

雕 刻 中 国 人 的 阅 读 地 图

01  
2023

# 中国图书评论

China Book Review

- 国家社科基金资助期刊
- 中国社会科学引文索引 (CSSCI) 扩展版来源期刊
- 中国综合性人文社会科学类核心期刊
- 国家新闻出版署 (中国期刊方阵) 双效期刊

中国图书评论

2023/01 总383期

## 阅读地图

2022年11月“中国好书”

## 概 念

林云柯 “概念”废墟中的拾荒 (主持人语)

朱恬晔 “算法”提喻的话语陷阱

章含舟 移情≠同情, 移情→同情

## 冷 眼

郭晓斌 也谈鲁迅与《Noa Noa》

## 社会关注

王 鑫 面向数据库的开放游戏



中國書評

創始於1986年



中国书协

## · 本期导读 ·

## 概念的边界

□ 周志强

顾问 鄔书林 高明光

编委 (按姓氏笔画排序)

鄔书林 刘小枫 刘东杰 杨平  
宋伟 张柠 周志强 孟宪实  
贺耀敏 高明光 郭义强 黄轩庄  
程巍

社长  
总编辑 杨平

副总编辑 刘苏哲

副社长 夏泽平

执行主编 周志强

编辑部主任 李文杰

编辑 管飞

陈琼娇 高宇 李昕 郎静

龚燕 刘英 王天孜 戴利明

封面题画 游江

封面设计 刘冰宇

主管单位 中国图书评论学会

主办单位 中国图书评论杂志社

出版单位 《中国图书评论》编辑部

杂志社社址 北京市西城区北三环中路6号京版大厦B座

电话传真 (010)64173406

邮政编码 100120

电子信箱 chinabookreview@163.com

杂志官微 中国图书评论杂志官方微博

制版 北京心水衡昌广告设计有限公司

印刷 沈阳市池陆广告印刷有限公司

国外代号 M1029 ISSN1002-235X

国内统一刊号 CN21-1035/G2

国内邮发代号 8-46

发行范围 国内外发行

国内总发行 辽宁省报刊发行局

订购处 全国各地邮局

国外发行 中国国际图书贸易总公司(北京399信箱)

出版日期 每月10日出版

定价 15.00元

广告经营许可证号 2017000029

理论界似乎一直存在“概念的边界”问题：一是西方的理论概念如何减少误读和误用；二是学者如何合理地“自创”概念——这两个问题似乎很令人焦虑（当然，也有人极力反对概念的滋生，认为通俗易懂才是理论的正统——这种市民主义的思想当然不能完全适合学术研究的专业化要求）。在这里，概念确实存在两个“边界”：已经具有内在特定含义的概念，要考量其学术共同体与理论史的语境；同时，概念的使用还要在继承中合理修改、错用与新造，与自身理论的现实语境相配合。历史语境与现实语境，分别构成概念的层级能指。

对古典概念和国外概念的多种借用形成了中国现代学术史的特点，也造成了概念使用的边界混乱。林云柯博士在评刊会上建议设置“概念”栏目，对多年来“重大概念”进行观念史的梳理和内涵的辨析。这是一件极其重要的事情，也是一个应该马上做起来和做好的事情。

我非常认同黑格尔或德勒兹的态度：概念是撬动现实的新可能性的东西。倘如此，对于概念的梳理和辨析，何尝不是对于新现实的发现和开掘呢？

2023年01月



中国好书二维码



中国图书评论二维码

本刊编辑部	2022年度好书评	3
阅读地图		
中国图书评论学会	2022年11月“中国好书”	4
概念		
林云柯	“概念”废墟中的拾荒（主持人语）	8
朱恬骅	“算法”提喻的话语陷阱	10
章含舟	移情≠同情，移情→同情 ——关怀伦理学中的移情与同情之辨	18

# “算法”提喻的话语陷阱

□ 朱恬骅

**【导读】**当前对于“算法”的文化批评聚焦于其社会效应中显现的非中立性。然而，就算法作为处理某个数学问题有限而确定的步骤而言，它具有数学上的中立性。科技巨头所代表的一众行为主体，试图借“算法”的名义占有中立性，以掩盖先于算法的人为决策（如对于问题本身的定义、数据来源和维度选择、标签类目和结果的使用方式等），逃避其社会责任。对算法中立性“神话”的批判，在无形中落入了这种假托“算法”之名的话语圈套。作为替代，可以使用“模型”的概念，凸显人为决策引入的非中立性，从而将对“算法”的批判重定向为对技术中起主导作用的“规范共识”的批判。

**【关键词】** 算法 模型 技术的社会建构

10

在谈论当今高新技术的社会影响时，“算法”无疑已成为一个关键词汇，以“算法”为定语的术语，如“算法社会”“算法治理”“算法规制”“算法意识”等纷至沓来，大有令人应接不暇之势。有国内学者认为，“算法社会”是雅克·埃吕尔（Jacques Ellul）“技术化社会”在当代的具体体现，“算法”已经成为数字时代的问题关键。<sup>[1]</sup>在文化批评的语境中，“算法”被视为技术决定论乃至“技治主义”的核心要素，是一系列社会现象特别是负面现象产生的直接原因，它放大社会偏见与歧视，激化社会矛盾，加剧社会不公。在这种观点看来，“算法”体现了某种价

值判断，如何种信息是值得推荐的而何种信息则应被阻止，而将“算法”视为一种中立性的事物完全是人为建构的神话。

但在另一方面，“算法”的本来含义同数字的计算息息相关，即便在计算机技术的语境中，它也被定义为一种单纯的数学过程。例如，在久负盛名的《计算机程序设计的艺术》（*The Art of Computer Programming*）中，高德纳（Donald E. Knuth）将“算法”一词界定为“一组有穷的规则，这些规则给出求解特定类型问题的运算序列”，明确它具有有限性、确定性、能行性等特征。<sup>[2]1:4-5</sup>而在同样被奉为经典教材的《算法导论》

(*Introduction to Algorithms*) 一书中, 托马斯·科尔曼等也将算法定义为“定义良好的计算过程”“用来将输入数据转换成输出结果”。<sup>[3]</sup> 这种数学意义上的“算法”似乎有着不言而喻的中立性, 从而与针对“算法”的文化批评形成了对垒之势。

## 一、“算法”(algorithm) 一词的源流

从词源上看, 英语的“algorithm”来自 algorism, 后者则来自 12 世纪欧洲翻译者对波斯数学家花刺子米 (Muḥ ammad ibn Mūsā al-Khwārizmī) 姓氏的拉丁化“algorizm”。在拉丁文译文残篇“花刺子米如是说”(dixit algorizmi...) 中, 详细记录了使用印度数字(今阿拉伯数字)进行四则运算的方法<sup>[4]</sup>, 使得这一原本意为“花刺子模人”的姓氏在欧洲获得了“计算方法”的意义。久之, 人们将“algorism”同古希腊语表示“数”的 ἀριθμός (arithmos) 一词相比附, “algorithm”的拼写方法由此产生, 尽管二者之间并无词源上的关联。这种比附体现了人们将“algorism”与“数”相关联的意识。

不过, 对于汉语读者而言, “算法”的历史实则更为源远流长。姑且不论中国古代数学中关于计算方法的记述存在以“(算)术”命名的

传统(如《九章算术》), “算法”二字就已经出现在《汉书·律历志》中: “其算法用竹, 径一分, 长六寸, 二百七十一枚而成六觚, 为一握。”此处的“算法”虽然是指算筹的制作之法, 但到唐代, 讲述计算方法的“算法”已作为书题出现。值得一提的是, 南宋末年流行的民间百科《事林广记》就设有“算法”一节, 所载“算法源流”条目语云: “夫算法者, 伏羲始画八卦, 周公叙述《九章》。至于玄元、益古、如积、细草, 其旨渊奥, 难可寻绎, 初学者无所措手。其加减因折乘除之法, 所以上揆星躔, 下营地理, 巨无不揽, 细无不规。”文中所列举之“玄元、益古、如积、细草”即为筹算天元术(解高次方程)的方法名称, 足见此时的“算法”已经固定为“计算方法”之意。

科学史学者指出, 宋元时期的中国数学发生了从“实物语言”主导到“文本语言”主导的转变, 也就是从使用算筹等实物计算工具进行操作, 转向由对算筹操作的文本记录而形成符号操作。<sup>[5]</sup> “算法”词意的变化恰好暗合这种转变, 从最初制作计算工具的方法到操作这些工具的方法, 再到将操作方法文本化而开始显现为书面符号的变换逐步走向抽象。

这就使得古代“算法”同今天

人们所谈论的计算机算法，并非只有语词上的关联。现代意义上的算法肇始于图灵（Alan M. Turing）的《论可计算的数》（On Computable Numbers）<sup>[6]</sup>，文中提出了著名的“图灵机”这一数学模型。不同于实际建造的计算机，这种设想中的机器可以拥有两端无限长的计算纸条，但每次只能对纸条上的一个符号进行操作，无论是读取（执行）还是擦除、改写。图灵机在数学上的应用为求解希尔伯特“判定性问题”（Entscheidungsproblem，即是否所有数学问题都是可判定的）提供了一种重要的等价形式，但主导其行文的精神则在于对计算本身进行了规则化的描述，以使之成为一系列关于符号的操作。它明确了计算机器所可能操作的“数”的性质，而这进一步来说又是通过对“计算过程”本身进行定义得到的，因而成为定义“算法”时参照的对象。

1954年，小马尔可夫（Andrei A. Markov, Jr.）将算法定义为数学中的精确操作，具有三个主要性质：其步骤不允许任意性，而且应当是普遍可理解的，因而具有确定性；算法接受一类数作为输入，而不只是在某个特定值上运作，此即群体性；算法应当倾向于获得确定的结果，而且对于正确的输入而言，输出也应当是符合预期的，此即结果性。<sup>[7]</sup>

不满于这种文字性的描述，小马尔可夫通过他所定义的“正规算法”（normal algorithms），将算法形式化为一种对于字母表（符号）和字串的操作，并指出其与递归函数、 $\lambda$ 演算和图灵机具有等价性。这一定义也被称为“马尔可夫算法”（Markov algorithm）。

沿着形式化的路径，数学家们对算法做出了更为精细的定义。2012年，《计算机协会通讯》（*Communications of the ACM*）在“编者按”中将其总结为“抽象状态机”（abstract state machine）和“递归器”（recursor）两条路径<sup>[8]</sup>，二者大致可分别追溯到图灵机和丘奇（Alonzo Church）的递归函数。另一方面，寻求算法哲学本体论意义的尝试并未中断。罗宾·希尔（Robin K. Hill）从不同的形式化定义出发，认为一种更为简洁而易于大众理解的文字定义仍然是有效的，并将算法归结为“一个有限的、抽象的、有效的、复合的控制结构，在特定限制下完成特定目的”<sup>[9]</sup>。

无论是在中西文语境中的原本含义，还是在丘奇、图灵等开创的现代意义上，“算法”一词都指向对一组有限符号进行有明确定义的操作。但是当代对于“算法”一词的运用，显然已经远远超出了上述数学意义的有限性和明确性。归结起

来，可以说“算法”成为了一系列技术手段和非技术决策的总体性象征，一种以部分代全体的提喻修辞法。

## 二、成为技术提喻的“算法”

在种种以“算法”为定语术语中，当属“算法社会”的覆盖面最为广泛。美国科罗拉多大学哲学教授约翰·加纳（John Carnes）可能是最早采用这一术语的学者之一。在1985年举办的国际哲学与技术学会双年会上，他在一篇探讨职业与教育的文章标题中亮出了“算法社会”（algorithmic society）之名。<sup>[10]</sup>2021年，里克·彼得斯（Rik Peeters）和马克·舒伦伯格（Marc Schuilenburg）在所编文集《算法社会：技术、权力与知识》中，将“算法社会”归纳为“一套实践和话语，牵涉到政府和私有部门之间的混合联系，它由一套相对较新的数据驱动技术支撑，通过自己的知识模式和形成新主体的特殊方式，为社会的治理增加了新的层次”<sup>[11]195</sup>。

值得注意的是，在上述界定中并未出现计算方法意义上的“算法”本身，而是借此指代了“一套相对较新的数据驱动技术”。尽管我们也可以将计算的方法视为一种“技术”，如维特根斯坦曾经将读图称为“技术”（Technik）<sup>[12]249</sup>；但是掌握

一种抽象的计算方法、一种操作符号的思想，毕竟无法比拟于建造一种自动运算的机器，将它们配置并部署到现实生活的某个情境中，切实地发挥作用。后者是今天的人们谈论“技术”时的主要意涵所在。而从针对“算法”中立性的批判来看，对于技术运行结果的评判往往构成了论者展开辨析的出发点。为人诟病的数据的采集、特征的选取、目标的设定等虽然也以各种方式在算法中得到体现，但它们本身并非算法运行的结果，而是在算法得到设计之前就已经决定了的事物。

这就意味着对于“算法”中立性的批判，实际并不是针对作为计算方法的算法本身。它的名称实际指代了诸多先于算法设计的决策和算法本体之外的数据。“算法”成为各类计算机技术特别是人工智能技术的提喻（synecdoche）。

技术的提喻此前多以实物的形态出现。西蒙栋（Gilbert Simondon）曾举例分析了仪表盘、尾气管等汽车部件对于汽车性能的暗示作用。通过这些部件夸张的外形设计，制造汽车的技术得到了象征性的显现，即所谓的“技术显像”（technophanie）。<sup>[13]39</sup>在他看来，技术显像的出现标志着一种将技术重新纳入“文化堡垒”中的努力，也就是为技术因素（如汽车的加速性能）赋予文化上的含

义（如与英雄气概相关联）。与此类似，“算法”这一名称在相当程度上为人工智能技术中多变、复杂而缺乏可解释性的技术构成赋予了一个概括性的名称。

算法所描述的计算过程之所以有意义、被采纳，有赖于一系列在此之前和在此之外施加的对输入—输出对应关系的规约，而这正是科技巨头企图利用算法作为数学过程的中立性所掩盖的。它们的逻辑是：由于算法是中立的，它本身无法做出价值的评判，而仅仅以计算的结果排列或推荐用户消息或新闻内容，社交网络、手机应用等技术产品也就是中立的。但是从设计的角度来看，如果广告的投放以转化率为唯一的衡量标准，在此规定下设计并运作的算法也就不可能将广告内容的合规与否纳入检测的范围。换言之，“算法”之所以被征用为人工智能技术产品的提喻，其目的不在于澄清或代表其中的作用机制，而是在于遮蔽算法以外参与人工智能技术建构的社会观念因素。

### 三、“模型”取代“算法”

作为一种技术构建，形形色色的人工智能技术应用从一开始就受到了社会因素的规定。芬伯格（Andrew Feenberg）对于技术的社会构建

论观点在此仍然适用：“技术包含着美学、伦理和文化领域中规范共识的成果，而不仅仅是纯粹的效率至上或用户至上渴望获取的狂热。”<sup>[14]</sup><sup>15</sup>作为诸种概念“物质化”（materialize）的结果，技术产品体现了这些“规范共识”的作用，并为之准备了自动的“次级施行者”（secondary agent）。

但是数学意义上的算法并不是这样的技术产品，二者间存在着巨大的鸿沟：一种“数学上可行”的算法在现实中完全可能是不可行的。算法所提供的是一种维特根斯坦所说的“象征的机器”，沿着逻辑的必然性“运作”，“从一开始就在自身中包含着它的作用方式……它将造就的各种运转，似乎已经完全决定好了”。<sup>[12]</sup><sup>89</sup>这也就意味着，算法和图纸一样，无法突破纸面而“生长”出现实的机器，后者由于物理的限制而可能产生出预料之外的结果。更不用说，算法对输入和输出数据所做的数学上的预设无法从现实中直接取得，而是需要经过技术上的转换。这些现实技术的成分是算法发挥作用的前提。

另一方面，从文化批评引入“算法”概念作为批判靶标的动机来看，其所关心的也不是作为数学过程的算法本身，而恰恰是将其近似、转变，由通用性的计算方法具体化

为满足外部特定要求的上述技术，也就是“算法”的提喻。问题在于，批评者对此往往并不自知。如同科技巨头企图将数学意义上的算法同他们的技术产品短接以赋予后者“中立性”，批评者也将他们从技术中发现的非中立性倒转为“算法”的性质。可见，“算法”作为技术的提喻并不恰当。人们需要一种更具概括性的话语，显性地揭示技术独立于数学过程并受制于社会关系的方面，才能让文化批评真正进入“规范共识”的层面，也就是现实中业已存在着的偏见、不平等或技治主义倾向，从而合法地发挥批判的作用。

“模型”提供了一种可行的替代方案。尽管“模型”不像“算法”那样存在较为公认或形式化的定义，我们仍可以从人工智能技术的语境中看到，模型突出的特点在于定义了算法所需求解的问题本身。如果从最初华伦·麦克库洛（Warren McCulloch）与瓦尔特·皮茨（Walter Pitts）对于神经元的数学建模算起<sup>[15]</sup>，“模型”在人工智能的领域中就不仅包含了对于某个问题的解决方法，而且首要地呈现了对模糊问题进行严格定义与描述的要求。特别是在当代以机器学习为主要路径、人工神经网络为范式的人工智能技术中，严格意义上的算法主要在数值优化

的基础环节发挥作用，而更为整体性的“模型”才与具体的用途相关联。

以艺术领域的人工智能技术应用为例，2016年，列昂·盖蒂斯（Leon A. Gatys）等人在提出有关“艺术风格转移”时，即通过一组矩阵来定义图像中的“风格信息”，并进一步将已有的模型同他们提出的新结构适配、嵌合。<sup>[16]</sup>他们用一组良好定义的（well-defined）矩阵替代了艺术上存在争议的“风格”概念，“风格”的接近则通过评估矩阵之间的差异得到。这样，模型确定了算法所要优化的目标函数有何意义，而所求得参数又应满足何种条件。最终，整体计算过程的性能或实用性则在相当程度上由模型所要求的训练数据（图像）决定。“风格”的模型是否有效并不是通过深入矩阵元素和某种关于“风格”的艺术理论之间的对应性上建立的，而只是研究者的一种人为规定，体现了他们所持有的、局部性的“共识”。

因此，如果说算法至多只是将与应用相关的信息隐含在对计算过程的描述之中，模型则将计算过程得以决定之前的先决条件和可能的输入范围，清晰呈现在人们的面前。即便是那些强调通用性的模型，如文本处理领域能在文本分类、自动摘要、生成等方面发挥作用的 GPT-3，

也有明确的作用范围（现代英语文本，特别集中在说明性文本）。而这又是因为模型对于数据的来源、性质和意义进行充分的考虑。由于英语维基百科的内容在 GPT-3 的训练过程中占有重要地位，其在处理说明性文本时显然就更可能存在优势。而当人们用 GPT-3 相同的算法甚至结构，用另一种语言的文本、另一组数据进行“训练”之后，所获得的新模型体现的显然已是另一种语言、另一些文体所具有的特征了，因而至多只是与 GPT-3 同构的另一个模型。遑论不同语言的文本涉及不同的具体处理方式，如中文文本是按照字符还是按照词为单位进入算法，都为模型引入了新的自由度。

当然，模型和算法一样，并不是具体的技术产物本身，它同样是对技术的一种抽象。但是，这种抽象仍比算法更为接近最终的实现；而且它给出了人们对于数据可能性范围和达成目标的认识和理解，从而直接地显现出技术之所受制与反映的“规范共识”。仍以前文提及的“风格转移”为例，当盖蒂斯等人以凡·高的《星空》为例证，说明其模型的有效性时，无疑就已包含了他们以凡·高的油画为一种典型绘画风格的艺术认识。这些作者显然也没有预想到，如果以水墨画为目标风

格，其输出结果并不能如应对油画时那般令人满意；这种“偏见”并不是由于算法本身所造成的，而是由于模型的设定和训练过程导致。例如，北京大学的段凌宇团队在此基础上重新设计了目标函数，并选取了徐悲鸿的画作图像作为水墨画的代表进行训练，取得了令人满意的结果。<sup>[17]</sup> 而由此取得的新模型，在应对水墨风格时就获得了一定的优势，尽管在底层上可以认为二者中起作用的“算法”是高度相似乃至相同的。

模型的改变同共识的改变相关。由于包含了更多具有现实意义的信息，模型成为数学意义上的算法和技术产品之间的桥梁，明确显现了技术的社会建构性质。以“模型”取代“算法”作为人工智能技术的提喻，有助于人们跳出“算法”提喻的话语陷阱，从对技术的间接批判更快地转入对技术的社会建构中“规范共识”的直接批判：如何将真正具有广泛代表性的社会共识通过技术得到表达，而不是那些只服务于少数人利益的“共识”。如是，从人工智能技术的种种社会后果出发的文化批评才能建立起合法性，免于陷入卢德主义的困境。

[本文系国家社科基金艺术学青

年项目“计算机艺术历史生成问题的人类学美学研究”(21CA169)的阶段性成果。]

### 注释

[1] 贾开. 算法社会的技术内涵、演化过程与治理创新[J]. 探索, 2022(2): 164-178.

[2] [美] Knuth D E. 计算机程序设计艺术[M]. 李伯民, 范明, 蒋爱军译. 北京: 人民邮电出版社, 2016: 1.

[3] [美] Cormen T H, Leiserson C E, Rivest R L, 等. 算法导论(原书第2版)[M]. 潘金贵等译. 北京: 机械工业出版社, 2006.

[4] Crossley J N, Henry A S. Thus Spake Al-Khwārizmī: A Translation of the Text of Cambridge University Library Ms. Ii. vi. 5 [J]. *Historia Mathematica*, 1990, 17(2): 103-131.

[5] 朱一文. 数学的语言: 算筹和文本——以天元术为中心[J]. 九州学林, 2010(7): 82-105.

[6] Turing A M. On Computable Numbers, with an Application to the Entscheidungsproblem [J]. *Journal of Mathematics*, 1936, 58: 345-363.

[7] Markov A A. The Theory of Algorithms [J]. *Trudy Mat. Inst. Steklov*, 1954, 42: 3-375.

[8] Vardi M Y. What Is an Algorithm? [J]. *Communications of the ACM*, 2012, 55(3): 5.

[9] Hill R K. What an Algorithm Is

[J]. *Philosophy and Technology*, 2016, 29(1): 35-59.

[10] Announcements: Federal Agency [J]. *4S Review*, 1985, 3(1): 42-48.

[11] *The Algorithmic Society: Technology, Power, and Knowledge* [M]. M. Schuilenburg, R. Peeters. London: Routledge, 2021.

[12] [英] 维特根斯坦. 哲学研究 [M]. 陈嘉映译. 上海: 上海人民出版社, 2005.

[13] Simondon G. *Sur La Technique* [M]. Paris: Presses Universitaires de France, 2014.

[14] [加] 安德鲁·芬伯格. 可选择的现代性 [M]. 陆俊译. 北京: 中国社会科学出版社, 2003.

[15] McCulloch W S, Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity [J]. *The Bulletin of Mathematical Biophysics*, 1943, 5(4): 115-133.

[16] Gatys L A, Ecker A S, Bethge M. Image Style Transfer Using Convolutional Neural Networks [A]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* [C]. 2016: 2414-2423.

[17] He B, Gao F, Ma D, et al. ChipGAN: A Generative Adversarial Network for Chinese Ink Wash Painting Style Transfer [A]. *2018 ACM Multimedia Conference on Multimedia Conference-MM'18* [C]. 2018: 1172-1180.

作者单位: 上海社会科学院文学研究所  
(责任编辑 郎静)