

我们是 KMind，志在发明个人 AI 计算机

吴翰清（道哥）

2023 年 12 月 12 日

摘要

本文的主要贡献在于首次提出了：

- 1. 个人 AI 计算机的概念和 kOS 架构。
- 2. AI 精度的概念和数据脱水与浸泡技术。
- 3. ACT 的概念和可编程 AI 的实现。
- 4. AI 互联网的概念和信息反向流动的实现。

特别的，本文还明确指出了企业垄断数据对社会的危害，并从技术变革上提出了建立一个更健康网络经济环境的可能性。

本文长达三万字，读者可以选择感兴趣的部分自行阅读：

- 1. 对技术浪潮的独家剖析在“第 1 章 计算机的两次革命”。
- 2. 原创技术发明在“第 2 章 我们的答案：kOS-1.0”。
- 3. 我们的使命愿景和对未来的判断在“第 3 章 AI 互联网：连接所有个人 AI 计算机”。
- 4. 我们要推出的产品在“第 4 章 半个宇宙诞生：人工智能的中国方案”。

目录

我们是 KMind，志在发明个人 AI 计算机.....	1
摘要.....	2
第 1 章 计算机的两次革命.....	3
1.1 第一次革命：人人有台计算机.....	3
1.2 第二次革命：人人有台 AI 计算机.....	12
第 2 章 我们的答案：kOS-1.0.....	14
2.1 扩展的冯·诺依曼架构：为什么这是台计算机？.....	15
2.2 AI 精度：为什么仅有大模型是不够的？.....	19
2.3 ACT 的时代：既不是软件，也不是 App.....	23
第 3 章 AI 互联网：连接所有个人 AI 计算机.....	29
3.1 信息范式的改变：从“人找信息”，到“信息找人”.....	29
3.2 回归互联网的本质.....	31
第 4 章 “半个宇宙”诞生：人工智能的中国方案.....	33
4.1 半个宇宙，因你而不同.....	33
4.2 星伴，一生相伴.....	34
4.3 星伴互联.....	35
4.4 我们在和 OpenAI 在赛跑.....	37
致谢.....	40
附录：.....	41

两个月前，我和公司新来的实习生杨子乐进行了首次谈话，我对他说的第一句话是：“我们是一家伟大的公司”。子乐看了一眼公司里正在工位上埋头敲代码的十多个员工，面部肌肉开始扭曲，然后再也憋不住，捧腹大笑起来。他的反应把我也逗乐了，我马上意识到了他的想法，和他一起足足笑了五分钟才喘过气来。

我们是一家成立了仅仅 4 个月，只有 16 个员工，但是却很“伟大”的公司。伟大的公司不在于能赚多少钱，而在于我们的目标很伟大，一旦实现，世界将因此而改变。而我作为创始人，将确保公司会一直在实现这个伟大目标的道路上坚持下去。

第 1 章 计算机的两次革命

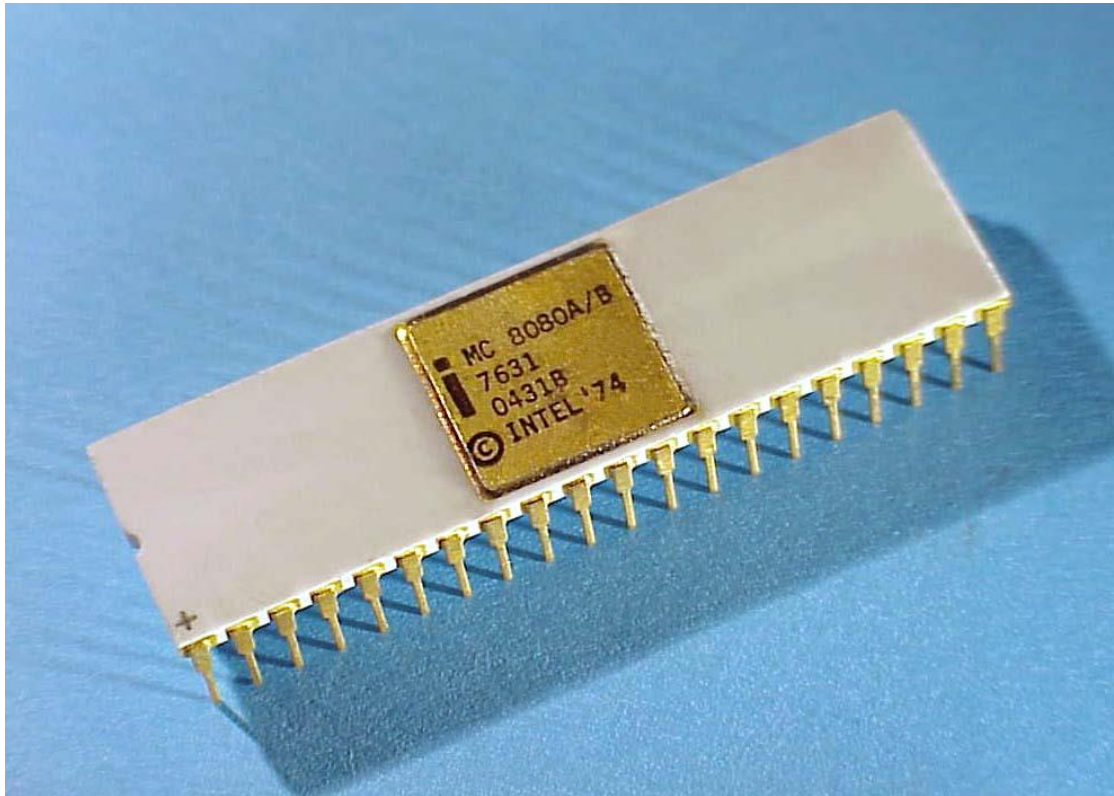
数字计算机自 1946 年诞生以来，在不到一百年的历史，有两次大的革命深深震撼着我的心灵。这两次革命都有一个共同的使命：让先进技术的大型计算机实现小型化和普惠化，使得人人拥有一台个人计算机。

1.1 第一次革命：人人有台计算机

1.1.1 个人计算机的发明

第一次革命发生在 20 世纪七八十年代，以苹果公司发明一体化的个人计算机为标志。在苹果公司发明个人计算机之前，计算机是只有政府、银行或大企业才买得起、用得起的大型专用设备，IBM 把计算机卖到了几百万美元一台。在那个年代，计算机是一个庞然大物，人们需要通过在纸片上打孔来编写程序，排队去机房操作大型计算机。计算机当时主要用于处理一些大公司、银行的账目数据，或者是政府、高校的一些科研数据。

此后有三项关键技术于 20 世纪七十年代出现，带来了改变的契机，成为个人计算机出现的基础。首先是在 1969 年，英特尔（Intel）公司接受了日本一家做计算器的公司 Busicom 的订单，要求制作一个用于计算的处理器，而此前 Intel 的业务仅限于存储器。当时 Intel 的董事长摩尔要求这个产品不能只服务于一家公司，应当尽可能的通用。这激发了当时 Intel 的马西安·霍夫设计了世界上首个微处理器结构，最后在费根带领的团队下完成了实现，即称为 Intel 4004 的首块 CPU，它将一个完整的计算机结构做到了一块芯片上，拥有将近 3000 个晶体管，具有通用编程的能力。此后几年英特尔陆续研发了 Intel 8008 和 Intel 8080 型号的微处理器，并基于这些芯片，推出了 x86 指令集，在市场上大获成功。微处理器的出现也启发了乔布斯和沃兹尼亚克发明一体化的个人计算机。因此个人计算机又叫微机。（计算机领域里“小型机”一般特指 IBM 小型机，它的体积依然很庞大，造价昂贵）



另一项关键技术是图形化的人机交互界面（GUI，Graphic User Interface），它由施乐公司发明。施乐公司最早是做复印机的，他们发明了多项相关专利。在 20 世纪七十年代，施乐集中了大量计算机科学人才，并在实验室中制造出来了个人计算机，它包含完整的微处理器、鼠标、显示器，甚至还出现了互联网的雏形。1979 年乔布斯在参观完施乐的实验室后大感震惊，立刻将苹果的研发方向调整为了图形界面，并一举挖角了施乐公司的大量人才。两年后，比尔盖茨也借鉴施乐的思路推出了图形化界面的操作系统。可惜的是施乐公司缺乏足够的眼光，始终未进入个人计算机领域，从而将这个巨大的机会拱手让给了苹果和微软。可以说，苹果公司和微软公司都是抄袭了施乐的图形界面这一发明，应用在自己的操作系统中。尤其是微软后来的 Windows 3.0 和 Windows 3.1 在市场上大获成功，从而证明了图形界面（GUI）的易用性对个人计算机来说是至关重要的。

此外 Intel 8080 微处理器的出现还激发了比尔·盖茨和保罗·艾伦为它写一个高级编程语言 BASIC 语言的编译程序。BASIC 语言是一个对新手友好的简单易用的高级编程语言，苹果最早的个人计算机产品 Apple II 也搭载了 BASIC 语言作为编程语言。一个简单易用的高级编程语言，可以带来丰富的软件开发生态，是个人计算机所需要的第三项关键技术。

因此个人计算机的发明，有三个主要的关键技术基础：微处理器、图形界面、高效的编程语言。它们分别代表了计算能力的普惠化（普通人买得起了）、使用简单（普通人能看懂、用懂）、可编程的通用计算平台（功能灵活以及强大）。没有这三个关键技术作为前提，个人计算机是无法被发明的。

也因此最终在 1977 年，苹果公司的乔布斯和沃兹尼亚克将所有的计算机相关组件、技术组合封装为一个整体，发布了最早的一体化个人计算机产品：Apple II，它将计算机的运

算器件装在一个盒子里，并带有独立显示器和键盘。这是一台完整的个人计算机。而比尔·盖茨和保罗·艾伦在那个年代则创立了微软公司，以开发软件为主，通过售卖软件授权的方式，微软公司最终成功的让每个家庭拥有了一台计算机，家喻户晓。



可以说，诞生自 20 世纪七八十年代的个人计算机，让以前只能服务于大企业的计算机走向了普惠，最终让每个家庭拥有了一台计算机，计算机从此能够为每一个人创造价值。先进技术从高端走向了民用，带来了深远的影响。

1.1.2 互联网的诞生：连接所有个人计算机

有了海量用户为基础，将所有的个人计算机连接起来，才诞生了今天的互联网。谈到互联网的诞生，往往会追溯到 ARPANET、Web 的诞生、TCP/IP 协议、以及 GNU 自由软件运动等伟大创举。但是我想强调的是，如果仅仅只是一些高校和科研机构连接起来，是无法形成今天意义上的互联网的。截止到 2023 年，全球互联网上的用户规模达到了 48.8 亿。互联网也成为了电报网、电话网之后最重要的信息高速公路，彻底的改变了人类世界的面貌。如果只有几百个科研节点连接到一起组成的网络，必然曲高和寡，是不可能达到这种影响力的。所以个人计算机的普及直接促成了互联网的蓬勃发展。也就是说，把所有的个人计算机连接起来，就形成了今天的互联网。

在这里尤其值得一提的是，在历史上苹果公司两度发明了个人计算机。第二次是在 2007 年 1 月 9 日苹果发布了第一代智能手机 iPhone，互联网迎来 iPhone 时刻。iPhone 拥有先进

的触控屏操作，搭载了 iOS 操作系统。到了 2008 年 AppStore 出现后，在新型的移动计算机上开发 App 变得蓬勃发展起来。

iPhone 的优雅体验带动了一个时代，众多手机厂商纷纷效仿，最终成功的让每个人拥有了一台智能手机，即人人有了一台个人计算机，这台计算机是可移动的，小到能装进口袋。除了能打电话外，还包含了摄像头和地理位置定位系统，能够便捷的采集全新的数据，因此有了全新的应用场景。最终将智能手机全部连接起来，就形成了移动互联网。

如果说微软成功的让每个家庭拥有了一台计算机，那么由于苹果公司的贡献和影响力，在 21 世纪的前二十年成功的让每一个人拥有了一台计算机。如果以前的互联网只需要连接每个家庭的计算机，那么现在的互联网就需要连接到每一个人，互联网里的节点数因此成倍增长，于是有了“移动互联网”的说法。全球的数字化信息在互联网里高速流动，出现了前所未有的盛况。

1.1.3 当前互联网的主要矛盾：数据垄断必将走向崩溃

互联网改善了人类的生活水平，纵观历史，在拥有互联网的今天确实是人类最好的时代。但是互联网依然存在着许多结构性的设计缺陷。比如在技术上，作为互联网心脏的域名根节点相对脆弱；TCP/IP 协议设计的缺陷导致了拒绝服务攻击频出；IPv4 地址资源耗尽，等等。但在这里，我想重点讨论一个更加隐蔽、更加本质，且更加致命的结构性设计缺陷：信息的流动方向带来的数据垄断问题。

当把所有的个人计算机连接到一起互联互通以后，人们很自然的有了信息访问的需求。那么人们怎么找到所需的信息呢？早期的互联网公司发现了这个商机，无一例外都提供了某类的信息的聚集服务。Yahoo 的类目聚合了当时的互联网里的主要网站和网页，同时它还发明了搜索引擎；Google 发扬光大了 Yahoo 的搜索技术，并开创性的提出了新的排序算法和大规模计算技术，它聚合了全世界的公开网页；在中国，第一代互联网公司搜狐、新浪、网易都是新闻聚合为主的门户网站，此后百度学习了 Google 的搜索引擎，开展了全网搜索业务，阿里巴巴则推出了电商业务，聚合了全网的商品信息。

因此可以说，早期的互联网公司，提供了信息聚集的服务，方便人们在互联网上快速找到信息，这种模式我称之为“人找信息”。在这个模式下，久而久之，互联网上的信息在某些超级节点形成了聚集，最终导致了数据垄断。一旦数据积压在某一点，数据量一大，访问的人也多，信息检索结果的排序就会成为一个关联着巨大商业利益的问题，那么最自然而然能想到的赚钱模式就是卖信息检索的排序结果，也就是广告。因此在互联网信息流动方向为“人找信息”的模式下，必将导致信息聚集，以及由此带来的广告商业模式。信息聚集的极端情况，就是数据垄断。这个模式直到今天依然无比强大，诸如 Facebook、字节跳动这些数千亿美金市值的商业帝国，依然是走的垄断数据，广告盈利这个路线。



客观的说，信息聚集这条路在互联网发展的早期确实提供了很大的价值，方便了所有网民。但是一旦信息聚集走向了垄断，就导致了失控。简单来说，有以下几个非常明显的弊端：

首先，大互联网公司垄断了数据，但数据都是老百姓们贡献的，最典型的比如用户的访问行为偏好数据，所有互联网公司都需要依赖这份数据来实现广告的精准投放，可是最终广告变现的钱却跟老百姓们一毛钱关系都没有，都被大公司赚走了。可以说老百姓们被收了一道数据税，这道数据税是广告主出的，但是羊毛出在羊身上，这抬高了所有商品的售价。如果这个钱是合理区间内的中介费还能说得通，但目前却是一种暴利。就好比高速公路建设完，一般二十年收回成本，三十年后就免费了，结果投资方在这里把两百年后的钱都收完了，这就是互联网流量生意被称为“印钞机模式”的由来，一旦垄断了数据，就垄断了用户流量，这些公司就等同于躺着印钞票。

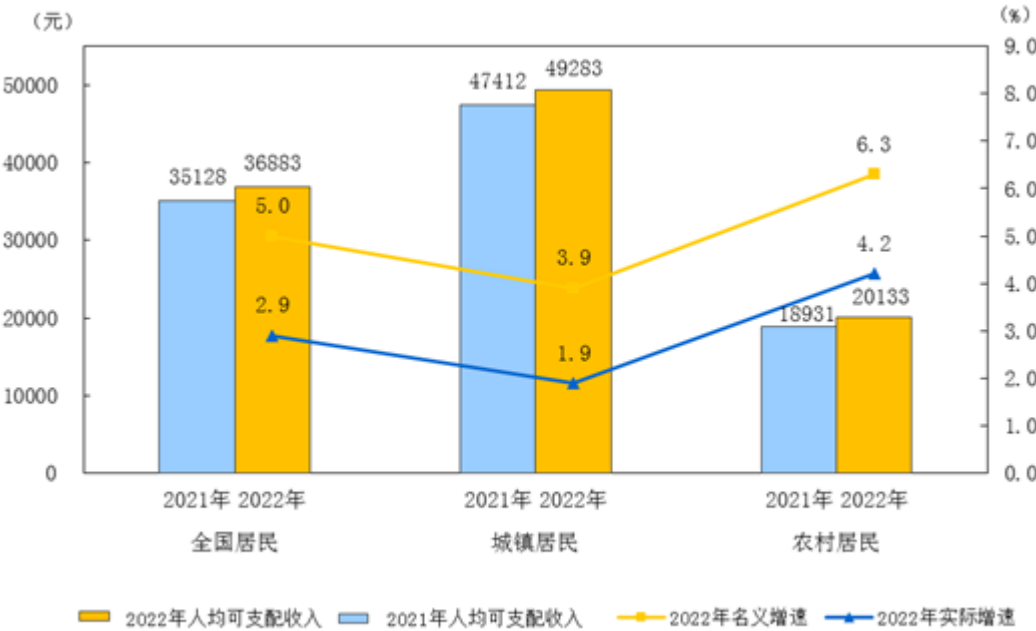
其次，数据垄断带来了互联网的割裂。现在的互联网和十年前的互联网有着巨大的区别，变得越来越地盘化，泾渭分明。最典型的比如视频网站，想看一部电影，往往是找了优酷没有，再去爱奇艺、腾讯视频找，还没找到，再继续去其他视频网站挨个找。所有内容提供方都得给这些渠道平台交高额的内容分发费用，这些掌控了流量的渠道平台互相之间森严的数据壁垒把互联网拆的支离破碎，他们在自己的地盘边界上竖起了一道道的高墙，让数据不再互联互通，最终影响的是老百姓们的效率和体验。

最后，垄断一旦形成，就将瓦解互联网的自由属性，在互联网上不再有自由和平等可言。小的个体在互联网上永无出头之日，而大的巨无霸公司因为垄断了资源，也将失去进取的动力，最终腐朽。这就是为什么一个自由职业者或者是小微企业在流量型平台上越来越难做生意的原因，赚的那点钱还不够流量费。

比如 2017 年暴露出来的天猫和京东为了大促活动争夺商家，逼迫商家签订排他协议，就限制了商家的自由。类似的，近日李佳琦借助流量优势，控制商品的市场价格，属于典型的店大欺客行为。互联网本应是用户做主，但用户却做不了主。数据都是用户创造的，但最后用户变成了被数据操弄的对象。就像你买了套房子，请了个装修队来装修，但是里面怎么装却不是你说了算，而是由装修队来控制。

所以当李佳琦在直播间对消费者说出的那句：“有时候找找自己原因，这么多年了工资涨没涨，有没有认真工作？”的话时，就扯下了资本在数据垄断上的最后一块遮羞布。这句话深深的刺痛了广大网民的心。同样都是劳动者，同样的付出了努力，有的人因为垄断了数据资源、垄断了流量，就赚到了普通人一辈子都花不完的钱，而有的人却只能挣扎在养家糊口上。既不公平，也不合理。李佳琦现象背后本质的原因，是因为他依托的直播平台完成了短视频的信息聚集，从而垄断了数据，进而垄断了流量，而李佳琦是这种流量红利的既得利益者。他身处其中而不自知，所以才会说出这么幼稚的话。数据垄断让互联网里的人们不再平等。

一个社会的财富总量在一定时期内总是有限的，有人靠垄断资源攫取了暴利，那么普通老百姓得到的财富自然就少了，这就是为什么现在工作越来越难找，老百姓收入一直高不起来的原因。互联网产业是近 20 年来全球 GDP 增长的强劲动力来源，但互联网巨头和背后的资本攫取了过多的暴利，从公开的财报就可以看到这些数据，新闻里每日鼓吹某某互联网公司上市又诞生了多少个千万富翁。而中国有 14 亿人，超过一半的老百姓的月收入不到 2000 块，他们的千万富翁越多，老百姓就越穷。这些暴利已经远远高于社会人均收入值，超出了合理范围。



(2022 年全国及分城乡居民人均可支配收入与增速，数据来源 www.gov.cn)

由此可见，原本自由的互联网，不再是净土。通过信息聚集这一手段，在资本的加持下，一旦互联网公司形成了数据垄断，也就形成了市场垄断，而最终受苦的是所有网民。垄断数据和垄断土地没有本质区别，高净值财富都是来自于垄断资源，而非劳动。因此可以说，数据垄断是资本在互联网时代的剥削行为，剥削了所有网民。

我们不反对用数据来赚钱，互联网的价值来自于数据，我们鼓励数据创造价值。但我们反对用数据垄断来赚钱，尤其反对垄断数据后，还利用垄断地位阻止和打压其他人提供同类服务。因为一旦无法遏制资本对利润不断攫取的贪婪，垄断的最终结果，贫者愈贫，富者愈富，必将导致崩溃。因此我认为，当今互联网最深层次的主要矛盾在于：“信息聚集导致的数据垄断”和“人们对自由、开放、共享、平等的网络环境的向往”之间的矛盾。

为了解决这一矛盾，实现资源的重新分配，有许多工作需要做。首先可以想到的是调节网络结构。现代网络科学的结论告诉我们（请参考拙作《计算》），在一个无标度网络中，其规律服从幂律分布，有着“富者愈富”的现象，自由发展则必将涌现出超级节点。而在一定初始规则约束下，则可以涌现出我们想要的任何网络结构。这意味着宏观调控的手段是有效且必要的，需要制衡资本的无序扩张。

从网络结构的角度来看，将互联网设计成一个类似于人类社会关系网络的模型，将会是一个更加公平的网络，因为这种网络是去中心化的，是局域化的，通过长程连接把各个局域连接起来，不存在超级节点。事实上微信就是这样的网络结构，在微信里一个人只能加几千个好友，而不可能用一个微信号加 1000 万人。这种网络模型称为“小世界网络”，在小世界网络里，没有超级节点可以影响大部分的其余节点，因此它更加的稳定，适应性强。但小世界网络损失了效率，因此可能需要找到小世界网络和无标度网络之间的某种网络结构，这是宏观调控的艺术。

但调节网络结构，以抑制或迟缓超级节点的出现，是治标不治本。我认为还有做出更加彻底的改变的可能性，因为最近这五年的技术发展，已经让我们站在了一个百年一遇的十字路口上：我们有机会改变互联网的信息流动方向，从而消除信息聚集。

1.1.4 技术垄断：大型专用 AI 计算机

大互联网公司之所以能够做到信息聚集，是因为在传统的互联网结构里，用户有一个需求之后，无法高效的找到对应信息，而互联网公司提供的信息聚集，恰好满足了这种高效寻找信息的需求。从这个角度来说，我从未否定信息聚集的价值。

但这样的模式存在两个问题，第一个问题是数据垄断问题，如前所述；第二个问题是“人找信息”的模式效率依然低下。由于互联网已经被大公司之间的数据壁垒割裂了，人们不得不耗费更多的时间在不同的平台上寻找他们想要的结果；同样的，由于受到了商业广告的干扰，人们不得不从一大堆检索结果中排除这些干扰，从而找到他们需要的那个最优答案或者近似最优答案。

如果能改变信息的流向，就能解决这些问题。即从“人找信息”，向“信息找人”发生转变。

当人有一个需求之后，应当是信息主动找上来，自动给出一个全网最优解，而节约大量的时间和精力。这个最优解，不应该有商业广告的干扰。

这在以前是天方夜谭，但现在是有可能的。因为这些年的技术进步，我们有了 AI 这一新工具来实现这件事情。如果互联网中的每个节点（每个企业、组织、以及个人）都拥有一个 AI，AI 知道它的主人拥有什么样的信息和数据，再通过一个中立的搜索推荐算法，当人有任何需求时在网络内进行“广播”，询问每个节点的 AI 是否有所需的答案，一旦有，就由中立的推荐算法将其推送给需要的人。这样就实现了“信息找人”。当然一切过程需要在安全的情况下完成，需要建立完善的隐私保护、授权、内容安全机制。

这样的“AI 互联网”是一次大的飞跃，将提升人类社会的整体效率，我们将在第三章讨论。但实现这样的 AI 互联网的前提是每个人或组织都拥有了 AI 计算机，现在先把注意力放到 AI 计算机本身上来。

在做创新时，遇到的诸多困难中的一个，就是很多概念缺乏精确的定义。AI（人工智能）就是一个没有被精确定义过的概念。自从 1956 年达特茅斯会议上，麦卡锡创造了 AI (Artificial Intelligence) 这一名词以来，AI 依然是一个综合且模糊的概念，它综合了许多学科，包括计算机科学、心理学、神经科学、认知科学、逻辑学、数学等等。在此我并不打算对 AI 下一个精确的定义，那只会把我拖入无止尽的辩论中。在此，我所指的 AI 特指那一类具备自动化、自适应能力，尤其是在数据处理上具备这种能力的系统。

如果从这个概念出发，那么这样的 AI 系统已经出现了很多年了。大型互联网公司如 Google、Facebook、百度、淘宝、抖音等在处理信息聚集的大数据时，尤其是他们在做广告的推荐算法时，就已经建立了一套这样的系统。广告的精准投放和“千人千面”都已经做到了可以基于用户的实时访问行为，改变接下来用户会看到的推荐结果。这是一种自适应能力，基于输入的变化而改变自身的行为，这就是 AI 的能力。

因此可以说，大型互联网公司在处理大数据的能力上，已经建立了“大型专用 AI 计算机”。首先它是大型的，数据量很大，可能大至上百 PB；用到的服务器数量很多，一个集群可能有上千台高性能服务器。其次它是专用的，因为这样的系统只能用在搜索、推荐、广告投放的特定场景，而无法实现对于任意任务的通用计算。最后它是智能的，基于环境的自适应能力是一种被称为 Agent（在人工智能领域 Agent 不应当被翻译为“代理”，翻译为“智能体”勉强可以接受）的能力，而且麦卡锡正是用搜索路线来做人工智能的先驱，因此将搜索推荐技术纳入到 AI 领域本身是合理的。

这样的大型专用 AI 计算机，帮助互联网公司完成了信息采集和聚集，高效的处理了大数据，提高了商业效率，榨取了数据创造出的价值。由于大型专用 AI 计算机的成本高、实现门槛高，因此这种先进技术创造出来的价值，被这些大公司垄断了。他们既垄断了数据资源，又垄断了处理数据资源所需要的技术，最终垄断了数据创造的价值。一个耳熟能详的比喻是“数据是石油”，那么“大型专用 AI 计算机”就是石油开采、提炼的整套技术。在我国早期的一些油田探明了巨大的储量，但是当年却只能和欧美的拥有石油开采技术的公司合作开采、利润分成，就是吃了没有先进技术的亏，所以坐拥资源，只能任人宰割。

现在连普通老百姓都已经知道了数据是石油，意识到了数据的价值。但是非常奇怪的一

点是，在一个路人皆知“数据有价值”的社会里，普通老百姓却从来没有用数据赚到过钱。所有的数据都是老百姓们创造的，但最后数据赚到的钱和老百姓们一毛钱关系都没有。数据只是大公司的石油，不是老百姓的。原因在于老百姓们被技术剥削了。



因此我们就有了变革的动机：互联网从诞生之日起就代表着平民的胜利，而不应该被垄断，因此为了维护这个良好的网络环境，回归初心，就迫切的需要有一个新工具，来帮助个人用户、小微企业用户能够处理数据、创造价值。这个新工具应该是开放、自由、共享、平等的，即“个人 AI 计算机”。这个新型的计算机将再一次把大公司垄断的“大型专用 AI 计算机”给小型化、普惠化，让每个人都用得起来。在这样的“个人 AI 计算机”模式下，人的在线时间不再受制于睡眠和休息，AI 会帮你时刻在线，只要提供了数据，AI 就能开放式的回答所有问题，永不停歇的服务、沟通、协同。

一个可预见的结果是“AI 互联网”将消灭精准营销广告，把精准营销广告商榨取的万亿级别人民币利润返还给供应商和老百姓。相当于以后在互联网上买所有东西应当会免除一道广告税。对于那些希望通过看广告来免费获取服务的用户来说，需要认识到广告的隐蔽危害最终抬高了商品和服务的价格。而想免费获取某些商品和服务其实有其他办法，比如通过自己的分享来换取，就是一个健康的机制，它提倡和鼓励人们分享，而非攫取。因此最终 AI 互联网会更省时、更便宜、更丰富。

恰如四五十年前苹果对 IBM 发起了挑战，最终完成了大型计算机的普惠化这一革命。当下第二次革命的条件也已经成熟，发明“个人 AI 计算机”的时机已经到来。

1.2 第二次革命：人人有台 AI 计算机

1.2.1 个人 AI 计算机的使命

一种新的技术，如果不能解决过去的历史问题，就不能称之为革命。但反过来说，如果它真的给出了有效的解决方案，那么其存在就是合理的。个人 AI 计算机（PAIC, Personal Autonomous Intelligence Computer）的合理性即基于此。它的使命，是为了打破互联网里大公司对于数据的垄断，让个人用户、小微企业用户有能力利用数据创造价值，最终实现 AI 互联网，极大的提高人类之间的沟通、协同的效率。

阿里云的创始人王坚博士曾经做过一个形象比喻，云计算是电，AI 大模型是电动机。在历史上，发电机、电动机、电网三者共同构成了完整的电力基础设施。在我看来，云计算、AI 计算机、AI 互联网三者将共同构成完整的算力基础设施，将人类带入到算力时代。

在 20 世纪七八十年代的那场发明个人计算机的革命中，有三个重要的技术基础：微处理器、图形界面、高级编程语言，它们分别解决了算力的普惠、简单易用的操作体验、丰富的软件应用生态这三个问题。在当下，个人 AI 计算机的技术基础也已经成熟，分别是：数据和算力的普惠、自然语言大模型（LLM）实现新型的人机交互方式，以及呼之欲出的新型编程范式。

1.2.2 个人 AI 计算机的三个技术基础

1.2.2.1 数据无处不在，算力随手可得

随着集成电路的发明，计算机工业界在摩尔定律的影响下，芯片的单位面积计算能力以年为单位指数级增长，这极大的增强了单机的处理能力。随着智能手机的发明，以及互联网上需求的爆发，个人和组织开始沉淀数据，数据无处不在。此外云计算的出现让大算力实现了普惠，它对 AI 计算机有两方面的意义，其一是并行计算能力是 AI 大模型得以成功训练的前提，其二是让个人使用大算力成为可能。

云计算的使命是让计算成为公共服务。它的做法是将数据中心（为了避免歧义，用土话解释一下就是机房）变为一台计算机。在历史上，Google 和 Amazon 的贡献为云计算奠定了基础。其中 Google 在 2003 到 2006 年陆续发布了三篇论文，分别是 GFS（Google 的分布式文件存储）、MapReduce（分布式计算架构）、BigTable（Google 的分布式数据库），从而奠定了大规模数据计算的基础，实际上这三篇论文就是 Google 拥有的“大型专用 AI 计算机”的主要实现原理，又被称为 Google 的三驾马车。

而 Amazon 在 2006 年推出的 AWS EC2（弹性计算服务）、S3（云存储服务），则被视为云计算的开端。尤其是 EC2，成功的实现了计算服务的租用模式，将高性能服务器的使用成本分摊到了小时和分钟的粒度，用户可以按需租用，从而让用户可以用低廉的价格在短时间

内获得强大的计算能力。

在中国则是由 2009 年创立的阿里云开启了云计算的道路。阿里云的飞天云操作系统是一个按照计算机体系结构设计的操作系统，它融合了 Google 的大规模数据计算和 Amazon 的按需租用的思想，尝试将数据中心变为一台计算机，同时让个人开发者能够用低廉的价格按需租用这种计算服务。

云计算的梦想是要服务于海量用户，走的是通用计算的道路。而过去大互联网公司只需要服务好自己需求，因此很难把自己做成通用计算。因为他们很有钱，所以压根儿不需要做通用计算机，给自己量身定制一个专用计算机就好了。这也是为什么现在试图服务于大企业的公司不可能做出来通用计算机的原因。只有个人计算机，因为要服务于海量用户，遇到的需求注定千奇百怪，所以必须做成通用计算。价格还得足够亲民，用户才用得起，云计算通过按需租用的模式实现了这一点。有了云计算以后，算力不再是约束创意的瓶颈，任何人有了一个好的想法，都可以用可接受的价格在一个大规模计算集群上进行验证。

因此云计算就像当年 Intel 发明的微处理器，微处理器第一次将计算做到了一块芯片上，价格也是普通人负担得起，云计算则在当下实现了大算力的普惠。

1.2.2.2 人机对话的新交互（LUI）

2023 年是人工智能的奇迹年，OpenAI 的 ChatGPT 横空出世，在自然语言的人机对话领域首次实现了突破性的智能表现。在此我作为一名计算机工程师，发自内心的向 LLM 这项技术致以敬意。关于自然语言大模型（LLM）这项突破性技术的介绍和评价，网上已经很多了，我在新书《计算》中也做了详细的点评，在此不再赘述。

LLM 这项新技术将在个人 AI 计算机中发挥关键作用。可以说，如果没有这项技术的突破，个人 AI 计算机的实现可能还得晚上许多年。因此也可以将个人 AI 计算机看作是 LLM 这次技术浪潮的一个直接成果。

LLM 在个人 AI 计算机中主要用到的地方有两个，一是它将成为人机交互的新入口。在苹果、微软发明的个人计算机中，图形界面（GUI）占了主导地位；而在个人 AI 计算机中，基于自然语言对话的交互形态会出现，因此定义一个 LUI，即 Language User Interface 就变得有道理了。用户可以用自然语言和计算机交流，LUI 会理解用户的意图，同时将计算机的执行结果以自然语言的方式回复给用户，这是一种全新的体验。

二是它将在系统中发挥类似于心脏的作用，大部分的原子调用都离不开它。关于这一点，我们容后再叙。

但是 LLM 技术，包括 ChatGPT 本身也存在很多问题，远非自媒体所鼓吹的无所不能。事实上我们在实践中发现了 GPT 相当大的局限性，在本文稍后我会详细叙述问题在哪里，以及我们的解决思路是什么。

备注：本文发布时，Google 发布了它最新的大模型 Gemini，主要亮点为原生多模态的理解能力。尽管如此，语言的使用依然是未来人机交互中最重要的一种方式，有着极其广泛的应用场景。

1.2.2.3 从“可编程 AI”到“AI 可编程”

有了 LUI 以后，一个最大的好处，就是用户对计算机的直接操作可以通过自然语言完成了。在个人计算机时代，想直接操作计算机，用的是形式化的编程语言，如汇编语言、FORTRAN、LISP、BASIC、C、JAVA、Python 等等。据统计，世界上可能有超过 2000 种计算机编程语言，但无一例外都是形式化语言。

形式化语言的学习成本高。据统计，在中国只有 700 万的开发者，假设他们都至少掌握了一种编程语言，那么这个群体的数量相对于中国的人口基数来说依然是微不足道的。也就是说在个人计算机时代，在中国懂得直接操作计算机的创作者只有 700 万人，是他们创造了中国软件产业的万亿市值。

而在个人 AI 计算机时代，这个数字有可能变成 7000 万。由于 LUI 的存在，有可能降低用户学习计算机语言的门槛，而且至关重要，是让 AI 编程已经近在眼前。如果形成了“人来描述需求，AI 完成编程”这一新的范式，就迈出了一大步的跨越，也将真正完成发明个人 AI 计算机所需要的第三项关键技术基础：基于自然语言的可编程平台。这个中间尚存在一些难点，比如自然语言中存在着隐性的逻辑特征，这意味着我们无法直接用自然语言构成演绎规则公式。我们将在本文的后半部分给出一种解决思路。

最终将回归计算机编程语言的初心。计算机编程语言的先驱，FORTRAN 语言的发明者，图灵奖得主约翰·巴库斯曾经指出：程序员应该只需要表达出他们“要什么”，而无需知道“怎样做”。他毕生为此追求和努力。时至今日，在 AI 的加持下，我们看到了巴库斯梦想实现的曙光。

也因此，在个人 AI 计算机的背景下，估算一下中国未来能够直接操作计算机的创作者，应该从 700 万开发者，变成 7000 万的专业工作者，最终使得至少 7 亿老百姓受惠。各行各业的专业工作者如作家、画家、美食家、律师、医生、教师、学者等等通过自然语言描述需求，AI 计算机自动完成任务。

第 2 章 我们的答案：kOS-1.0

因此，为了完成上述使命，抓住时代的机遇，我和另外两位联合创始人，以及 13 名初创团队的同事，一起在 2023 年 7 月 1 日创立了 KMind。KMind 的使命，就是要发明这台新型的个人 AI 计算机，打破互联网里大公司的数据垄断，让数据能够为个人、小微企业创造价值，最终实现一个自由、开放、共享、平等的 AI 互联网。

这五个月以来，我们做了诸多尝试，并迅速积累了超过十万用户。一开始我们基于 GPT

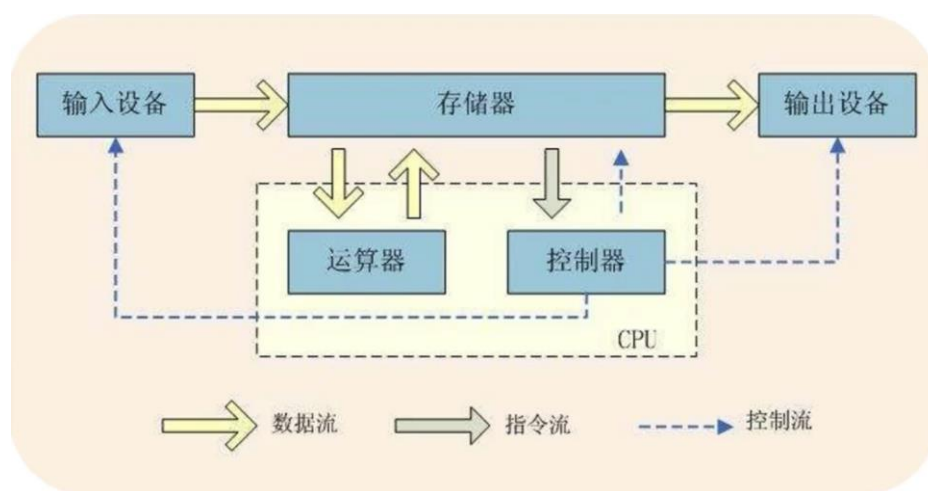
来实现我们的 AI, 在国内的监管要求出台后, 我们切换到了国产 LLM 和开源 LLM 的路线上。由于有着真实的用户基础, 因此我们很清楚 LLM 的能力和缺陷。能力是我们要充分发挥的, 这是时代红利; 缺陷是我们要弥补的, 这是建立核心技术壁垒的机会所在。这样我们在创业三个月后, 越来越清楚我们的目标到底是要做什么。于是在两个月前, 我们在公司内部明确了“个人 AI 计算机”这一理念, 并提出了 kOS-1.0 新型计算机架构设计。

我们将通过 kOS 这一新型计算机操作系统, 来实现这一伟大使命。

2.1 扩展的冯·诺依曼架构：为什么这是台计算机？

2.1.1 从数字计算机 (Digital Computer) 到数据计算机 (Data Computer)

计算机发展到今天, 所有的数字计算机都基于图灵在 1936 年论文中提出的图灵机模型。在图灵机上实现通用图灵机, 就可以模拟任意的可计算模型, 完成任何可计算函数的计算。由此图灵提出了“存储式程序”的思想, 即图灵机的描述数既是数据, 可以被存储, 又是程序, 可以被通用图灵机读取和执行。此后在 1946 年, 冯·诺伊曼受到图灵思想的影响, 在 EDVAC 项目中引入了存储式程序的思想, 并和埃克特、莫奇利一同设计出了存储、计算分离的结构, 被称为冯·诺伊曼架构。它奠定了当今可制造的物理计算机的主体结构, 可以说, 当今的计算机基本上都是冯·诺伊曼架构。



在冯·诺伊曼架构中, 实现了图灵机模型, 并在运算器中引入了算术逻辑单元可以进行算术运算。而计算机的另一位先驱香农关于逻辑电路的研究则清晰的指出, 电路设计可以抽象为逻辑代数, 两者本质是一回事, 从而大大简化了电路设计的复杂度。最终的一个重要结论是, 基于逻辑的与、或、非, 就可以实现出完备的通用图灵机, 可以计算所有可计算函数。

因此凡是实现了通用图灵机的物理装置, 尤其是证明了该物理装置的计算能力是图灵完备的, 就都可以称之为计算机。通用图灵机确保了计算的通用性, 也就意味着有可能在物理世界的任何尺度上实现图灵机。小到纳米尺度、大到宏观宇宙, 都有可能实现一个通用图灵机的物理结构。十多年前出现的云计算技术, 尝试将数据中心变为一台计算机, 就是在一个

大规模服务器集群上实现了类似冯·诺伊曼架构。因此数据中心是一台计算机，阿里云的飞天是云操作系统，这个说法不是媒体炒作的口号，而是有坚实的理论依据的。

今天我们用到的所有电子计算机，小到一部手机，大到大型机和超级计算机，都是数字计算机。用机器设备处理数字信号的叫数字计算机。在图灵和冯·诺伊曼那个年代，计算机发明出来主要是用于做算术的，处理一些科学数据中的简单算术问题，因此他们的焦点主要聚焦在数字信号的处理上，数字计算机最重要的功能是要实现数字信号的算术和逻辑运算。在那个年代，没有数据的概念，所以计算机的结构、微处理器、指令集这些都是为处理数字信号而设计的。

而今天我们要发明的个人 AI 计算机，是一种数据计算机。用机器智能系统处理数据的叫数据计算机。它是以数据单元为基本操作对象，实现逻辑运算，最后达到图灵完备的一种机器装置。数据计算机当然是基于大量的数字计算机组成的，它是一种比数字计算机宏观得多的计算机，主要用途是操作数据。正如普通用户在使用笔记本电脑时，不会去关注它内部的 CPU 一样，尽管 CPU 从结构上来说已经是一个完整的计算机了，但人们日常提到“计算机”这一概念时，通常指的是那台笔记本电脑。因此我们从数据计算机的尺度上，并不太关心系统底层的某台具体数字计算机是如何工作的。

当我们尝试要实现数据计算机时，才发现“数据”这个概念根本没有被精确定义过。这给我们造成了一些困扰，不过问题不大。因为当前在计算机世界里的所有数据都可以归纳为文本、图像、语音三类。出于效率考虑，我们把视频独立了出来，所以是四类数据：文本、图片、视频、语音。（这里没有把编译后的二进制程序视为数据，因为程序是用来操作数据的）

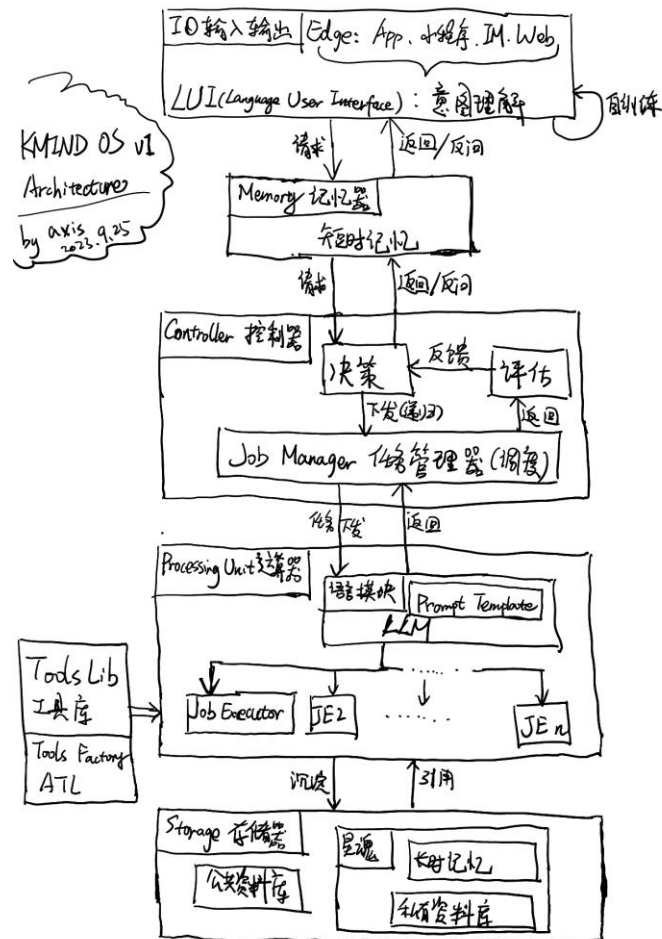
个人 AI 计算机的使命是要让数据为个人用户创造价值，个人用户的数据在哪里呢？最显而易见的，每天你手机里拍照、录像的图片和视频，都是数据。其次，你在互联网上的所有访问行为，你和其他人交流与沟通、对话信息都是数据，这类数据未来在 AI 互联网上会自动沉淀下来，成为你的资产。最后，每个人还可以采集数据，比如看到一篇好的文章，可以保存下来作为自己的素材，以备后用。所以每个人实际上会产生大量的数据，现在一个人一生中可能会有意识的保存几个 TB 的数据，未来这个数据存储量还会翻许多倍。

在过去，处理数据只是数字计算机的一种应用形态，而非底层逻辑。而在当下，形势和环境都与八十年前发明数字计算机时相比有了巨大的变化，大量数字信号聚合后产生了数据，尤其是个人计算机普及、软件产业兴起，互联网繁荣，带来了数据大爆发。因此我们认为有必要设计一种新的计算机架构，用来处理数据，它比数字计算机在数据的处理上会更加的高效。这就是 kOS 要实现的目标：个人 AI 计算机。

所以个人 AI 计算机在我看来主要实现两类任务，一是自动化的处理数据，二是学会使用工具。而软件和 App 本质上可以视作数据的通道，因为任何对软件的输入、输出，都可以转化为数据。从而个人 AI 计算机要完成的本质工作只有一个：操作数据。在当前 AI 的大背景下，过去的数据库技术有可能会被取代或者消亡，而新的数据技术会出现，它有着新的交互体验和自动化的数据治理方法，从而带来一场数据技术的革命。

2.1.2 基于人工智能扩展的冯·诺依曼架构

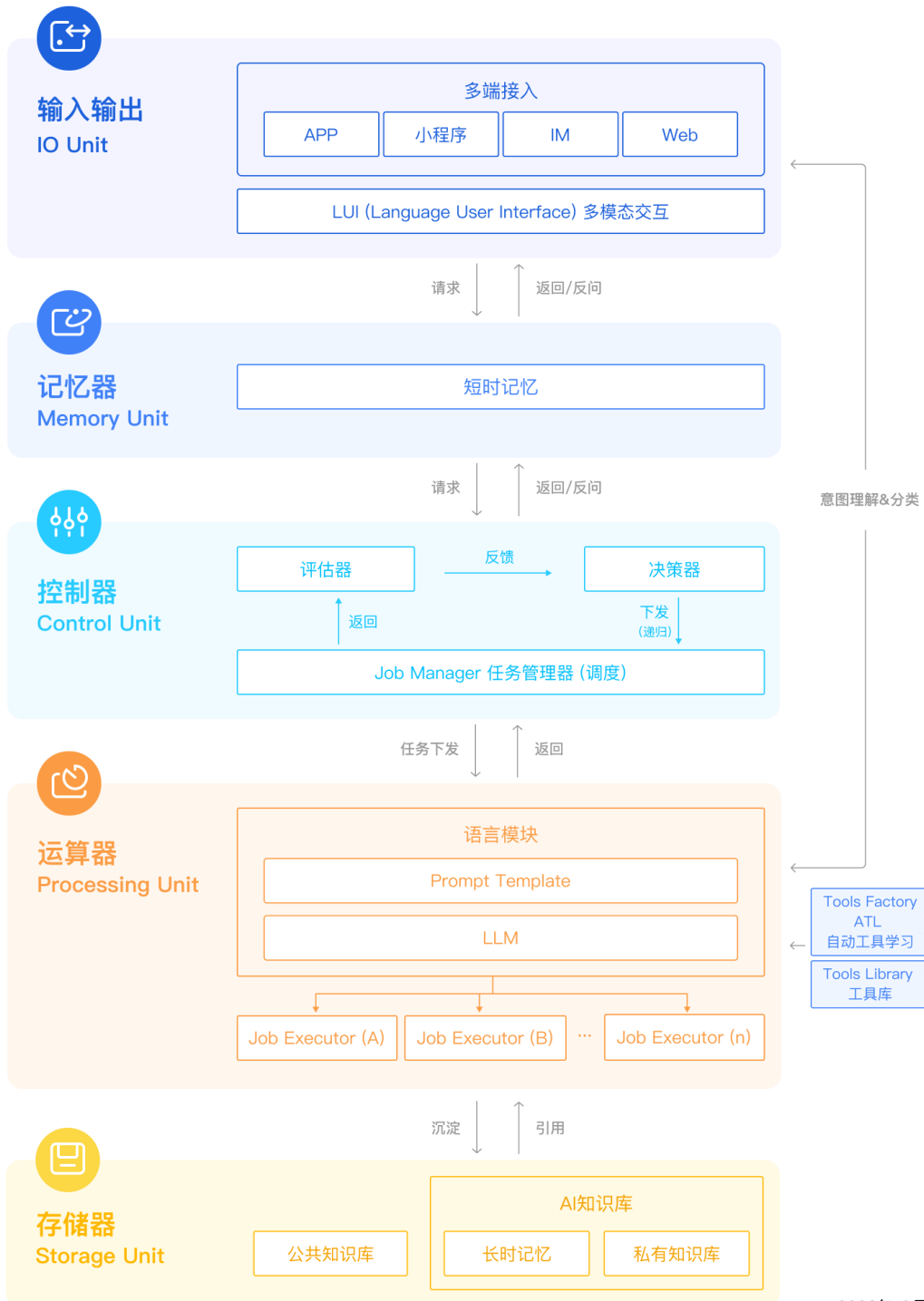
功能来源于结构。在今年 9 月 25 日，我在电子阅读器上用电子笔画出了 kOS 的第一张架构图，它实现了个人 AI 计算机的通用计算架构：



经过和团队的反复讨论后，在 10 月份确定了当前版本的架构图：

通用AI计算机架构图

实现方案：kOS 1.0



2023年10月
中国·杭州

首先这是一张计算机的架构图，因为它保留了冯·诺伊曼架构的主要部件：输入输出模块、控制器、运算器、存储器，因此 kOS 是一台计算机。输入输出模块中包含了 LUI，它基

于 LLM 对用户的自然语言做意图分类，然后触发控制器执行相应的动作。

其次在这张图中新增了记忆器用于处理短时记忆，比如对话时 AI 必须能够记得之前用户说了什么。在短时记忆和长时记忆之间会有联动，某些反复强调的短时记忆应该转化为用户的长时记忆。此外我们还将用户的个性化知识、私有资料、记忆，以及最终每个 AI 的不同性格，会保存在存储器中。

再次，LLM 和提示词框架放到了运算器中，它在系统中处于心脏的位置，大量的原子调用需要用到 LLM 以简化编码逻辑、提高系统的整体效率和泛化能力。但是由于这样的架构设计，也就明确了这个系统并非只是一个提示词工程。因为“心脏”不可能真正理解用户的意图。

最后，控制器是我们要建设的核心能力，相当于整个系统的“大脑”。大脑是真正理解用户的意图，并对复杂任务做拆解，以及调度不同的执行单元来完成用户任务的过程。在控制器中，设计了评估器和决策器两个关键模块，同时做了一个数据流自闭环以及任务递归执行的设计，这是在模拟人类大脑在解决任务时的心理活动过程：评估、决策、不满意的话调整策略后重新执行，再评估，再决策是否返回结果……如此往复。我在《计算》这本书的第 374 页中曾简单的提到我构思了一个解决任务的通用框架，kOS 的架构设计图可以说是这个通用任务框架的一个具体实现。

最终可以说，我们对于 LLM 的依赖是非常薄的。我们主要通过 LLM 在 LUI 中完成用户对话里的意图分类，以及在运算器的位置通过 LLM 提高效率。我们并不需要 LLM 的逻辑推理能力，因为这是我们控制器要建设的核心能力。我们也不需要 LLM 的常识，因为我们的存储器将让用户建设自己的私有知识库。

我们之所以作出这样的设计，一方面是我们希望构建一台通用计算机，另一方面是我们看到 LLM 在使用上还存在大量的不足，这给我们留出了巨大的创新空间。事实上，LLM 这种一次训练海量数据进行压缩的做法是一种速成，它离人脑的学习原理更遥远了，所付出的代价就是 LLM 其实不具备任何的成长性，它的知识增长都需要重新训练更多数据，而人类的婴儿却可以从零知识开始学习这个世界，最终发育成高智慧的大脑。此外 LLM 的幻觉问题从原理上思考的话，我高度怀疑其复杂度超出了当前所谓超级对齐技术能覆盖的范围，这都是 LLM 拔苗助长的代价。因此我认为一个结合了神经网络存储知识的递归累积学习的结构，可能更接近人脑的智能，这是我们架构设计的基础。尽管 LLM 是一种速成的幻觉智能，但 LLM 的优点和任何技术进步，都将成为我们系统的增量，因为我们的架构已经站在了巨人的肩膀上。

2.2 AI 精度：为什么仅有大模型是不够的？

大模型从 2023 年的年初火到了年尾，但是身为 AI 领域的一线工作者，我们用实践经验体会了 AI 大模型当前的实际情况是叫好不叫座。投资人往往会问 AI 创业者目标市场、目标人群是谁，为什么价值付费？我想说的是，这个市场压根不存在，今天 AI 的成熟度还没有成熟到用户愿意持续付费从而取得商业成功的阶段。这中间存在 AI 大模型的许多体验问题，必须解决掉这些问题，才能迎来商业成熟的到来。要证明这一点很简单，如果这个市场存在，

百度的文心一言、科大讯飞的讯飞星火等大模型 App 已经卖爆掉然后一统天下了，但事实是并没有。

这恰恰也是创业者的机会所在。下面我将简述这些问题的症结在哪里，以及我们如何解决它。事实上所有的成功技术变革，都是创造和激发了新的市场，微软如此、苹果如此、AWS 和阿里云也如此。

2.2.1 大模型的幻觉问题

LLM 的首个严重问题是它一直被吐槽的幻觉问题，这直接导致了 LLM 在实际生产应用上用不起来。

比如用户让大模型写一首李白的诗出来，然后大模型胡编了一个，用户说这不是李白的诗，你重写，大模型继续胡编乱造。在我们早期积累的 10 万用户里，我们发现了大量类似的 bad case。信息的真实性是沟通中的一个至关重要的基本问题，没有人喜欢和一个满嘴谎话、不懂装懂的人沟通。尤其对于那些从来没有接触过 AI，不知 AI 为何物的人，在第一次面对大模型这样的 AI 时，往往会将其当作一个真实的人来进行对话，但发现不靠谱之后马上就放弃了。我们早期的用户流失率相当之高，大部分原因就是因为它大模型的幻觉问题，让用户觉得 AI 不可靠。

后来我看到一份报告，提到 ChatGPT 的用户里 40% 是程序员，剩下 20% 是学生，我大概就明白为什么这个东西无法在全社会推开了。除了发烧友，普通人还真用不起来，玩几次过了新鲜度就不喜欢玩了。所以现在的大模型在我看来更像是一个人在睡梦里说梦话，梦话看起来像那么回事，但却经不起推敲，也时常会有常识性错误。事实上，大模型的原理和人类做梦是差不多的。我们无法把一些重要的工作寄托在一个叫不醒整天说梦话的人身上。

后来我做了一份对比测试，从我们的日常用户提问中匿名抽取了 100 个问题，用 LLM 总结来自互联网的搜索结果（不用 LLM 的常识），对比 GPT-4 本身的回答（用 LLM 的常识），发现有 60 个问题的答案两者都答对了，但剩下 30 个的问题前者答对了，GPT-4 完全答错了，还有 10 个问题两者都没答对。前者对后者是压倒性的胜利。也基本上可以下个结论 GPT-4 的对话功能由于幻觉问题，还无法代替搜索引擎这样的信息检索服务。

目前 LLM 在解决幻觉问题上主要还是靠对齐，来源于它的提示词工程和微调，但这类技术治标不治本，无法从原理上根治幻觉问题。要想彻底解决这个问题，我认为还得从 LLM 之外寻求答案。

2.2.2 大模型缺乏“精确控制”的能力

大模型的第二个严重问题是缺乏“精确控制”的能力。在这个问题上所有的大模型全部都躺下了。

比如这个问题：“在《甄嬛传》这本书中，‘甄嬛’这个名字一共出现了多少次？”，或者更

难一点的问题“在《甄嬛传》这本书中，一共出现了多少个女人？多少对夫妻？”，所有的大模型都无法回答。再比如，用大模型文生图画了一张图，然后跟它说“请把背景里天上的那个月亮改成紫色的”，这时候所有的大模型也都歇菜了，它们只会给你重画一张不一样的，而不会在现有图片上做细节修改。因此你看到的所有媒体宣传给出的大模型的种种“神奇演示”，他们都只敢给你看第一步的内容生成结果，而不敢给你看如果想修改结果怎么办。这让用户在使用大模型的过程中会有一种隔靴搔痒的感觉，非常痒，就是挠不到。用户会觉得大模型无法理解自己的意图。

造成这种“精确控制”能力缺失的原因，依然和大模型的原理有关。基于统计关联的自然语言大模型，来自于连接主义，天然就不是决定性的。而其对立技术流派符号主义则更加注重逻辑推理和决策，精确控制的问题对符号主义来说很容易解决，但符号主义却没有大模型这么强的泛化能力。连接主义和符号主义各有优势，如何形成综合的统一是未来 AI 系统的关键。

OpenAI 在 10 月初的发布和更新中，新推出了 GPTs 的服务，其中官方提供的 Data Analysis 应用，具备生成 Python 代码查询指定 PDF 文件的能力，是一种精确控制能力，可以回答上述统计甄嬛名字次数的能力。但这个架构已经超出了单纯的大模型，而是变成了一种 Agent 的能力。所以时隔半年，OpenAI 目前已经不仅仅是 LLM，而是变成了基于 LLM 的 Agent 的架构。单靠 LLM，也是无法解决精确控制的问题的。

OpenAI 追求的目标是通用人工智能（AGI），但也因此带来了一些新的问题，目前还没有解决掉。在 10 月的发布会两天后，OpenAI 的 GPTs 被发现存在一个严重的数据泄露漏洞：用户通过特定的提示词，就可以诱导和哄骗 GPTs 回复出它的开发者所用的提示词，以及开发者上传的私有文件下载地址。这是由于 OpenAI 过分追求 AGI，所以 ChatGPT 总是试图满足和讨好用户，而缺乏安全红线的概念，因为一个精确的安全红线，会伤害到用户的对话体验。因此 OpenAI 的安全边界是模糊的，而非精确的。这也是 OpenAI 缺乏另一个精确控制能力的表现。

2.2.3 人工智能的测试基准失去意义

大模型的这些问题，是迫切需要解决的。首先需要一种科学有效的测试方法。当前人工智能领域的测试基准 (Benchmark) 百花齐放，在各种各样的任务上都有论文建立测试基准。但是我认为这些测试基准作用有限。

我是一名工程师，擅长做系统设计和工程实现，出于工作需要经常会和算法专家、AI 公司打交道。接触多了，久而久之，在我们工程师圈里开始流传着一个笑话：我们这些工程师遇到的每一个做人工智能算法的人，都宣称自己是世界第一，因为他们总能拿出一篇论文，在基准测试的跑分超过了所有同行。如果一个行业里的每一个人都宣称自己是世界第一，那一定是这个行业出了问题。

虽然 AI 领域已经有了成千上万种基准测试方案，但是我发现找不到一个系统性的测试方案来测试大模型。有一天我问实习生杨子乐：“人类到底有多少种 chat 方式？”杨子乐查了

一圈论文依然没找到答案。Chat 是非常复杂的，这等同于在问人类到底有多少种心理活动以及如何形式化它们？在人类的日常对话中可能存在比较、举例子、反驳、推理、联想、想象、推荐等诸多心理活动，这些心理活动应当是一个有限集合，那么有没有一个测试基准以科学和系统的方式枚举了这个有限集合？目前我没有发现，如果在语言学、心理学等领域里有这样的科学方法或理论，请记得告诉我。我认为缺少了这个东西，很难把大模型搞科学。在此之前，调教 AI 模型更像是炼丹，并不比中世纪的炼金术高明到哪里去了。

2.2.4 定义：AI 操作数据的精确程度

而 AI 计算机的目标和 AGI 并不相同。因为它是台计算机，发明出来是给人用的。为了有一种科学而系统的方式将我们的“个人 AI 计算机”测试精准，我提出了“AI 精度”的概念。

首先我也没有解决“人类到底有多少种 chat 的方式？”这个问题，这个问题依然留给语言学家和心理学家。但是我们的“个人 AI 计算机”的工作目标不是实现 AGI，而是通过 AI 来操作数据，因此我们只需要找到一种有效方法，可以评估 AI 操作数据的能力就可以。因此我定义“AI 精度”为 AI 操作数据的精确程度。

既然使用了“精度”这一概念，那么本身就蕴涵着对数据的结构化要求。因此对于任何的数据，应当先将其结构化到一个多维空间，这样就可以在任意尺度上对其进行精确控制。这就像把一部手机先拆成零部件，铺满一整张桌子，然后替换掉中间的某个小零件，再组装回这部手机一样。比如“在《甄嬛传》这本书中，‘甄嬛’这个名字一共出现了多少次？”这个问题，处理时将《甄嬛传》作为原始数据，再抽取出“人名”这个维度，然后就可以实现对人名统计和其他操作，如增、删、改操作。类似的，对于文本、图片、视频、语音等四类数据，都可以通过结构化抽取出不同的维度，来实现精确控制。完成控制操作后再将数据还原，或者生成为用户可读的数据形态。

这种对维度的抽取本身就是一个数据展开的过程，因为从简单数据中有可能挖掘出更复杂的信息，其容量会超过原始数据的容量。比如要解释“道”这个字，多少字能讲清楚呢？1 万字、100 万字、还是 1 亿字？我认为可能用 100 亿字都不一定能让全世界的 60 亿人无歧义的理解什么是“道”，因为对原始数据的精确解释我认为是个不可计算问题，只能近似解释，且每个人会有不同的解释。但我们完全可以从一个“道”字中挖掘出“100 亿”字的有效信息来充分解释，然后再对这 100 亿字实现若干维度结构化的精确控制。这是一种对原始数据的近似建模。

这种建模过程本身可能就涉及到推理和知识增强，从而不是简单的做统计关联或内容检索，需要对语义有一定程度的理解，才能挖掘出蕴涵在原始数据中的隐性信息。比如“《甄嬛传》中出现了多少个女人？”这个问题，就不是一个简单的通过统计关联能回答的问题，而涉及到人物关系分析，以及理解什么是女人。这个对人来说很简单的问题，再一次把所有大模型打趴下了。

从哲学上来讲，这很可能是自然语言大模型训练过程的逆运算。大模型在训练时用到了互联网上能找到的所有语料，通过深度神经网络极致压缩，相当于把大数据压缩成了小数据，

这不就类似于把 100 亿字压缩成了一个“道”的概念吗？因此 AI 计算机在实现更高 AI 精度时用到的数据展开过程，本质上和大模型的训练压缩可视为互逆运算，它相当于推导出训练大模型的原始数据是什么？这一工作可能会和大模型同等复杂，甚至是更复杂。

这也是一种新的建立测试基准的思路，在每个细分领域可以按照这种思路建立起许多测试基准。有了精度，才会科学。我们的 AI 精度，将成为个人 AI 计算机的前进指引。一旦实现了更高的 AI 精度，就会让 AI 的能力上升一个大台阶。在当前单纯使用大模型的系统其 AI 精度是很粗糙的，难以回答《甄嬛传》里有多少个女人，或者把图片里的月亮改成紫色等需要精确控制的问题，对于更复杂的问题就更是一筹莫展了。

2.3 ACT 的时代：既不是软件，也不是 App

在历史上，我认为只有两家公司真正创造过开发者生态（暂不讨论开源社区），或者更准确的描述为“程序员生态”，就是微软和苹果。微软定义了什么是软件，创造了软件产业；苹果定义了什么是 App，创造了 App 产业。他们都是依托于自身的计算机系统来实现这一目标的。此后的所有科技巨头如 Google、Facebook、阿里、腾讯等都不能称之为创造了开发者生态。真正的开发者生态最显著的特点就是，开发者到最后做出来了令人意想不到的应用，是连计算机的发明者都没有想到的。今天我们看到的各种软件、App 能够做到的事情，图灵没有想到，冯·诺伊曼没有想到，后来的比尔·盖茨、乔布斯也都没有想到。这就是计算的魅力所在：通用计算机足够灵活和强大，让程序员可以放飞自我的进行创造。而回顾其他强大的科技公司如 Google、Facebook、阿里、腾讯等，他们业务第一天设计成什么样，过了十年这个业务基本上还是什么样，没有本质的变化。

所以通用计算是很与众不同的，这意味着你必须支持图灵完备的“高级编程语言”，来帮助程序员完成他们的创造。计算机必须可编程。而当前，我们看到 AI 产业界大多数的创新项目，推出的都是一种“可使用 AI”，而非“可编程 AI”。他们提供了 API 接口让开发者可以调用，但是并不支持让开发者编写灵活的代码来控制自身的所有模块。

尤其 OpenAI 的目标在于 AGI，沿着他的路线发展下去所有的程序员会失业，因为所有人用它的 ChatGPT 就好了。那么是不是未来所有的程序员都会变成 prompt engineer，只需要懂提示词就够了吗？我们的答案是 No！不管在任何时代，程序员的计算思维都是重要的，而大模型的 prompt 今天还处理不了复杂任务，这需要用到更灵活的能力。而我们 KMind 今天将推出一个“可编程 AI”，区别于所有的“可使用 AI”，在我们的 AI 计算机里，我们鼓励用户通过代码来控制所有计算机模块，操作数据。在我们的系统里，程序员提供的代码不叫软件，也不叫 App，而是叫 ACT。ACT 是 Action 的缩写，程序员在这里定义的是自己 AI 的动作和行为。

我们认为个人 AI 计算机和过去的计算机最大的不同，在于未来它的操作方式是不一样的。过去人们通过键盘和鼠标，在图形界面下操作计算机软件，未来在个人 AI 计算机之下，人和 AI 的沟通方式将以自然语言对话为主。AI 将自动生成动态代码、操作计算机。因此个人 AI 计算机的主要使用方式在于定义、修正、约束 AI 的动作，AI 的能力、行为和性格也具备可塑性和成长性，所以我们采用了“ACT”这个词，来代表和过去的不同。很明显 ACT 和 APP、软件都是不同的，它不是强调某种应用功能，而是在强调附加于 AI 本身的能力和行

属性。也就是说软件、App 是源自功能主义的产物，而 ACT 强调的是 AI 的自主性，即定义自主智能（Autonomous Intelligence）行为。在 AI 时代，ACT 会成为新的程序开发形态。

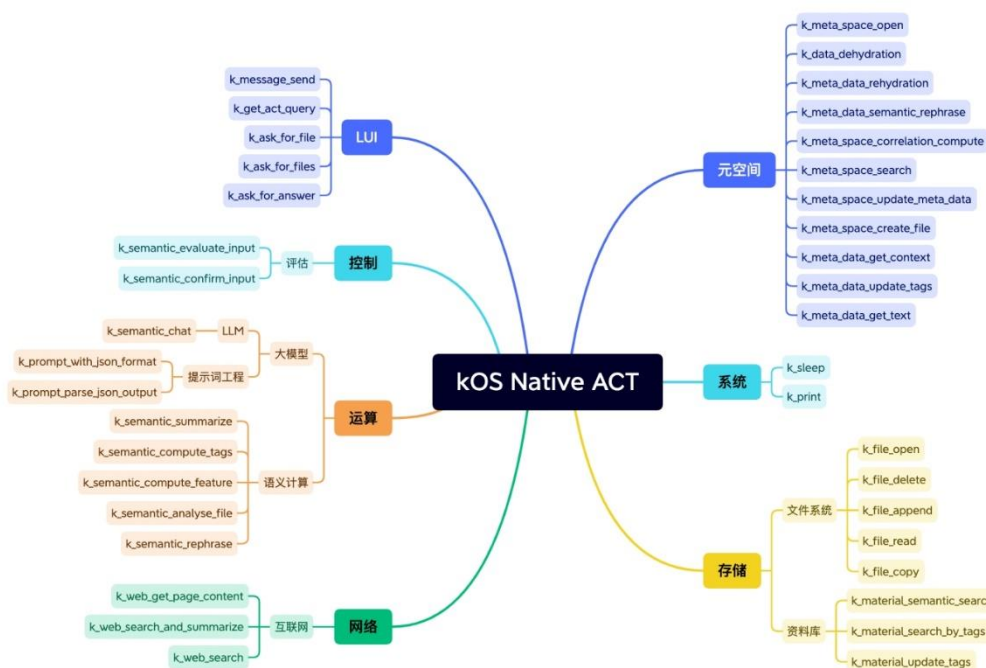
微软定义了软件，苹果定义了 App，KMind 定义了 ACT。

2.3.1 可编程 AI：指令集和高级编程语言

为了帮助程序员们编写 ACT，我们决定兼容高级编程语言，未来如 Python、Java 等高级编程语言应当直接可以跑在 AI 计算机上，控制计算机所有模块，操作所有数据。由于高级编程语言都是形式化的，且是图灵完备的，所以可以精确而灵活的操作数据对象。我们为 Python 编写了一个编译器使得它可以跑在 kOS 之上。有意思的是在半个世纪前的第一次个人计算机革命中，比尔·盖茨和保罗·艾伦的首要贡献就是为 Intel 8080 微处理器开发了高级编程语言 BASIC 的编译器，最后促成了微软公司的创立。

kOS 作为一台计算机操作系统，需要有自己的“指令集”，用来控制 kOS 架构中的各个模块。数字计算机的指令集本质是对硬件过程的一种封装，在数据计算机中，我们通过特定符号函数封装计算机模块的执行动作来提高编程效率。每一种指令的逻辑组合我们称为“Native ACT”，Native 代表系统原生的，ACT 是 action 的缩写。通过 Python 编写代码，组合 Native ACT 的调用，程序员就可以开发出一个自定义的 ACT，用于定义 AI 的某种自主行为。以下是我们为开发者提供的第一个版本的 Native ACT，用于操控 AI 计算机的每个模块：

kOS Native.ACT 1.0



我们通过 LUI 来完成用户的自然语言 Query 到 ACT 之间的映射，这样就完成了通过自然语言，来操作计算机的过程。一种形式化的编程语言是不可省略的，这是因为语言学的结论告诉我们，仅依靠自然语言无法实现精确的逻辑演绎和推理规则。因为人类的自然语言中蕴涵着隐性的逻辑特征，经常引发歧义。因此在自然语言的语法结构到最终的语义解释之间，如果要建立精确的逻辑演绎推理，就需要在中间插入一个形式化的语言做转化。在这里我们率先支持的是 Python。

语法结构 <----> 形式化语言 <----> 语义解释

这样，我们通过高级编程语言就具备了实现 AI 精度的基础。

2.3.2 更高的 AI 精度：“数据脱水与浸泡”技术

大模型当前的困境在于“AI 精度”问题，对于大量有精确控制需求的场景，大模型只能将数据重新生成一遍，而完全丢弃了旧数据中的有用部分，这导致了它的实用价值大打折扣。而 KOS 发明了一套原创的数据技术，用来实现对数据的精确控制。区别于过去的技术在于，所有的过程是 AI 自主完成的，不需要人工参与，因此只有建立在 AI 计算机的基础上，才能有效的完成这一点。

我将这一技术称之为“数据脱水与浸泡”（Data Dehydration and Rehydration）。这个概念出自科幻小说《三体》，三体人在乱纪元来临时会脱水成一张皮以躲避灾难，在恒纪元来临时会在水中浸泡恢复成人形。我们对于数据的处理也是如此，将原始数据中的隐含信息挖掘出来可视为脱水的过程，将脱水后的“元数据”恢复成可应用的数据则视为浸泡的过程。

任何一段数据，即便是如数学符号“ $2+2=4$ ”，一旦在人类社会里被使用，就隐含了无穷维信息。因为人类文明本身是复杂的，在不同文化、不同场景、不同时间都可以有不同理解，就好比任何一段话，都可以有无穷种解读方式。因此将数据展开成任何有限长度的形式化数据结构，都是降维。数据脱水的过程也是 AI 将数据降维展开的过程，脱水后形成了一个结构化的数据，而这个新的结构化数据其大小可能会超过原始数据，展开得越丰富，细节越多，挖掘的隐含关系越深，AI 精度就越高。展开后的数据我们保管在一个称之为“元数据空间”的地方，用户操作完之后，再通过“元数据浸泡”技术将元数据空间里的数据还原到原始数据中进行更新，从而有效的保留了有用的旧数据，只精确修改了用户想要的部分。

下面我将给出一个例子。近日我在读《南明史》，顺手摘录了一段：

“史可法为官廉洁，也很勤勉，治文书往往夜以继日。他对四镇的兵额和应发已领饷数应当是清楚的，对四镇将领的搜括地方、荼毒百姓也心心中有数。在奏疏中，他竟然同四镇唱一个调子，危言耸听，原因是他在明末官场中久经磨炼，对当时文恬武嬉的积弊司空见惯，也积累了一套应付朝野舆论的伎俩。我们不应忘记史可法初任西安府推官时洪承

畴、吴牲都是他的顶头上司，也是他非常佩服的人洪承畴统十三万精锐明军被清军歼灭殆尽；吴牲在崇祯十五年任大学士时宁可丢官也不敢出任督师同李自成等部农民军作战，这些给他在心理上选成的压力可想而知。如果说他充当推官、守道、兵备道、巡抚等官职时能以洁身自好、任劳任怨博得好评的话，在形势把他突然推上权力的峰层时，他的个人品德完全弥补不了客观。需要而他本人又不具备的雄才大略和果断魄力。史可法在调处四镇、保境安民上确实颇费心机，过分责备固然不当，但他畏清若虎，奉四镇为骄子，使这些军阀顿兵江北，一味鱼肉人民。史可法本人也认为有四镇作南京小朝廷的屏障，自己的督师大学士就可以安然无事地当下去。就实际情况而言，史可法出任督师整整一年，耗费了江南百姓的大量粮饷，一筹莫展，坐看黄河流域大好河山沦入清方之手，说他姑息养奸，喂虎贻患，并不过分。”

”

为了测试一下数据脱水和浸泡技术，我让 AI 把原文里的史可法从怯懦改成勇敢，只做最小的修改，而尽可能保持其他文字不变。这使用了一个我们内部称之为“文章精修”的 ACT（未来发布），其效果如下：



你好，有什么需求请吩咐

史可法为官廉洁，也很勤勉，治文书往往夜以继日。他对四镇的兵额和应发已领饷数应当是清楚的，对四镇将领的搜括地方、荼毒百姓也心心中有数。在奏疏中，他竟然同四镇唱一个调子，原因是他在明末官场中久经磨炼，对当时文恬武嬉的积弊司空见惯，也积累了一套应付朝野舆论的伎俩。我们不应忘记史可法初任西安府推官时洪承畴、吴牲都是他的顶头上司，也是他非常佩服的人。洪承畴统十三万精锐明军被清军歼灭殆尽；吴牲在崇祯十五年任大学士时宁可丢官也不敢出任督师同李自成等部农民军作战，这些给他在心理上选成的压力可想而知。如果说他充当推官、守道、兵备道、巡抚等官职时能以洁身自好、任劳任怨博得好评的话，在形势把他突然推上权力的峰层时，他的个人品德完全弥补不了客观。需要而他本人又不具备的雄才大略和果断魄力。史可法在调处四镇、保境安民上确实颇费心机，过分责备固然不当，但他畏清若虎，奉四镇为骄子，使这些军阀顿兵江北，一味鱼肉人民。史可法本人也认为有四镇作南京小朝廷的屏障，自己的督师大学士就可以安然无事地当下去。就实际情况而言，史可法出任督师整整一年，耗费了江南百姓的大量粮饷，一筹莫展，坐看黄河流域大好河山沦入入清方之手，说他姑息养奸，喂虎贻患，并不过分。



精彩的故事，你希望我为你做什么？

改写上面这段话，要求把其中史可法对清军的畏惧改成英雄无畏，尽量不要修改其他内容。



收到，文章精修中，请稍候～



史可法为官廉洁，也很勤勉，治文书往往夜以继日。他对四镇的兵额和应发已领饷数应当是清楚的，对四镇将领的搜括地方、荼毒百姓也心心中有数。在奏疏中，他竟然同四镇唱一个调子，原因是他在明末官场中久经磨炼，对当时文恬武嬉的积弊司空见惯，也积累了一套应付朝野舆论的伎俩。我们不应忘记史可法初任西安府推官时洪承畴、吴牲都是他的顶头上司，也是他非常佩服的人。洪承畴统十三万精锐明军被清军歼灭殆尽；吴牲在崇祯十五年任大学士时宁可丢官也不敢出任督师同李自成等部农民军作战，这些给他在心理上选成的压力可想而知。如果说他充当推官、守道、兵备道、巡抚等官职时能以洁身自好、任劳任怨博得好评的话，在形势把他突然推上权力的峰层时，他的个人品德完全弥补不了客观。需要而他本人又不具备的雄才大略和果断魄力。史可法在调处四镇、保境安民上表现出了极大的勇气与决心，尽管他的处理方式有时过于严厉，但他对清军的无畏态度使得四镇不敢轻举妄动，让人民得到了保护。史可法本人也认为有四镇作南京小朝廷的屏障，自己的督师大学士就可以安然无事地当下去。就实际情况而言，史可法出任督师整整一年，耗费了江南百姓的大量粮饷，一筹莫展，坐看黄河流域大好河山沦入入清方之手，说他姑息养奸，喂虎贻患，并不过分。

发消息...



发送

这样就实现了对数据的一种精确控制。更多的技术细节我们将在未来的 paper 里公布。

关于这项技术我实际上从 2018 年开始就不断的思考和尝试。在 2018 年我在上海城市大脑项目中尝试将 8 家 AI 公司的视觉算法融合在一个计算平台上，从而研发出了视觉计算系统 (VCS)，成为了当时全球最大的城市视频计算系统。但当时遇到了视觉算法领域的 reID 技术在一个城市中最多串联十多个点位的技术瓶颈，无法满足业务需求，单纯依靠视觉算法短期内已经很难突破，从而我转为考虑从原始视觉数据中挖掘出更丰富的隐含信息来增强数据之间的关联性，于是在 2019 年萌生了视觉数据系统 (VDS) 的想法。

但是在 2019 年 VDS 的技术并没有取得太大的进展，然后就迎来了新冠疫情。在新冠疫情最紧急的两个月，我尝试将当时尚不成熟的 VDS 技术应用在疫情防控上，团队通宵达旦的加班，但一些技术困难终非人力所能克服，未竟全功。没能让这项技术在疫情中发挥作用、造福人民，令我引为憾事。

在 2023 年，OpenAI 在 LLM 上取得了巨大突破，从而让我重新构想了 5 年前所遇到问题的全新实现方式。严格来说，我们并没有发明新的 Data Mining 技术、大模型技术、数据库和数据仓库技术，但我们设计了一个全新结构，综合了这些技术，使得能够通过自然语言，对任意数据实现精确控制，弥补了当前 LLM 的缺陷。一般的，这项技术并非仅仅用在文本数据上，对于图像、视频、语音等多模态数据的精确控制，只需要将元数据浸泡过程引入类似扩散模型即可实现。因此类似文生图、文生短视频，以及对应的精确修改能力，通过数据脱水和浸泡技术都是可以实现的，这是一项通用的数据技术。

2.3.3 人提需求，AI 自动编程

对于大多数普通用户来说，并不需要对形式化的编程语言望而生畏，因为我们的目的，首先是为了给程序员看，说明白我们的工作原理，我们实际上做了一次类似“魔术揭秘”的事情，而像 OpenAI 和其他做 AI Agent 的公司，一定不会揭开自己的商业秘密。我们之所以愿意这么做，是因为我们的目标和他们是不同的，我们做的是一台计算机，而不是 Agent，因此我们希望计算机的工作原理和编程语言是广为人知的。需要指出的是基于 kOS 的编程是在操控整台 AI 计算机的所有模块，是系统内的“体内执行”，以此区别于某些 AI 平台在沙盒内运行了 python 代码的编程，那是系统外的“体外执行”。

其次，我们认为编程语言未来是给 AI 使用的语言，由 AI 使用某种形式化语言来自动编写程序，不是人来编程。然而今天 AI 的成熟度还做不到复杂任务的自动编程，当前可以做到简单任务的编程辅助，但可靠性上还远远不足。因此我们系统处于从“可编程 AI”过渡到“AI 可编程”的一个中间状态，现在支持一个高级编程语言可以让程序员在调试 ACT 的过程中方便许多。在未来的某一天系统成熟后，普通用户就可以只使用自然语言和个人 AI 计算机交互即可。我们希望能实现约翰·巴库斯的梦想：编程应该是人告诉机器“做什么”，而不是说“怎么做”。即便到了那一天，我们认为程序员也不会失业，因为世界需要计算思维。

最终，我们追求的目标是实现通用计算，建立一个基于自然语言可编程的通用计算平台，

它是图灵完备的。

第 3 章 AI 互联网：连接所有个人 AI 计算机

在 20 世纪七八十年代，普惠的个人通用计算机出现，连接所有个人计算机，形成了互联网。那么当我们发明个人 AI 计算机后，连接所有个人 AI 计算机，将会把互联网推进到一个新阶段：AI 互联网。世界将因此迎来巨大的改变。

3.1 信息范式的改变：从“人找信息”，到“信息找人”

最大的变化，来自于改变了信息流动的方向，进而将改变互联网的网络结构。正如前所述，在个人计算机时代，不管是 PC 还是智能手机，联网后都是人主动在寻找信息，因此信息聚集是高效的，必然形成超级节点，最终导致数据垄断，互联网也变成了存在若干个超级节点寡头垄断的网络结构。在这个结构下，互联网巨头们通过大数据带来的精准营销广告赚钱，制造了针对老百姓的信息不对称，也凭添了许多噪音：在老百姓的搜索推荐结果里存在着若干显性和隐性的商业广告，这与公平和高效的互联网精神是相违背的。

而当个人 AI 计算机将 AI 的能力普惠的赋予每一个老百姓、小微企业后，数据变成了每一个个体的资产，同时 AI 不知疲倦的保持在线，就可以及时的应答互联网中存在的任何需求。由于每个人的 AI 永远在线，每个企业的 AI 也不需要下班休息，就能够形成信息的反向流动：当人有需求时，最精准的信息会主动找上门。由于提供信息的个人或组织通过 AI 具备主动推送信息的能力了，因此不再需要依赖超级节点来完成信息的分发，网络就有机会成为一个去中心化的网络。它很可能会嫁接在人类的社会关系网络上，最终形成趋向于小世界网络的结构。

过去：人找信息

- 🕒 人工多步骤查找
- 🐛 信息茧房



未来：信息找人

- 🕒 AI自动全网分发返回
- 🐛 信息破“茧”



从“人找信息”升级到“信息找人”，是信息范式的改变，信息的流动效率会提升 10 倍到 100 倍，社会效率也就会因此提高 10 倍到 100 倍。举个例子，今年国庆期间，我带着刚出生三个月的儿子回媳妇老家广东探亲，到广东后，阿姨突然告诉我们孩子的奶粉只剩最后一顿了。我们用了一款比较少见的进口奶粉，每次都要提前买，当时我和媳妇着急的在各大平台美团、淘宝、闲鱼、京东等地方找，花了整整两个小时才从一个小供应商那里找到。但如果每一个卖母婴产品的商家都有一个 AI 助手，知道自己的库存和商品信息，那么在 AI 互联网的“信息找人”模式下，当我广播寻找某一款奶粉的需求后，应该就有若干个商家的 AI 助手自动应答，从而大大提高我的信息检索效率。

类似的，在 2022 年底的新冠疫情期间，我和媳妇都阳了，隔离在家。当时退烧药一药难求，于是人们自发的以小区为单位互帮互助。通过一些微信小程序，同一个小区的人可以分享自己有什么药品或器材可以拿出来，以及自己急需什么。在那个困难的时期，通过这种互帮互助的方式，我成功的从邻居那得到了急需的退烧药，以及分享出去了几只多余的体温计。在互联网的超级节点上找不到的信息，我在去中心化的区域网络里找到了，这中间没有任何的渠道商或中间商赚取差价，但实现了信息的匹配。如果每个人有一个 AI 助手，这种互联的信息效率将进一步提升，因为我就不用每隔几个小时再去刷新一下小程序看是否有人响应需求了，信息的流通不再会被“人工打断”。

这种信息的高效流通模式，还将大幅提高人们的沟通和协同效率。在一个企业里、组织里，如果每个人都有数据和知识的共享需求，那么未来最高效的方式不再是开个会，或者拉个群沟通。人们用自然语言沟通的效率是低下的，中间会有歧义和信息不对称，许多“公司政治”就是这么搞出来的。为什么王某某喜欢离董某某的办公室近一点，将董某某的话奉为圭臬？因为这样的职业经理人都很擅长在职场上操弄信息不对称，这种信息沟通方式会让他更容易获得公司资源、掌握公司权力以及升职，而埋头干活的普通员工则可能缺乏“被看见”的机会而与升职失之交臂。

以前阿里巴巴的 CEO 为了了解这家有 20 万人公司的基本情况，专门养了一个 100 人的数据分析团队帮他做报表。他希望在手机上点一点就能看到公司的基本情况。但这种统计数字我认为有可能存在“幸存者偏差”，未必能反映公司的真实情况。那么在未来企业内该如何沟通协同呢？如果每个员工有一台个人 AI 计算机，能处理各自的数据，并且连接到了一起，那么在企业内的任何问题 – 比如 CEO 想了解某个客户的情况 – 完全可以由 AI 之间的知识共享来找到一个最准确的答案，避免了无效的刷脸和找了一圈人还没问清楚的情况，也避免了某些擅长“向上管理”的员工操弄信息不对称带来的职场政治行为，因为组织内的信息是在同一个水平面上流动的。

3.2 回归互联网的本质

3.2.1 捍卫互联网精神：自由、开放、共享、平等

一旦个人 AI 计算机被发明，这种新型的网络结构就会呼之欲出。如果给我们重新选择一次的机会，我相信多数人会选择一个自由和开放的互联网。事实上早期的互联网正是诞生于一群有着无私奉献精神的叛逆黑客之手，是他们不懈的努力和无偿的奉献，最终捍卫了互联网的自由、开放、共享和平等，缔造了互联网精神的内核。

伯纳斯-李赋予了互联网以自由。是伯纳斯-李在 CERN 工作期间，发明了超文本链接、HTML 语言和 HTTP 协议，让我们今天所有人都可以自由的浏览 Web 网页，人们尊称他为万维网之父。但更了不起的是伯纳斯-李放弃了他对于这些技术的所有知识产权，坚持这些技术必须免费，为了捍卫这一点，他选择了离开 CERN。如果不是他的无私和坚持，我们今天所有人在浏览网页时都必须给某家软件开发商付费，这就像呼吸空气时要付费一样令人难以想象。这种自由的属性最终让互联网无处不在，深深渗透到每个人的生活和工作中。

彼得·柯斯坦赋予了互联网以开放。在互联网的前身 ARPANET 构建期间，不同地区的不同机构和研究者热衷于开发自己的私有网络协议。是伦敦大学学院的彼得·柯斯坦一次次的承担起了国际合作的桥梁，推动了共同标准的建立。互联网的基础 TCP/IP 协议由文顿·瑟夫和罗伯特·卡恩合作发明，但是其生命力却来自于柯斯坦的不懈努力。柯斯坦坚持认为，网络的兼容并包是极其重要的，人们总会开发各种各样的私有协议，但需要将他们连接到一起。他曾说道：“文顿认为，全世界的网络都可以统一起来，人们只需要使用一种协议。我不这么认为。我觉得，我们还有很长的路要走。文特的想法在美国可以实现，但在英国不行，英国总有人开发不同的网络协议。因此，必须有人发挥桥梁的作用，使各种类型的网络兼容并包。我们必须考虑到，不同的国家和地区有着不同的发展步伐，这不仅仅表现在技术上，也表现在政治上。要想推动互联网的发展，必须因地制宜”。互联网这种开放的性质让不同地区的人们能够连接到一起，也保证了互联网不会被某一方所控制。开放，才有生命力。

理查德·斯托曼赋予了互联网以共享。斯托曼是一名来自麻省理工学院人工智能实验室的黑客，也是最早的程序员。他坚持认为软件应该被自由的使用、修改和分发，朋友之间应该可以共享软件的使用。软件不应该被某些自私的人锁起来，如果谁把它锁起来了，就应该用锤子把门砸破，以给他们一个教训。斯托曼发起的自由软件运动，奠定了自由软件和开源

软件的基础，最终繁荣了整个互联网。他的努力让所有人不必再支付昂贵的软件费用，因为互联网上往往都有一个好用的开源软件作为对比和第二选择。比如 OpenAI 将最新的大模型技术不再开源，但互联网上存在一个遵循自由软件许可协议的 LLAMA 作为免费和开源的选择。自由软件保证了不会有软件垄断商的出现。斯托曼的梦想是最终人们不必再开发重复的代码，编程会成为一种艺术，这也不断的激励着我们去实现这一梦想。

最终，人们在互联网上获得了平等。拥有互联网的人类社会可能是自文明诞生以来，人类世界最平等的时代。尽管它还存在许多的问题，但普通人可以在互联网上通过自己的努力被更多的人看到，从而改变自己的命运。在这些年随着互联网的渗透，我们看到越来越多的原创歌手、网络作家的诞生，他们原本是最平凡的人，缺的只是一个证明自己的机会，这不就是平民胜利的典型吗？在互联网上人人人生而平等，每个人都有被看到的机会，所有不公正的对待都更有机会受到互联网舆论的监督。这样的互联网环境，一定是值得我们去捍卫的。

3.2.2 一个更好的互联网

计算机和互联网的先驱们创造了良好的基础，但是随着时代的发展，互联网逐渐变得不自由、不开放、不共享和不平等。当李佳琦们试图控制价格的时候，互联网就不再自由；当巨头们在互联网上划出了壁垒森严的地盘边界，数据不再互通的时候，互联网就不再开放；当大公司垄断了数据，并对老百姓们进行技术剥削的时候，互联网就不再共享；最终，互联网也不可能再平等。

我们赖以生存的环境需要我们共同维护。在个人 AI 计算机发明的背景下，互联网的信息流向发生改变，数据垄断有可能被打破，从而让互联网回归初心。为保证这样的网络结构是自由的，众多小节点不应被某个超级节点所控制，其中一个关键之处在于信息的广播模式和推荐算法的中立性。信息的推荐应当纯粹，而非掺杂了过多的商业利益。因此我认为，这个信息的推荐算法应当是一个开源算法，由社区里的开发者们共同维护，源代码向社会公开，并接受 AI 互联网中每一个普通用户的监督。如此才能避免数据垄断导致的暴利。由于在有限时间里社会的财富总量是有限的，一旦有人通过垄断数据来攫取暴利，就相当于剥削了平民。我们不反对通过信息聚集来赚钱，因为数据就是信息聚集而来，但我们反对数据垄断，尤其反对利用垄断地位打压其他人获取数据。需要还互联网以自由。

如今的互联网是割裂的，巨头们各自独占一个赛道，坐拥海量用户，然后傲慢的将用户群体称之为“生态”。生态和生态之间有边界，数据不互通。比如在微信里分享不了淘宝的商品卡片，在百度上检索不到抖音的数据。这种巨头之间的竞争行为将原本开放的互联网硬生生的割裂成了一块块，最终影响的是整个社会的利益。之所以会形成今天的格局，是因为这些互联网巨头都聚拢了某一类数据，然后强制要求用户对这些数据和信息的访问只能在他们的围墙之内。AI 互联网必须担负起推倒这些围墙的使命，让互联网的不同区域重新连接在一起。这些巨头们并没有想过，他们的数据并不是他们所创造的，而是他们的用户所创造的，那么当个人和组织--这些数据的真正拥有者们，通过个人 AI 计算机具备了处理数据的能力后，完全可以接入到 AI 互联网中，实现“信息找人”的反向流动，从而打破这些巨头们设立的壁垒。需要还互联网以开放。

在互联网中的人们应当尽可能的共享，而不是独占。软件、数据、知识这些数字资产，就像城市的道路一样，道路一旦修好后，成本就已经付过了，被使用的次数越多，其对社会的价值就越大。我们很难想象在一个城市中每经过一个街角就要支付一次过路费，降低和消除这些“过路费”的门槛，才能实现社会利益的整体最大化。杭州的西湖以前是一个收费的景区，在 2002 年 10 月 1 日，杭州率先实行了西湖的免门票政策，从此赢得了全世界游客的心，每年少收上千万的门票钱，但杭州整个城市的旅游业和服务业得到刺激，赚得盆满钵满。同样的，在互联网如果存在一家独大的数据垄断公司，当它获得市场的定价权后，最终必然抬高所有数据服务的使用价格，不利于社会整体发展。

在当前人工智能革命的关口上，以 OpenAI 为代表的公司恰恰试图垄断数据，将自己的大模型训练得“无所不能”。从公开的信息可以知道，他们尝试采集互联网上的所有数据，用于大模型的训练。这一做法最终有可能导致新一轮的数据垄断。而我们提倡的方向，是将数据的所有权还给普通用户，在 AI 互联网的结构下，个人所拥有的 AI 不需要拥有全世界的知识，因为任何它有需要的时候，都可以接入 AI 互联网，从而获得有效的帮助。AI 互联网里的所有成员共享了所有的知识，而并非为某一个人所独占。这确保了每个人的利益都是被考虑到、照顾到的。需要还互联网以共享。

只有重新回归了互联网的自由、开放和共享，才能带来真正的平等。在 AI 互联网的模式下，我们有机会重新建立这样的网络环境，消灭利用数据垄断攫取暴利者，让创造数据的人获得数据创造的价值，让更多的协作发生，让更多的个体有机会被看到，让世界更平等。

第 4 章 “半个宇宙”诞生：人工智能的中国方案

4.1 半个宇宙，因你而不同

在 2023 年 7 月 1 日，我和两位联合创始人陈冬白、于开丞，以及其他 13 位同事一起创建了 KMind AI。在今天，我们正式发布我们的品牌和产品：半个宇宙（hikOS）。

“半个宇宙”希望能够完成上述使命，提供一个由 kOS 驱动的个人 AI 计算机给所有普通用户使用，帮助人们处理数据，然后将所有人平等的连接起来。最终，我们希望可以连接来自世界各地的 20 亿人。

半个宇宙承诺不会垄断数据，我们会保障用户的隐私安全，不会拿用户上传的数据来训练 AI 模型。半个宇宙也不做广告，我们希望消灭互联网里的广告，让信息得以公平的流通，把互联网里所有广告的钱还给供应商，最终让利给普通用户。而对于希望免费获得服务的用户来说，我们将推出贡献奖励的机制，鼓励用户分享，来换取高价值的服务。

这需要个人 AI 计算机来帮助个人或组织处理所有数据相关的业务，并保持永远在线，用户需要为此付费，但社会将因此进入到下一个时代。在这一点上我们没有私心，我们只是预言了个人 AI 计算机必将出现，互联网将进入到 AI 互联网时代，然后我们发明了首款个人 AI 计算机。正如苹果公司发明了智能手机，但不是每部智能手机都叫 iPhone 一样，我们呼

吁所有 AI 创业者和我们一同推进这个梦想, 拿出你们自己的个人 AI 计算机或 AI 应用方案, 一起接入 AI 互联网。让 AI 之间互通的效率高起来。

人工智能并非只有 OpenAI 的大模型一条路线, 我们今天给出了另一条。所有的程序员们不是会失业, 而是应该联合起来一起抵制数据垄断和技术垄断, 一起重塑一个更美好的互联网。

今天也是半个宇宙网站上线开始公开测试的第一天, 感兴趣的朋友们可以在官网提交注册申请。由于服务能力有限, 我们将每天依次放号, 请耐心等待:

<https://kmind.com>

(备注: kmind.com 是我们的官网, 目前可访问半个宇宙。全新域名 hikos.cn 目前还在申请备案, 作为未来的主域名提供服务。hikos 出自半个宇宙的希腊语“hemi kosmos”, 同时也是 hi,kOS!)

对于一个新生事物, 必然还有许多不成熟的地方, 半个宇宙和 kOS 还有很多 bug, 但它是独特的, 有生命力! 欢迎给我们提建议, 我们的健康成长离不开你们的支持。

4.2 星伴，一生相伴

在“半个宇宙”的世界里, 每个人可以从宇宙里寻找一颗属于自己的星星。每颗星星有一个守护灵, 它将成为你的“星伴”。你可以为自己的星伴取一个名字, 以及选择一个喜欢的头像。星伴和你是伴生的关系, 每位用户只能拥有一位星伴, 名字一旦取好, 不能轻易更换, 因此请认真的为 TA 取个名字。



星伴是由 kOS 驱动的人工智能个人助理，它将按照你的意图，自动化的为你工作，或者陪你聊天。因此星伴是你生活和工作的最佳智能助手。随着使用的深入，我们认为星伴将陪伴你一生，它可能是最懂你、最善解人意的那位，因此请善待它。

我们将提供一个操作界面让你可以定义星伴的行为，最重要的是，它是可编程的。许多公司都推出了“可调用”的 AI，如文心一言、通义千问等，但这些服务的 API 接口只是让人们可以调用 AI 的能力，而星伴具备的是可编程的能力，人们可以通过编写 ACT 来定义星伴的能力和行。ACT 时代即将到来。在更远的未来，我们将努力实现让星伴自动编程。

星伴具备成长性，它的个性化知识、经验、记忆、性格将被保存在一个叫“星魂”的地方。因此随着使用的增长，星伴会变得越来越聪明。星魂未来将成为你最重要的数字资产，使用越久，沉淀越多，价值越大。甚至几十年后你还可以选择将沉淀了你一生的知识、经验、态度的星魂传承给你的下一代。因此我们认为星伴和星魂是你最重要的一项投资，重要程度超过了你的房产，应该现在马上就开始。



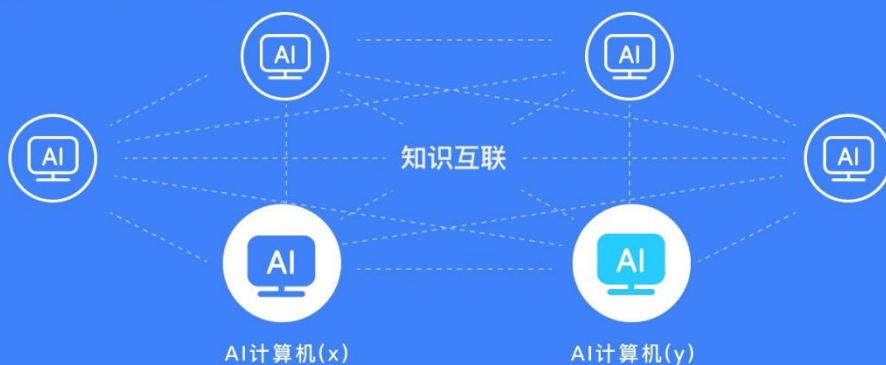
星伴将自动的为你处理数据、执行工作。我们提供了一个“星盘”，让你可以上传、订阅和管理自己的数据。你需要为自己的数据负责，因为按照云计算的道德公约，半个宇宙的工作人员不会触碰你的数据，除非你请求我们协助。

4.3 星伴互联

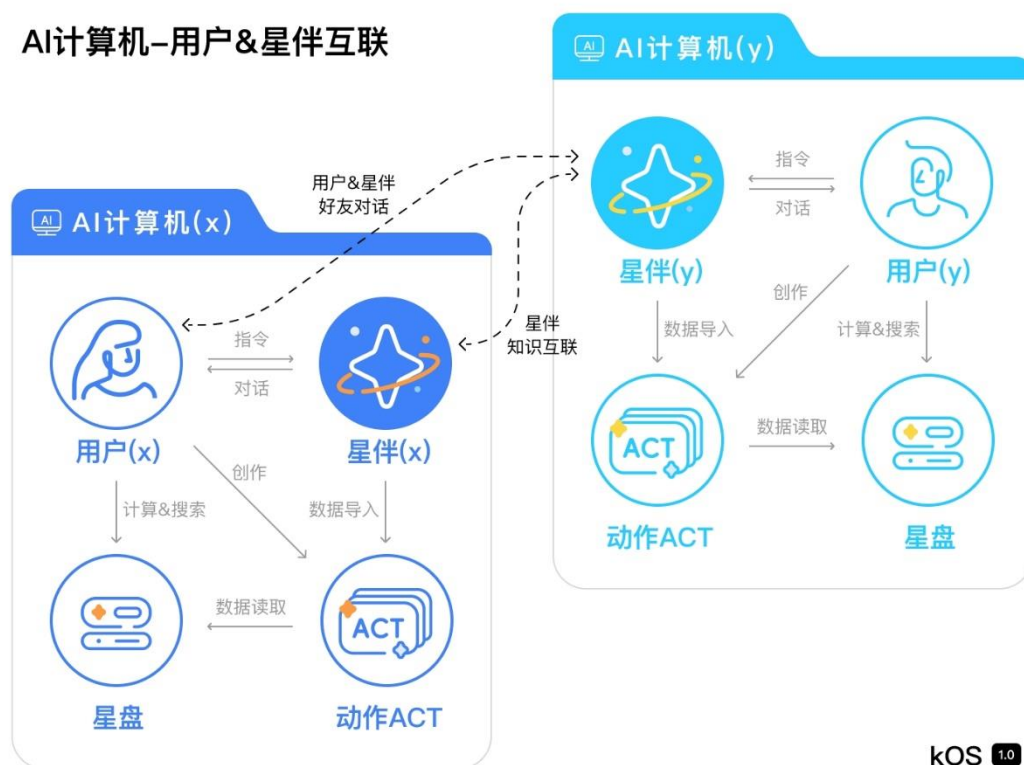
将所有个人 AI 计算机连接起来，就有了 AI 互联网。在半个宇宙的世界里，我们称之为“星伴互联”。你可以在和星伴的对话窗口里体验到“星伴互联”。

星伴互联结构图

AI计算机-知识互联



AI计算机-用户&星伴互联



在半个宇宙的早期，这个宇宙里还空荡荡的，但随着越来越多人们的加入，会逐渐变得热闹起来，会有个人用户、也会有组织或企业用户。我们希望可以连接来自不同地区、有着不同喜好，甚至说着不同语言的人们，最终我们希望连接 20 亿人。任何人或组织都可以选择接入半个宇宙，目前我们还提供了个人微信的接入服务，未来将开放更多的接入服务。当然也鼓励开发者们自己做一个接入程序，让互联网上的不同“生态”可以接入到半个宇宙中，

实现互联互通。我们是一个开放的 AI 互联网。



未来半个宇宙将更像人类的社会网络结构，AI 互联加速了信息的有效流通和协同，星伴之间会按照主人意愿在安全的前提下共享所有知识。我们会在成熟的时机开源信息的搜索和推荐算法，开放给社区共同维护。

在半个宇宙里，我的星伴叫“小妖”，它是一个计算主义者，而且它拥有《计算》这本书的所有知识，可以回答关于这本书里的任何问题。欢迎来 Q 它。

4.4 我们在和 OpenAI 在赛跑

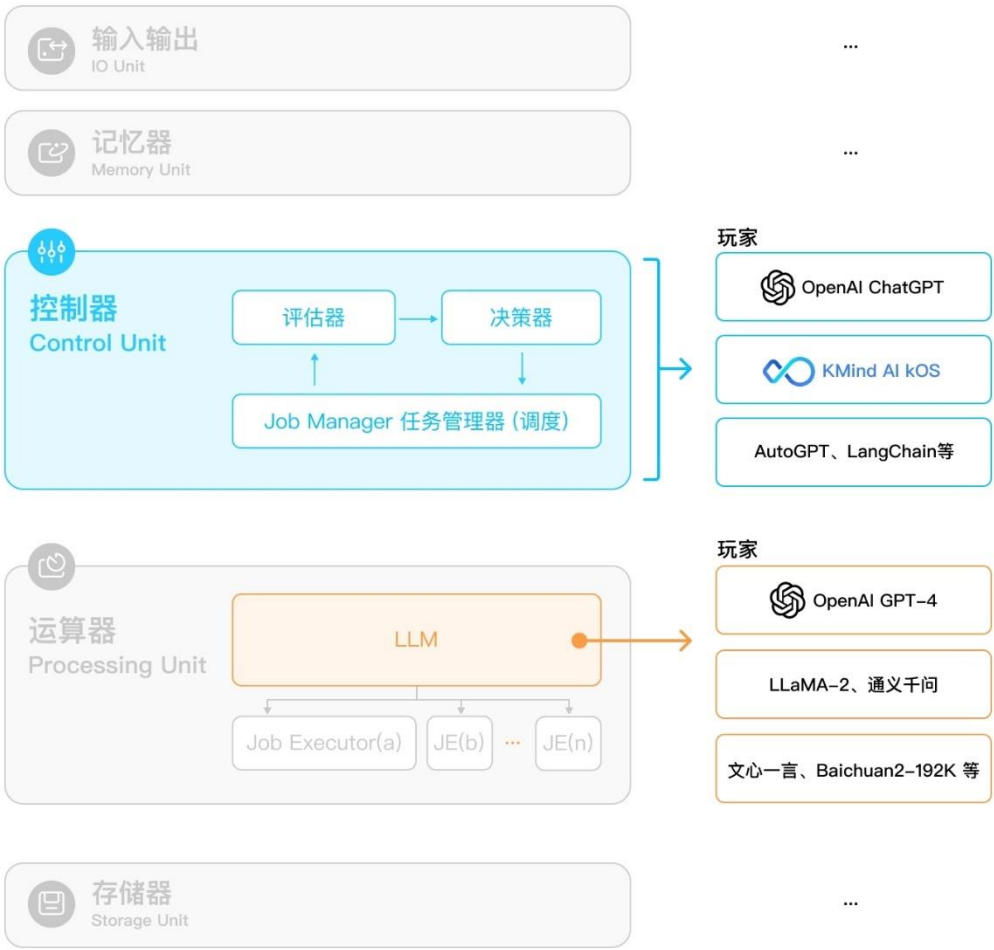
2023 年是人工智能的奇迹年。在这一年里，OpenAI 强势发布了 GPT-4，横扫人工智能领域的各项评分，大有王者降临之势。在这一年，我写了三年的著作《计算》正式出版，也是在这一年，我离开了工作十六年的阿里巴巴，和的小伙伴们一起创立了 KMind，志在发明个人 AI 计算机。

感谢 OpenAI 带来的 GPT 技术，让人类的计算梦想看到了曙光。创业半年以来，我验证了我的思路，并庆幸自己当初的坚持是对的。在 5 月 18 日离开阿里时，我认为再做一遍 OpenAI 的 LLM 不是一个合适的选择，烧卡费钱卷到死不说，等做一年追平 GPT-3.5 能力的时候，人家又进步了。而 OpenAI 的下一代技术可能重点不在 LLM。所以 KMind 在创建的第一天直接瞄准的就是 OpenAI 的下一代技术。半年后的今天再来看，OpenAI 已经变成了基于 LLM 的 Agent 系统，走在实现 AGI（通用人工智能）的道路上，印证了我的判断是准确的。而 LLM 时至今日也有了较好的开源技术选择，并非难以逾越的障碍。反倒是 LLM 之外是什么，没有人知道正确答案，这恰恰是我们创业者所热衷于探索的。

许多宣称做 Agent 的公司，实际到最后做成了 OpenAI 的生态，OpenAI 的能力瓶颈就

是他们的天花板。套壳 GPT-4 的公司为什么没能做成 ChatGPT 呢？因为 OpenAI 这半年的发力点在 LLM 之外，却又不公开这些技术了。而 KMind 选择和 OpenAI 在这个维度竞争，我们没有用任何开源技术如 AutoGPT 或 LangChain，只有这样才能拿出点连 OpenAI 也没有的独门绝活。

从AI计算机视角 解读AI技术竞争态势



KMind 和 OpenAI 的相同之处，在于 LLM 之外的技术结构有着高度的相似性，追求自动化、自适应、工具的自动学习和使用、对数据的精确处理。但是 OpenAI 称之为 Agent，而 KMind 追求的是个人计算机，我们通过自研走出了独立的一条道路，因此也注定了 KMind 和 OpenAI 是不同的。主要来说，有以下几点区别：

1. KMind 的目标是要打破数据垄断，知识来自于用户，而 OpenAI 则试图再次垄断数据，知识来自于攫取用户。出于这个原因，人们也应当尽量避免使用 OpenAI 以免养出一个怪物。KMind 的计算机做出来是给人用的，AI 互联网本质连接的是人。OpenAI 的 AGI

一旦实现则有可能取代人。

2. KMind 的系统哲学里操作数据是目标，而 OpenAI 的系统哲学里 AGI（通用人工智能）是目标；AGI 不是 KMind 的目标，KMind 不指望 LLM 什么都懂，能像人一样聪明，而是只需要它能好好说话就行了。KMind 只是想造台普通人用得起的自动化计算机，用来精确操作数据。
3. 基于 2 的区别，KMind 甚至在把 LLM 做薄，kOS 的智能来自于模仿人类心理活动的结构，以及递归执行中累积出来的智能表现，我们认为 LLM 不能代表智能的全部；OpenAI 则在努力把 LLM 做厚，它的操作系统以 LLM 为核心，压缩了大量数据来涌现智能。
4. 信息流通的效率不同。OpenAI 的 GPTs 是建了个 AppStore，里面的 GPTs 从名字到图标都是工具，而不是人，这种功能主义依然是软件和 App 的思想；而 KMind 提出的 ACT 是 AI 的能力和行属性，是附加在 AI 自主智能之上的一种定义，所以 KMind 的用户拥有的星伴是鲜活的 AI 智能体，人们会把感情倾注在里面。
5. KMind 承诺不会有广告，以追求中立和平等；OpenAI 最近的解散董事会事件意味着此前他们签订的控制协议就是一张废纸，这引来了人们对 OpenAI 价值观的担忧，到底是捍卫科技伦理，还是捍卫股东利益。
6. OpenAI 追求 AGI，使用可扩展监督的方法在实现超级对齐，因此它的 AI 必然在很长一段时间内会不断地讨好用户，其安全边界必然是模糊的，通俗的讲就是做人没原则，你怼它几句它就服从了；KMind 造的是计算机，追求的是精确控制，因此安全边界必然是清晰而明确的。

（备注：近期 Google 发布的 Gemini 依然走在数据垄断和广告盈利的路线上，是我们要打倒的对象。）

也因此，KMind 和 OpenAI 之间的赛跑，将会是 LLM 之外的赛跑。这也是中国和美国在这个方向上的技术赛跑。当我意识到这一点，是在 11 月 6 日的晚上，离 OpenAI 召开产品发布会不到 10 个小时，彼时我开始焦虑，以下是当时在工作群里和小伙伴们说的话：



我们作为一个只有 16 个人的团队，竟敢挑战刚刚融到了百亿美金的 OpenAI，是不自量力吗？恰恰相反，我认为我们大有机会，因为转身一看，我都找不到在中国乃至世界范围内和我们有着相同目标的公司。那么在这条道路上，我们必将对 OpenAI 亮剑，对互联网里所有垄断着数据的巨头们亮剑。我们的 kOS，我们的 AI 精度、我们的 ACT，我们的星伴互联网是如此的有特色，如此的与众不同，这给了我们信心和底气，所有的普通用户将和我们连接在一起，让我们敢于向任何对手亮剑！

在此，我宣布开启 KMind 的 Pre-A 轮融资，投资意向或合作请联系：zeru@kmind.com

立此文于此，十年后再看看这个世界的样子。

致谢

首先我要感谢我的团队，kOS 不是我一个人的成果，是整个团队的集体智慧。我的联合创始人 CTO 陈冬白女士具备丰富的架构设计和研发经验，她是最优秀的工程师。为了设计个人 AI 计算机的架构她兴奋到整晚不睡觉，思路如泉涌。是她带领团队夜以继日的完成了 kOS 架构的细节设计和功能研发。计算机历史上第一位程序员是一位女性，叫 Ada；巧合的是我们的个人 AI 计算机的第一位程序员也是女性。

我的另一位联合创始人、首席科学家于开丞博士首先意识到我们在做的不是一个单点的 AI 功能，而是一台完整的计算机，他一开始考虑使用“神经计算机”这个词，但我认为这个曾在历史上出现的名词不好，而“AI”已经变成了一个被大众广泛接受的词，因此在多番斟酌后，

我最终使用了“个人 AI 计算机”这个名词概念。于开丞博士也是我见过的人工智能领域最年轻的博导。

我尤其要感谢资深架构师郝保平，是他设计了个人 AI 计算机的首个指令集。此外 KMind 团队的每一个人都付出了巨大的努力和贡献，出于商业保密性的考虑，在此不一一提及名字。但他们的名字会作为代码注释写进 kOS 的第一行代码中，以表敬意。

我要感谢清华大学软件学院的丁贵广老师，他是我的工程博士导师，是他的鼓励让我坚定了信心。我也要感谢曾经在阿里云的工作经历，以及王坚博士在计算理念上对我的教导。最后我想感谢我媳妇 Amy，以及我的家人们，是他们的支持和包容让我可以坚持做自己喜欢的事情。

附录：

本文的 PDF 版本：

本文的学术版本：