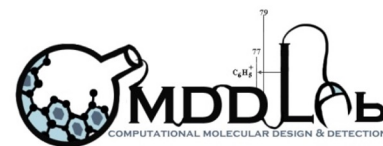




Computational *Molecular Design & Metabolomics*

Y. Jane Tseng

Graduate institute of Biomedical Electronics and Bioinformatics,
Department of Computer Science and Information Engineering,
School of Pharmacy,
National Taiwan University

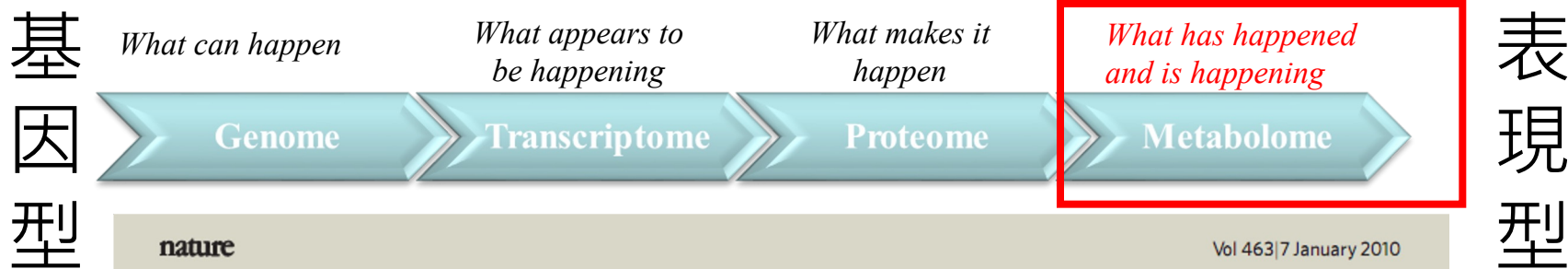


研究

*AI in drug discovery and Biomedical
advancement*

研究主題一：診斷

代謝體學



OPINION

nature publishing group



代謝體

未來十年會改善人類生活的14學科之一 (Nature's 2020 vision)

14學科 — 搜尋、個人化醫療、藥物開發、微生物學、精神健康...

14學科中之個人化醫療、藥物開發、微生物學、精神健康認為自己的領域要突破需要代謝體的協助

研究主題一：代謝體學

圖譜計算、系統生物網路計算

Software

3Omics: a web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data

Tien-Chueh Kuo^{1,2}, Tze-Feng Tian^{2,3} and Yufeng Jane Tseng^{1,2,3*}

Highly accessed

Open Access

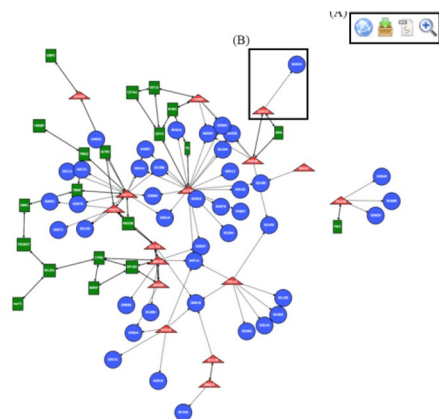
analytical
chemistry

Article
pubs.acs.org/ac

Batch Normalizer: A Fast Total Abundance Regression Calibration Method to Simultaneously Adjust Batch and Injection Order Effects in Liquid Chromatography/Time-of-Flight Mass Spectrometry-Based Metabolomics Data and Comparison with Current Calibration

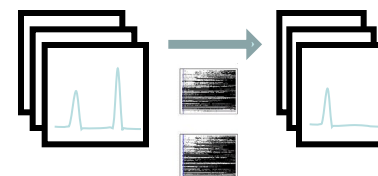
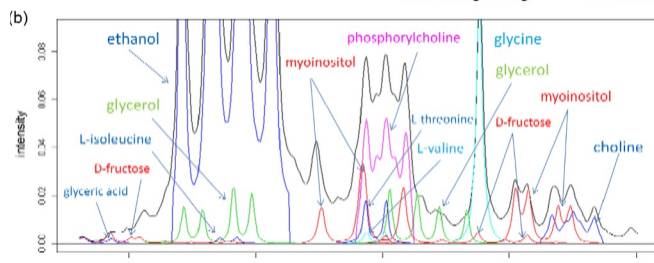
analytical
chemistry

Article
pubs.acs.org/ac



Distribution-Based Classification Method for Baseline Correction of Metabolomic 1D Proton Nuclear Magnetic Resonance Spectra

Kuo-Ching Wang,^{†,||,L,†,V} San-Yuan Wang,^{‡,||} Ching-hua Kuo,^{§,||} and Yufeng J. Tseng^{*,†,‡,§,||}



- 發展可分析混和物的質譜與核磁共振光譜演算法
- 2011國際代謝體年會最佳論文獎
- 最多理論計算發表到分析化學最好的期刊「Analytical Chemistry」
- 系統生物網路計算 3Omics – **Highly Accessed Article** [BMC Systems Biology (2013)]

研究主題一：代謝體學

鑑定臨床上診斷、治療的生物指標

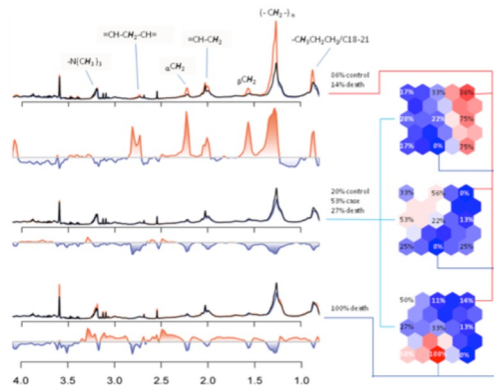
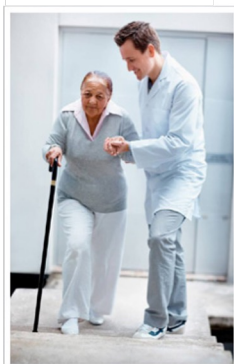
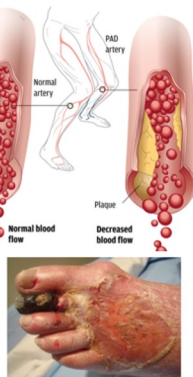
Peripheral arterial disease

Peripheral arterial disease (PAD) is a common circulation problem: arteries that carry blood to the legs or arms become narrowed or clogged.

Main cause of PAD

Narrowing of arteries
Plaque builds up in arteries, narrows, blocks blood flow (also called atherosclerosis)

Risk factors
- Smoking
- Diabetes
- Obesity (a body mass index over 30)
- High blood pressure (160/90 millimeters of mercury or higher)
- High cholesterol total



周邊動脈疾病

- 三成死亡率
- 全球最大的周邊動脈疾病血液樣本
- 全球唯一團隊

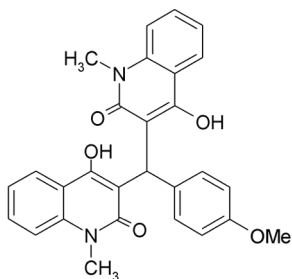
- **FDA**同級的台大代謝體核心實驗室
- 世界上第一個可預知一年後周邊動脈疾病死亡率的生物指標
- 開啟台灣學界代謝體本領域研究

Lab Focus



診斷

- 找尋生物指標
- 圖譜、系統生物網路計算
- 精準醫學



分子結構計算



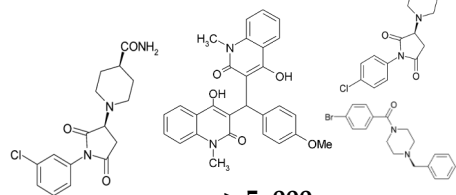
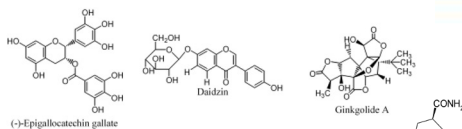
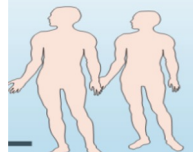
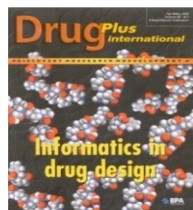
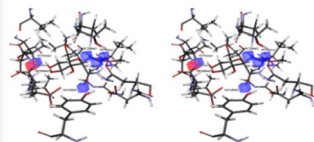
治療

- 電腦藥物設計
- 電腦藥物篩選
 - 計算模型取代生物實驗標準
- AI應用於藥物開發

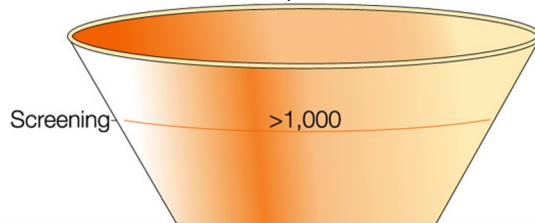
電腦輔助藥物設計

10 years

7 Years



> 5,000



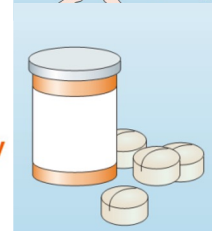
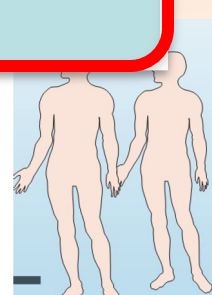
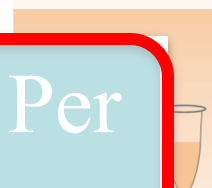
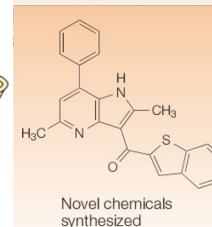
US \$ 1.7 Billion Per Drug

Phase II 2.4

Phase III 1.2

Launch of NCE

Nature Reviews | Drug Discovery



20 years

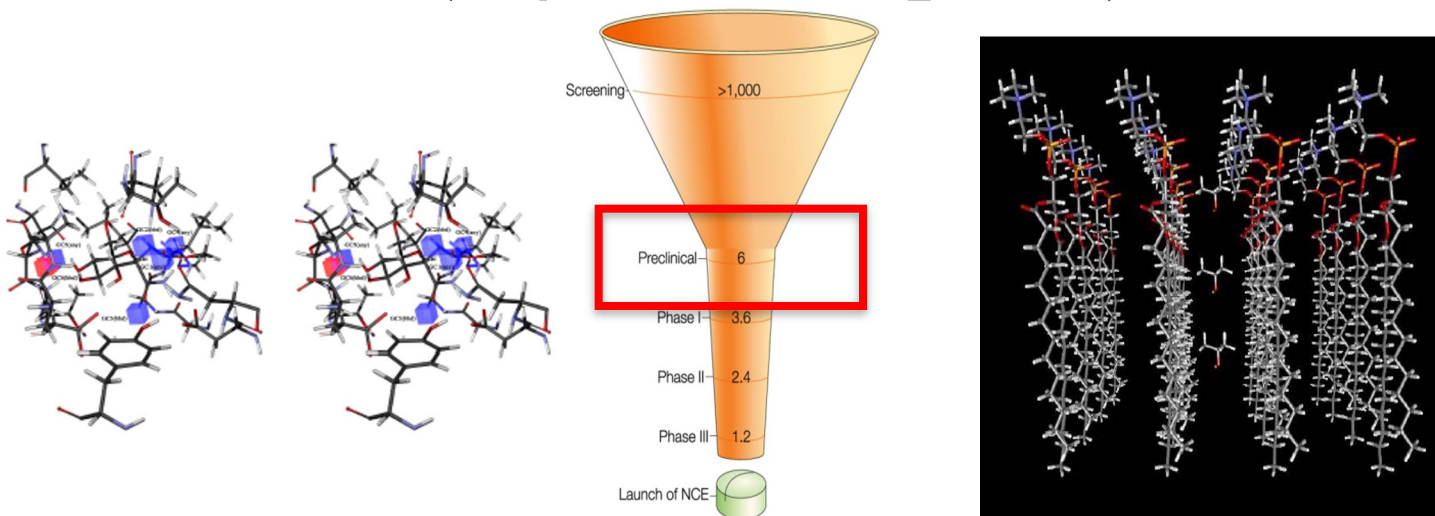
7 Years

傳統藥物開發

研究主題二：電腦藥物設計

減少毒性

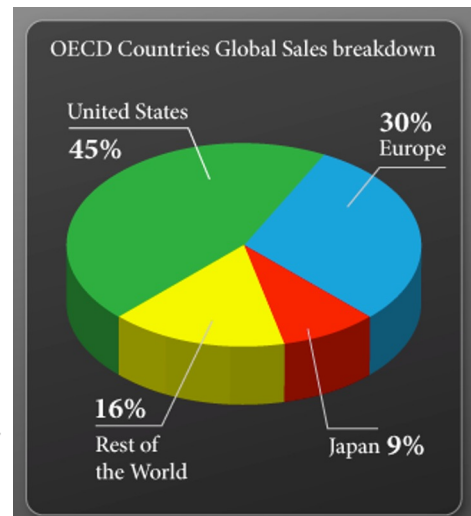
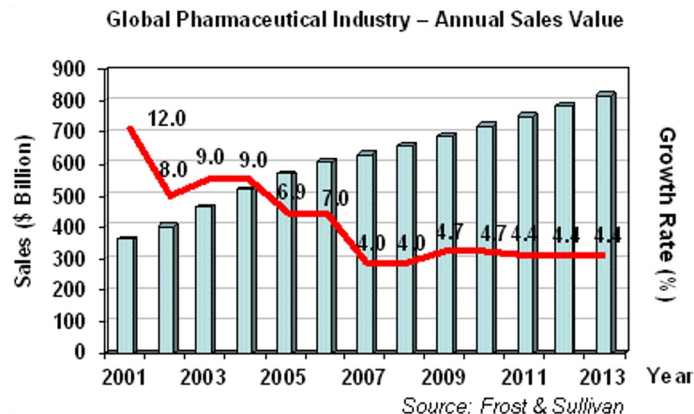
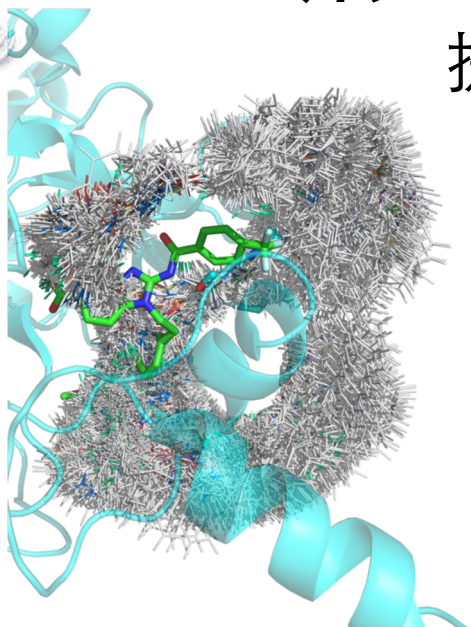
美國國家衛生研究院唯一臨床前電腦預測藥物篩選之計畫
(2006-2009，總共10個領域得到支持)



- 已成為國際業界標準：
 - 美國知名製藥公司採用的電腦藥物篩選模型(2007-)
- 計算模型取代生物實驗
 - 口腔腸道藥物吸收、皮膚致敏毒性、皮膚吸收、皮膚滲透的計算模型
 - 歐盟在2012年啟用禁止生物測試化學物質經由皮膚吸收毒性法規

研究主題二：電腦藥物設計

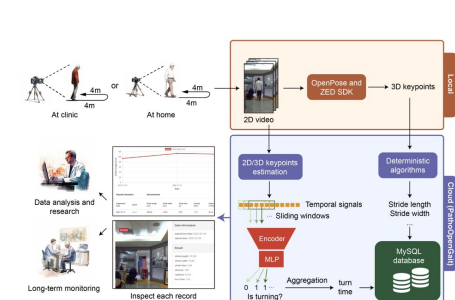
提高結構活性、獲得藥物專利



- 藥物要能獲利是要有藥品結構專利
- 一年全球獲准上市的藥物不超出17種
 - 在一個治療類別裡第一個上市的藥物 (**first in the class**) 可以拿下一半的全球市場15至20年

研究主題三：AI for Healthcare

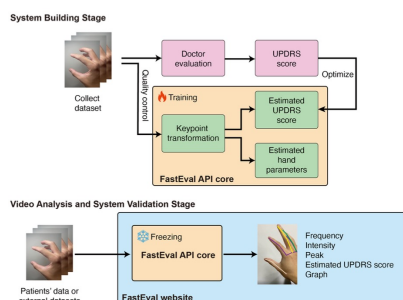
We leverage computer vision, nature language processing, and machine learning methods to understand human movement and behavior, toward scalable and accessible healthcare solutions.



PathoOpenGaIt

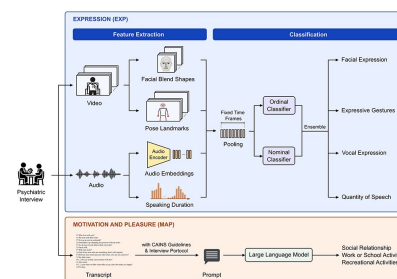
(IEEE JBHI 2024)

Open-source cloud-enabled platform for pathological gait analysis.
— Semi-supervised learning + cloud inference for scalable gait evaluation.



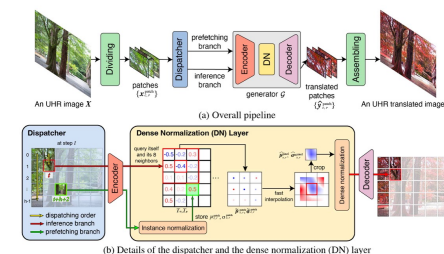
FastEval Parkinsonism (npj Digital Medicine 2024)

Instant video-based online assessment system for Parkinsonian motor symptoms.
— First real-time, deep-learning-assisted evaluation platform.



AI for Schizophrenia (Schizophrenia Bulletin 2025)

Gen AI + ML for automated symptom assessment.
— First attempt to quantify negative symptoms using LLM/AIGC methods.



Dense Normalization (ECCV 2024)

Ultra-high-resolution unpaired image-to-image translation via dense normalization.
— Enabling data augmentation for large images in healthcare scenarios.

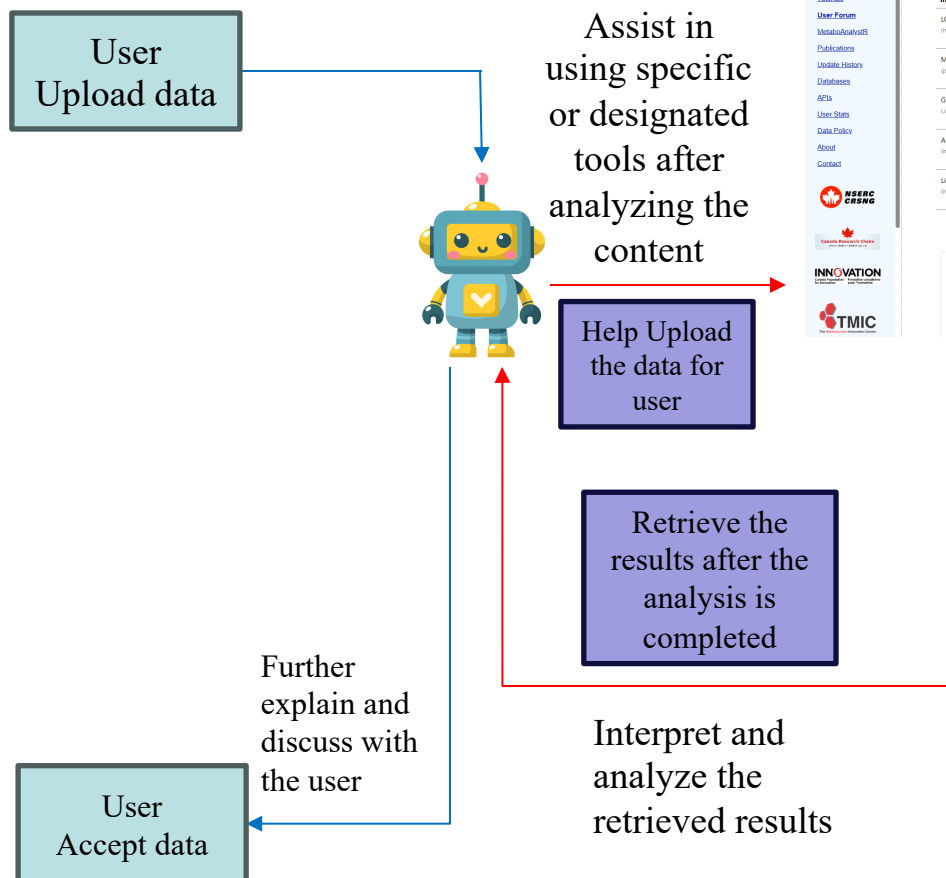


Computational Molecular Design
& Metabolomics Lab



研究主題四：AI Agent

We employ AI agents to assist users in analyzing mass spectrometry data through reliable databases, reducing complex manual procedures and facilitating deeper data interpretation.



MetaboAnalyst 6.0

MetaboAnalyst 6.0 - from raw spectra to biomarkers, patterns, functions and systems biology

Module Overview

Input Data Type	Available Modules (click on a module to proceed, or scroll down to explore a total of 18 modules including utilities)				
LC-MS Spectra (mzML, mzML, or mzML)			Spectra Processing (LC-MS w/MS/MS)		
MS Peaks (peak list or intensity table)		Peak Annotation (MS2, DDA/MS1)	Functional Analysis (IC-MS)	Functional Meta-analysis (IC-MS)	
Generic Format (CSV or TSV table file)	Statistical Analysis (one factor)	Statistical Analysis (metadata table)	Biomarker Analysis	Dose Response Analysis	Statistical Meta-analysis
Annotated Features (metabolite list or table)		Enrichment Analysis	Pathway Analysis	Network Analysis	
Link to Genomics & Phenotypes (metabolite list)			Causal Analysis (Mendelian randomization)		

Spectra Processing (LC-MS w/MS/MS)

This module allows users to upload raw LC-MS spectra (mzML, mzML, or mzML) to be processed using our optimized workflow based on MetaboAnalyst 6.0 or the latest gpl algorithm. Users can also include MS2 spectra (both DDA or SSW/MS1-MS2 are supported) for peak annotation.

Peak Annotation (MS2 DDA/MS1)

This module performs MS2 peak annotation based on a comprehensive list of public databases. Users can either directly enter a two-column peak list containing m/z and intensity values (DDA) or upload a .mgf file produced by MS/MS or MS1-MS2 after the spectral deconvolution (MS/MS+Peak).

Functional Analysis (IC-MS)

This module accepts high-resolution LC-MS spectral peak data to perform metabolic pathway enrichment analysis and visual exploration based on the [gpl](#) or [gpl](#) algorithms. It currently supports 20 organisms including Human, Mouse, Drosophila, C. elegans, and other species.

Choose the analysis approach according to the data

Pathway Analysis

Enrichment Analysis

Dose Response Analysis



研究主題四：AI Agent

We use AI agent to help users explore and understand toxic chemicals they are curious about, by gathering and interpreting information from academic literature and chemical databases, enabling deeper insights into the substances of interest.

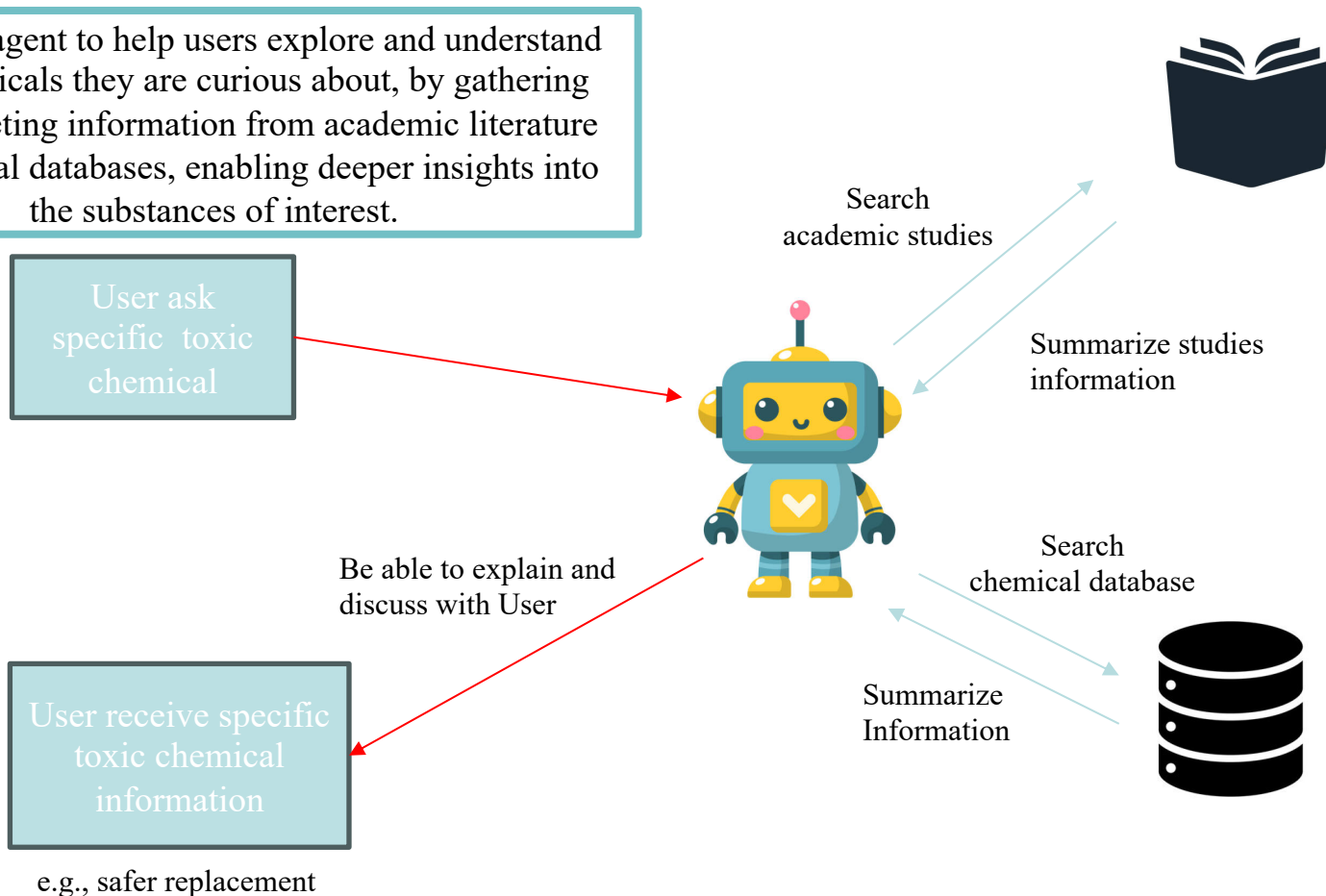


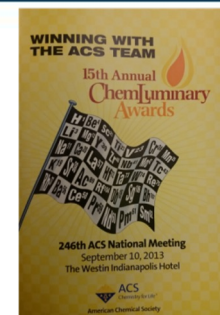
Figure resource :

<https://share.google/images/thOzC0Te9PccKQHFD>

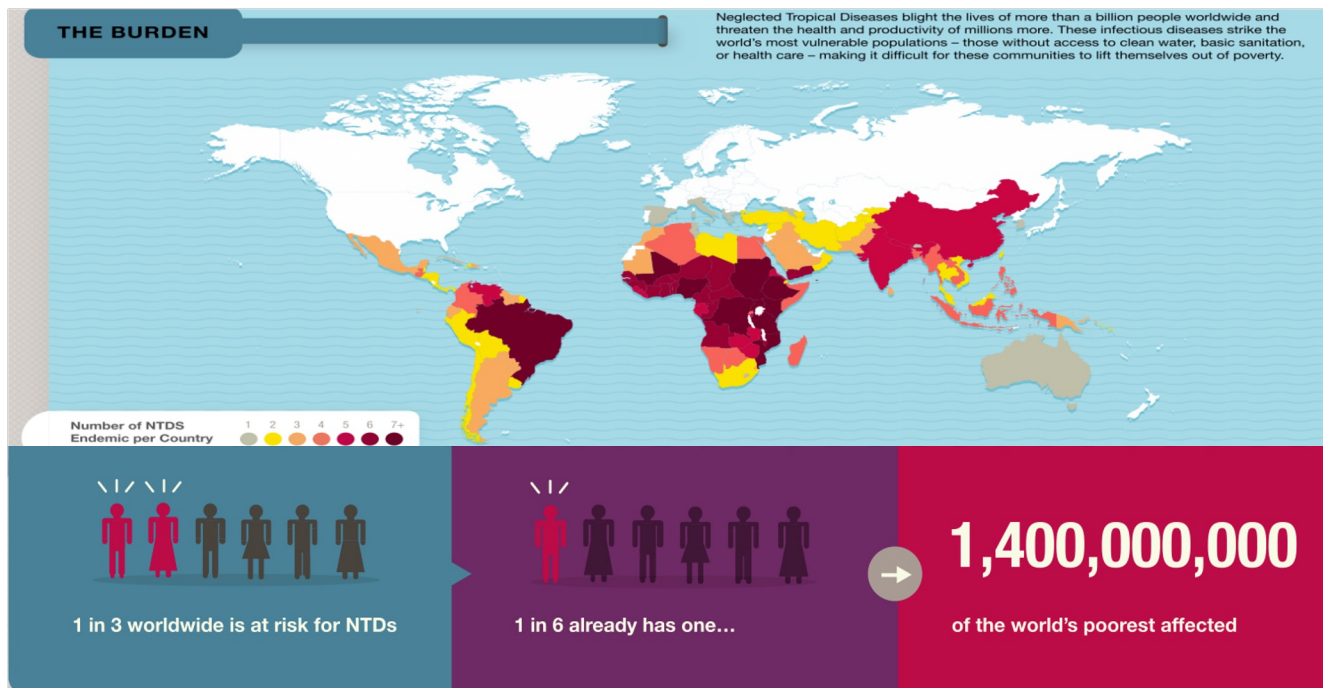
<https://share.google/images/iY2BLqx1KsKWUQukw>

國際服務

「Teach Discover Treat (TDT)」 initiatives



- 集合**全球學界**開發**忽略性**疾病藥物 (neglected disease) since 2010
- 病人眾多，但集中在藥廠不關心的非歐美地區，無現有有效的藥
- 「**台大**，Novartis，Merck，UC」 Founder (4)，Scientific Board Member
- **2012 美國化學協會創新研究獎 ACS IPG innovation award**
- **2013 美國化學協會創新服務獎 ACS Chemluminary award**



跨國跨院服務

跨院系服務

- 藥物研究中心
- 基因體醫學研究中心代謝體核心
 - 設立台大醫學院代謝體核心實驗室
 - 國科會代謝體核心實驗室

跨國服務

- 跨國頂尖癌症研究中心 (i-Rice) [2012年迄今]



WELCOME

資訊工程學系, Room 529

yjtseng@csie.ntu.edu

<http://www.cmdm.tw>

<http://www.csie.ntu.edu.tw/~yjtseng>