



# CCL2018 学生研讨会

---

## 如何应对科研中的困难与挑战 -兼科研体会杂谈

万小军

北京大学计算机研究所

<http://www.icst.pku.edu.cn/lcwm/wanxj>

2018年10月19日



# 如何应对科研中的困难与挑战

---

## ➤ 简单的答案

重在坚持、努力克服



# 影响科研成败的因素

---

- 环境
- 能力
- 投入
- 运气



# 影响科研成败的因素

---

## ➤ 环境

- “老板”的水平和态度
- 实验室氛围
- 硬件设备



# 影响科研成败的因素

---

## ➤ 能力

- 调研能力
- 学术思维能力
- 编码实现能力
- 沟通能力
- 写作表达能力



# 影响科研成败的因素

---

- 投入
  - 时间
  - 效率
  - 思想



# 影响科研成败的因素

---

## ➤ 运气

- 课题的选择
- 实验结果
- 论文投稿



# 困难无处不在

---

- 导师不靠谱，关系不融洽
- 选题偏差
- 基准方法不能复现
- 想不到创新点 / 创新点冲突
- 实验结果不理想
- 写作太烂
- 评审人不靠谱
- 遇到不正当竞争
- ...





# 困难无处不在

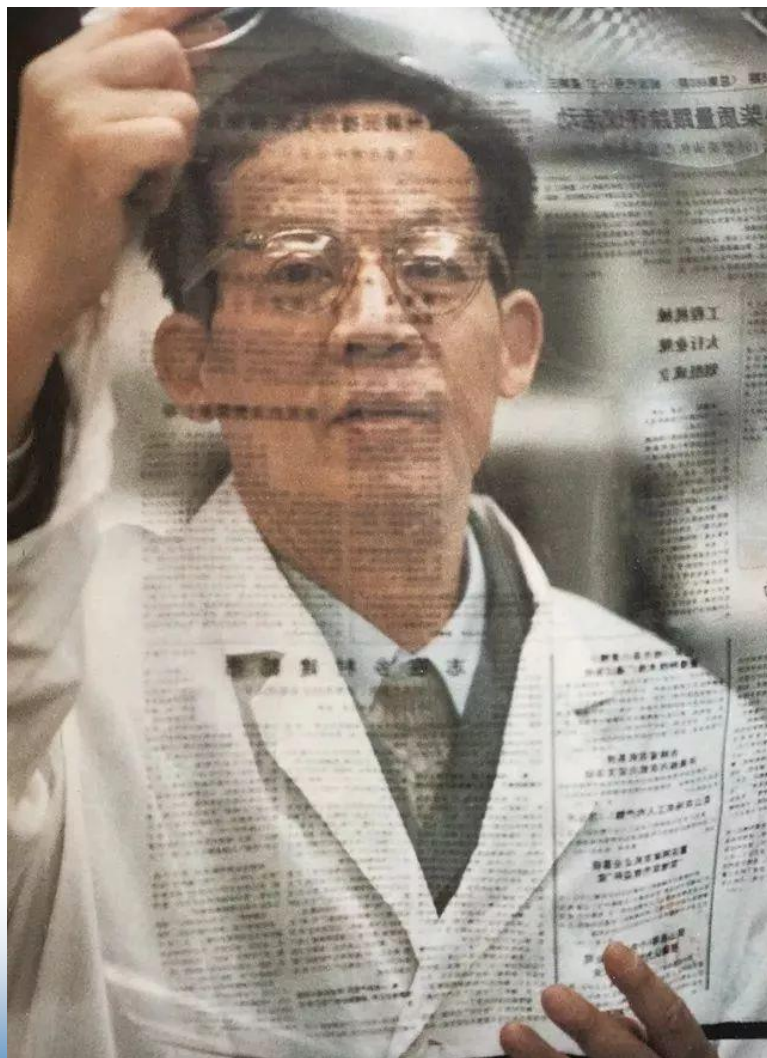
---

- 战略上藐视困难
- 战术上重视困难



# 举例 - 王选先生

---

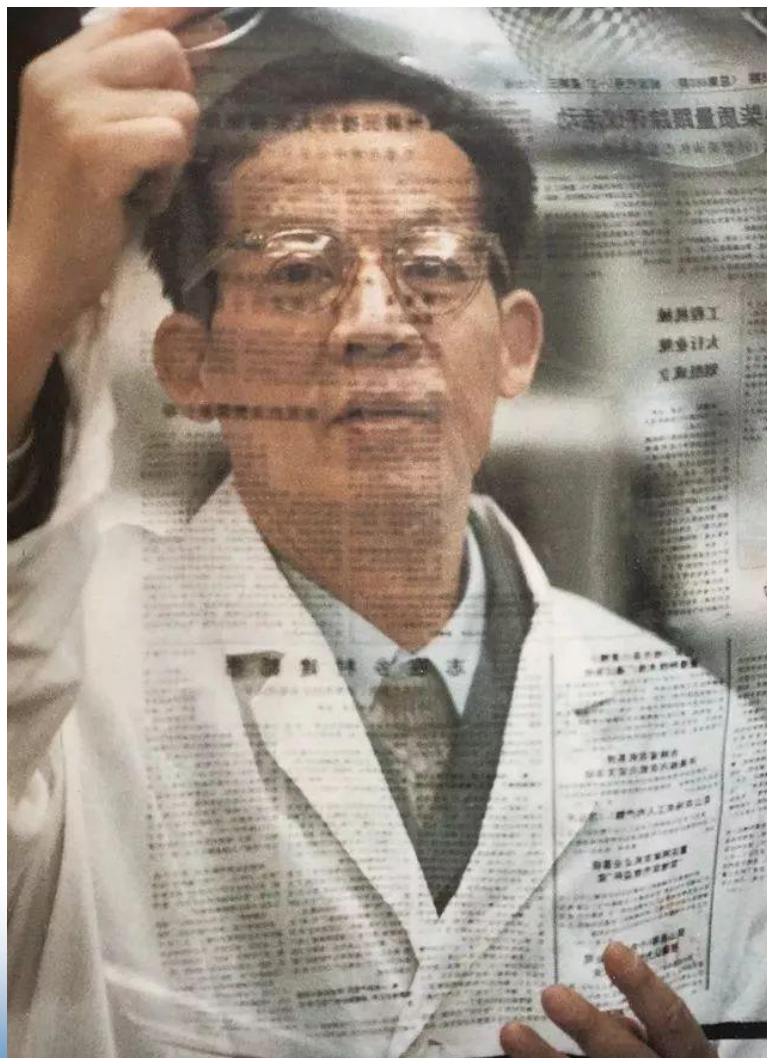


代表性成果：汉字激光照排技术与系统

使汉字印刷术——“告别铅与火，迎来光与电”



# 举例 - 王选先生

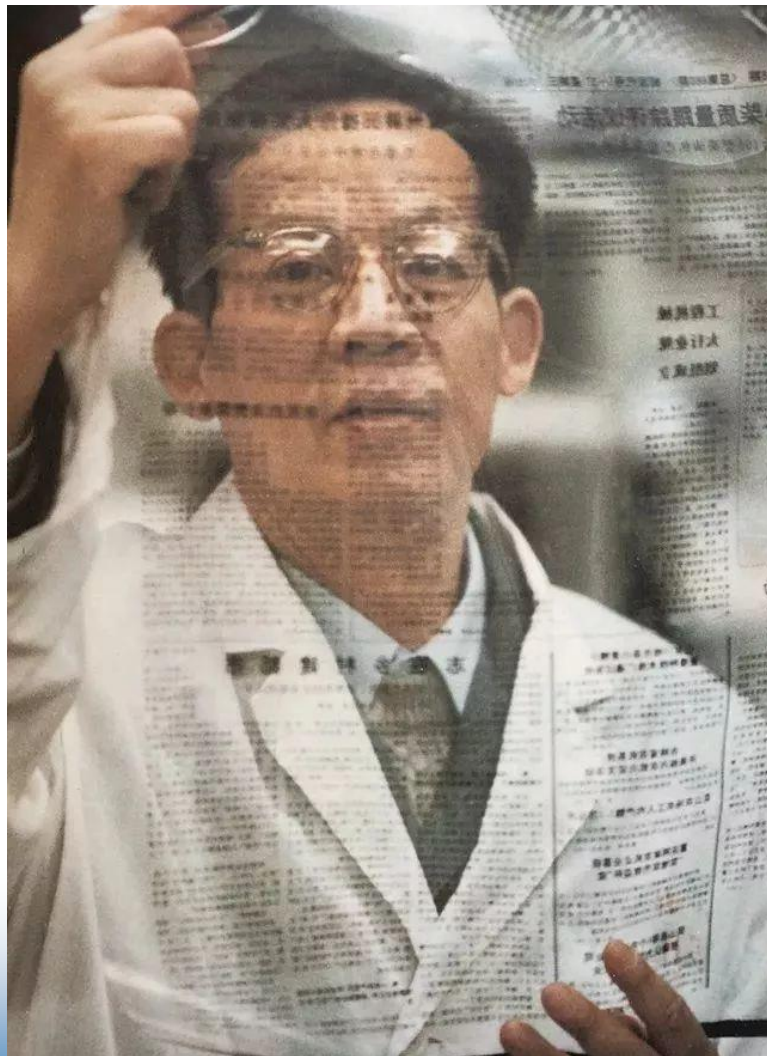


困难与挑战:

- 1975年38岁时才开始研究



# 举例 - 王选先生



## 困难与挑战:

- 技术难点：汉字字形庞大的信息量

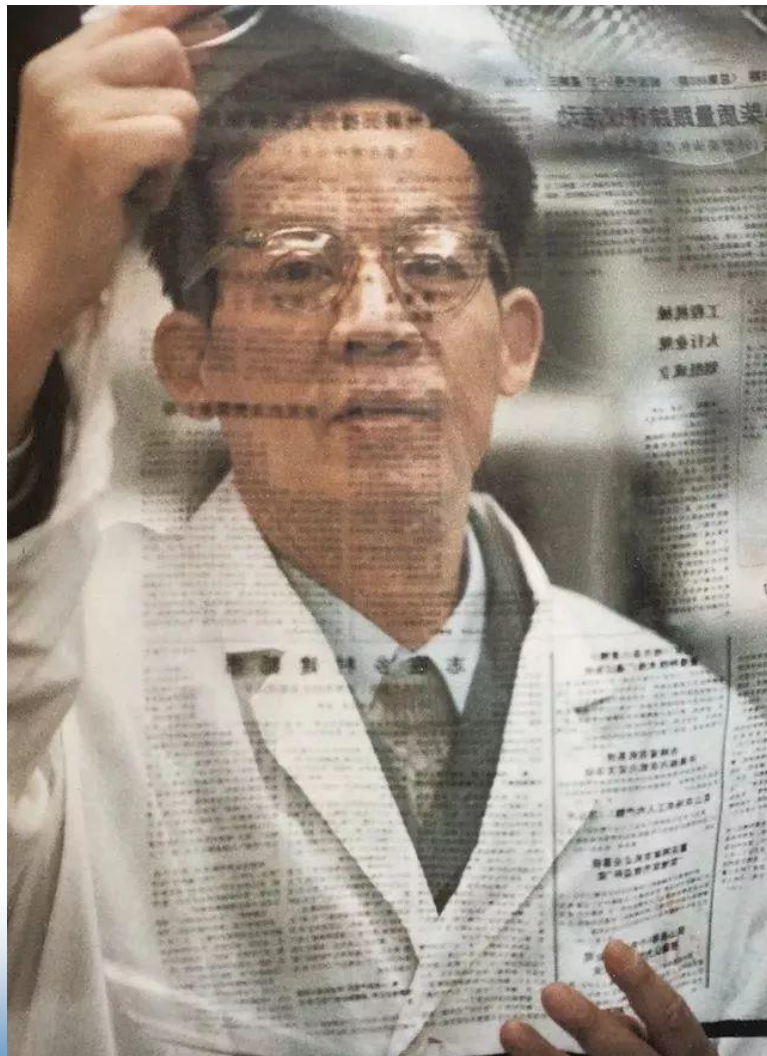
英文只有26个字母，而汉字的字数是英文的数百倍，还有多种字体，如宋体、黑体、楷体等十余种。每种字体又有十余种不同大小的字号。全部算起来，存数量要达到数百亿位，而当时能够使用的DJS-130计算机的磁心存储器，最大容量只有64KB。

“轮廓+参数方法”





# 举例 - 王选先生



## 困难与挑战:

- 项目获取

“748工程”办公室主任郭平欣指定了10个字，还特别加了个“湘”字，请王选和同事们做信息压缩还原点阵的模拟实验。

王选、陈堃铎用一个月的奋战，顺利完成了任务，不久，由郭平欣签署的“汉字精密照排系统”的研制任务，正式下达给北京大学。

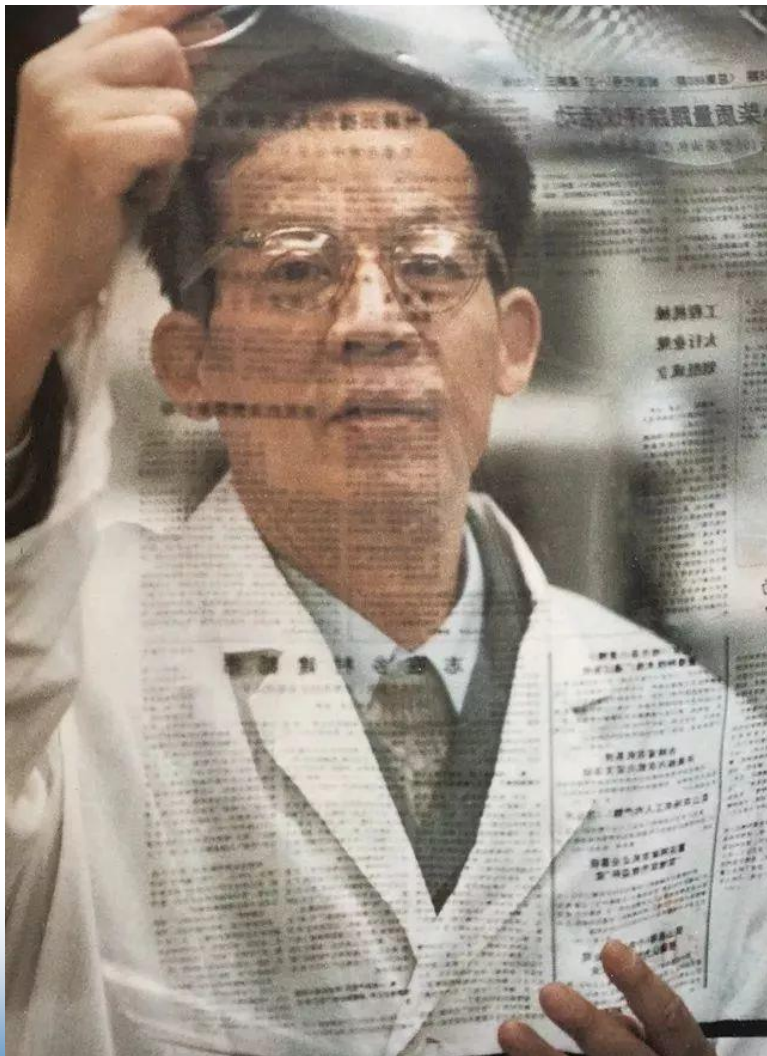


# 举例 - 王选先生

困难与挑战:

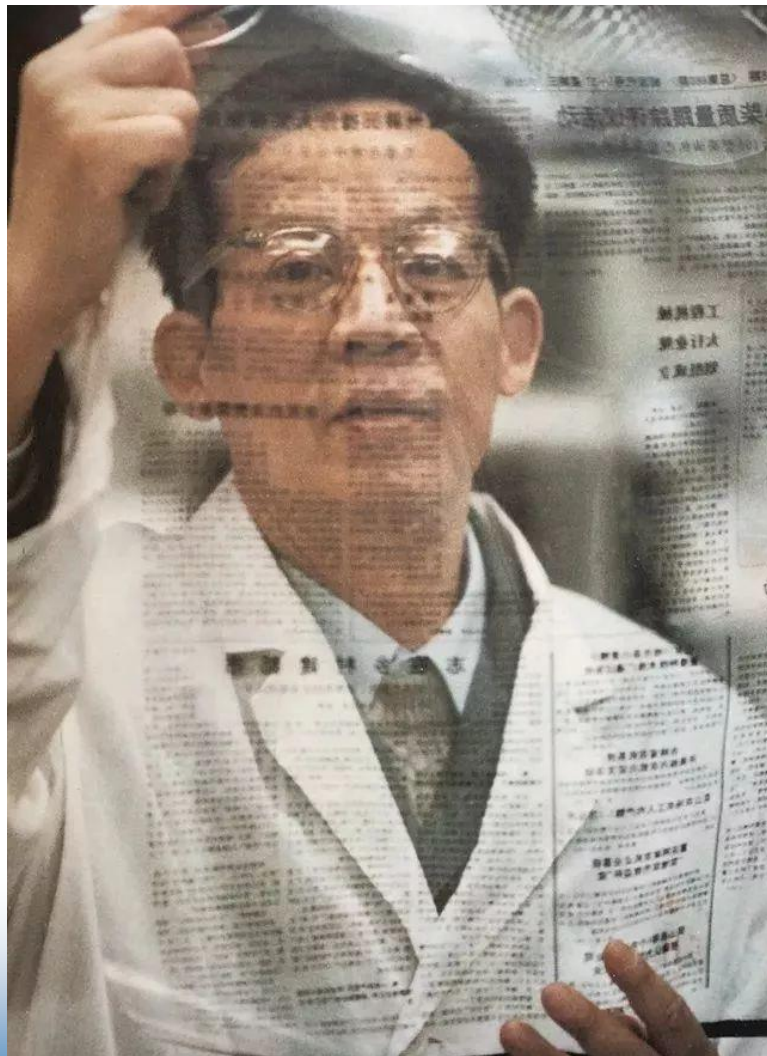
- 1979年报纸样张输出

工作了近两年，期间经历人员流失，任务繁重、机械故障等等问题。海外公司准备大举进华。





# 举例 - 王选先生



困难与挑战:

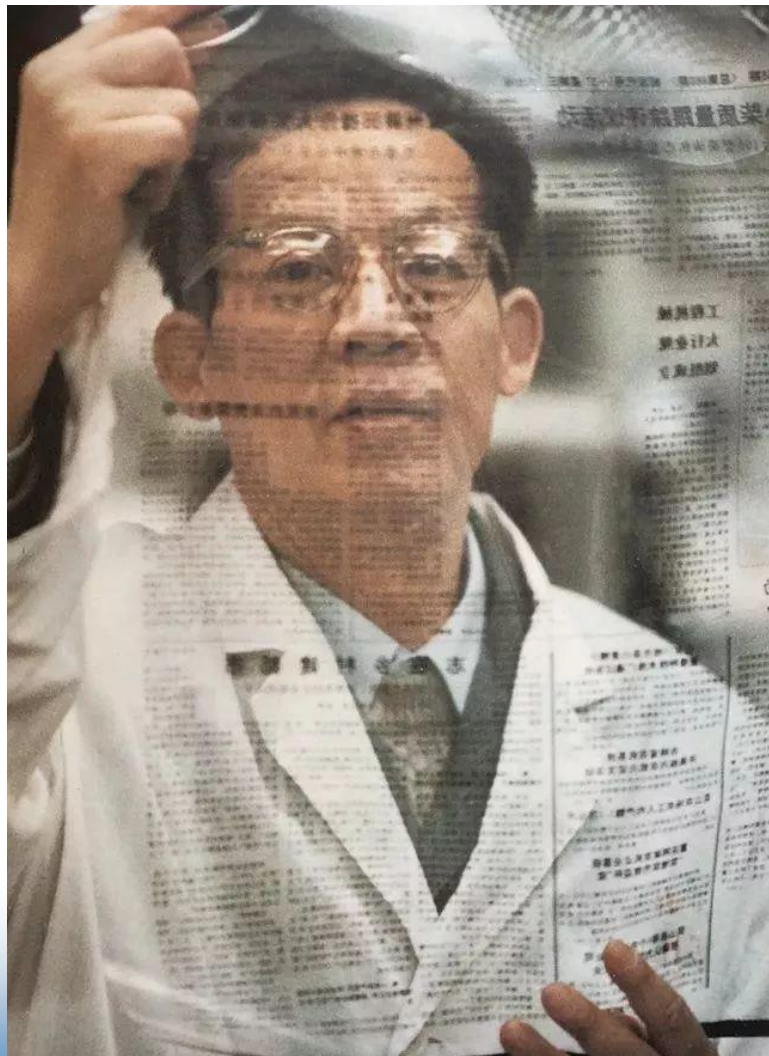
- 应用不顺利

1984年8月，国产激光照排系统在新华社进行了试用，由于问题重重，频繁出现故障，引起了许多人的质疑。业内人士纷纷引进外国产品，国产系统遭遇冷嘲热讽，“搞出来也是落后的”“‘748’不如叫‘气死吧’”





# 举例 - 王选先生



困难与挑战:

- 应用不顺利

王选和同事们又花费了5年的时间，不断攻克更新，于1987年，装备两块专用芯片WA、WI的华光IV型系统问世。此后，王选又先后设计出第五、第六代照排控制器，使来华销售照排系统的外国公司全退出了中国。





# 杂谈 - 科研选题

---

## ➤ “热门” vs. “冷门”

- 例如 社交媒体挖掘 vs. 指代消解
- “热门”题目易出小成果，但存在盲从的现象
- “冷门”题目需要坚持，耐得住寂寞

## ➤ 科研选题 vs. 科研项目

- 无强相关性
- “国家需求”不完全代表科研方向

## ➤ 自己感兴趣、愿意钻研的题目就是好选题



# 杂谈 - 研究方法

---

- 不存在对每个人都有效的研究方法
- 每个人要根据自己的特点找到适合自己的方法
  - 多交流 vs. 独立思考
  - 先读大量文献 vs. 自己先思考问题
  - 先思考再动手 vs. 边动手尝试边思考



# 杂谈 - 良好科研素质

---

- 兴趣
- 善于学习
- 积极思考
- 勤于动手
- 不怕失败 (信心)
- 诚实公正



# 杂谈 - 几点建议

---

- 思考无处不在
- 随时记录“灵感”
- 拒绝无休止的思考
- 不耻下问
- 追求应该越来越高
- 努力克服不利环境
- 客观看待同行成果



# 杂谈 - 研究成果

---

- 通过研究某个问题，得到以下研究成果
  - 高质量的学术论文
    - 问题+数据+方法+结果
  - 有意思的原型系统
  - 代码和数据



# 杂谈 - 关于论文

---

## ➤ 论文目的

- 为进一步的研究提供基础，一般情况下非完全解决某个问题；

## ➤ 论文和实用有一定的差距

- 追求目标不同：效果 vs. 效率
- 测试数据 vs. 现实数据

## ➤ 论文和研究的关系

- 论文为研究服务，反之不成立：论文是研究的一个重要步骤

## ➤ 对论文的认识是逐步提高的

- 数量=》档次=》影响力



# 杂谈 - 论文创新性

---

## ➤ 创新性从何而来

- 提出新问题并初步解决：思考+灵感
- 揭示某个问题的基本规律：灵感+深入分析总结
- 提出新的解决方案/模型
  - 相似问题映射：偷天换日，移花接木
- 改造原有方案/模型
  - 分析原有方法的缺点，对原有模型进行合理改造
- 以上都离不开经常性的深入思考

## ➤ 通过交流、分析加以确认



# 杂谈 - 高质量论文

---

- 新问题，合理的方法及结果
- 老问题，新颖的方法及满意的结果
- 老问题，对已有方法的独到性的比较分析





# 杂谈 - 对论文的几个误解

---

- 论文成果一定能实用
- 论文完整解决了问题
- 论文是写出来的
- 论文被拒全是审稿人的错
- 论文等于科研



# 杂谈 - 国内优秀学生水平与趋势

---

- 研究生、本科生发顶级会议论文已成常态
- 论文奖项的获得：SIGIR、CIKM、ACL、COLING、AAAI、IJCAI的最佳(学生)论文奖
- 今后将更关注研究工作的系统性与影响力，而不是论文数量，学术评价更加注重代表性工作的价值与影响力



---

不要迷信权威！  
“小人物”也能做出大成就！

谢谢