

ChatGPT 时代我们的“脸面”何在

——多维视野下的颜面现象学

徐英瑾

内容摘要：人脸在人类的社会生活中扮演了双重角色：一方面，人脸如同一扇门，向我们展现了他人的部分心理活动；另一方面，人脸又如一道墙，向我们遮蔽了他人另外一部分的心理活动。人类所有的社会交往活动都有赖于人脸的这种微妙平衡机制。不过，人脸的这种微妙作用，却在刘慈欣的科幻小说《三体》的“思维透明人”假设中被无视。根据这种假设，任何一个三体人都能绕开“人脸”这一媒介而对他者的颅内信息进行透视。无独有偶，时下流行的网络聊天软件 ChatGPT 所依赖的超级语料训练模型也是“无脸”的，因为该模型已经隐去了对于个体知识提供者的“脸”的标注。然而，无论是刘慈欣的科幻假设还是 ChatGPT 的操作模式，都会对以科学精英的“脸面”为核心的人类传统科研模式的可持续性构成威胁，特别会对使科研创造性得以可能的底层逻辑构成戕害。

关键词：《三体》 思维透明人 ChatGPT 人脸 进化论 颜面现象学

作者：徐英瑾，复旦大学哲学学院教授。（上海 200433）

基金项目：国家社会科学基金项目“对于通用人工智能与特定文化风土之间关系的哲学研究”（22BZX031）；国家自然科学基金项目“探索研究 AI 伦理对科研环境的影响”（L2124040）

人脸在人类的社会生活中扮演了双重角色：一方面，人脸如同一扇门，向我们展现了他人的部分心理活动；另一方面，人脸又如一道墙，向我们遮蔽了他人另外一部分的心理活动。人脸的存在，在人类的社会生活中平衡了“公开性”与“隐私性”这两个貌似冲突的特征。人类所有的社会交往活动——当然也包括科技活动——都有赖





① 刘慈欣：《三体》（全三册），重庆：重庆出版社，2010年。

② Emmanuel Levinas, *Ethics and Infinity*, Richard Cohen trans., Pittsburg: Duquesne University Press, 1985, pp. 85-92.

③ 和辻哲郎：《面とペルソナ》，《思想》1935年6月号。此文还有一个线上阅读版本：https://www.aozora.gr.jp/cards/001395/files/49911_41926.html。本引文直接翻译自该网络电子版。

于人脸的这种微妙平衡机制（详后）。不过，在刘慈欣的著名科幻小说《三体》^①中，作者却构想了一种完全抛弃人脸的上述平衡机制的社会组织方式。具体而言，作为外星人的“三体人”的思维是彼此完全透明的，张三完全知道李四在想什么，李四亦然。因此，在三体人的社会组织里，不可能存在欺骗活动，也不可能有什么马基雅维利式的诡诈——当然，更没有任何隐私。不过，在《三体》构造的这个幻想世界里，一个内部完全不存在欺骗的社会团体却能够爆发出惊人的科技生产力，并因此获得远超地球人的科技成就。

《三体》中所做的上述设想虽是科幻文学的一部分，却也具有哲学思想实验的特征。我们不妨将该思想实验的场景挪移到地球上，并沿着刘氏的思路抛出下面这个问题：如果人类现有的“读心”（mind-reading）技术发展足以让我们通过特定设备全面监测到他人内心活动的水平的话，那么，这究竟会促进科技的进步，还是会起到相反的作用呢？对于这个问题，笔者给出的答案是否定的。笔者借以支撑上述答案的论据，一方面来自日本哲学家和辻哲郎（1889—1960）对于人脸的现象学分析，另一方面则是来自进化论的考量。同时，笔者关于“透明人假设”的这种批判意见，也可以被延伸到对于时下正火热的 ChatGPT 技术的批评上去。

和辻哲郎论“脸”

谈到与人脸现象学有关的哲学资源，不少读者可能会想起法国哲学家列维纳斯在《伦理与无限性》中的相关讨论。^②不过，列维纳斯主要是在马丁·布伯式的“吾-汝”（I-Thou）关系中讨论人脸所揭示的伦理学意义与神学意义的，与我们这里所聚焦的日常人际协作关系并不直接相关。相比较而言，日本哲学家和辻哲郎在1935年发表的短文《面与假面》（「面とペルソナ」）中对于“脸”的讨论，则具有更丰富的现象学细节，对于日常人际协作活动的指涉性也更强。他写道：

我们在可以不知道对方“脸”的情况下和他人交往。书信、口信等语言的表现都可以作为交往的媒介发挥作用。但即使是在这类情况下，我们并不会因为不知道对方的脸，而认为对方没有脸。多数情况下，或是基于对方通过语言所表现的态度，或是基于文字所蕴含的情绪，我们都在无意识地想象着对方的“脸”。这虽然往往是被大家忽视的一件事情，但在有时候却也能给人强烈印象……除了脸，同样联结于人们记忆的，还有诸如肩膀、背影、步履等其他特征。但是，即使我们将这一切特征都加以排除，我们依然能想起此人；唯独只有脸，是绝不可以被排除出去的。即使是在我们怀想一个人的背影之时，脸庞也正对着我们。^③

现在我们使用现象学的语言将和辻的这段话再“翻译”一下。在日常的劳动协作中我们将不得不面对大量来自他人的信息，而其中很大一部分信息（如书信、电报、电邮等）都是以“无脸”的形式出现的。但是，要将这些信息在脑海中激活，我们依然需要构建出一个现象学界面，以便赋予每一个他人的名字以一张活生生的脸，然后

才能决定与这些他人打交道的具体策略。甚至在我们所面对的他人是素昧平生的前提下，我们也会根据其文笔与字迹构想出一张他的脸，以便其能够成为现象学界面上的一个操作按钮。之所以唯有脸——而不是手或者背影——才能成为对于他人的最集中的现象学概括方式，则是因为人脸能够承载人的其他身体部位所难以承载的如下核心信息：他人的年龄、性格、态度甚至是德性状态。因此，脸便成了我们得以窥见他人心灵活动的窗户。

然而，和辻又在同一篇小文中将笔锋一转，谈到了日本能乐表演中的假面的作用。具体而言，这些假面将演员的脸完全遮蔽起来，而该面具的表情则是纯然僵死的。因此，演员只能通过丰富的肢体表演来赋予自己所塑造的人物以情感，而观众也只能通过自己的“脑补”为僵死的面具添加上自己想象的表情。非常有意思的是，这种将人类最能表现其情感的身体部位加以遮盖的表演方式，恰恰带给了观众巨大的审美享受。

和辻讨论能乐表演形式的真正哲学目的，是为了展现“无”（即对脸的遮蔽）对于“有”（即被展现出来的脸）的巨大提示作用，由此再迂回到对于一种基于“无”的佛教式审美哲学的辩护上去。不过，若将此类讨论的美学意蕴过滤掉，我们也能从中提炼出一些对一般意义上的“自-他”关系均有效的哲学推论。概而言之，一个人的脸并不是在任何时候都带有表情的，毋宁说，人有时候会面无表情，有时候则会强装欢笑（或者悲痛）。因此，虽然世界上的大多数人都不是能乐演员，但也都会间歇性地戴上假面。从这个角度看，脸就具有了“既展现情感，又遮掩情感”的辩证性。很显然，和辻哲郎本人并不对脸的这种辩证性感到不安，反而对这种辩证性表示欢迎，因为在他看来，这种辩证性本就是植根于“有”“无”之间的生死依存关系的。

由此，从对于脸的现象学分析出发，我们就能获得某项对人类社会普遍有效的人际交往规则：人必须通过人脸将一部分内部情感信息表达出来，同时，脸部肌肉的可控性又使得我们随时能将内心的一部分遮掩起来。需要指出的是，在和辻哲郎看来，现代通信技术的日益进步，其实并无法真正动摇这种基于脸的“揭示-遮蔽”辩证法的有效性。换言之，“公开性”与“私密性”之间的张力是无法被任何技术进步所克服的。首先，这是因为小团队的既有成员未必想将小团队内部的事情全部予以公开化。

“对于吾辈心中深深的苦痛，我们只能向亲友或者爱人吐露，却不会向无利害关系的人或向非亲非故者敞开心扉”。因此，对一个不想对圈外人表示痛苦的人来说，他依然会“满脸笑容”。其次，由于各小团体各自的风土差异所导致的兴趣差异，外人也未必会对一件本地新闻具有深入了解的意向。因此，“那些负责向国际社会传送本国新闻的人，会基于国际读者的兴趣而选择本地事件的一小部分予以报道，并不会真想将本地发生的一切让世界知道”。^①据此，一个明知听众对某事不感兴趣的电视记者，亦会在脸上佯装出自己也对此漠不关心的表情。

在笔者看来，脸之所以有这种遮蔽作用，在相当大程度上是为了使得社会网络中的每一个主体都能以更快捷的方式进行信息处理，以免出现“信息过载”的问题。

① 和辻哲郎：《面とペルソナ》，第224页。



① 和辻哲郎:《面とペルソナ》,第225页。

② Daniel Dennett, *Darwin's Dangerous Idea*, New York: Simon & Schuster. Illustrated ed., June 12, 1996.

③ Jeremy M. DeSilva James F. A. Traniello, Alexander G. Claxton, et al., "When and Why Did Human Brains Decrease in Size? A New Change-Point Analysis and Insights From Brain Evolution in Ants," *Frontiers in Ecology and Evolution*, vol.9, 2021, <https://doi.org/10.3389/fevo.2021.742639>.

④ 对于认知谬误的详细讨论,参见徐英瑾:《认知成见》,上海:复旦大学出版社,2015年。

⑤ Nelson Cowan, "The Magical Mystery Four: How is Working Memory Capacity Limited, and Why?" *Curr Dir Psychol Sci.* vol.19, no.1, 2010 Feb 1, pp. 51-57, doi: 10.1177/0963721409359277.

这一过滤机制非但不会随着信息技术网络的扩大而被削弱,反而会由此得到加强,这又是因为:人脑信息处理能力的上限是一个生物学定量,它不会随着外部信息处理技术的增强而增强。因此,外部信息量的激增,反而会进一步凸显人脑的信息处理能力与有待处理的海量信息之间的矛盾。在这种情况下,海量的信息本身必须经由某些代理者(如媒体管理者)的运作才能够以恰当的方式得到筛选,以便顾及每一个信息接受主体的信息处理能力的上限。从这个角度看,我们也就能理解和过下述评论的含义了:“所谓的‘诸个体的国际社会’一语,仅仅是指那些各自代表各自民族的个体所构成的社会,而不是指那种全然脱离了各民族特性的世界公民所构成的社会。”^①这也就是说,只有通过各个民族的政治精英所实现的纵向管理以及对于相关民族的“代表”机制,诸民族之间的国际信息联通才会成为可能。因此,对于特定民族共同体的成员来说,国际新闻也往往是通过一张张“脸”(特别是新闻主播的脸)来加以传达的。

在这种情况下,如果我们强行按照刘慈欣的假设撤去每一个人的脸,并设法使得每个人的内心活动都全部公开化,又会发生什么事情呢?难道真会发生他所说的科技大爆发吗?下面我会从进化论的角度说明:这是不可能的。

进化论视野中的“无脸人”

在这里笔者首先假设丹尼特(Daniel Dennett)在《达尔文的危险观念》^②中表述的上述观点是对的:“进化论是腐蚀宇宙一切角落的超级强酸”——也就是说,自然选择的道理,即使对于地外生命都是适用的。而自然选择的一个基本原理就是“吝啬性”,换言之,如果演化出特定生物学性状的生物学投入高于其生物学产出,那么,该性状就不会被自然选择所偏好。鸵鸟飞行能力的退化便是一个明证。

而在人类的各种器官中,作为中枢神经系统的脑所消耗的生物学资源是非常可观的。没有明显的进化论证据说明人脑的生物学机制在比较晚近的历史中在持续进化,正如生物学家德希尔瓦(Jeremy DeSilva)所说:“人类的大脑的现有尺寸,是六百万年前的人类的四倍(彼时人类与其他猿类刚刚分离),但从上一次冰期开始,人类的脑容量却缩小了。”^③因此,按照上述的“吝啬性”原则,我们大致可以认为:人类现有的大脑已经达到了灵长类所能达到的增长极限,因为人脑容量的进一步增长所带来的生物学副作用(如庞大的头颅给妇女分娩所带来的危险)可能是人类所无法承担的。然而,即使在如此奢侈的生物学配置的支持下,人脑依然会倾向于犯下各种各样的推理谬误,如在逻辑或概率演算中系统违背形式规律的要求,等等。^④此外,工作记忆时间的有限性又使得我们在一时间段内只能处理几项数量有限的认知任务。^⑤由此不难推出:人脑现有的信息处理能力的上限,已经是生物学演化所能提供的“天花板”。

那么,在别的星球上,外星智慧生物的大脑是否能够将这一天天花板抬得更高呢?

我们当然不能排除这种可能性，但从哲学角度看，一个被抬得再高的天花板毕竟也有自己的限度。外星大脑与人类大脑之间唯一不同的便是：一个外星超级大脑会有一个更大的工作记忆池来一下子记住更多的当下任务，会有一个更敏捷的联想机制以便将不同的知识模块联系起来。然而，这样的大脑依然会遗忘不少事项以便聚焦当下的任务，依然会有“锚定效应”或“刻板成见”以防止自己在无穷无尽的归纳活动中损耗自己的生命。因此，他们依然需要一个“现象学界面”以便控制从“内”流向“外”或从“外”流向“内”的信息总量——而这个界面就是“脸”。我们在此当然无法预报外星人的“脸”的具体物理形式（假若有人告诉我某个星球上的人的脸是以等离子体出现的，我也不会太吃惊），但只要外星人也是哲学意义上的“有限存在者”，他们定然是“有脸”的。

需要指出的是，就地球上的科技文明发展而言，“脸”的存在构成了科技发展的基本先验条件。我们都知道，伽利略-牛顿的科学革命之所以会在西方引发巨大的反响，便是因为“科学共同体”在近代欧洲的出现。此类共同体通过科学协会、科学杂志等建制化设施的作用，频繁地进行科学交流，由此使得学科进步的速率大大提升。而从和辻的“颜面现象学”的角度看，所谓“科学共同体”，其实也就是科学家“露脸”的舞台——也就是说，每一次公开的科学报告与论文发表，都构成了一位科学家的“脸”。这些“脸”当然也承载了和辻所说的“遮蔽—去蔽”的辩证法——也就是说，科学家既通过这些“脸”表述了自己想说出来的科学观点，也遮蔽了那些科学共同体所不关心的信息，诸如自己最近的感情生活细节。由此，科学共同体才能以更高的信息处理效率进行科学理论的迭代。从这个角度看，中国古代科技之所以落后于西方，根本原因也不是因为古代中国人“好面子”，而恰恰是因为古代中国人没有在原始意义的“面子”上构建出虚拟意义上的“面子”，即公共的科学信息交流平台。

而从历史唯物论的角度看，“脸”的两面性也使得生产力发展的一个关键性要素——分工机制得以迅速发展。很明显，分工协作活动本身就是一种对于生产信息的“提取”与“过滤”的双重机制：一方面，协作方当然要知道关于另一方的生产技能的信息；另一方面，彼方与生产协作无关的那些信息则是此方所不关心的。因此，参与公司面试的求职者就需要以“求职者”的脸面示人，当他在周末进入一家餐厅吃饭时，他立即就会换上一张“消费者”的脸。而从“颜面现象学”的角度看，市场经济超越自然经济的一个关键点，便是其所能提供的“颜面界面”的丰富程度与替换频度都有了长足的进步。与这种进步相伴而生的，则是人类工艺能力的日益精进以及商品的极大丰富。因此，脸或面具的丰富化其实本就是科技树增长的重要指标。

由此反观刘慈欣的“无脸人”或“思维透明人”假设，我们不难发现其与我们已知的自然规律与社会规律的冲突之处。

其一，该假设破坏了进化论的基本预设，即吝啬性原则。换言之，如果另外一个智慧生物的所有内部信息都会事无巨细地涌入“我”（假设“我”也是一个三体人）的



脑海的话，我是否有足够的心理学资源与生物学资源来应对这些信息呢？假若与我进行联系的其他三体人不是一个，而是一百个，那我又怎么可能处理那么多信息呢？除了对那么多“他者”的信息进行抽象，并由此构造出一张张关于他们的“脸”之外，我还有别的应对方法吗？答案显然是否定的。但如果三体星系的自然选择过程偏好那种能够将“脸”抽象出来的认知机制的话，那么按照吝啬性原则，该选择过程为何又要偏好那种演化出“无脸人”的选项呢？请别忘记了，按照《三体》小说的原始设定，三体人所处的生态环境本就是十分恶劣的，因此，一种需要更长演化时间的认知机制，本就不可能胜过一种只需要更短演化时间的认知机制。既然一种不需要“脸”做中介的全知性认知机制的出现显然会消耗更多的演化时间，这种机制自身又怎么可能在三体星的恶劣环境中应运而生呢？

其二，即使我们悬置上述第一点批评，刘慈欣的“无脸人”假设也会使得有效的社会分工成为不可能。前文已说过，有效的社会分工需要社会成员在不同的工作场域戴上不同的面具，以便将那些与当下工作无关的信息遮蔽掉。但在这些面具本身不存在的前提下，那些与当下工作无关的信息又如何被遮蔽掉呢？而假若这些无关信息不被及时遮蔽掉的话，工作效率的提高又如何得以可能呢？

面对上述批评，一种同情“思维透明人”假设的修补方案或许是这样的：一个三体人固然没有一张固定的脸，但是他依然会在信息负载太大的情况下启动遗忘机制，以便将注意力聚焦当下的任务。不过，这样的三体人依然具备一项地球人所不具备的超能力：一旦其认知资源累积到足够的程度，就能随时对他者的心理活动进行福柯式的全景式扫描。

不难看出，上述修补方案其实是将原始版本的三体人具有的“思维透视力”从一种现实变成了一种潜能。换言之，在这个方案中，三体人能够在不需要使用这种穿透力的前提下节省自己的认知资源，由此使得自身的认知架构被自然选择所偏好的概率上升。然而，在笔者看来，即使是这种弱化的“思维透视力”，也会与科技发展的一项关键性内在要求发生冲突——此要求即对于个体的知识产权的保护。

众所周知，工业革命在欧美的勃发，一项关键性的制度性因素就是对于发明家的知识产权的保护。此类保护措施使得发明家的科研或创业热情不至于因为盗版者的大量出现而被打压。但在任何一个发明家的内部科研思路都在原则上可以被他人看透的前提下，对于剽窃者的甄别就会在原则上成为不可能了，遑论对其进行惩罚。在这种情况下，谁又来保护少数天才的科研热情，由此推动社会的进步呢？

“思维透明人”假设的支持者或许会反驳说，“三体人”的社会道德已经摒弃了对于“私人利益”的任何执念，正如昆虫世界中的工蚁那样没有任何私念。然而，在笔者看来，这种辩驳本身其实已经误解了“智慧”一词在进化论语境中的本义。“智慧”的生物学含义恰恰是为了提高特定基因的传播效率，基因的传播又总是借助于个体的生存才是可能的。因此，从进化论的角度看，一个生物学个体如果能在生存斗争中发展出更为丰富的斗争策略来增大其基因得以被传播的概率，其智慧程度也就越高。因

此，在生物学家的词汇里，“马基雅维利式的智慧”其实并不算是一个贬义词。^①而对于一种内部完全没有马基雅维利式竞争的生物学种群而言，它又是通过什么机制来使得其基因组合方式能够不断得到优化呢？对于个体生物学利益的摒弃，本就是与智慧的增长不相容的——而这一点也就解释了蚂蚁的昆虫世界为何只能具有这么一点可怜的智慧。同时，这一点也能有力地解释：为何那些对个体的产权利益能够做出更有效保护的人类社会，其科技的发展水平也相对较高。

① Richard W. Byrne and Andrew Whiten, *Machiavellian Intelligence*, Oxford: Oxford University Press, 1988.

ChatGPT 时代“脸”的模糊化

上述基于进化论的考察结论，其实也是适用于对 ChatGPT 技术的讨论。ChatGPT 的全名是“Chat Generative Pre-trained Transformer”（含义是“预训练的聊天生成转换器”），其本质是一种人工智能聊天机器人程序。ChatGPT 目前主要以文字方式与用户交互信息，能够完成诸如自动文本生成、自动问答、自动摘要等在内的多种任务。目前美国一些高校的学生已经在该软件的帮助下完成学术论文，甚至还能获得不错的成绩，这就使得一些学校开始出台新规定，以禁止学生使用该软件完成作业。现在我们就将在“颜面现象学”这一新的维度中评估 ChatGPT 对科技发展的意义。

笔者总的诊断意见是：ChatGPT 是一种将人脸加以模糊化的新技术，其会因为这种模糊化而对人类科技研究的分工体系构成破坏。因此，尽管这种技术貌似既时髦又有趣，但实际上却将在总体上对人类科技发展的速率构成消极影响。

为了让读者能接受这种貌似非常“非主流”的意见，可以先回顾一下笔者在前文中所给出的对于人类科技协作过程的“颜面现象学”的分析结论：在特定的科学共同体中，专家们通过专业学术发表这一“颜面”来展现自己的学术成果。“颜面”在这里扮演了几重关键性角色：第一，屏蔽与当下的学术讨论无关的信息；第二，展现特定科学家的个性，并由此确定科研成果的责任人；第三，由于学术发表机会的相对稀缺性，能够在高端平台上“露脸”这件事本身就意味着一种学术资质。因此，传统的学术协作机制本身就带有一种精英主义色彩：特定领域内的优秀科学家与工程师作为新知识的贡献者，带动整个人类的科技进步。基于如下理由，笔者认为这种传统的精英主义的科学知识生产方式是基本合理的：第一，优秀的科学家与工程师的培养成本极高，因此，从教育经济学的角度看，重大的科研进步就只能是精英的事业；第二，成熟的社会分工方式能够使得科研成果以“涓滴”的方式惠及社会各个角落，因此，大多数受众只需要在承认科学权威的前提下接受他们的结论，而不必参与对于新知识疆土的开拓；第三，相对高门槛的科学知识创造准入资格，也能使得同行评议的工作量得到控制，最终使得新知识的确认得到足够充足的学术理由的背书；第四，与精英学术体制捆绑在一起的则是复杂的文责追认机制，这就使得新知识创制过程中出现的谬误能够得到及时的追溯与修正。总之，因为传统的科学知识创制方式是预设大家都“有脸”，为了保住“颜面”，科学共同体大概率就不会做“丢脸”之事。



这种传统的知识生产模式当然并不会在原则上拒绝现代信息技术。实际上，在 ChatGPT 问世之前，学术界早就开始使用搜索引擎来提高知识生产的效率。需要指出的是，基于如下理由，搜索引擎技术并没有构成对于传统知识生产方式的真正威胁：

第一，搜索引擎所获得的信息依然是科学共同体的可追溯的“颜面”——比如，如果你搜索到你需要的某类信息出现在某个期刊的网络版上，该期刊网站自身的权威性就能够为其内容的权威性背书。由此被搜索到的信息也是公开的，这就使得一个建立在搜索引擎技术上的科学共同体依然还是一个“人脸矩阵”。

第二，搜索引擎技术的使用本身具有一定门槛，需要使用者灵活替换关键词组合方式才会搜索到自己所需要的信息。当被搜索的信息涉及某些未被翻译为主流语言的内容时，搜索者也应当具备一定的外语素质，或至少是活用机器翻译软件的能力。因此，一种基于搜索引擎技术的知识探索方式依然能够磨砺探索者的智性能力，并由此促成其具有个性的“科学颜面”的形成。

第三，搜索引擎最后所定位的知识表述未必是“用户友好”的。以笔者所熟悉的哲学研究为例：业内人士一般都知道，很多哲学家思想的介绍可以在“斯坦福百科全书”上找到，而且利用搜索引擎找到特定哲学家的介绍页面也几乎是举手之劳，但即使对于以英语为母语的学生来说，一般也需要本科水平的哲学训练基础才能真正读懂这些词条。因此，搜索引擎技术的出现，几乎对学术生产的精英主义传统无甚影响。

ChatGPT 技术与搜索引擎不同。这是一种彻底颠覆精英主义传统的知识获取方式。根据现有的报道，使得 ChatGPT 得以运作的语言生成模型是通过大量数据标注工的人工劳动而获得海量的训练数据的。这些数据标注工来自肯尼亚等欠发达国家，他们拿着比发达国家劳工低得多的薪水，在特定数据处理政策的指导下对海量数据进行人工标注，以便由此间接地调节系统的输入-输出映射模式。^①不过，基于如下几重考量，上述数据训练模式在原则上就无法使得高质量的科学知识在用户界面上出现。

第一，由于新科学知识（比如最新医疗期刊上出现的新疗法）的出现往往很难被既有的训练数据所覆盖到，因此，现有的 ChatGPT 无法对最新发生的科技新闻进行回应（笔者曾就俄乌战争中的一些新技术趋势询问 ChatGPT，都被对方以“不在训练范围内”为由而“打发”）。

第二，由于数据标注工本身不是各行各业的专家，很多专业的学术信息在互联网上也并非是可被轻易获取的（特别是那些通过小语种表述的信息），因此，ChatGPT 无法应对用户询问相对冷门的知识的情况（笔者曾就几位冷门的日本哲学家的情况询问 ChatGPT，对方都无法应对）。

第三，由于基于大数据训练的任何模型反映的都是一般人对某个问题的普遍看法，因此，ChatGPT 本身就是“无脸”的：其回答没有个性，不能反映其在特殊的社会分工中所扮演的特定角色，更谈不上成为特定的责任者，以便为错误的知识表达负责。更糟糕的是，与传统的搜索引擎技术不同，ChatGPT 无法为其知识来源进行精确的标注，这就使得我们甚至不能以 ChatGPT 为媒介去寻找可能的错谬信息传播源。很显然，

① Julia Zorthian, *OpenAI Paid Kenyan Workers Less than \$2 an Hour to Make ChatGPT Less Toxic*, <https://time.com/6247678/openai-chatgpt-kenya-workers/>.

ChatGPT 为可能出现的虚假信息的广泛流传创造了技术土壤。

第四，ChatGPT 不但消除了自己的脸，而且会慢慢模糊化用户自己的脸。前文刚提到，搜索引擎技术的出现并不会影响科研者学术个性的形成，因为对被搜索的关键词的设置方式就是对于搜索者自己的科研洞见力的考察。但 ChatGPT 却提供了一种新的知识提供方式，即用户可以通过简单的自然语言应答而获得相应的答案。在系统实际提供的答案质量其实并不算高的前提下（理由请参看前述三点），这样的做法其实就等于剥夺了优秀的知识索取者获取更优质答案的机会——这就好比某个国王强迫自己所有的臣民只能通过外卖点餐。此做法虽然免去了不少人亲自下厨的麻烦，却剥夺了某些烹调爱好者自己琢磨厨艺的机会。很显然，这样的做法迟早会降低整个王国的厨艺水平——因为这个国家将不再会有值得被记住的大厨的脸（同样的分析显然也适用于对于该国科研水平的分析）。

第五，正是因为 ChatGPT 的“无脸性”，此技术平台甚至也不能承载“脸面”在传统科研成果发表平台上所具有的另外一项功能——遮蔽对科研无意义的各种废话。与之相反，现有的 ChatGPT 简直就是“正确的废话”的制造机制。譬如，当用户问该软件是否同意“在国外销售的茅台酒比国内便宜”的时候，该软件只会说出这样的一些“车轱辘”话：“这一点得通过具体国家的税费与物价水平来确定，而不能一概而论。”虽然在日常生活中我们有时也会说这些“正确的废话”，但机器通过惊人的复制力，却能迅速制造出一个个体很难与之抗衡的语言垃圾场，由此反而使得那种单刀直入的科学文风成为“不合时宜的少数派”。此类软件在教学场景中的广泛应用，也会使得学生们失去对于语言海洋中核心信息的把握力，由此成为对于网络信息的无灵魂——自然也是无脸的——“人肉复读机”。

从上面的分析来看，不但科幻小说中的“无脸人”假设无法支撑外星的发达科技，甚至就地球当下的科技发展水平而言，已经快要造就“无脸人”之横行的 ChatGPT 技术也迟早会对人类的科技进步构成戕害。

反思与总结

最后，我还想就《三体》中的“思维透明人”假设与以 ChatGPT 技术为代表的聊天机器人之间的关系，发表一点管见。尽管从表面上看来，后一技术所代表的技术水平是远远不及三体人文明的，但两者的哲学本质却依然有不少类似之处。两者都试图在一个超级的平台上将一个文明的所有信息都铺设开来，由此实现一种信息场的“大一统”。只不过在三体文明中，这个信息铺设的平台是该文明的“元首”（他被假定为一个几乎不会犯错的外星政治领导人），而在我们这里，该信息的铺设平台则是 ChatGPT 背后的超级语料训练模型。两者之间的隐蔽平行关系还体现在：外星元首被设定为对于该文明的所有的政治权力的掌握者，而 ChatGPT 则被设定为对于地球上的充分的语料知识的掌握者。换言之，在这两种思维模型背后，都有一种对于全知全能



者的诉求。

然后，即使是中世纪的那些痴迷于讨论针尖上天使数量的经院哲学家，都能一眼看出上述诉求有多么荒谬。在基督教哲学的背景中，真正全知全能的只能是上帝，而无论是外星人还是人类编制的任何软件，都只能是“有限存在者”。作为无神论者的我们当然不会接受基督教哲学家关于上帝的设定，但是他们对于人的有限性的认定却完全可以在无神论的框架中得到进化论的背书：自然选择没有理由偏好那些全知全能的认知架构。同样的道理，仅仅是作为有限的人类个体的谈话方式之“统计学平均机制”的 ChatGPT，又如何可能做到全知全能呢？

不过，《三体》的透明人假设也好，时下对于 ChatGPT 的热捧也罢，颇能反映社会上相当一部分人对于科技发展的误解：如果我们能够让每个人的智慧都毫无挂碍地联结到一起，由此形成的超级智慧就能用来攻克一切科技难题。在这种思维模式的驱使下，集中大量资源、攻克科技难题的思路就会占据优势。但人类科技发展的真实情况恰恰是：大量的科技发明都是少数精英在缺乏社会资源支持下独立研发的产物（如巴斯德对于细菌的发现、基尔比对于集成电路的发明），因此，对于科研个体创新热情的保护反倒是促进一国科技发展的题中应有之义。同时，现代科技产品（特别是芯片）的制造所需要的超级复杂的社会分工体系，也需要一套同样复杂的“颜面现象学”与之配套，而“颜面现象学”的这种复杂性，又需要社会管理的权威在每一层次的“颜面”上被分解成不同层次的亚权威甚至亚-亚权威。因此，与刘慈欣的预想相反，高度发达的科技协作系统恰恰不需要信息的高度集中，而需要信息的高度分散，甚至需要强大的产权保护机制以使得这种分散性被建制化（譬如，若某类生产工艺的机密的确属于某公司的话，一个合理运作的国家机器就应当竭力保护这种秘密性）。

这种分散式的科研资源配置方案的合理性甚至可以得到概率论上的支持。不同的科研方向就类似一个将领在不同方向上派出的侦察兵——他派出的侦察兵的路向越多，探听到敌情的概率也就越大。反之，对于科研资源的过分集中，反而会使得被派出的侦察兵的可能路向被减少，由此亦会减少实现科研突破的机会。此外，对于上述分散性的资源配置方案的支持，亦是演化的吝啬性原则的题中应有之义。既然演化进程大概率不会演化出一个全知全能的超级信息处理者，那么，将全社会的信息处理权上交给任何一个个体信息处理者的做法都会是不合理的。从这个角度看，假若三体人的社会架构与信息交流方式果真是如《三体》所描述的那样，那么我们这些地球人反而不需要担心他们对我们的威胁了——除非我们已经预先通过 ChatGPT 将自己训练为“无脸人”了。

概而言之，人类的科技发展是需要英雄的，而所有的英雄是需要“脸面”的，这一点甚至对外星球的科技发展也是适用的。因此，以精英科学家之脸面为基础的人类科技传统发展模式依然是我们需要小心呵护的文明瑰宝——如果我们失去了这个模式，我们将真正失去人类的未来。

编辑 张 蕾 特约编辑 阮 凯