



香港中文大學(深圳)  
The Chinese University of Hong Kong, Shenzhen

# AI-Supervisor in Chemistry and Materials Science Research

朱熹

香港中文大学 (深圳)

ACS Taiwan Library Event  
2023/09/21

# CONTENTS

## 01. 传统的科研模式

The Traditional Research Paradigm

## 02. AI-Supervisor

Prosperous AI Publishing Service

## 03. 推动机器人科学研究

Towards the Sustainable Prosperity

## 04. 总结和致谢

Conclusion and Acknowledgement





# CONTENTS

## 01. 传统的科研模式

The Traditional Research Paradigm

## 02. AI-Supervisor

Prosperous AI Publishing Service

## 03. 推动机器人科学研究

Towards the Sustainable Prosperity

## 04. 总结和致谢

Conclusion and Acknowledgement

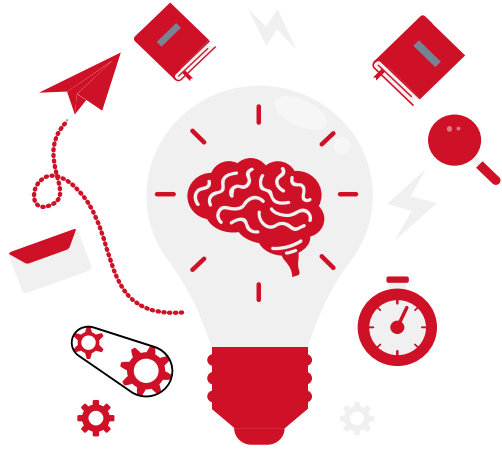


# 传统的科研模式

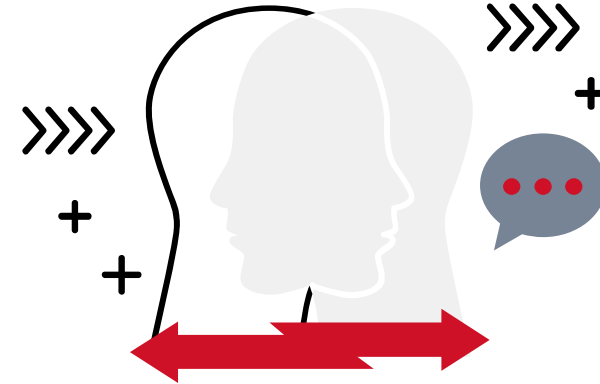


香港中文大學(深圳)  
The Chinese University of Hong Kong, Shenzhen

## Knowledges, Ideas



## Communications

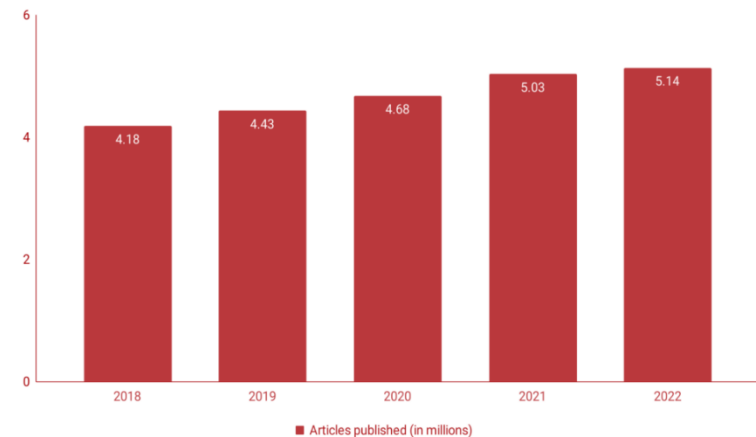


**Sci Disco  $\propto$  manpower**

## Experiments



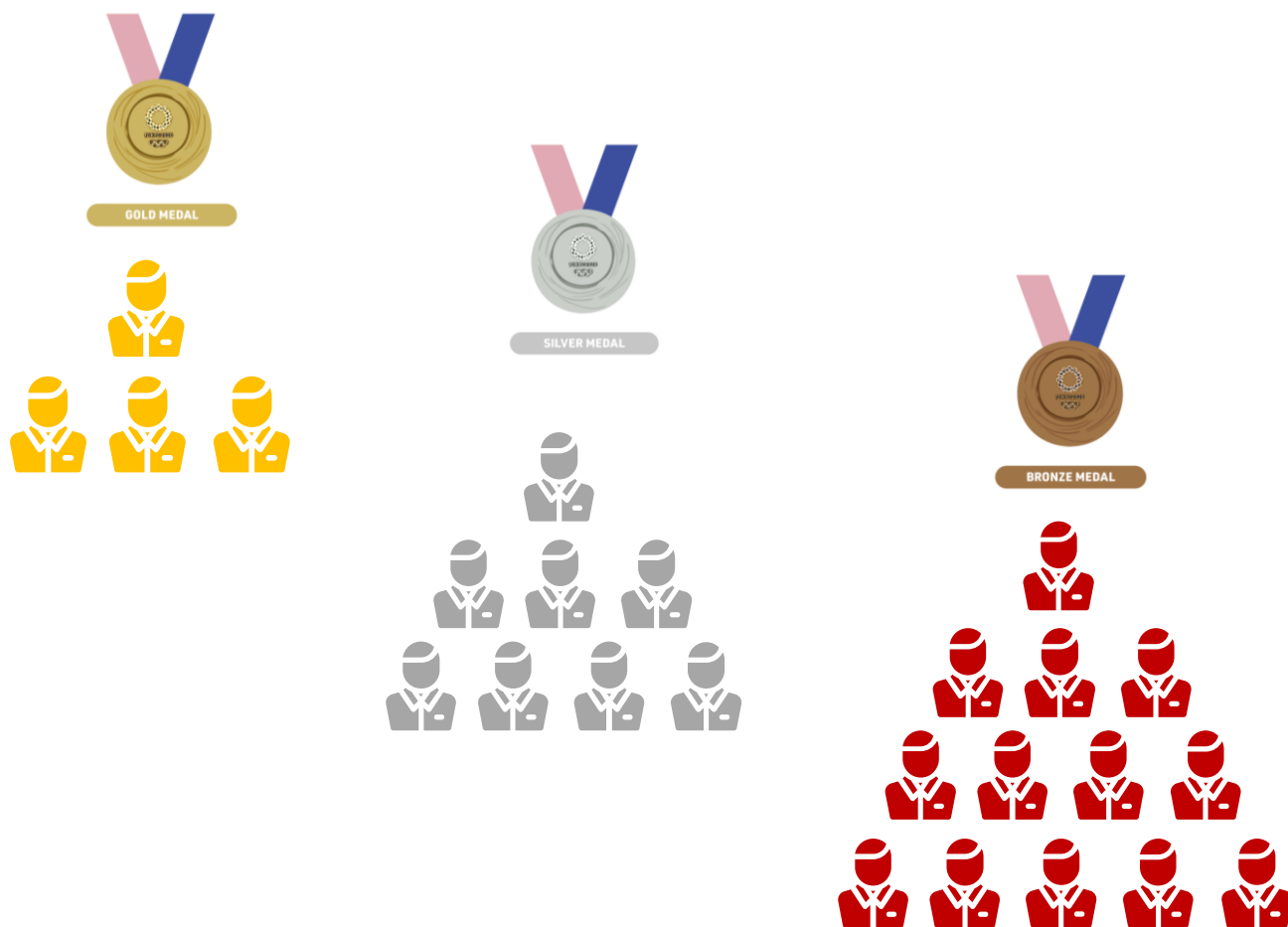
Number of academic papers published by year (in millions)





# 传统的科研模式: 学术壁垒

## ● 较封闭式的知识传承体系



2015-16 Rank	Institution	Country
1	California Institute of Technology	United States
2	University of Oxford	United Kingdom
3	Stanford University	United States
4	University of Cambridge	United Kingdom
5	Massachusetts Institute of Technology	United States
6	Harvard University	United States
7	Princeton University	United States
8	Imperial College London	United Kingdom
9	ETH Zurich-Swiss Federal Institute of Technology Zurich	Switzerland
10	University of Chicago	United States

□ 普通大学学术资源有限

# 传统的科研模式

打包全历史数据



M

Tell me how ChatGPT works.



ChatGPT is a large language model that uses deep learning techniques to generate human-like text. It is based on the GPT (Generative Pre-trained Transformer) architecture, which uses a transformer neural network to process and generate text. The model is pre-trained on a massive dataset of text, such as books, articles, and websites, so it can understand the patterns and structure of natural language. When given a prompt or a starting point, the model uses this pre-trained knowledge to generate text that continues the given input in a coherent and natural way.





# 传统的科研模式

## 6 Amazing Things ChatGPT Can Do

ChatGPT is a powerful artificial intelligence (AI) system developed by OpenAI

- Automate simple tasks and workflows
- Generate viral social media content
- Generate content ideas
- Write goal-oriented articles and blog posts
- Can help you find the best way to write your code
- Explain complicated concepts, processes, and phenomena in easily understandable terms



**Can AI breaks Scientific knowledge barrier ?**

pakwired.com

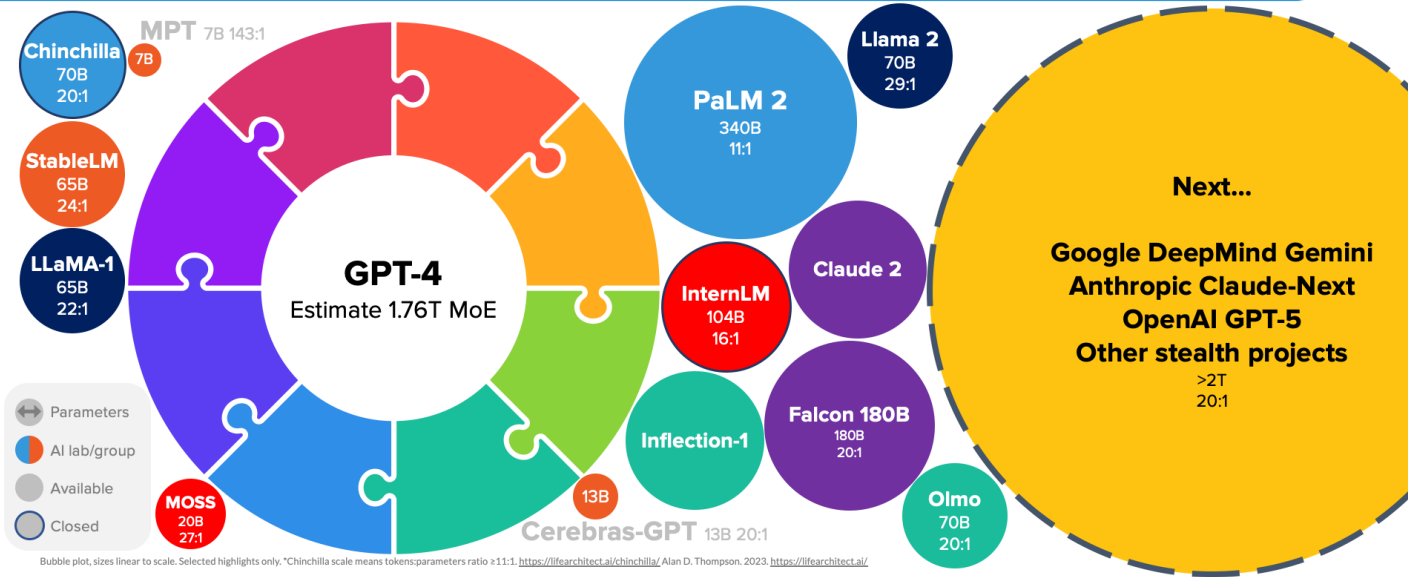
<https://pakwired.com/what-is-chatgpt/>

# 繁荣的AI出版服务

Prosperous AI Publishing Service

## 国际上大模型现状 (2023年6月)

### 2023-2024 OPTIMAL LANGUAGE MODELS SEP/2023



Bubble plot, sizes linear to scale. Selected highlights only. \*Chinchilla scale means tokens-parameters ratio  $\approx 11:1$ . <https://life architect.ai/chinchilla/>, Alan D. Thompson, 2023. <https://life architect.ai/>

[LifeArchitect.ai/models](https://life architect.ai/models)



- May**  
 CodeT5+ 16B (model)  
 Salesforce  
[Read the paper.](#)
- May**  
 Formosa 176B (model)  
 Asus  
[View the announce.](#)
- May**  
 LIMA 65B (model)  
 Meta AI  
[Read the paper.](#)
- May**  
 Guanaco 65B (model)  
 UW  
[Read the paper.](#)
- May**  
 Falcon 40B (model)  
 TII  
[View the repo.](#)
- May**  
 GPT-4 MathMix (model)  
 OpenAI  
[Read the paper.](#)
- June**  
 DIDACT (model)  
 Google DeepMind  
[Read the announced.](#)
- Next...**
- TBA**  
 Gemini 1T (model)  
 Google DeepMind  
[Read more,](#)  
[Alan's report](#)
- TBA**  
 Claude-Next 1T (model)  
 Anthropic  
[Read more,](#)  
[Alan's report](#)



# 繁荣的AI出版服务

Prosperous AI Publishing Service

## 国内大模型数目 > 100

<b>科大讯飞</b> <b>1+N认知智能大模型</b> 未发布 预计2023年5月6日发布	<b>昆仑万维</b> <b>天工3.5</b> 未发布 预计2023年4月17日开始测试	<b>华为</b> <b>盘古NLP模型</b> 未发布 2023年4月10日举行发布会但未公布上线时间
<b>达观数据</b> <b>曹植</b> 未发布，可试用 2023年3月18日公布研发进度可申请试用	<b>网易</b> <b>玉言</b> 未发布 发布时间未知	<b>阿里巴巴</b> <b>通义千问</b> 2023年4月11日发布 将接入所有阿里产品
<b>商汤科技</b> <b>日日新</b> 2023年4月10日发布	<b>360</b> <b>360智脑</b> 2023年4月10日发布	<b>百度</b> <b>文心</b> 2023年3月16日发布 关键产品“文心一言”
<b>智谱AI</b> <b>ChatGLM-6B</b> 2023年3月14日发布	<b>澜舟科技</b> <b>孟子</b> 2023年3月14日发布	<b>复旦大学</b> <b>MOSS</b> 2023年2月21日发布
<b>腾讯</b> <b>混元</b> 2022年12月发布，预计关键产品“混元助手”近期上线	<b>中科院自动化所</b> <b>紫东太初</b> 2021年9月27日发布	<b>智源研究院</b> <b>悟道2.0</b> 2021年6月1日发布

### IDC Market Glance: 中国AI大模型生态图谱 (V1.0)



来源: IDC, 2022 对于已公开发布市场份额的市场, 市场份额前3-5的厂商已经被展示; 对于IDC未公开发布市场份额的领域, 所展示的厂商是由分析师自行决定的。



# CONTENTS

## 01. 传统的科研模式

The Traditional Research Paradigm

## 02. AI-Supervisor

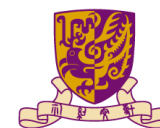
Prosperous AI Publishing Service

## 03. 推动机器人科学研究

Towards the Sustainable Prosperity

## 04. 总结和致谢

Conclusion and Acknowledgement





# AI-Supervisor

AI-Supervisor

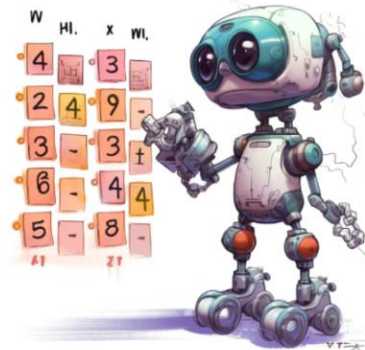
## AI-Supervisor

### Primary Features



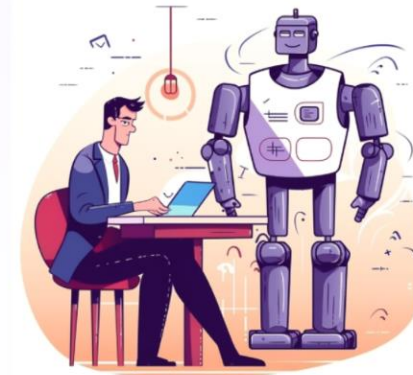
#### Research Trending

Summarizes previous hotspots in the form of a word cloud and provides analysis results in natural language. Predicts potential hotspots for the next one, two, or three years and beyond.



#### Necessity Rating

KnowledgeWorks can analyze papers and provide scores across different criteria.



#### Paper Discussion

Highlight the key contributions of the paper. Suggest possible points for improvement. Propose potential areas for further research, including interdisciplinary topics.

# AI-Supervisor

AI-Supervisor

## 全周期智能科研服务

### 论文知识



On top

AI 指导下的科研CRO

AI 内容总结

创新价值评判

学术知识图谱

论文传承分析

ScholarNet

张量网络

API 接口

多模态模型

跨领域文本数据

大型自然语言模型



## 大型语言模型基础



# AI-Supervisor

AI-Supervisor

支持使用多种模式

# KnowledgeWorks

Model Version

- Fast
- High Performance

# KnowledgeWorks

KnowledgeWorks

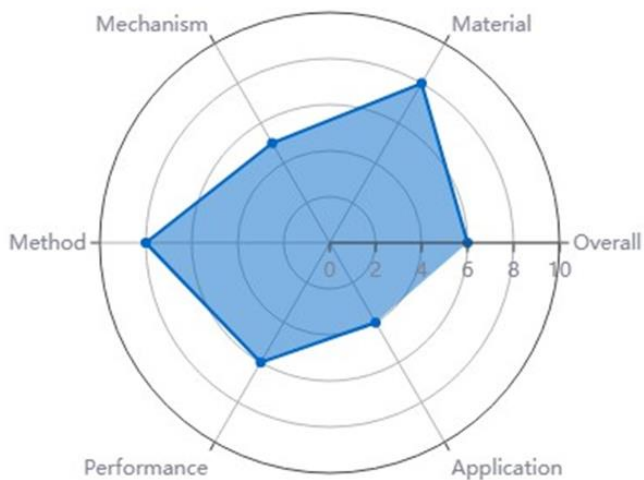
## 论文鉴定

KnowledgeWorks

Model Version  
● Fast  
○ High Performance

Home Trending **Rating & Analysis** Paper Discussion Teaching & Generation

### Paper Rating and Analyzing by AI Supervisor



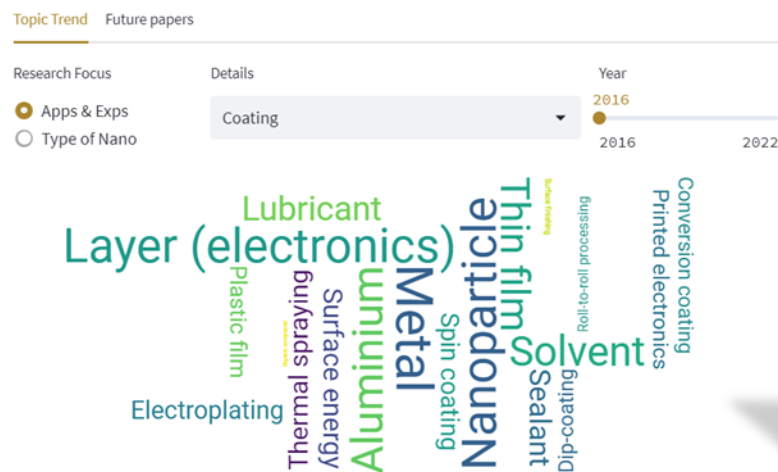
## 知识总结

KnowledgeWorks

Model Version  
● Fast  
○ High Performance

Home **Trending** Rating & Analysis Paper Discussion Teaching & Generation

### Research trending analysis by AI Supervisor



## 内容讨论

KnowledgeWorks

Model Version  
● Fast  
○ High Performance

Home Trending Rating & Analysis **Paper Discussion** Teaching & Generation

### Talk with AI Supervisor

Input DOI Upload file

Chat content

Please provide the DOI links for the papers you're interested in. AI supervisor can analyze the papers' limitations and offer suggestions for future research. Alternatively, feel free to ask about specific scientists or concepts directly.

<https://doi.org/10.1039/CTTA00203C>  
<https://doi.org/10.1002/adma.201102306>

The information for paper 10.1002/adma.201102306:  
**Title:** Two-Dimensional Nanocrystals Produced by Exfoliation of Ti3AlC2  
**Abstract:** 2D Ti3C2 nanosheets, multilayer structures, and conical scrolls produced by room temperature exfoliation of Ti3AlC2 in HF are reported. Since Ti3AlC2 is a member of a 60+ group of layered ternary carbides and nitrides, this discovery opens a door to the synthesis of a large number of other 2D crystals.  
**Publication Date:** 2011-10-04  
**Journal:** Advanced Materials  
**Publisher:** Wiley-Blackwell  
**Area:** X-Ray Diffraction, Temperature, Nitrogen, Carbon, Aluminum Compounds, Titanium, Carbon Compounds, Inorganic, Photoelectron Spectroscopy



**KnowledgeWorks**

KnowledgeWorks

**How**

**KnowledgeWorks**

**Works for**



**ACS**  
Chemistry for Life<sup>®</sup>

## nature

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾ [Subscribe](#)

[nature](#) > [nature index](#) > [article](#)

NATURE INDEX | 03 May 2023

## Researchers who agree to manipulate citations are more likely to get their papers published

Data suggest that these researchers are more willing to publish in journals that participate in such coercion.

[Dalmeet Singh Chawla](#)



先发表，再判断。



KnowledgeWorks

## 文章必要性和创新性评分

当用户将论文提交给AI Supervisor时,

我们的系统可以分析它并提供不同标准的分数, 包括五类评分标准:

机理

方法

材料

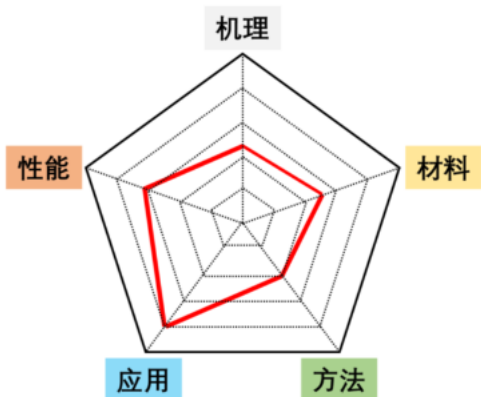
性能

应用

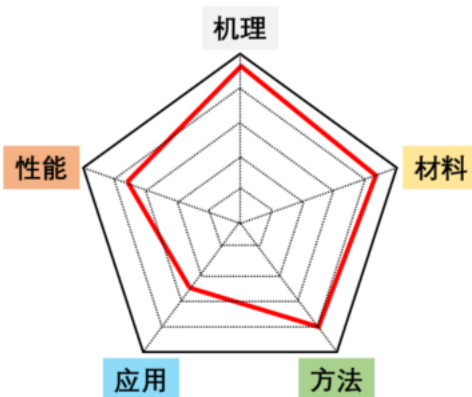
AI智慧导师的评分来源于基于历史文献数据训练的大模型, 以及其对应的数据库。通过对论文的拆解和语义的分类标注, AI智慧导师可以判断相似关键概念在历史文献中的出现频率, 从而得出评分, 相似概念出现频率越高, 创新度评分就越低。

需要注意的是, 我们的评级是完全基于语料的, 可以在论文发表前就给出客观的评分, 而不需要使用人为创建的评估参数, 如H指数、影响因子或引用数。

中国材料化学论文平均创新评分



高创新度论文案例: 单层石墨烯的首次制备  
Novoselov, Science 306.5696 (2004): 666-669.



## 先判断, 再发表





# 学者分析

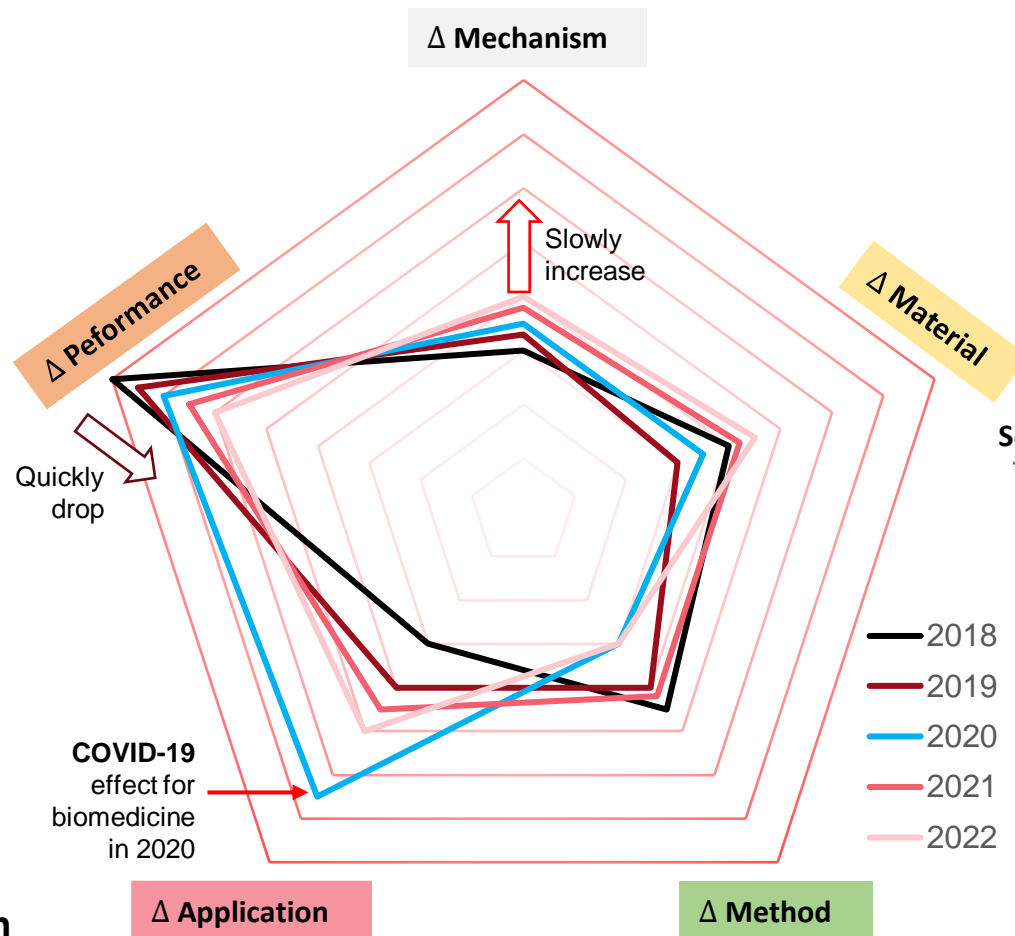
Scholar Profile

The variation of publication key concepts (delta  $\Delta$ ) in 5 dimensions in the past 5 years

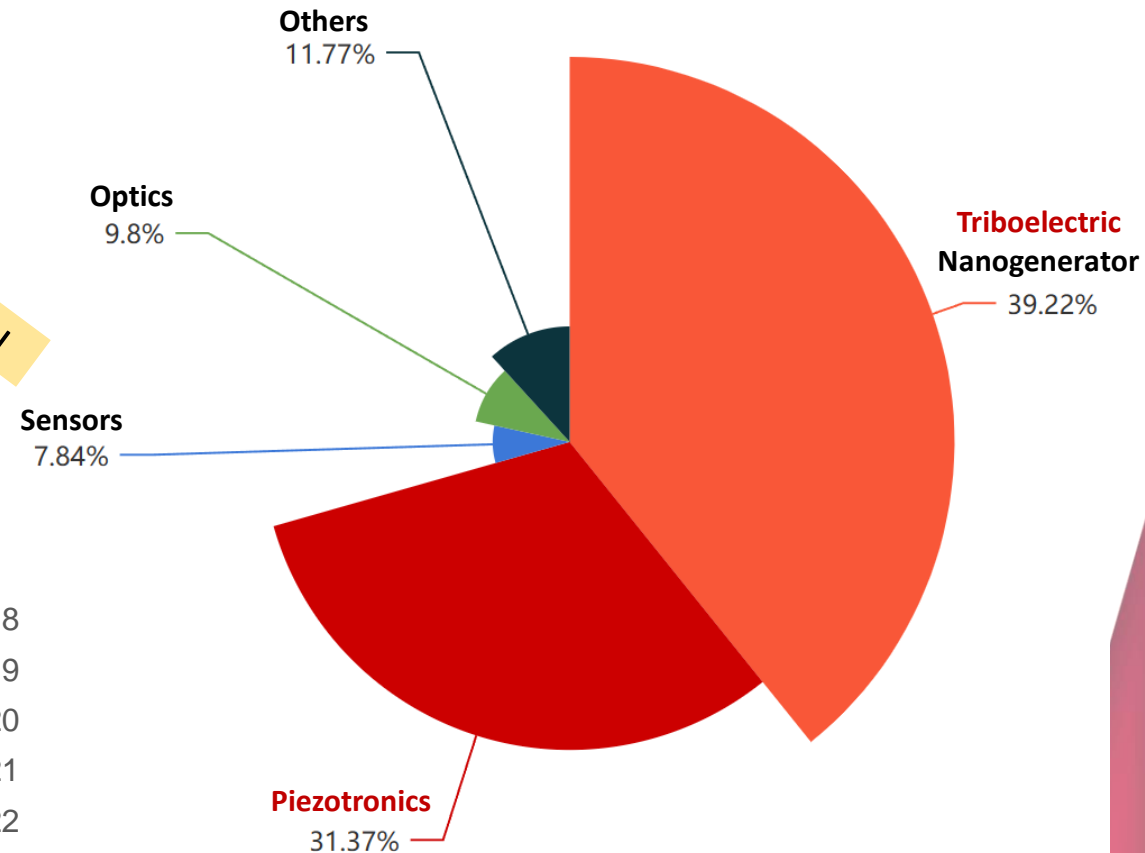


Focus:  
Nanomaterials

Score: 0-10, higher means more variation



Wang's Publication Categories in the past 5 years



**Conclusion:** Very limited invention change in most of the aspects. Keep refresh records in the same application field. Stay in the same research area (>70%). Suitable to be replaced by AI & automation!

# 设计新的论文

Design for New Papers

## 建议的新步骤

### Suggested methods

precursor can be used to grow CsPbBr<sub>3</sub> nanocrystals, which are known to exhibit excellent optoelectronic properties.

Next,  $x \mu\text{L}$  ( $0 \leq x \leq 800$ ) methanol was added into 10 mL toluene, referred to as the "bad solvent". It is crucial to optimize the concentration of methanol for obtaining high-quality nanocrystals with optimal optoelectronic properties.

Then, 1 mL of the precursor solution was quickly injected into 10 mL of the "bad solvent" to trigger supersaturation. This process can be controlled at different rates to modify the growth of CsPbBr<sub>3</sub> nanocrystals and investigate its impact on their optoelectronic properties.

After the synthesis, the as-prepared CsPbBr<sub>3</sub> nanocrystals can be subjected to various characterization techniques to study their optoelectronic properties, such as:

1. UV-Visible absorption spectroscopy and photoluminescence (PL) spectroscopy to investigate their absorption and emission

## Method Query

[Suggest](#) [Ask & modify](#)

### Chat content



Please tell me what you would like to synthesize with detailed demand.



**用户输入现有实验**

First, the precursor solution was prepared by mixing 0.4 mmol CsBr, 0.4 mmol PbBr<sub>2</sub>, 0.5 mL OLA, and 1 mL OA in 10 mL DMF (dimethylformamide, the "good solvent"). Then,  $x \mu\text{L}$  ( $0 \leq x \leq 800$ ) methanol was added into 10 mL toluene, namely the "bad solvent". Finally, 1 mL of the precursor solution was injected into 10 mL of the "bad solvent" to trigger supersaturation.



**回复初步建议：可能的修改优化路径**

1. Phase Transition Analysis: Investigate the effects of varying methanol concentration on the formation and stability of perovskite nanocrystals, including crystalline phase, size, and morphology.
2. Optoelectronic Properties: Study the impact of the different methanol concentrations on the optical and charge transport properties of the formed nanocrystals and their potential use in optoelectronic devices.
3. Surface Chemistry Modification: Explore ligand exchange strategies and their influence on the stability, solubility, and optoelectronic performance of the perovskite nanocrystals, aiming for improved device fabrication and performance.
4. Environmentally Friendly Alternatives: Research non-toxic alternative materials to replace lead-based perovskites in order to minimize potential environmental and health hazards.

## Idea Suggestions

[QUAINT suggested ideas](#)

### Suggested directions

Phase Transition Analysis

Optoelectronic Properties

Surface Chemistry Modification

Environmentally Friendly Alternatives

AI-Assisted Material Discovery

**选择框，用户可选进一步探索方向**

# 设计新的论文

Design for New Papers



## QUAINT

### Method Query

[Suggest](#) [Ask & modify](#)

Chat content



Please tell me what you would like to synthesize with detailed demand.

Your Input:

First, the precursor solution was prepared by mixing 0.4 mmol CsBr, 0.4 mmol PbBr<sub>2</sub>, 0.5 mL OLA, and 1 mL OA in 10 mL DMF (dimethylformamide, the "good solvent"). Then,  $x \mu\text{L}$  ( $0 \leq x \leq 800$ ) methanol was added into 10 mL toluene, namely the "bad solvent". Finally, 1 mL of the precursor solution was injected into 10 mL of the "bad solvent" to trigger supersaturation.

Submit

### Idea Suggestions

[QUAINT suggested ideas](#)

Suggested directions

The QUAINT suggestions are shown here.

You may choose by clicking the buttons.

Suggested methods

Please choose a suggestion above.



# CONTENTS

## 01. 传统的科研模式

The Traditional Research Paradigm

## 02. AI-Supervisor

Prosperous AI Publishing Service

## 03. 推动机器人科学研究

Towards the Sustainable Prosperity

## 04. 总结和致谢

Conclusion and Acknowledgement





🔒 | SPECIAL ISSUE ARTICLE

## Automated Synthesis of Peptides: Solid-phase peptide synthesis, a simple and rapid synthetic method, has now been automated.

R. B. MERRIFIELD [Authors Info & Affiliations](#)

SCIENCE · 8 Oct 1965 · Vol 150, Issue 3693 · pp. 178-185 · DOI: 10.1126/science.150.3693.178



Robert Bruce Merrifield

1921-2006

Nobel Prize in Chemistry (1984)

for the invention of solid phase peptide synthesis

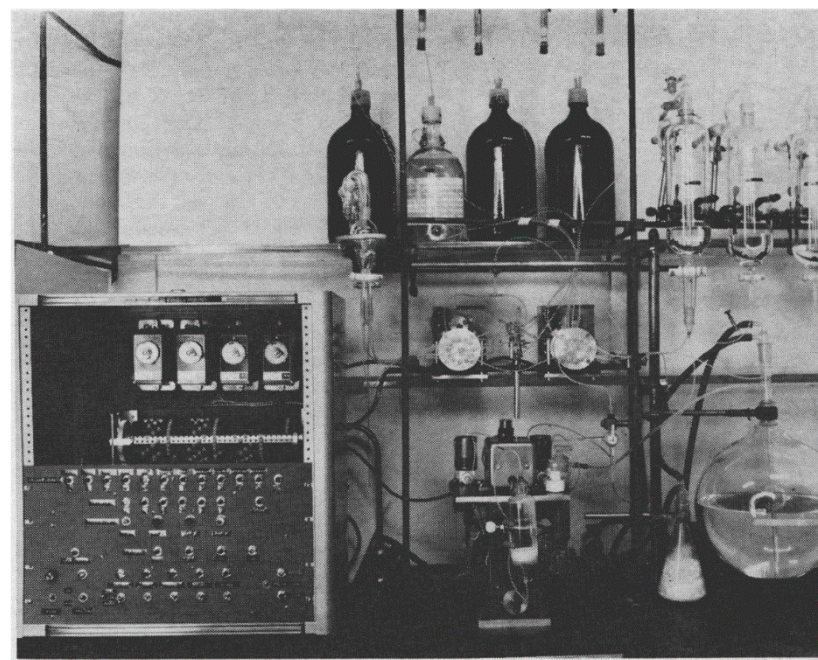


Fig. 7. The instrument designed for automation of peptide synthesis.

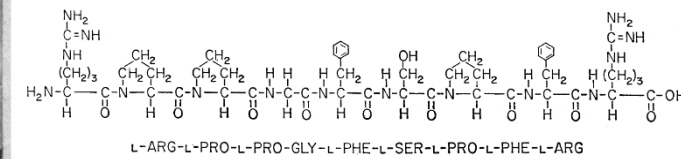
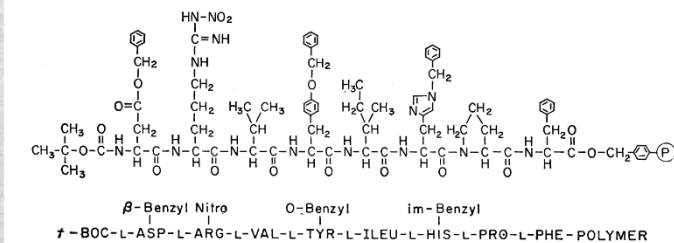


Fig. 5. Bradykinin.

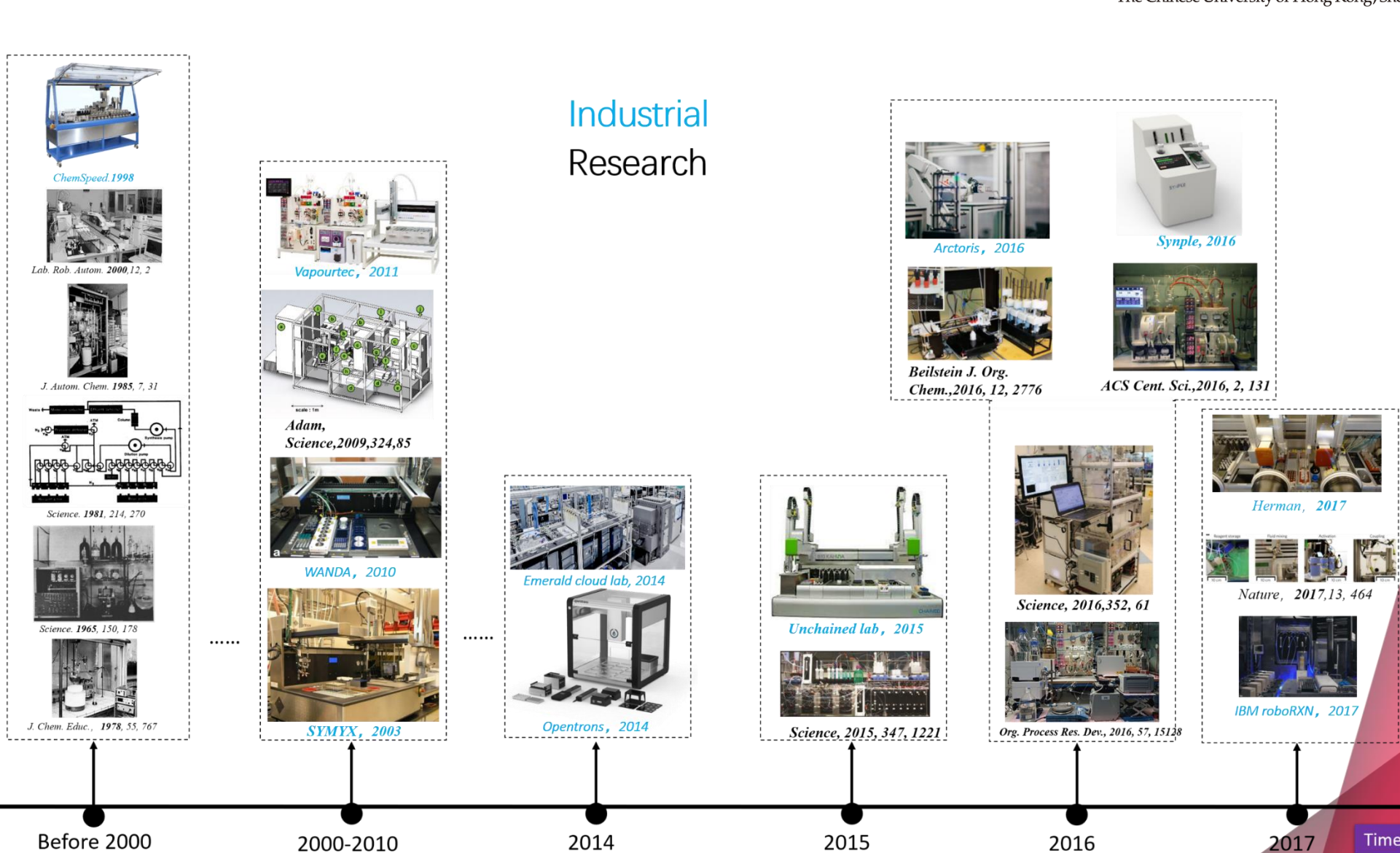
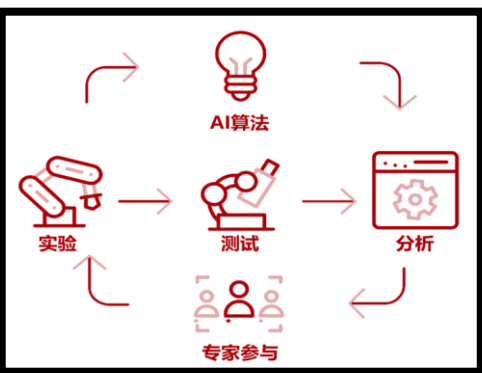




# AI和机器人的化学实验室：2017年以前

- 以自动化设备实现实验过程为主;
- 包括了常见的无机和有机反应;
- 主要应用于大药厂和生物医药行业。

## Industrial Research



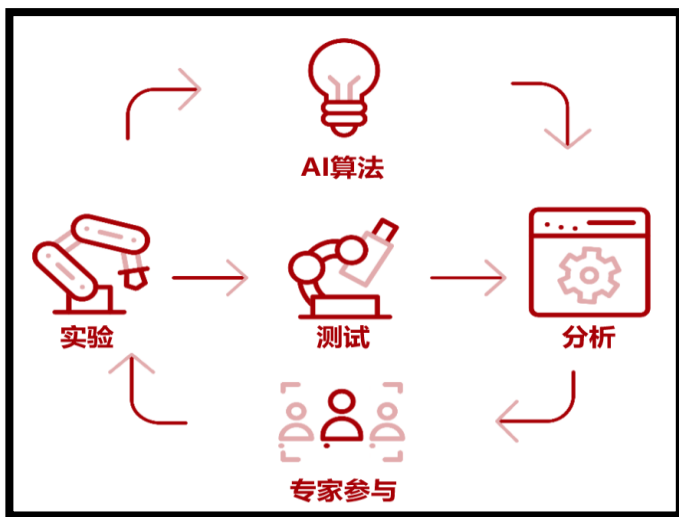




# AI和机器人的化学实验室：2017年以后

- AI技术大幅应用于机器;
- 覆盖更多的材料化学实验;
- 面向高校的课题组实验室。

Industrial  
Research



**AIR-Chem**, *J. Phys. Chem. A* **2018**, 122, 9142

**Project Ada**, **2018**

*Science*, **2018**, 361 (6408), 1220

*Adv. Sci.*, **2018**, 7, 1901957

*Science*, **2019**, 365, eaax1566

*Science*, **2019**, 363, eaav2211

**MAOSIC**, *Nat Commun* **2020**, 11, 2046

**BiacP**, *J. Phys. Chem. Lett.* **2020**, 11, 9995

**Artificial Chemist**

*Artificial Chemist*, *Adv. Mater.* **2020**, 32, 2001626

*Science* **2020**, 370, 101.

*J. Am. Chem. Soc.* **2020**, 142(19), pp.8561

*Nature*, **2020**, 579, 379

*Nature*, **2020**, 583, 237

*FineFanta*, **2021**

**VR<sup>2</sup>C**, *J. Chem. Educ.* **2021**, 98, 2720.

*Joule* **2021**, 5, 495.

**AI-EDISON**, *Sci Adv.* **2022**, eaab2626.

*Nat. Sci. Rev.* **2022**, 9, mvac190.

*Chem* **2022**, 8, 3202.

*AIM*, *ACS Nano* **2022**, 16, 1600.

**HAORI**, *Digital Discovery*, **2022**, 1, 375

*Nature Chemistry* **2022**, 14, 1311.

*FineFanta* **2023**

*Beam Therapeutics* **2023**

*Strateos & Lilly* **2023**

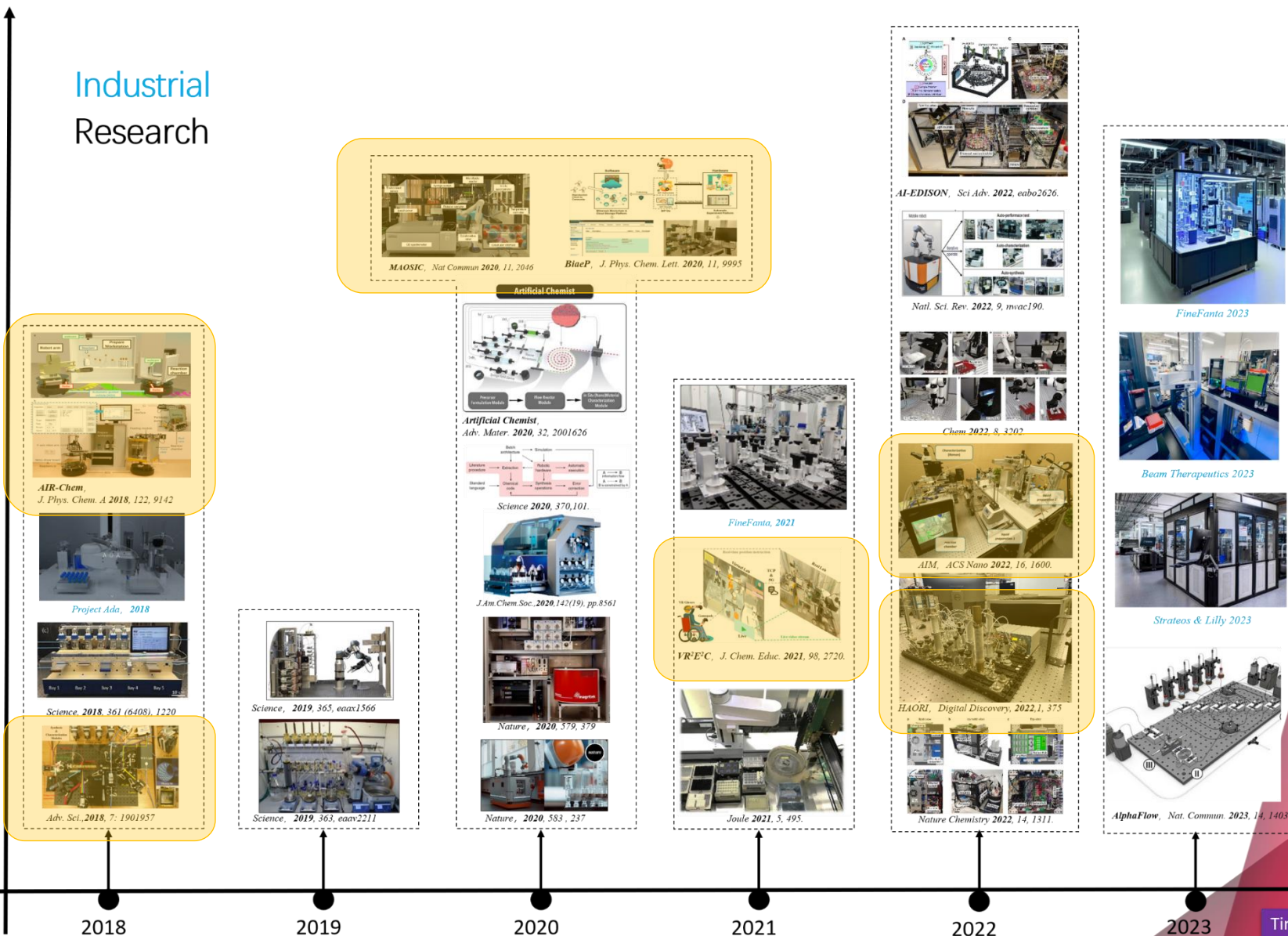
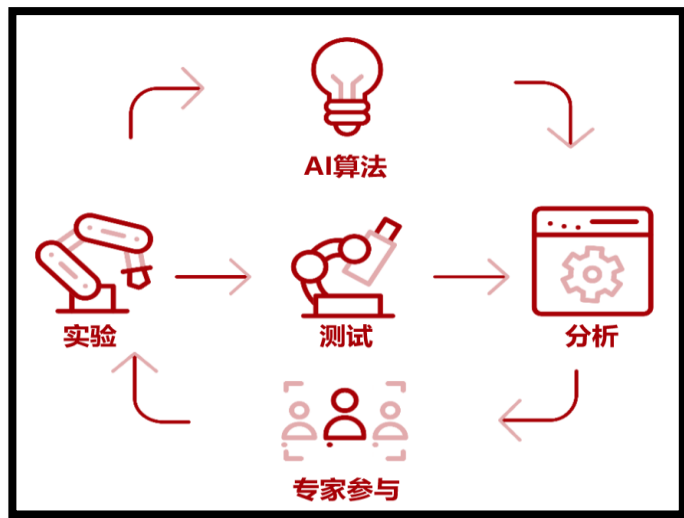
*AlphaFlow*, *Nat. Commun.* **2023**, 14, 1403.



# AI和机器人的化学实验室：2017年以后

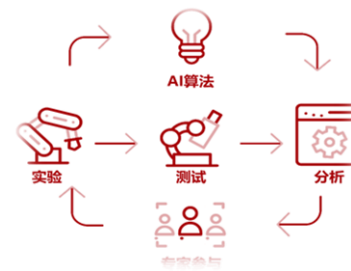
- AI技术大幅应用于机器;
- 覆盖更多的材料化学实验;
- 面向高校的课题组实验室。

Industrial  
Research



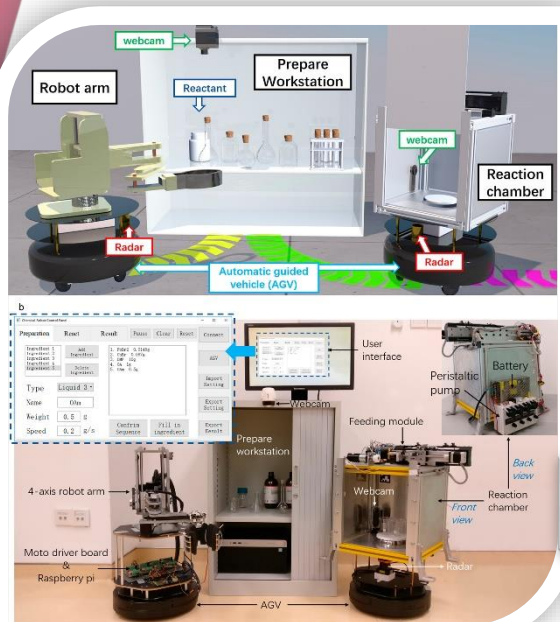


# AI和机器人的化学实验室



香港中文大學(深圳)  
大湾区国际碳中和研究院

## AIR-Chem (2017)



THE JOURNAL OF  
PHYSICAL CHEMISTRY A

Cite This: *J. Phys. Chem. A* 2018, 122, 9142–9148

Article

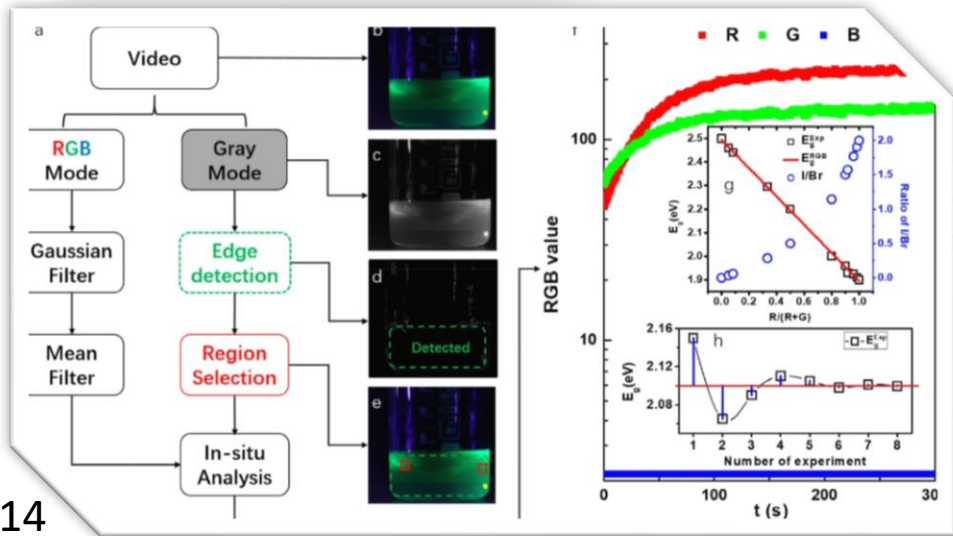
pubs.acs.org/JPCA

## AIR-Chem: Authentic Intelligent Robotics for Chemistry

Jiagen Li,<sup>†</sup> Yihua Lu,<sup>†</sup> Yanheng Xu,<sup>†</sup> Chongfeng Liu,<sup>†</sup> Yuxiao Tu,<sup>†</sup> Shuqian Ye,<sup>†</sup> Haochen Liu,<sup>‡</sup>  
Yi Xie,<sup>\*,§,¶,||</sup> Huihuan Qian,<sup>\*,†,||</sup> and Xi Zhu<sup>\*,†,||,||</sup>

Received: November 2, 2018

Published: November 5, 2018



Setup the parameter used  
in experiment on interface





# AI和机器人的化学实验室

AIR-Chem(2017)

MAOS (2018)

BiaeP (2019)

MAOSIC (2020)

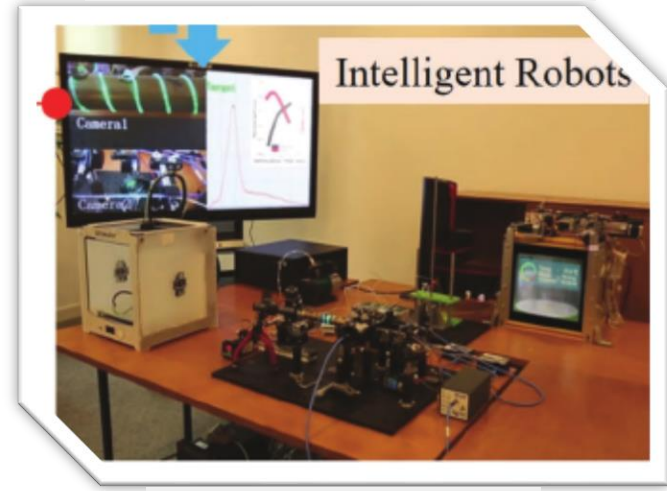
VR<sup>2</sup>E<sup>2</sup>C (2020)

AIRS-Lab (2021)

AIM (2022)



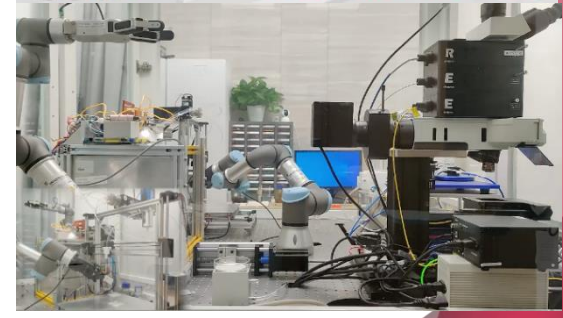
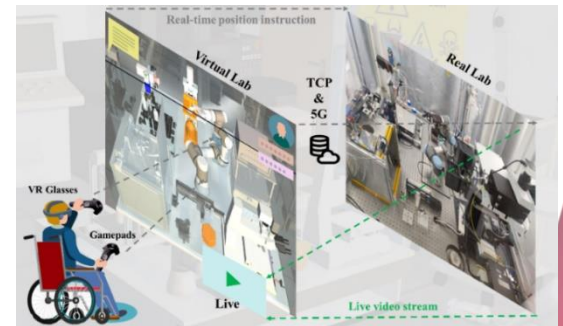
2017 AIR-Chem



2018 MAOS



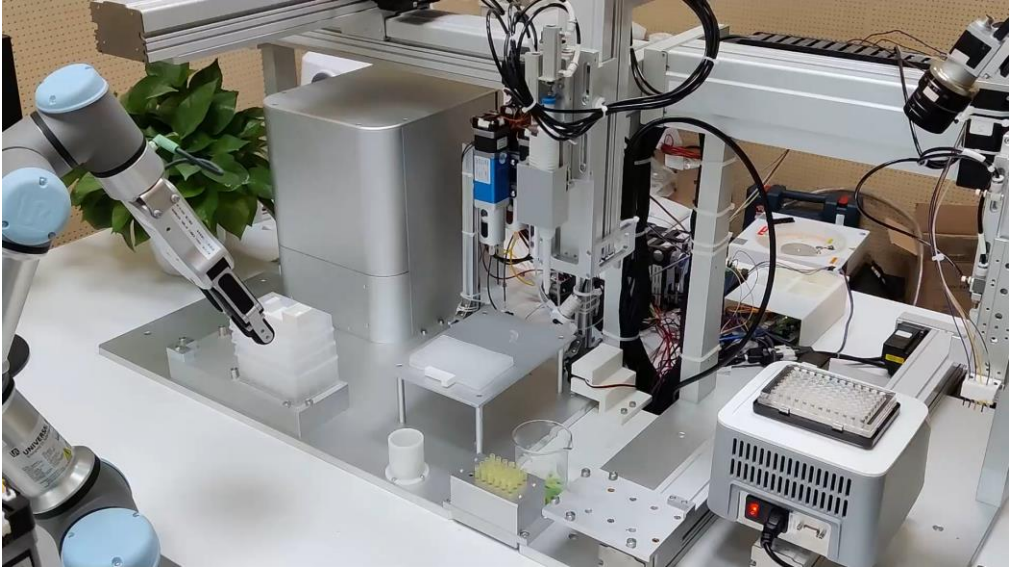
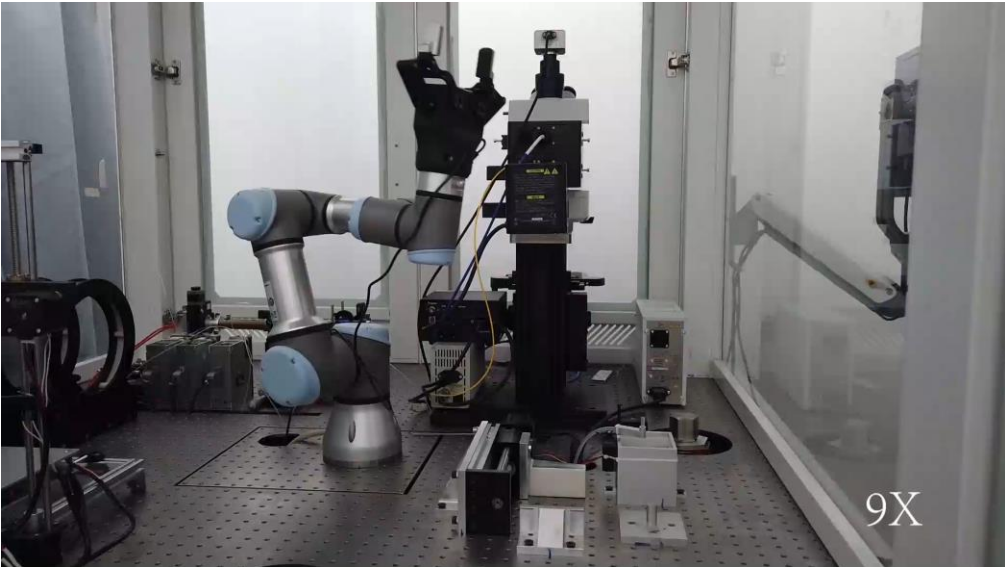
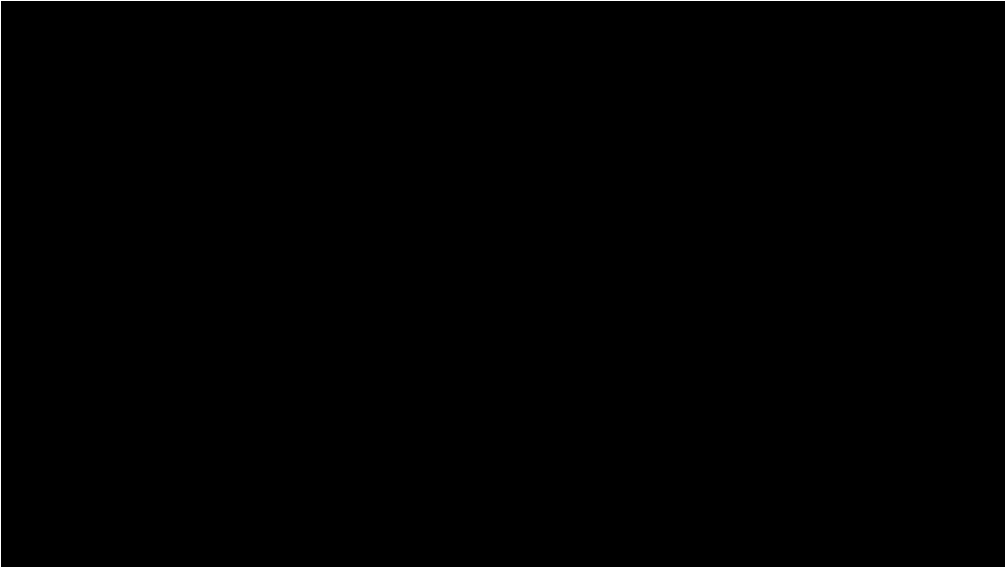
2019 BiaeP



2020 VR<sup>2</sup>E<sup>2</sup>C



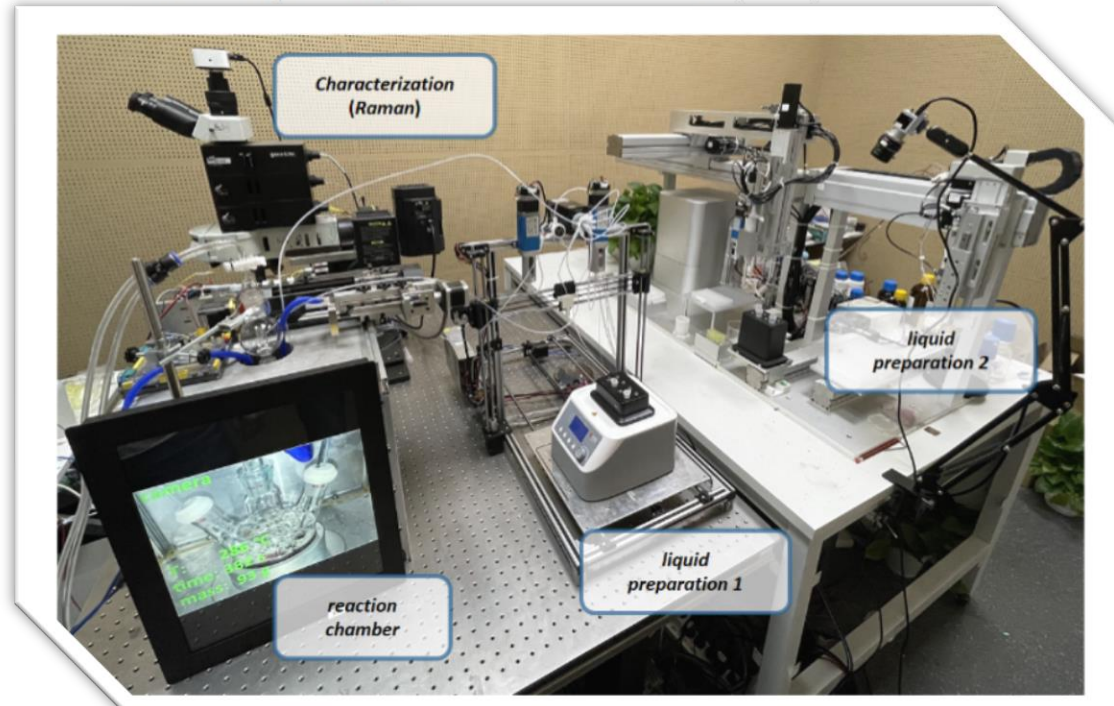
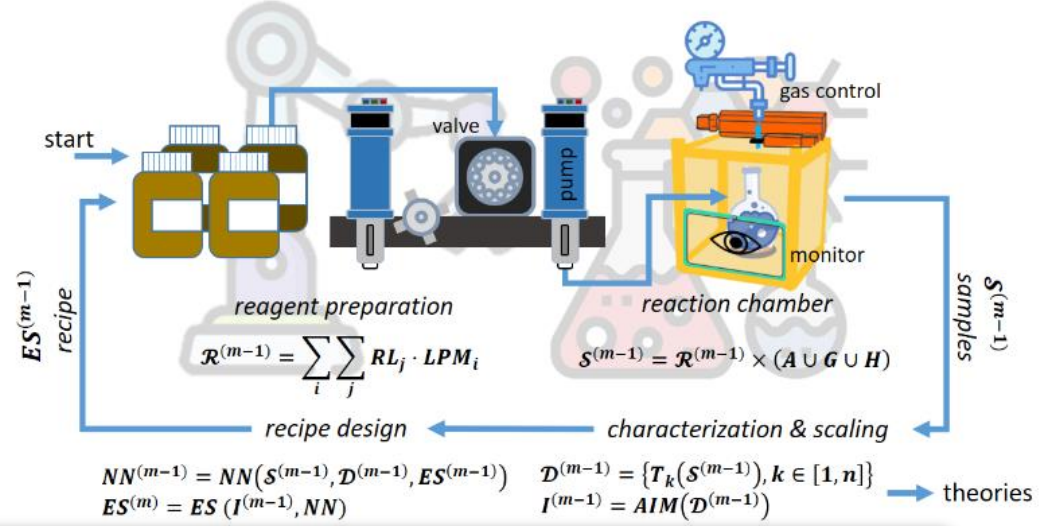
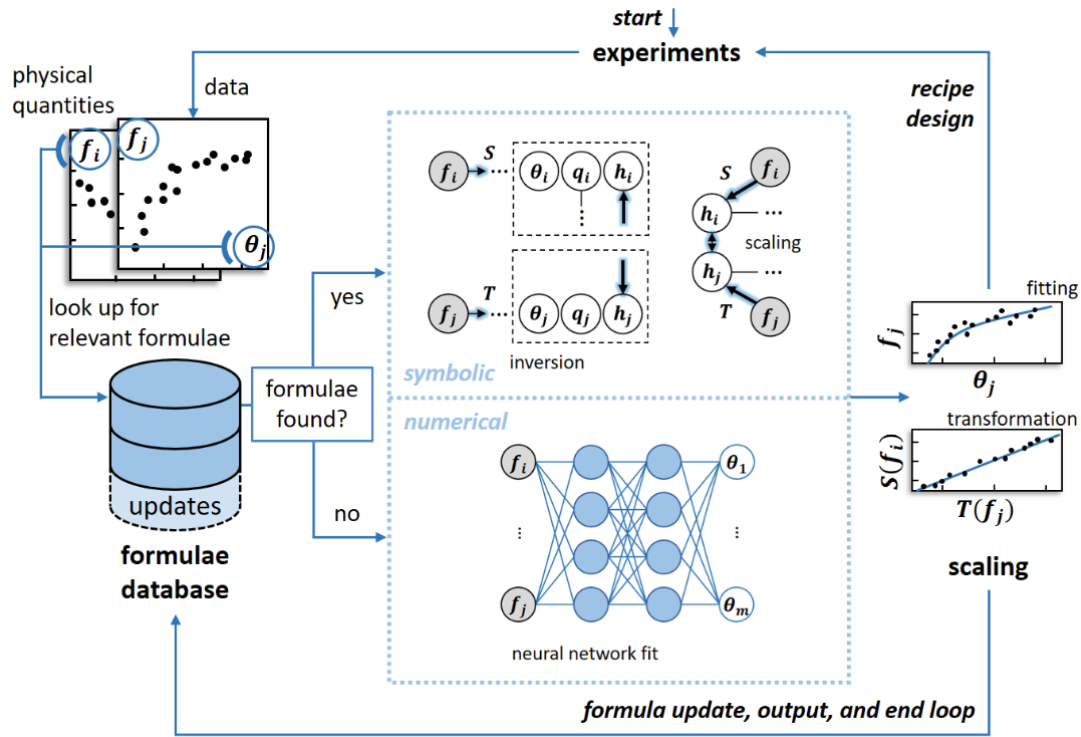
# AI和机器人的化学实验室(2020)







# AI和机器人的化学实验室 (Authentic Intelligent Machine)





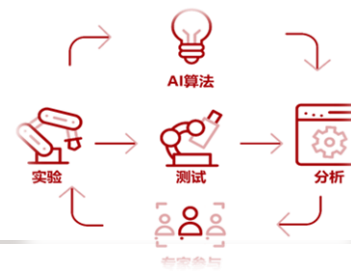


# AI和机器人的化学实验室(2021)





# AI和机器人的化学实验室(2021)

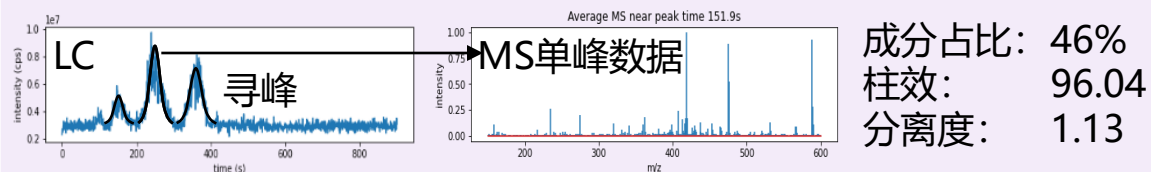


## LC/MS分析产物结构

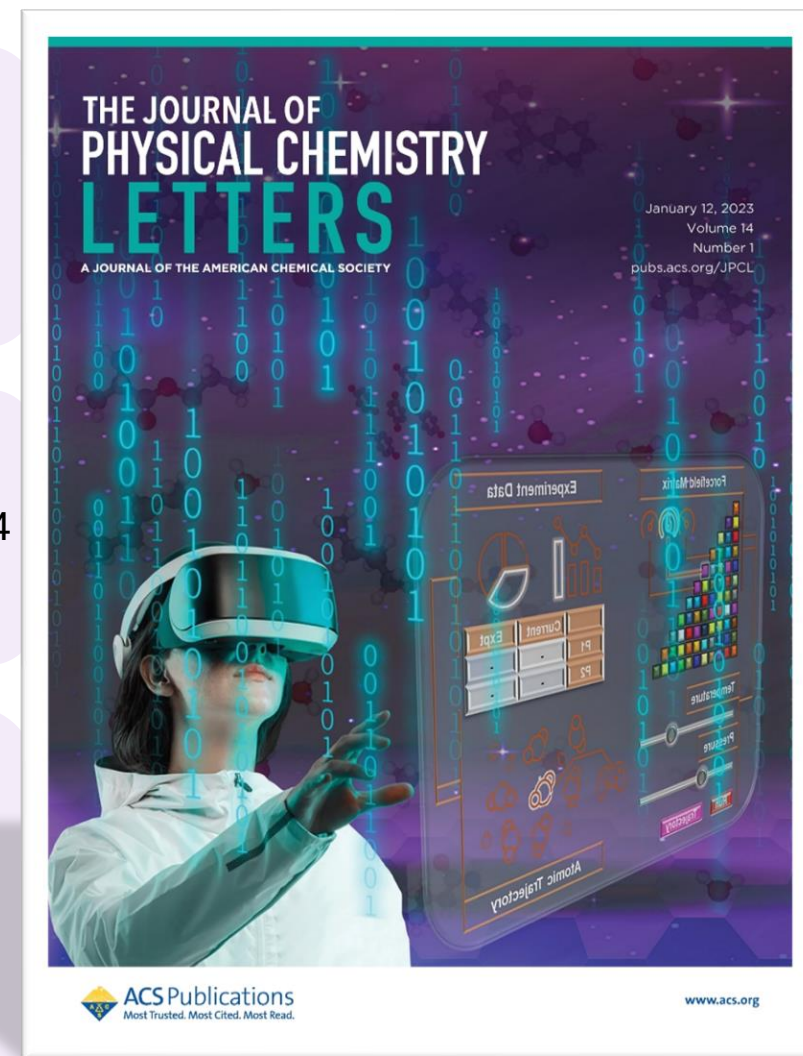
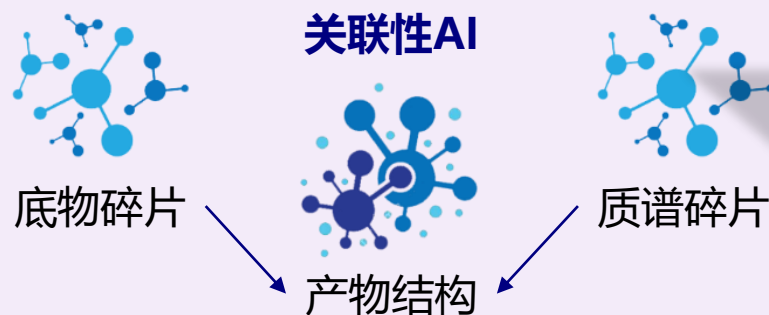
- ✓ 高通量实验数据生成  
快速测试实验方案  
数据集可用于模型更新



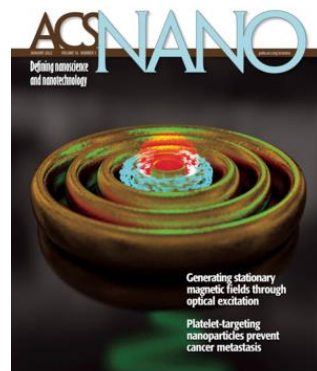
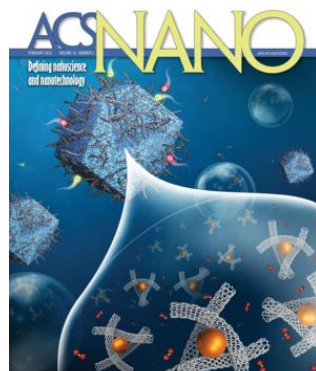
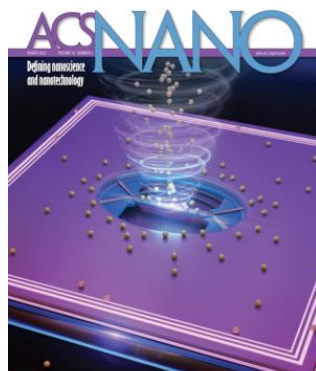
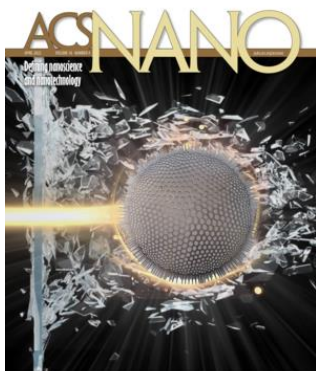
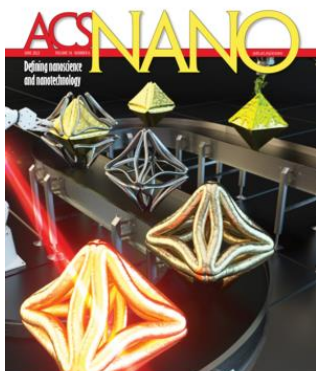
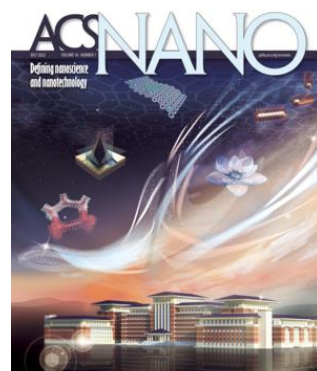
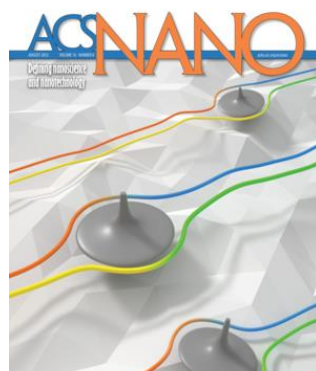
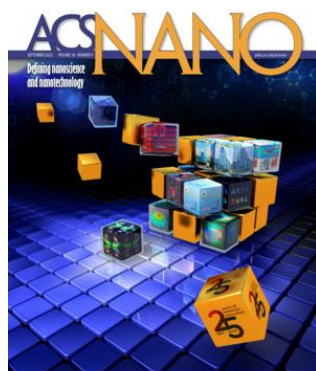
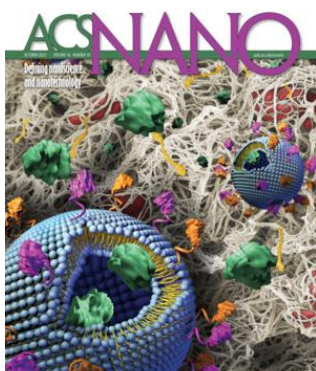
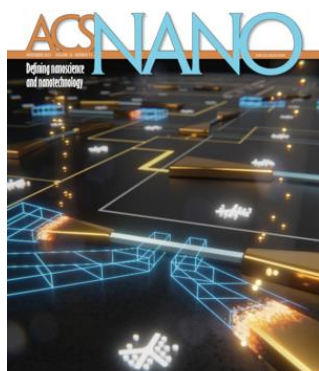
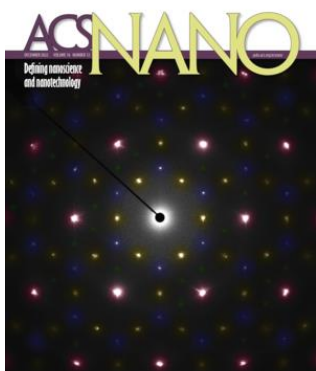
- ✓ 自动计算谱图指标  
判断数据质量  
分析单峰碎片构成



- ✓ AI模型预测产物结构  
关联性张量神经网络



## ACS Nano 2022 讲了什么?



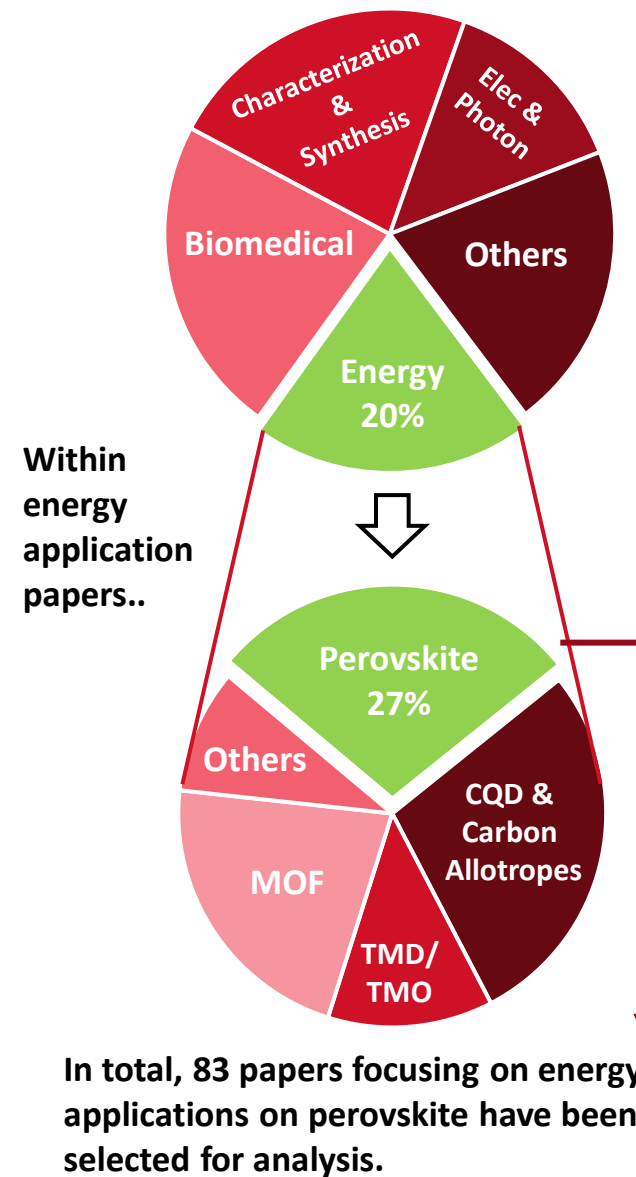


# 期刊合订本阅读

Reading in Bundle Volume

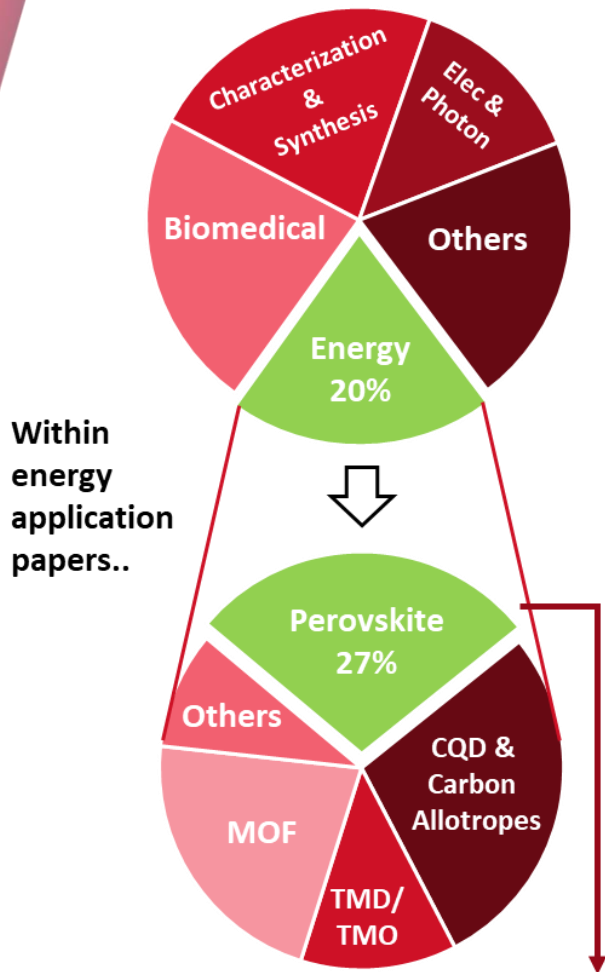


1724 review / research papers



# 实验验证

Data validation



Within energy application papers..

In total, 83 papers focusing on energy applications on perovskite have been selected for analysis.

## Idea Generation

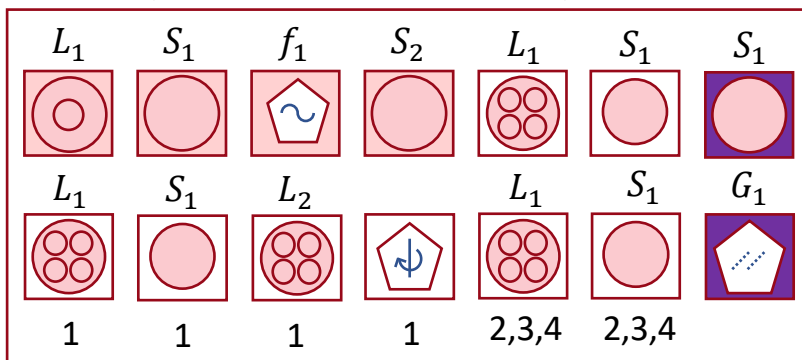
Merge similar cases

Precipitation Method in ACS Nano:



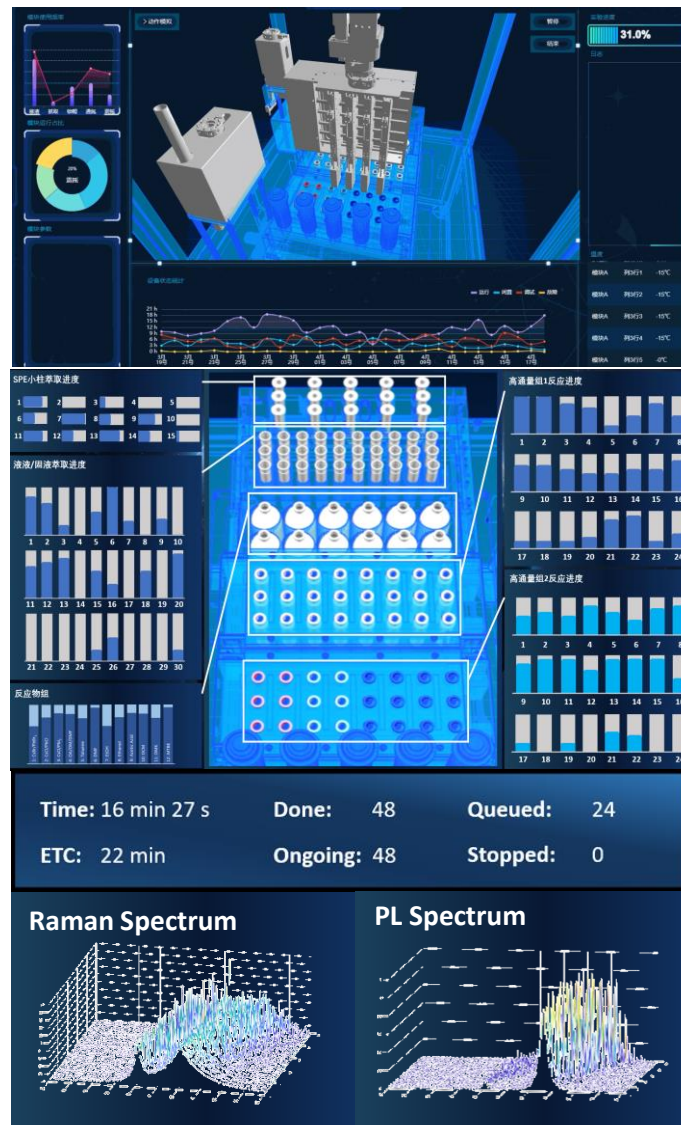
Precipitation for high-entropy (FAMACPX) perovskite

Extend parameters in various spaces



One Idea include all cases  
One paper > All papers in this field

## Idea Execution



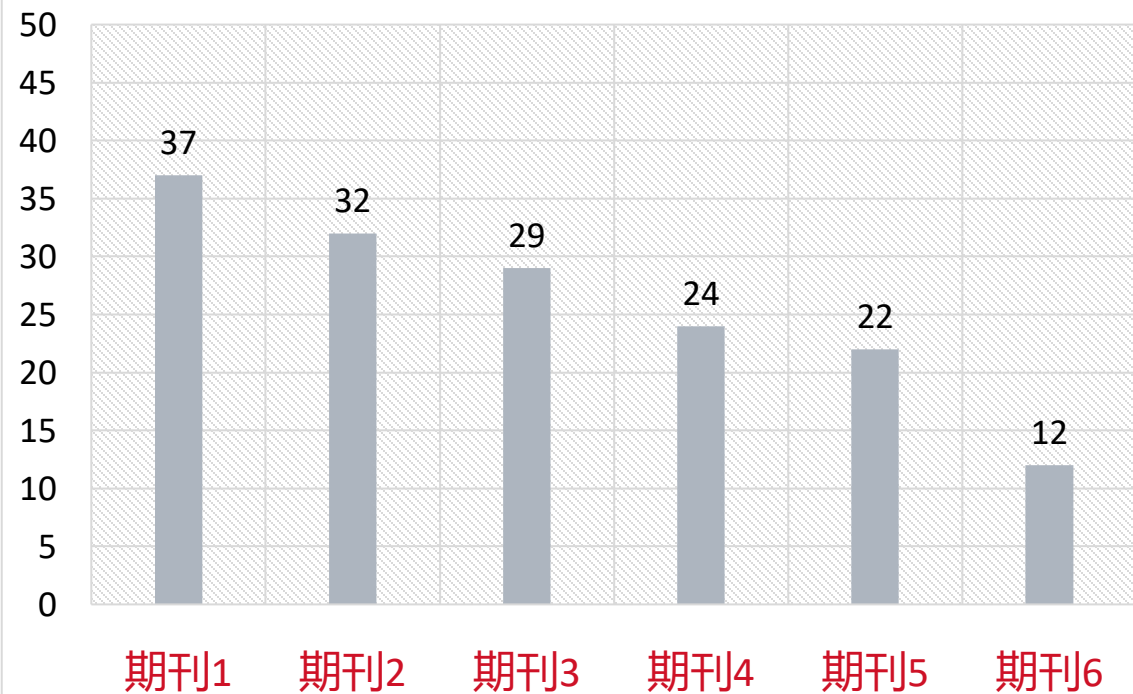
Automation high throughput Experiments  
Testing all the missing cases in idea generation



# 实验验证

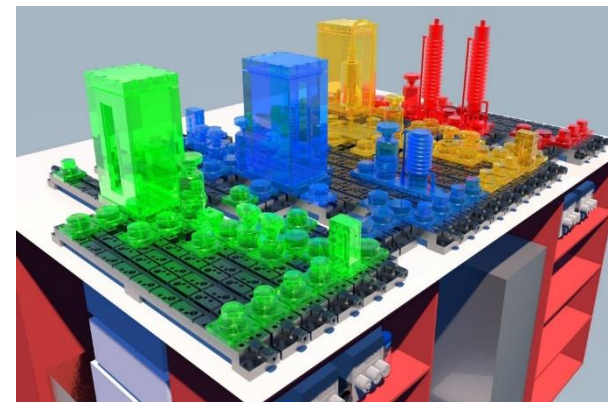
Data validation

## 相关期刊中量子点论文复现成功率 (%)



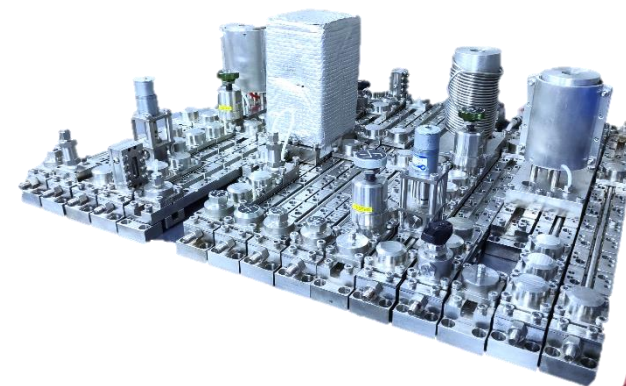
复现过程:

实验步骤输入数字孪生实验平台



↓ 控制

现实实验平台合成并表征测试





# 总结

Conclusion

**KnowledgeWorks 为读者提供了全新的服务：**

- 1. 广泛的学科资讯；**  
(学科日报，周报，月报，合订本)
- 2. 深度的知识挖掘；**  
(论文引用传承算法，AI问答系统)
- 3. 客观的论文评定；**  
(基于论文内容的评定算法，发表前评定)
- 4. 完整的实验验证**  
(驱动化学实验机器人的功能，验证/产生实验数据)
- 5. 独有的创新想法**  
(根据大量论文数据进行猜想假设，辅助创新)
- 6. 低碳的科研方式**  
(定量计算论文的碳排放量，通过AI机器人提供低碳解决方案)

## KnowledgeWorks

# CONTENTS

## 01. 传统的科研模式

The Traditional Research Paradigm

## 02. AI-Supervisor

Prosperous AI Publishing Service

## 03. 推动机器人科学研究

Towards the Sustainable Prosperity

## 04. 总结和致谢

Conclusion and Acknowledgement



# 致谢

Acknowledgement



张甲 馆长,  
香港科技大学 (广州)



香港科技大学 (广州)  
THE HONG KONG  
UNIVERSITY OF SCIENCE AND  
TECHNOLOGY (GUANGZHOU)

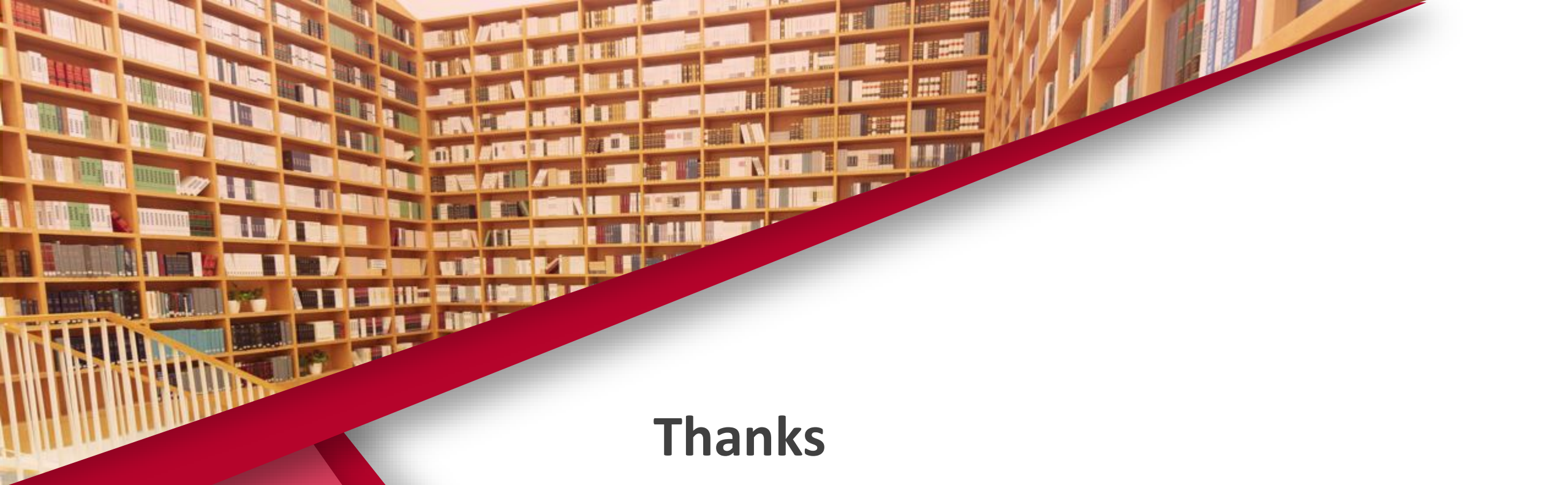
图书馆  
LIBRARY



香港中文大學 (深圳)  
The Chinese University of Hong Kong, Shenzhen

图书馆  
Library





Thanks  
**希望对您能有帮助**

Hope it can be helpful



香港中文大學(深圳)  
The Chinese University of Hong Kong, Shenzhen