算法作为文化：算法系统的民族志方法

# 零、术语焦虑

大数据（Big Data）是模糊的，而算法（algorithms）在计算机科学中有精确而直接的定义。在研究者眼里，它似乎只是理性化、量化方法和程序逻辑的产物——通过这种组合式的理解，“算法”的概念就解体了，在某种程度上成为了一个重要、强有效而含义模糊的“现代迷思”（modern myth）。

当我们谈“算法”的时候，我们不谈冒泡排序、并查集或快速幂，我们谈百度的搜索、抖音的视频流、饿了么和美团的交通情况预测，以及淘宝、京东和拼多多的广告推荐……这些**系统**（systems）都比计算机科学所称的**算法**更复杂、更难以界定。“算法”在人文社会科学与计算机科学的不同范畴之间，有“令人不安”（unsettling）的间隙。术语是学科内部的话语，对它的理解和使用区分了“圈内”和“圈外”。但这是否意味着人文社会科学研究者完全不了解所谓的算法呢？Nick Seaver在将算法作为文化的考察中给出了答案，通过经验研究与对日常体验的强调，来重新审视算法在日常生活中的意义。

# 一、文化中的算法

解决算法定义危机的一种方法，就是照搬计算机科学家的“精确定义”，但算法是一个流动的概念，不同历史时期、社会环境中的不同人，都可能对算法的内涵与外延做出不同的理解。对精确定义不加批判的依赖，会遮蔽知识前沿正在发生的微妙变化。

Seaver认为，“算法”实际上是一个经验问题：人们是否在根据算法的定义来实践（do）算法？Seaver在一家提供推荐算法的企业中展开了研究，他采访了该企业一半的雇员。所有被访者都或多或少地参与到算法的生产中。为了寻找算法实际的内涵，Seaver要求被访者辨明他们所摆弄的“算法”究竟是什么，而被访者往往难以回答这一问题。现实中的算法难以精确定位，“算法”实际上成为了一个集体的产物，每个人都对它感到陌生。



现实中的算法有着模糊而非技术性的含义，指向一个更广的体系。“技术人员”不是唯一被卷入算法的人，各色人等都在使用不同的技术、操着不同的理解来塑造“算法”这一松散而模糊的体系。算法的“正确”定义让它从人文社会科学研究者的关切中脱离出来，让算法过程从文化过程中独立出来。澄清“算法”的概念，就产生了算法和算法似的他者（algorithmic others）——算法不是自动化、不是代码文本，它不是它的实现形式（materialization），算法处于一种非社会的抽象空间里，因而应该也能够脱离算法实践与社会生活的纠缠。

不论是强调技术没有嵌入政治因素的算法科学家，还是将推荐算法当作社会生活的外源性威胁来防备的批评者，都仰仗一个共同的前提，即文化与技术的分离。对文化内的算法的强调将算法看作河流中的石砾，是共存而相互影响的两个独立体系。

# 二、算法作为文化

和算法类似，“文化”一词也曾在学界与社会遭遇不同命运，结果许多人类学家从对作为一致连贯的符号规则的探求，转向了对文化生活实践的探索。文化与其说是一套行为特征，不如说它更可能是人们行为的结果。在摩尔（Annemarie Mol）看来，“实践的本体论”（practical ontology）意味着行动者不是依照特定对象来行动，而是促使它诞生、将其带到现实中去的——这一过程就是制定（enactment）。因而，行动的对象在很大程度上成为了“多重的”（multiples），不是一致的、持续的“一”，也不是分散的、相对的、独立的“多”。从这一角度思考，算法就不再是一个稳定的互动对象，而是人们不同实践造就的多样结果。没有一个算法的“内在真理”（inner truth）决定这些互动行为。

算法作为文化，不是文化河流中的一个个技术石砾，而而是掺杂进河流的河水。算法是由行为塑造的，它并不在意技术和非技术关切之间的区别，相反，它能够实现二者的混合。在这一意义上，算法不是作为单一的技术对象进入到文化互动中，而是由人们的关乎文化的实践制定的。算法的文化属性并不来源于它的对象（音乐、电影、视频……），而是来源于它背后的人类实践。

# 三、方法的制订

算法的审计方法将算法当成一个黑箱，通过改变输入对比不同条件下算法在结果上的差异，进而讨论算法可能产生的影响。这种方法是一系列实践的一部分，这类实践将算法理解并塑造为保密的过程。当许多批判性的思路将算法黑箱的保密性当作算法的主要特征时，Seaver提供了更新颖的接近算法的民族志策略。

## 1. 拾荒（Scavenge）

即在许多不同的地点、场景尝试与研究对象互动，不拘一格地收集资料。民族志研究者一直致力于收集多种来源的信息。如果我们的兴趣不在于特定算法的具体型构，而在于算法所属的更具有持续性的文化世界上，可供研究的材料不再局限。“算法”在不同的场景显示出不同的特性。

## 2. 留意进入方式

日常的社会互动是保密的。尽管研究者可以自由地出入和进行采访，但总会有更多的障碍需要逾越。“知识的不对称”（asymmetries of knowledge）是田野研究的一部分而不是研究的障碍。算法的复杂性和分散性，人类社会生活的整体性让“内部人士”也不能一窥全貌。人类学的研究，譬如算法系统，是由有限的存在、不完整的信息和不确定的联系形塑的。

## 3. 将访谈当作实地调查

“拾荒”的民族志学家的工作大都是以采访为中心的。访谈不是参与观察，只能反映人们**说**了什么而非**做**了什么。但将采访看作文化行动发生，而非人为建构的场景，是有必要的。在算法系统内或周围工作的人们生活在“访谈文化”（interview culture）中，他们熟悉采访和访谈的过程，甚至亲身实践过。采访对于他们来说是已知的互动环境，有既定的规则和意义，是日常生活的一部分。

## 4. 解析公司杂语（heteroglossia）

巴赫金（Mikhail Bakhtin）用杂语描述掺杂多种用语（voices）的文本。公司式的语言往往是杂糅的。对于公司而言，杂语揭示了公司内部的相互作用，就“算法”而言，它可能是法律意义上的知识产权、计算方法或是一簇新颖、复杂的技术丛。民族志研究者需要拒斥一种表面性的理解，即公司是有公开声明的直白意图的单一的行动者。



## 5. 提防反语（irony）

计算机文化的民族志研究者指出，程序员群体倾向使用反语和玩笑，在分析他们的话语时，对深层意义的理解能力显得十分重要。只有通过深入参与和丰富的语境描述，民族志研究者才能区分这种变化。例如，一位被访者常开玩笑称，数据“强迫”他做出明显带有个人色彩的决定。但这并不是一种数据原教旨主义的崇拜，相反，是对数据原教旨主义这一说法对使用数据的复杂方式的揶揄。算法工作者作为个体必然是程序化的，这一看法过于简化和天真。重新理解算法系统内的社会动态，与置身于系统中的人沟通，是必要的，也是负责任的做法。

# 四、结语

Seaver指出了算法在本质上的文化属性，以及这一理解能带来的优势。民族志地理解算法，将其作为文化的一部分，不仅包括理性的过程，也包括制度、群体、交错的背景语境和日常文化生活中的释义。民族志为进入和理解意味深长的实践世界提供了方法。Seaver所提供的策略将算法系统看作异质而弥散的社会技术系统，有超越软件和产权边界的深层纠缠。尽管如此，因为科技公司在法律上和技术上的保密，特定算法在特定时间、环境下的特定作用，仍是无法回答的问题。

原文：Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. Big Data & Society. <https://doi.org/10.1177/2053951717738104>.

李凌浩

中国社会科学院大学社会学院社会学系