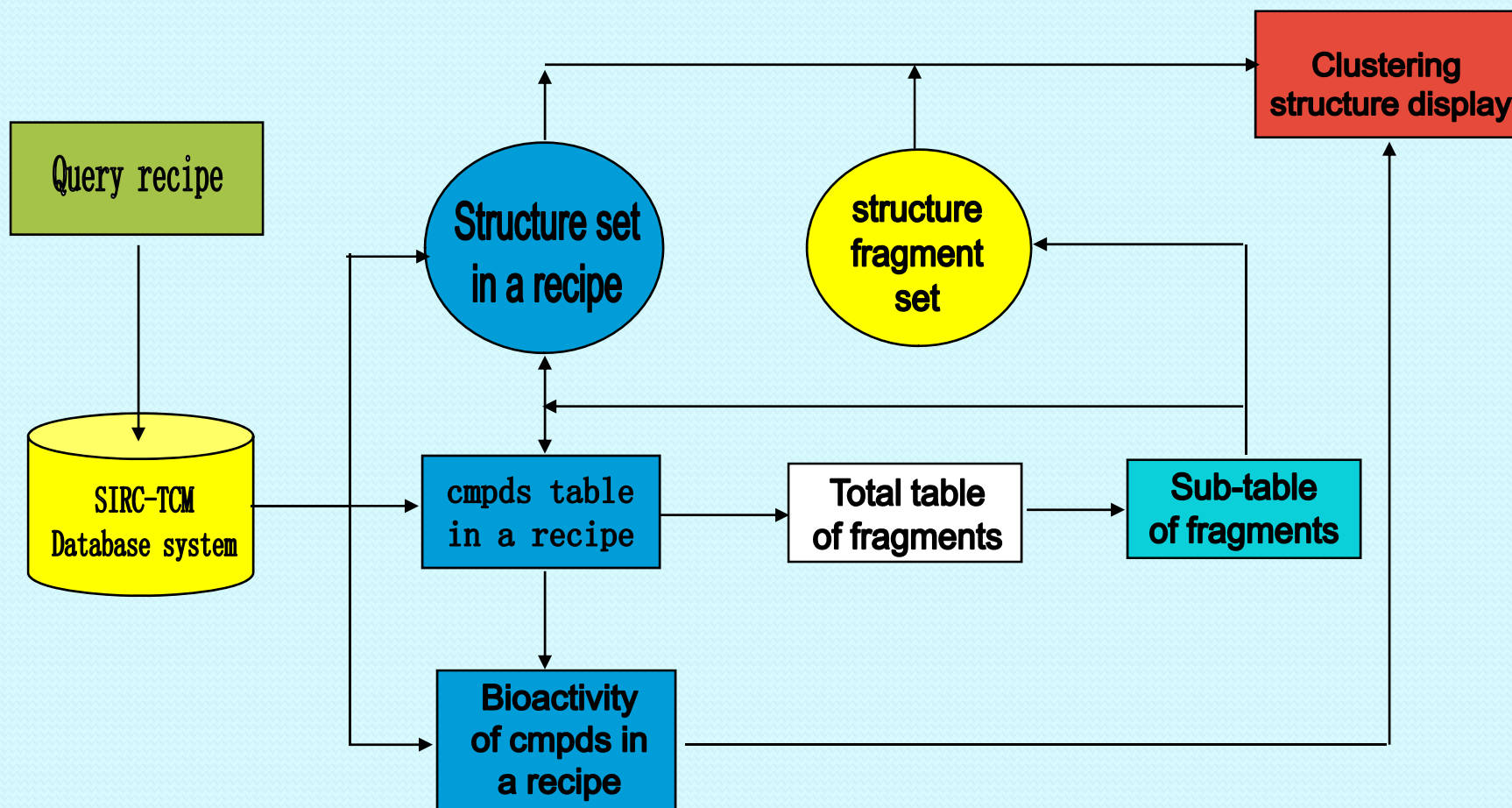
Flowchart of substance basis of complex recipe



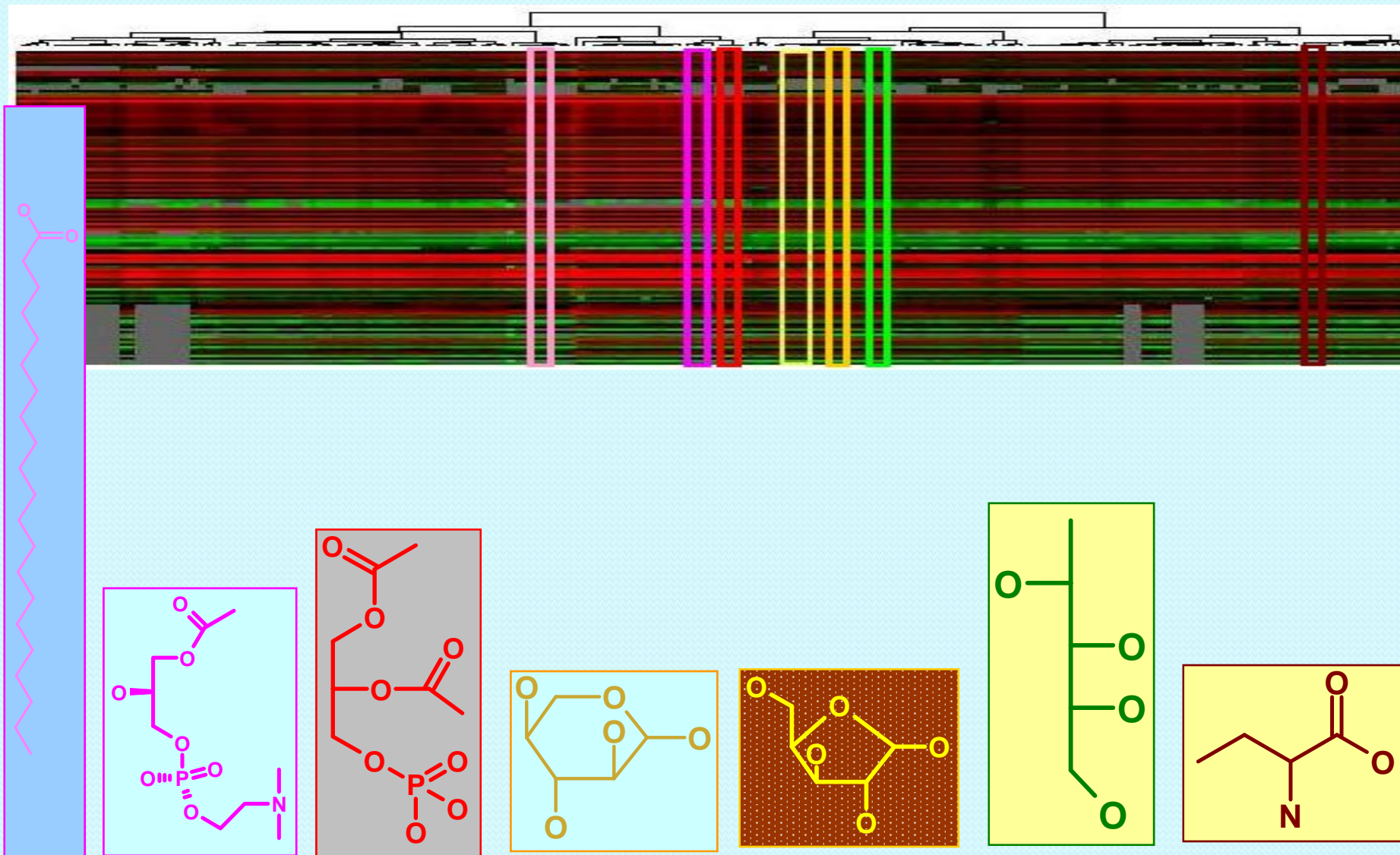
Clustering the chemical structures in a recipe

- **This is a demand by users**
- **They want to know the compound's type in a recipe and to study the relationship between structure and their bioactivity**

Flowchart of structures cluster in a recipe

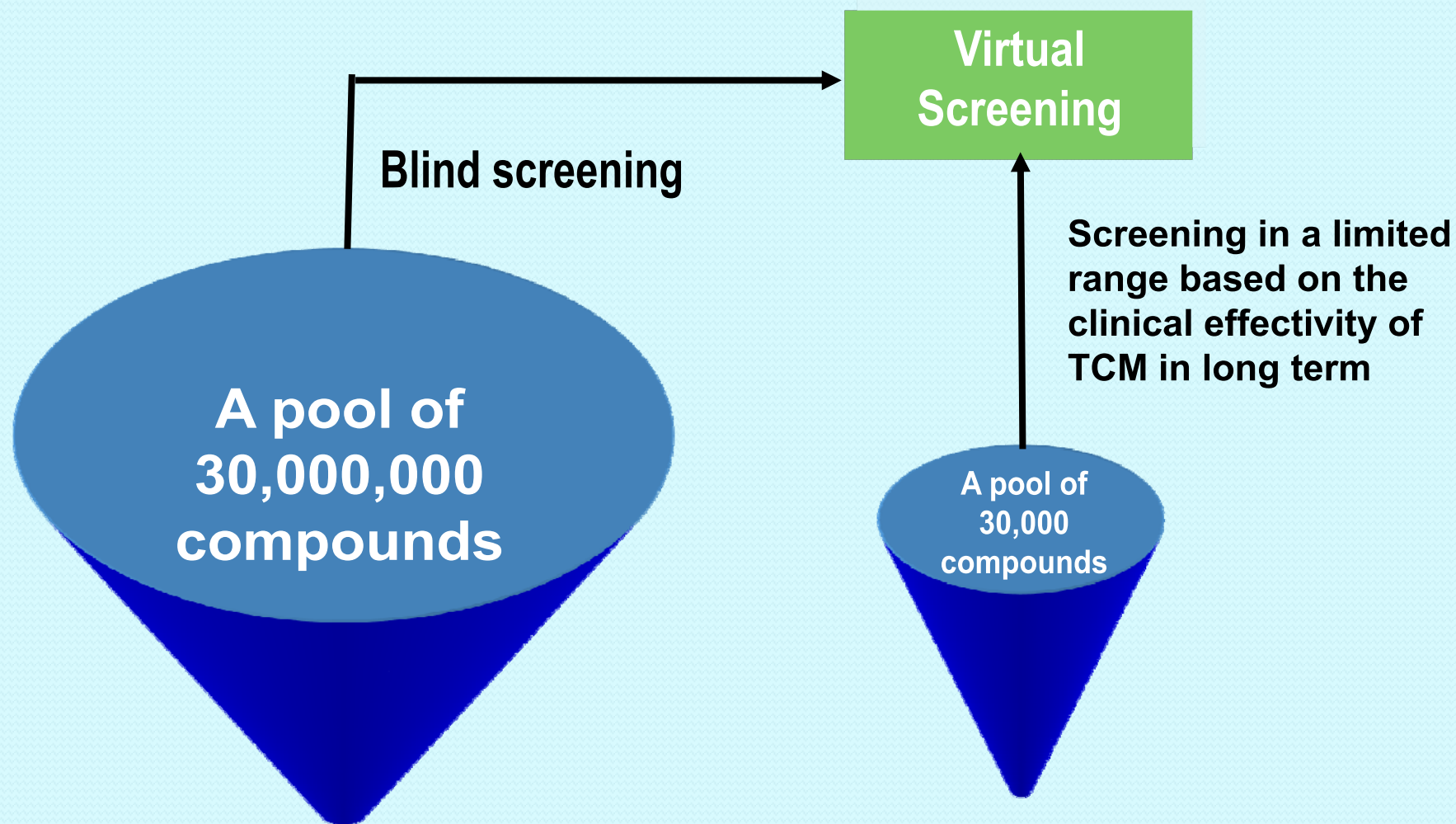


Cluster of core structures of Si Wu Tang



Role in the virtual screening of medicines

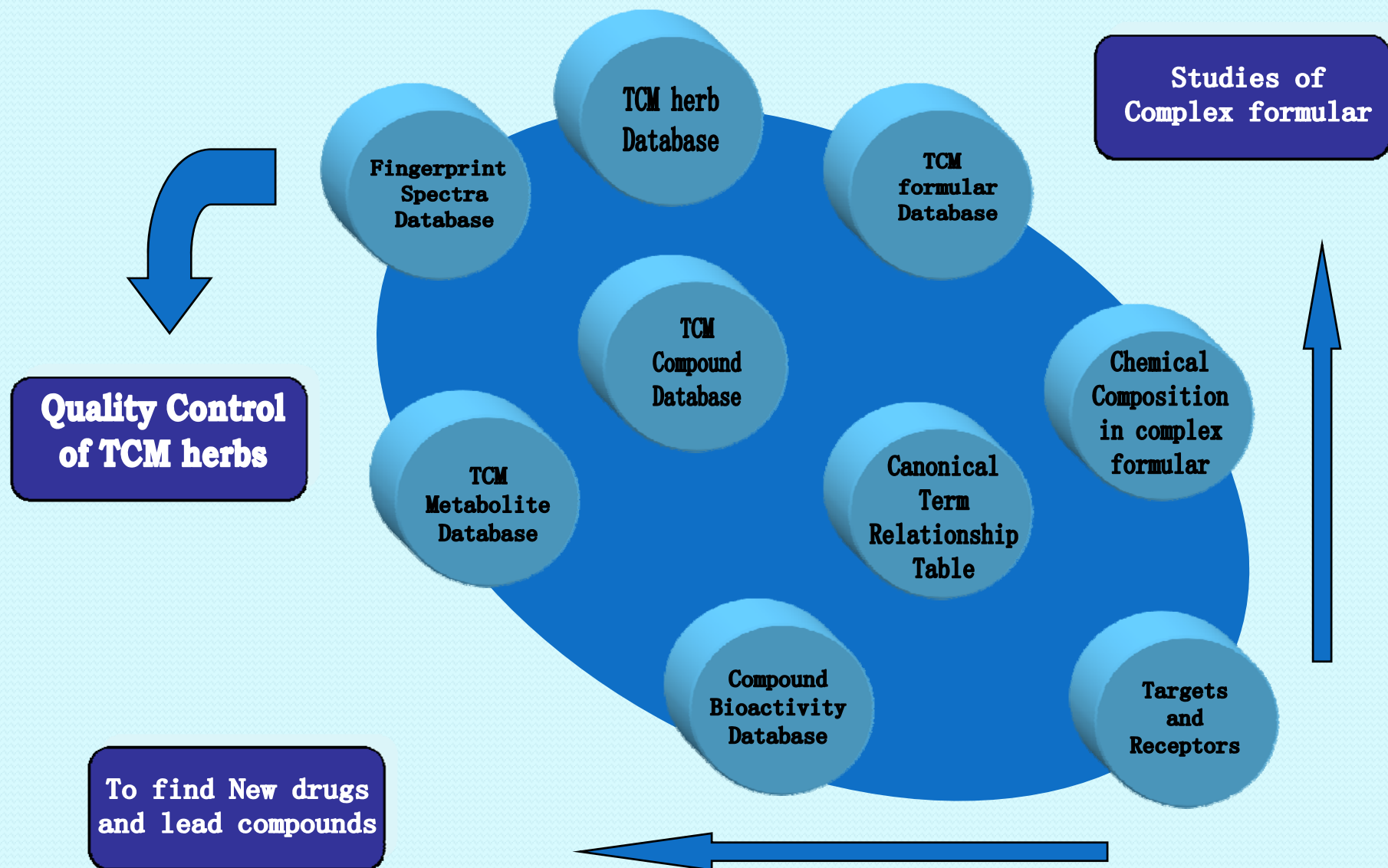
The effect in finding leading compds



The example of Innovation Services

- “Geng Shu An” granule for postmenopause
- “Nao Ling Qing” for Alzheimer
- The suggestion for R&D of designated function drugs
 - DSM、NESTLE、SK、NIVEA、AMWAY
- Kupio University used for screening of bioactivity compds
- Information services for SARS and bird flu

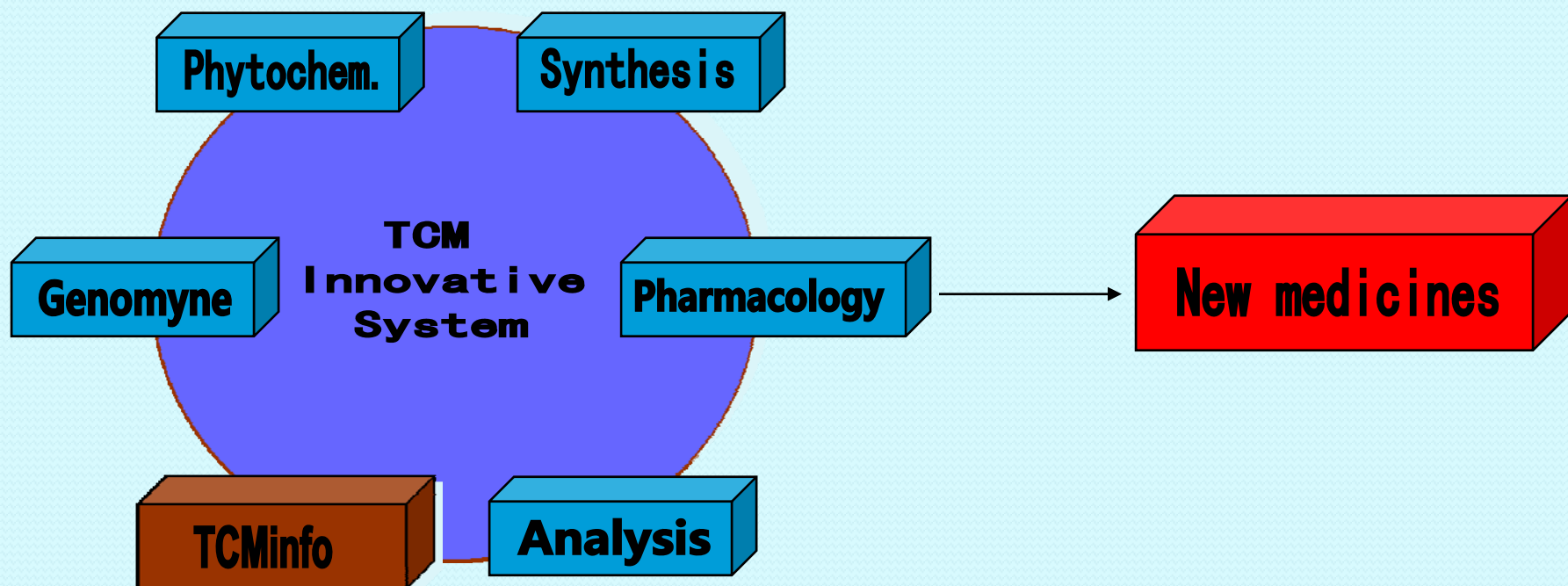
New Developments



Some inspiration for Development of Library and Information

- Is this work belong to Library or Information?
- Librarians and Informatists have made great contribution to information technology
 - Information theory have permanent value
 - Thesaurus
 - Punch card
- To emphasize the role of consults expert based on database
- It is the development trend of information field

Always remember that library and information are the component part of scientific research





Thanks for your attention

<http://www.tcm120.com>