

使一份信念破界而行

○ 矣晓沅（2012 级计算机）



矣晓沅（左 1）与合作者香港中文大学 Huang 教授（右 1）、矣晓沅团队两位科研实习生白钰卓（右 2，2018 级计算机）和清华硕士留学生 Pablo 合影

一种理想沐光而生

十年说来也短，短到不过几番朝暮、一程攻坚、些许论文。回顾十年前，许多画面奔涌而来，2016年发生了许多事。前一年底我刚获得本科生特等奖学金，成为当时唯一获此殊荣的重度残障学生。那年我选择在清华自然语言处理实验室（THUNLP）继续读博，导师孙茂松老师对我说，“做学问不能只为论文，应兼济天下，做对社会有意义的研究”；那年建筑学院的邵磊老师与我长谈，告诉我全球有约10亿残障者，中国有8500万；我也第一次理解到包容性设计与理念将如何帮助构建一个更自由、更开放、更包容的理想社会。正如《世界人权宣言》所言，“人人生而自由，在尊严和权利上一律平等”。这些微小的愿景汇聚成光，照亮了一个遥远却又似乎触手可及的“理想

国”，而它成为了我在接下来10年砥砺前行动力。那时我从没想过，这份理想会与日后我所从事的研究工作不期而遇。

一怀坚守免于彷徨

十年间，我目之所及的风景在不断变迁。从清华园到微软，从线下学生组织到前沿AI模型。“90后”的幸运来自于赶上了一个巨浪奔涌的时代，而不幸又起始于光阴流转间的彷徨。身边有太多的人囿于相互比较，从GPA到论文数量，从科研奖项到年薪等级，其间的挣扎逐渐剥蚀了我们的主体性。我的幸运在于对消弭偏见、构建公平的执着追问。

最初的实践便来源于校园内的亲身体验。2016年，伴随着清华大学无障碍发展研究院的成立，我们在其指导下同时发起成立了清华大学学生无障碍发展研究协会，发动同学们一起做调研、办活动、写倡议，开展清华校园无障碍地图的绘制，帮助残障学生在校园里通行；将“观念无障碍”的微光，从清华照向公益组织、北京的中小学以及更广阔的边界。直至今日，越来越多的高校能为残障学生提供支持，让他们有尊严地求学。我曾将那段时光的努力称为“微小而伟大的事”——每一分努力如雨滴，汇聚可成江海。我坚信改变始于具体的行动，始于对每一个“人”的看见。

2021年我博士毕业加入微软亚洲研究院后，这份信念也随我继续迁徙。彼时AI研究的热点是预训练，但我们却看到了

另一堵亟待打破的“高墙”。AI在变得更大、更快、更强的同时，也对包括少数群体在内的人类产生了不可避免的风险：歧视性语言、仇恨言论、虚假信息、隐私泄露、欺骗与谄媚……这些风险如暗礁潜藏，威胁的不仅是AI的技术迭代，更是涉及整个社会的平等与和谐。于是，我的研究焦点将无障碍从改善具体的物理环境延伸至数字世界。在过去五年中，我们聚焦于AI安全、AI对齐、风险评估等方向，开展结合心理学、社会科学、法律等多领域的交叉学科研究，致力于构建公平、包容、能提升人类福祉的AI，我们称之为社会责任人工智能（Societal AI）。这项工作的本质，与当年在清华推动“观念无障碍”一脉相承：都是试图在存在巨大差异的主体间，即残障人士与环境、人类与AI间，建立理解、尊重与和谐的协同机制。

在本职的研究工作之外，我也加入了微软亚洲研究院多元与包容（D&I）委员会，继续以志愿者的身份延续我对那个“理想国”的信念。我们发起了“AI for D&I”项目，鼓励用技术打破障碍；我们在微软全球黑客马拉松组织“Hack for D&I”挑战，见证了数百个充满善意的技术创意诞生，我们也将这些技术向善的努力传到了外部，在计算机领域、大学里、中小学之间，为他们埋下向善而行的理念。无论是打破物理空间中的偏见“高墙”，还是驯服数字世界里的算法“恶魔”，其内核都是对“人”的关怀，是对“何为美好、公正、包容社会”的求索。AI不应是加剧不平等的工具，而应成为普惠的赋能者。正如我在TEDxBeijing演讲中所言，要以科技之光照亮人机共生之路。这不仅需要算法创新，更需要我们将

人类历史中积淀的关于平等、尊严、互助的价值观，注入技术的血脉。

我们这一代人，有幸站在AI浪潮之巅，目睹它重塑一切。但人生的意义不仅是逐浪，还有航行的方向。我的十年，是从轮椅上的方寸之间，走向价值星海的探索之旅。真正的无障碍，终极目标是心智与价值的联通；而负责任的智能，是懂得对齐于人类共同福祉的智能。也正是这份坚守让我在过去十年没有被所谓的浪潮裹挟，免于彷徨于物质的索求，得以始终不违内心所向，去做自己相信正确的事。

一份信念破界而行

十年说来也长，长到坚守心中所信需要巨大的毅力与担当。这十年，是一个从接纳自身到关怀他人，再到思考未来、不断向外探求又向内深耕的过程。我时常自问，将如此多时间投入公益与非主流的研究，是否耽误了所谓主业？但当看到基于我们努力的一点点改变时，我确信这两条轨迹是共鸣的弦。现实是理想主义的坟场，妥协是它最廉价的悼词，但我却愿将它谱成一首盛大的挽歌。如今回首过去十年，我很欣慰自己最大的成就不是发表的那些论文、拿到的若干奖项，而是自己能在一定程度上放下个人得失，为让这个世界变得更美好一点点，贡献的微薄之力。

我由衷地感谢过去十年所有因共同信念而汇聚的朋友，是你们让我相信即使生活充斥着遗憾与无奈，也总会有勇敢的人坚守内心闪烁的理想，用真诚让生命变得更加绚烂。这些微小的努力让一份信念破界而行，驶往我们心中的那个“理想国”。

（部分内容源自2012级建筑学院校友唐义琴收集整理）