

# 质性知识构造与量化知识分析综论：兼论学术对标法理论基础

叶 鹰

**摘 要** 文章在简要综述知识基因、知识单元、知识表示和知识图谱的基础上,探讨知识的质性构造和量化分析基础,指出知识具有表型(外型)、构型(内型)和量型(量值)三型特征。用逻辑分析法从现象到本质揭示知识表型特征,包括语言文字型、公式符号型和图表图谱型;知识构型可用学科类别和特征概念标识;知识量型可基于数据或信息计算知识量。进一步提出用知识花图式和知识花程式展示质性知识构造与量化知识分析。由知识三型特征引出知识变换三假设以及学术评价“对标法”;揭示质性知识构造适合解析静态知识特征,量化知识分析可延伸处理动态知识演化。质性知识构造与量化知识分析同时适用于剖析知识客体和知识主体,并独具学术评价应用价值。倡导通过知识分析创新图书情报学,同时也通过图书情报学促进知识之花盛开。图9。表1。参考文献38。

**关键词** 知识构造 知识分析 知识基因 知识单元 知识表示 知识图谱 学术对标法  
**分类号** G250.2

## A Synthetic Probing into the Qualitative Knowledge Construction and Quantitative Knowledge Analytics: Concerning the Theoretical Foundations for Academic Benchmarking Method

Fred Y. Ye

### ABSTRACT

Based on a brief review of knowledge genes and knowledge units, knowledge representation and knowledge graphs, this paper explores the qualitative structure of knowledge and the basis of quantitative analysis, and points out that knowledge has three characteristics, namely phenotype (exo-type), configuration (endo-type) and sum type (quantity value).

This paper uses logical analysis to reveal knowledge phenotypes from phenomenon to essence, including language and text type, formula symbol type, and graph chart type. Knowledge configuration can be identified by subject category and feature concept, and knowledge value can be calculated by and based on data and information.

It is proposed to use knowledge floral graph and knowledge floral formula as a comprehensive representation of qualitative knowledge structure and quantitative knowledge analysis, and it is suggested to adopt three hypotheses of knowledge transformation and academic evaluation “benchmarking method” from the three characteristics of knowledge to analyze static knowledge characteristics and process dynamic

通信作者:叶鹰,Email:yeye@nju.edu.cn,ORCID:0000-0001-9426-934X (Correspondence should be addressed to Fred Y. Ye,Email:yeye@nju.edu.cn,ORCID:0000-0001-9426-934X)

knowledge evolution.

Qualitative knowledge construction and quantitative knowledge analytics are both suitable for analyzing knowledge objects and knowledge subjects, and possess unique academic evaluation application value. This paper advocates to innovate library and information science through knowledge analytics, and promote the blossom of knowledge flowers through library and information science. 9 figs. 1 tab. 38 refs.

## KEY WORDS

Knowledge construction. Knowledge analytics. Knowledge gene. Knowledge unit. Knowledge representation. Knowledge graph. Academic benchmarking method.

## 0 引言:论题缘起及简要综述

当世间所有知识皆已化为存储于图书馆、数据库中的文献信息、数字资源时,图书情报工作者无不管理着人类知识而自豪。然而,我们对知识的认识还远远不够,仅仅是把知识管理类同于数据管理或信息管理。实际上,知识是有结构的,并非简单的数据信息堆积。知识区别于数据、信息,重在知识构造,需要知识分析。图书情报学若要以研究知识为己任,就应当对知识分析有所贡献。因此,辨析知识构造和推进知识分析意义明显,是为本文立题探讨之缘起。

从费希特的《知识学》<sup>[1]</sup>到波普尔的《客观知识》<sup>[2]</sup>,作为哲学知识论的知识学概念由来已久,知识学作为认识论已永存于哲学。然而,科学的知识学却存在概念混沌、理论缺失、技术多样等问题,致使知识组织<sup>[3]</sup>、知识技术<sup>[4]</sup>、知识工程<sup>[5]</sup>、知识学<sup>[6]</sup>等术语所指内涵皆在数据、信息与知识间徘徊,不能有效区分知识与信息,难以体现知识特质。究其原因,根源在于没有深入到知识内部揭示知识本质,因此需要反思知识构造和知识分析之概念与方法。

涉及知识构造与知识分析的概念,国内以知识基因(Knowledge gene)和知识单元(Knowledge unit)居多,国外则以知识表示(Knowledge representation)和知识图谱(Knowledge graph)为主。

知识基因概念可追溯至美国学者 Dawkins 类比生物基因而提出的文化、思想、观念等在传承

过程中存在思想基因的想法而提出的模因(Meme)概念<sup>[7]</sup>,并将模因定义为具有稳定性、再现性和逐渐变异性的基本思想单元。经印度学者 Sen 转用作情报基因<sup>[8]</sup>后,我国学者刘植惠系统地引进并发展了知识基因说,并将知识基因定义为知识遗传与变异的最小功能单元<sup>[9-10]</sup>。

知识单元概念可追溯至“数据元”并与知识基因概念交织发展<sup>[11]</sup>。我国学者赵红州在1986年较早提出“知识单元”概念,认为任何科学知识都是由相对不变的知识单元构成<sup>[12]</sup>。国外学者 Bergocist 在1987年将问题求解系统看作是一个知识元系统,提出了建立基于知识的问题求解系统和面向类型编程(TOP)的方法,认为知识元(KU)是用于模拟现实世界的知识构件<sup>[13]</sup>;Alani 等则集成通用句法分析模块,采用本体技术研究全文本分析抽取知识元技术<sup>[14]</sup>。国内在相关研究中也提出过类似定义,例如:知识元=信息元+经验+智慧+解决问题<sup>[15]</sup>;知识元是知识分解成的可独立使用的由名称、属性、操作和导航四个要素组成的最小单位<sup>[16]</sup>;知识元是文献中最小的、不可分割的、相对独立的一个元素<sup>[17]</sup>;知识元是指语义上相对完整地表达特定知识的最小内容单元<sup>[18]</sup>;知识元是在一定语境中相对独立的、具有完备知识表达的、最小粒度的语义单元<sup>[19]</sup>,等等。概括起来,多数学者趋向于认为知识单元是具有完备语义结构的、独立的、不能分割的最小知识单位,这与知识基因异曲同工。

知识表示和知识图谱主要用于智能信息处理<sup>[20]</sup>,从 Vickery 在1980年代的知识表示综

述<sup>[21]</sup>,到 Fensel 等在 2020 年的知识图谱著作<sup>[22]</sup>,知识表示和知识图谱交织发展,并结合知识表示学习共同深化<sup>[23]</sup>。关于知识表示方法,目前研究和使用的较多的主要有谓词逻辑表示法、产生式表示法、框架表示法、语义网络表示法、面向对象表示法、基于 XML/本体的表示法等。基于三元组的知识表示形式在知识图谱应用中得到广泛认同<sup>①</sup>,但其在计算效率、数据稀疏性等方面却存在诸多问题。近年来,以深度学习为代表的知识表示学习取得了重要进展,可以将实体的语义信息表示为稠密低维实值向量,进而在低维空间中高效计算实体、关系及其相互之间的复杂语义关联,对知识库的构建、推理、融合以及应用均具有重要的技术意义<sup>[24]</sup>。

知识表示是知识图谱的前端技术,知识图谱则是涉及知识表示等的综合应用,知识表示在知识图谱构建、推理、融合中具有重要作用。目前存在的表示方式仍是基于三元组形式完成的语义映射,在面对复杂的知识类型、多源融合的信息时,其表达能力仍然有限。但知识图谱发展迅速,其本质是一种揭示实体之间关系的语义网络,可对现实世界的事物及其相互关系进行抽象化、概念化、形式化描述,从而在技术意义上成为大数据分析的语义桥梁<sup>[25-26]</sup>。

如今在深度学习热潮中,知识表示和知识图谱正向人工智能(AI)时代迈进,虽然图书情报学也关注信息可视化和科学知识图谱研究,但越来越技术化的知识表示和知识图谱研究却基本归属于计算机科学范畴。于是,知识表示研究更多关心机器理解和处理知识,知识图谱研究也偏重以算法展示知识演化,两者远离感性直观,偏离了图书情报知识体系。

知识基因和知识单元概念盛行于国内却未

成国际热点,而知识表示和知识图谱则有国际燎原之势。究其原因,知识基因和知识单元概念没有系统发展支持技术,而知识表示和知识图谱则大力发展计算机支持技术,这在数据驱动研究盛行、大数据风靡天下的当今学界分别演化成弱势和强势,成为本土学术和国际学术之间差异的一种体现。知识基因和知识单元没有统一规范的国际技术支持,只似萤火闪烁;知识表示和知识图谱则在计算机科技的强势支撑下开疆拓土。不过,知识基因和知识单元虽然技术支持不足,却是合理概念,而知识表示和知识图谱虽有技术处理优势,却也有破碎知识之嫌;此外,重视内涵的知识基因和知识单元略显外在呈现不足,面向技术实现的知识表示和知识图谱也在直观性上有所欠缺。因此,如何架设知识表示、知识图谱技术与知识基因、知识单元概念之间的直观通道,也是本文的探讨方向。鉴于当前数据积累尚难以完全对有史以来的知识进程进行量化分析,故以质性表征为基础,兼顾量化分析,创制具有科学哲学特色的知识分析学,将会有利于图书情报学视角的研究深化。

## 1 思想方法:知识构造及学术定义

思考知识构造和知识分析首先需要确认有关知识的基本概念,而基本概念源于客观认知。为达成共识,用逻辑构造法按定义、公理、定理、推论层次展开如下:

定义一:学术客体和学术主体<sup>[27]</sup>:学术客体是学术的载体,包括论著(论文、专著)和刊物(期刊、数据库)等;学术主体是学术客体的创造者,包括学术个体(学者、个人)和学术团体(机构、大学)等。

学术客体和学术主体互为表里,构成第一

① 知识图谱描述真实世界中存在的各种实体或概念及其关系或属性,构成语义网络图,其节点标识实体或概念,边标记关系或属性,构成“实体—关系—实体”或“实体—属性—属性值”等三元组形式,从而使基于三元组的知识表示成为知识图谱的通用表示方式。

对概念范畴。

学术主体创造知识,具有创造性;学术客体记录知识,具有新颖性。明确学术主体和学术客体的概念有利于区分创造性和新颖性,也有利于学术评价;学术主体的顶端是学术大师,学术客体的高端是经典著作。学术大师是经典著作的作者和主干知识的创立者。

定义二:经典著作和主干知识<sup>[28]</sup>:经典著作是代表一个学科或学科群知识来源的知识客体,具有不可替代性。主干知识是一个学科或学科群的根本知识,具有不可颠覆性。由经典著作奠定的主干知识是一个学科异于其他学科的结构化、体系化内核,包括特征概念和重要规律,因而具有不可替代性和不可颠覆性。

经典著作与主干知识互为表里,构成第二对概念范畴。

学术主体创造知识的过程,也是将知识记录进学术客体中得以保存的过程,这一过程需用语词、符号、图谱等表达形式来呈现知识,从而形成知识表型和知识构型。

定义三:知识表型和知识构型:知识表型是知识的外观呈现形式;知识构型是知识的内在学术属性。

知识表型和知识构型互为表里,构成第三对概念范畴。

知识表型多样,但因人类知识表达主要依靠语言文字而存在知识表达第一公理:

公理一:语言文字是表达知识的基本型式。

在此公理上,所有知识皆可用语言文字表述,故语言文字表型是基本的知识表型。然而,伴随人类知识的发展、深化,尤其是科学知识的积累,现有知识也非语言文字所能完全覆盖,因而存在知识表达第二公理:

公理二:公式符号和图表图谱是语言文字的扩展。

由上述定义和国家标准《文献著录总则》<sup>[29]</sup>中的定义,可以证明以下定理:

定理一:文献必有知识表型。

证明:因为知识必有知识表型,而文献是知

识的载体,所以文献作为载体化的知识也必有知识表型。【Q. E. D.】

这样,图书情报学研究关注的文献就自然而然与知识表型相关联,并可推论:

推论一:知识表型集中体现在文献里。

依据公理一和公理二,可以证明以下定理:

定理二:必有非语言文字型知识存在。

证明(反证法):若没有非语言文字型知识存在,则公式符号、图表图谱承载的知识将不存在;而公式符号、图表图谱承载的知识存在,故必有非语言文字型知识存在。【Q. E. D.】

于是,语言文字、公式符号和图表图谱成为表达知识的三种基本型式。若将语言文字纳入广义的符号,则卡西尔哲学“人是符号的动物”<sup>[30]</sup>与“人是语言的动物”同构。但因语言文字不足以表达所有人类知识,故有推论:

推论二:语言文字不能穷尽所有知识。

在科技和艺术领域,用公式符号和图表图谱表达的知识更加“硬核”。

若以学术符号代指所有表征学术知识的语言文字、公式符号和图表图谱,则可用学术符号概念定义知识基因和知识单元:

定义四:知识基因和知识单元:知识基因是学术符号组合而成的表达知识的最小独立功能单位,在同类知识中具有永不改变的涵义。知识单元是知识基因的应用组合。

由此定义,知识基因具有涵义不变性,虽有可能伴随“知识遗传和变异”而进化,但作为知识的最小独立功能单位,其意义相对稳定,可表现为语句、公式、图表等。知识单元可以是一个或多个知识基因,为便于应用而构成组合。

若将不同语言文字的同义转换称为翻译,则有如下定理:

定理三:翻译不改变知识基因。

证明:世界上多种语言文字同义转换(翻译)时必须保持知识涵义不变,而知识涵义由知识基因确定,所以翻译不改变知识基因。

【Q. E. D.】

因此可以推断:

推论三:不改变知识基因的翻译才是合格的翻译。

由此,“信、达、雅”确为翻译的好标准和高标准。

知识表型和知识构型合成质性知识构造,其感性直观呈现详见下文。

## 2 知识表型(外型)及外在呈现

知识的现实直观呈现必然是厚重的印刷书

刊和浩瀚的数字资源,印刷在书刊上的学术符号和记录成数字资源的信息编码就成为体现知识的外在型式,即知识表型。

根据上节对思想方法的阐述,知识除语言文字表型外,也有公式符号表型(见图1)和图表图谱表型(见图2),并成为科学知识表达所必需。

知识用公式符号表型能有效提升逻辑思维及推理的效率,其重要性恰如一句名言“万物速朽而公式永恒”(everything is ephemeral while formula is eternal)<sup>[31]</sup>。

$$e^{iz} = \cos z + i \sin z$$

$$E^2 = m^2 c^4 + p^2 c^2$$

$$\int_a^b f'(x) dx = f(b) - f(a)$$

$$H(p, q, t) = p_k \dot{q}_k - L$$

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

$$\frac{\partial H}{\partial p_k} = \dot{q}_k, \frac{\partial H}{\partial q_k} = -\dot{p}_k$$

$$\int_S d\omega = \oint_{\partial S} \omega$$

$$\delta S = \delta \int_S d^l x L = 0$$

图1 数理知识的公式符号表型示例

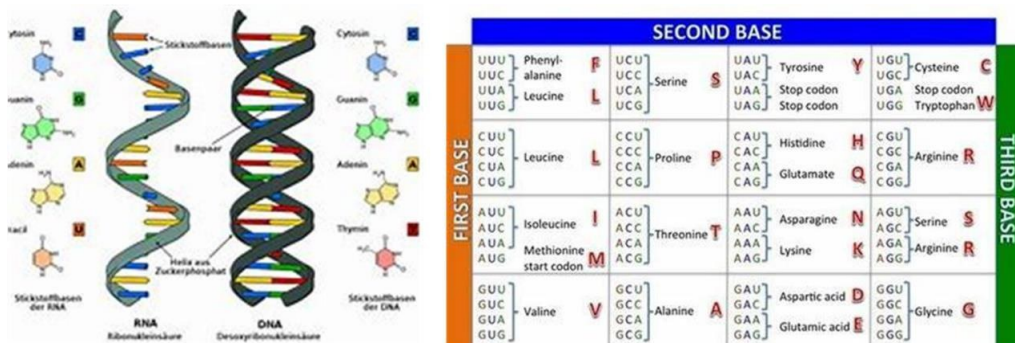


图2 科学知识的图表图谱表型示例(公有知识)

知识的图表图谱表型则能具象清晰地表明知识特质。如DNA结构、原子行星模型等,见图即明,其意义真可谓“一图值千言”(A figure is worth of thousands words),如图3所示原子模型的发展。

同理,科学方法中的光谱图(波谱图)也是超越文字符号的知识表型,从可见光谱到红外光谱(IR)、紫外光谱(UV)等(见图4),都构成

独特的知识,成为科学中的必要知识存在。

综上,在现代科学中,没有公式图表是不可想象的。公式符号表型、图表图谱表型的知识往往非语言文字所能取代,而且常常是“硬核”知识的表达型式。

艺术图谱也是特殊的图表图谱,包括图画、照片、乐谱等(从广义上说,音频和视频也是图谱)。如图5所示。

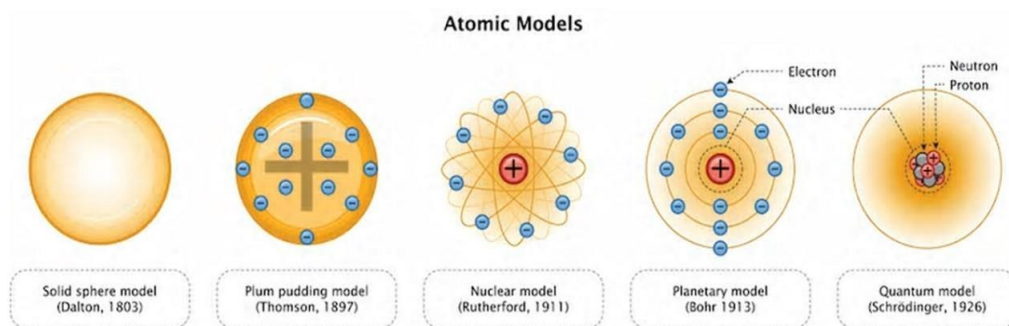


图 3 图表图谱表型之代表性原子模型(公有知识)

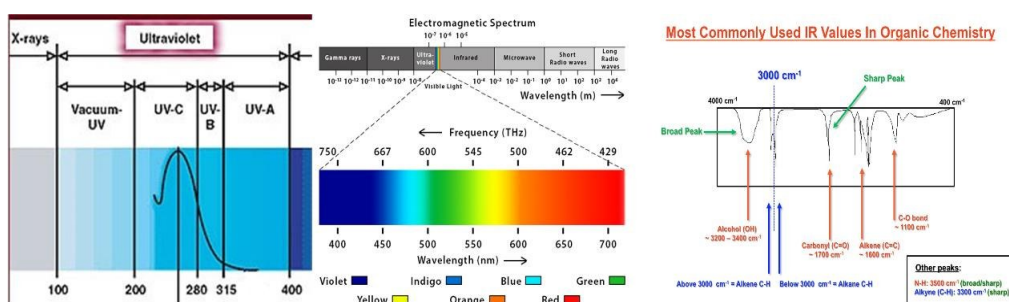


图 4 图表图谱表型之波谱图示意(公有知识)

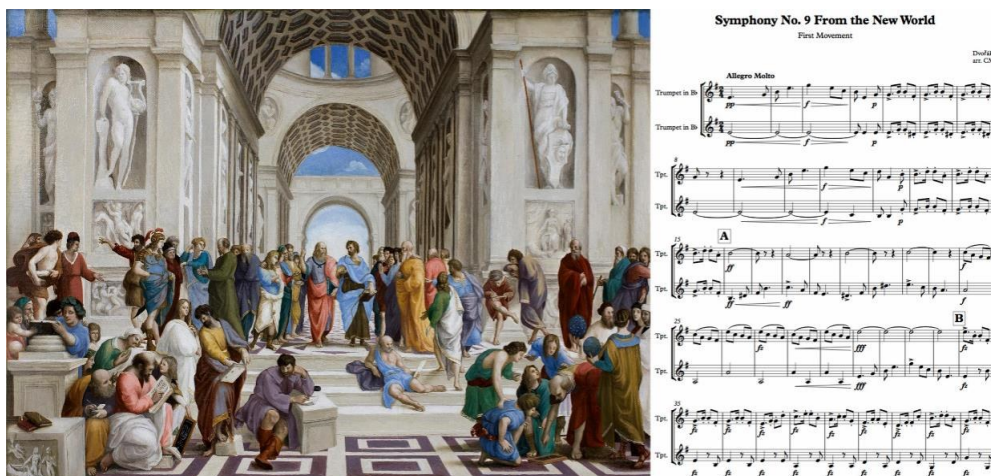


图 5 作为图表图谱的艺术图画和乐谱示例(公有知识)

由此可见,知识的图表图谱表型也是艺术知识表达的必需方式。没有图画,就没有美术;没有乐谱,就难以记录音乐。与科学知识的量化图表图谱特征不同的是,艺术知识的图表图谱没有量化约束。

### 3 知识构型(内型)及分类编码

当我们深入知识内容时,首先面临的是学科差异。

伴随知识的发展和深化,人类积累的庞大知识体系不得不分科而治,犹如经济社会发展形成社会分工一样。面对浩瀚的“知识海洋”,知识分类不可避免,其结果就是学科划分。不同学科从基本概念开始就存在差异,导致表述概念的语言文字必然也存在差异。

知识分类也是辨识知识基因所需。但无论是人为分类还是自动聚类,学科划分都不完美,造成知识分类的不完备性<sup>[32]</sup>。知识领域作为知识分类的结果,通常可以无歧义地分为人文(哲学、文艺等)、社科(经济、管理等)、科技(科学、技术等)三大领域。

而当分类继续深入时,分类编码就显得十分必要。这正是图书情报学必须研究知识分类的起因。结合语言文字编码,可设计简单组配编码分类标识以表征知识构型。

例如,学科分类码可用大写拉丁字母表示主类(留用 I、J、K):A-艺术;B-哲学、宗教;C-语言、文学;D-历史、社会;E-经济;F-法律、政治;G-管理、教育;H-人类知识综合;L-理学综合;M-数学;N-天文;O-地学;P-物理;Q-化学;R-生物;S-医学;T-工业技术;U-农学。小写字母表示亚类,如 Ta-、Tb-、Tc-等表示工业技术亚类。

语言文字则以 Ix 编码,用 Ia 和 Ib 分别指代人造语言文字和自然语言文字,x=c、d、e、f、r、ar、es、gr、it、la……分别指代汉语、德语、英语、法语、俄语、阿拉伯语、西班牙语、希腊语、意大利语、拉丁语等语言文字;公式符号以 IIy 编码,y=a、b、c……可用于区分专业符号、公式方程、特殊编码等;图表图谱以 IIIz 编码,z=a、b、c、d……可

分别指代线图、数表、图画、图谱等。统合 IxIIyIIIz,既体现了语言文字,也编码了知识外型。

综合成 Kc = {X; Y}, X = 学科分类码, Y = IxIIyIIIz 知识外型码,就能简洁地揭示知识构型内核。于是可用 Kc = {C; Ic} 表征汉语文学, Kc = {A; IIIId} 表征音乐乐谱, Kc = {P; IIb} 表征物理公式, Kc = {R; IIIb} 表征生物数据, Kc = {D; IeIIIc} 表征英语表达的图画历史,以此类推。

在已有知识中,哲学和文学拥有语言文字优势,科学和艺术则分别具有公式符号优势和图表图谱优势;以公式符号和图表图谱见长的科学和艺术并非排斥语言文字,而是在语言文字中渗入了高效有力的符号图表元素;若把符号和图表渗入古老的哲学、文学论述中,也有望拓展独特创新。人们通常只用语言文字表达思想,形成所谓“语言的界限即世界的界限”<sup>[33]</sup>,而公式符号和图表图谱则能为语言文字插上双翼,让思想展翅飞翔。

因此,既不应低估语言文字的力量,也不应排斥公式符号和图表图谱的运用,图书情报学在核心知识或主干知识表达上完全可以充分综合语言文字、公式符号和图表图谱的优势而有所建树。这也是知识分析对图书情报学知识创建的启示。

深入分析质性知识构造还需要量化知识,所以需要考虑知识量值。

#### 4 知识量值及公式刻画

知识量值可定义为:知识 K 是主观信息 J 对客观信息 I 的积分<sup>[34]</sup>,如公式(1)所示:

$$K = k \int J dI = k \int v \ln I dI = kvI(\ln I - 1) + K_0 = K_0 + \Delta K \quad (1)$$

其中 k 是信息的知识转化系数(单位 kit/bit);  $v \in [0, 1]$  是价值系数,  $K_0$  是代表原有知识的积分常数,  $\Delta K = kv(\ln I - 1)$  就是增加的知识。

公式(1)表征了静态知识测度,也连通了数据与信息。若附加动态参数 t 作为时间变量,则可对音频、视频等动态知识形态进行量化。

再以 Top 1% (顶端百分之一)、Top 0.1% (顶端千分之一)等来量化知识的品质优异程度,即可构成知识量值从数量到质量的通量。

综上,知识总特征可概括为三元组  $K = (X; Y; Z) = (Kc; Z)$ , 其中  $Kc = \{X; Y\}$  就是知识构型, X 表征学科类别, Y 揭示知识表型,  $Z = \Delta K$  标

示知识增量,这是与“实体—属性—属性值”类似的三元组知识表示。

## 5 知识花喻:知识花图式与知识花程式

常言“知识之花”,用花朵比喻知识,的确是一种可行和有用的方式。

前述知识的学科领域属性也可用花形表征内核:以圆及椭圆表征人文学科知识核心,譬喻其核心信念圆融;以多边形表征科学技术,譬喻其核心轮廓分明;以星形表征社会科学,譬喻其核心条理交融;不规则几何形状则可用于交叉学科。几何图形直径可表征知识贡献大小。各学科领域花形约定见表 1。

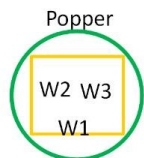
表 1 学科领域代码及花形表征

人文学科		社会科学		科学技术	
哲学 Ba	正圆	经济学 B	五角星	数学 M	正三角形
宗教、神学 Bb	斜椭圆	管理学 G	六角星	天文、地学 N, O	倒三角形
语言、文学 C	竖椭圆	社会学 De	四角星	物理 P	正方形
艺术、美术 A	横椭圆	政治、法律 F	三角星	化学 Q	苯环形
音乐 Ac	音符	教育、国情 Gg	对开书	生物 R	五瓣形
人文综合 Ah	球体	社科综合 Gh	星体	医学 S	十字形
	不规整圆	历史 Da *		农学 U	三叶形
			不规则星	心理 Rs *	
			八角星	工程 T	八边形
				科技综合 Th	多面体

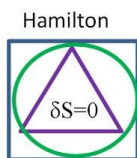
\* 历史按传统文史哲观属于人文学科,但按其现代理论方法假设应归社会科学。心理学兼有社会生物医学属性。

两两组合构成的交叉学科,也能赋予其几何形象。如数学哲学——三角形内接圆,哲学物理——圆内接正方形,物理化学——正方形内接五边形,等等。

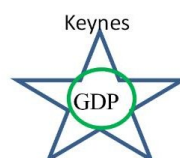
仿照植物花图式(Floral graph),将知识表型和知识构型结合的特征概念标记用花形图示,即构成知识花图式。用表 1 中几何形状为核心构造知识花图式,以象征知识构造,如图 6 所示。



人文知识之花:  
波普尔三世界论



科技知识之花:  
哈密顿原理



社科知识之花:  
凯恩斯经济思想

图 6 知识花图式示意

将知识花图式的核心展示于知识 K 和时间 t 构成的坐标系里并确定关联,就能得到如知识发展树似的知识花迹。图 7 和图 8 分别给出了物理知识主干树和哲学知识主干树的主体、客

体知识花迹。

图 7 中可见物理理论多有数学支点和实验物理来源;图 8 则提示哲学进步(P)从源头起皆有思想与方法、先验与经验之交融。



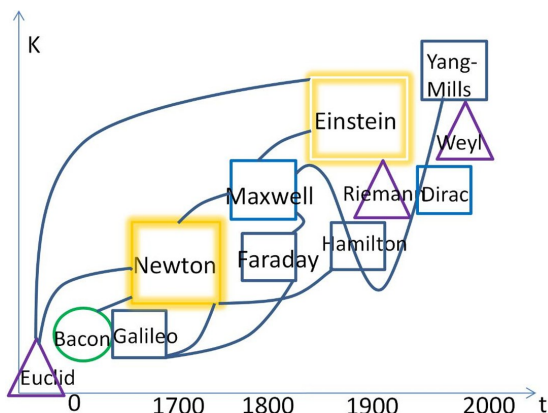


图7 物理主干知识发展示意:主体创造花迹

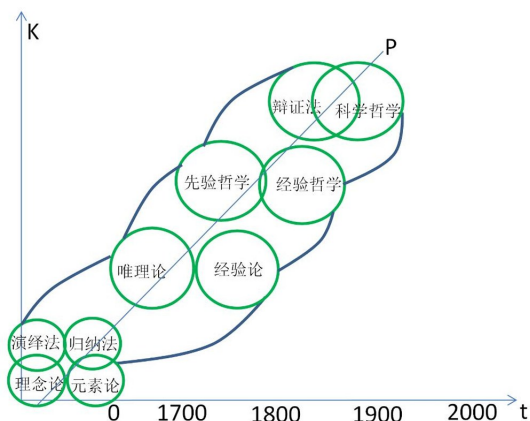


图8 哲学主干知识发展示意:客体学说花迹

同理,仿照植物花程式(Floral formula)<sup>①</sup>,可用知识花程式表征知识构造。以知识三元组定义其一般形式,见公式(2):

$$\text{知识花程式 } F[K] = [X; Xx(\dots); Y; IxHyIIIz; Z; [\Delta K]] / [\text{Range}] \quad (2)$$

【】内是知识三元组(X;Y;Z),Xx(…)是学科编码(特征概念提示),/[ ]界定可比范围。F[K]约定直指知识客体,若用F[S]表征知识主体,则有:

$$F[S] = [X; Xx(\dots); Y; IxHyIIIz; Z; [\Delta K]] / [\text{Range}] \quad (3)$$

① 花程式是用特定符号和数字来表征花的组成、排列、位置及其相互关系的一种形式,一般格式为: X Km; Cm; An; Gx; y。X为符号: \*表示辐射对称花, ↑表示两侧对称花, ♂表示雌花, ♀表示雄花; 符号 K、C、A、G 分别表示花萼、花冠(K、C 合并是花被 P)、雄蕊群和雌蕊群; m、n 为数字,通常写在相应字母右下角表示该结构的数目,数目很多时用 ∞,花的某部分相互连合时加( ); G 上划线代表下位子房, G 下划线代表上位子房, x 为心皮数, y 为子房数。如百合花程式为: \* P3+3; A3+3; G(3:3)。参见金银根主编《植物学》,科学出版社,2017 第 3 版, 266-267 页。

相对知识增量也可用 Top10%、Top1%、Top0.1%等定性表述:各学科学术经典是塑造人类知识的源泉依据,应是学术客体精华,故可设定其属于哲学经典:

$$F[K] = [X; B(\text{TaoTeching}); Y; Ic; Z: [Top 0.1\%]] / [Classics]$$

$$F[K] = [X; B(\text{Metaphysica}); Y; Igr; Z: [Top 0.1\%]] / [Classics]$$

$$F[K] = [X; B(\text{Tractatus}); Y; Ila; Z: [Top 0.1\%]] / [Classics]$$

社会科学经典:

$$F[K] = [X; E(\text{GeneralTheory}); Y; Ie; Z: [Top 0.1\%]] / [Classics]$$

$$F[K] = [X; D(\text{LawSpirit}); Y; If; Z: [Top 0.1\%]] / [Classics]$$

科学经典:

$$F[K] = [X; M(\text{Elements}); Y; IgrIbIIIa; Z: [Top 0.1\%]] / [Classics]$$

$$F[K] = [X; P(\text{Principia}); Y; IlaIbIIIa; Z: [Top 0.1\%]] / [Classics]$$

$$F[K] = [X; R(\text{OriginSpecies}); Y; IeIIIc; Z: [Top 0.1\%]] / [Classics]$$

也可以选列以下学术主体标杆:

哲学:

$$F[S] = [X; B(\text{Confucius}); Y; Ic; Z: [Top 0.1\%]] / [Thinkers]$$

$$F[S] = [X; B(\text{Aristotle}); Y; Igr; Z: [Top 0.1\%]] / [Thinkers]$$

$$F[S] = [X; B(\text{Kant}); Y; Id; Z: [Top 0.1\%]] / [Thinkers]$$

社会科学:

$$F[S] = [X; E(\text{Smith, Adam}); Y; IeIIIc; Z: [Top 0.1\%]] / [Thinkers]$$

$$F[S] = [X; E(\text{Keynes}); Y; IeIIIc; Z: [Top 0.1\%]] / [Thinkers]$$

科学:

$$F[S] = [X; P(\text{Newton}); Y; IeIbIIIb; Z: [Top 0.1\%]] / [Thinkers]$$

$$F[S] = [X; P(\text{Einstein}); Y; IdIIIbIIIb; Z: [Top 0.1\%]] / [Thinkers]$$

$$F[S] = [X; R(\text{Darwin}); Y; IeIIIc; Z: [Top 0.1\%]] / [Thinkers]$$

任何能为人类知识系统贡献  $\Delta K/K > 0$  的客体论著或主体学者都是杰出的,列出花程式便知概貌,例如:

$$F[K] = [X; P; Y; IxIIyIIIz; Z: [Top0.1\%]] / [10 Classics] \quad (4)$$

标示对物理学贡献知识 Top0.1% 的十大学术经典之一。又如:

$$F[S] = [X; Q; Y; IxIIyIIIz; Z: [5]] / [100 Thinkers] \quad (5)$$

标示对生物学贡献知识相对增量 5 的百位思想家之一。

知识花图式与知识花程式的优点是简洁明确、一目了然,缺点是蕴涵有限、信息不足。与知识的公式、图谱表型需用语言文字补充说明一样,知识花图式与知识花程式也需要用文字进一步解释。

由此,知识花图式与知识花程式相结合,可形象地表示“知识之花”,犹如定性知识图谱。与定量知识图谱相比<sup>[35-36]</sup>,定性知识图谱可以揭示知识的本质关联和概貌。定性图谱与定量图谱若能相互印证,符合双证原理<sup>[28]</sup>,即是真实的知识图谱。

## 6 知识分析:从知识变换三假设到学术对标法

如果把改变知识特征的过程称为知识变换,则知识变换是知识分析的重要形式,由知识三型特征可引申提出三假设供讨论。

假设一:在准确和精确操控下,改变知识表型的操作可以维持知识构型基本稳定。

由假设一可以推论:

- (1) 精准的翻译可以基本保持原作恒定。
- (2) 同一知识构型可以有不同知识表型。

假设二:改变知识构型的操作可能引起知识内容失真。

由假设二可以推论:

(1) 对原作附加任何操作皆可能导致原作失真。

(2) 原作具有不可替代性。

假设三:改变知识量值必以改变组成知识构型和知识表型的数据信息为前提。

由假设三可以推论:

(1) 知识贡献取决于知识表型、知识构型和知识量值的集成。

(2) 知识量值难以单独测算。

质性知识构造与量化知识分析相结合的启发是可以创立学术评价“对标法”:以历史上的知识成就为标杆,把后人成果按知识主体或知识客体分别与历代名家或名作进行对比,对标已有学术成就,即可对知识创新程度进行认知认定。图9给出了经济学知识对标示意,新的经济学作者或作品可通过同行对标或公众对标获取评价。

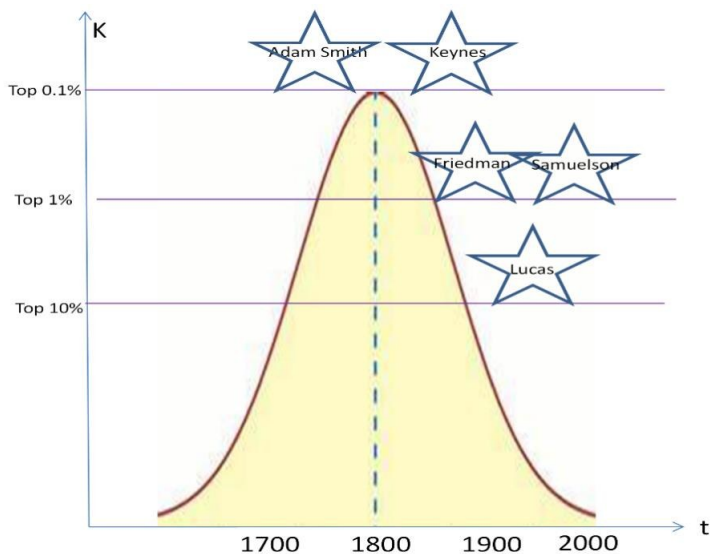


图9 经济学知识对标示意

如果有经济学新人或新著在图9对标体系里能跃进Top 10%(顶端十分之一)、Top 1%(顶端百分之一)、Top 0.1%(顶端千分之一),就必然成为不朽名家或名作,皆为不凡成就。从统计分布角度看,任何学科中的作者和作品的品

质高低都应服从正态分布,于是可以通过建立交互网站,以智能算法汇集同行和大众的开放评价数据,将动态结果作为测度质性知识成就和量化知识贡献的客观依据,从而以学术对标法创新学术评价<sup>①</sup>。

① 笔者已以“一种针对主体或客体进行评价和对标的方法”为题登记中国专利优先权。登记号:202110723425.5

过去学术评价的一个惯例是以产出论著的数量及其引用作为评判学术客体及其主体的主要依据,虽然就学术影响而言有一定判识作用,然而也可能导致严重失误。典型实例是黎曼(G. F. B. Riemann, 1826—1866),其毕生仅有10篇论文发表,至今引用也不高,但其知识贡献却遍及数学、物理、哲学思想等领域,成为人类有史以来稀有的杰出学者,为人类“硬核”知识做出了优异贡献。因此,需要改良学术评价。学术对标法即是以历代知识成果为参照的学术评价方法,这里仅简要明确其理论基础。

## 7 讨论:以图书情报学促进“知识之花”盛开

质性知识构造与量化知识分析相结合,可望贡献揭示核心知识的思想方法。在大数据时代,这是一种逆向思维:不追求数量众多而关注品质优良。

本文贡献两点原创设想:①可通过质性知识构造揭示知识本质,以少胜多,以质胜量地实现知识品质发现及提升,并连通知识基因说;②可通过量化知识分析桥接知识图谱,成为大数据知识图谱中小数据核心发现的桥梁,实现核心知识集萃。

“知识就是力量”,培根这句代表着西方科学精神的名言曾激励人们求知奋进。的确,知识既能充实人的精神,增强人的智慧,改造人的心灵,也能化为改造世界的力量。人类累积知识,集成在经典著作里,代代相传,构成质性知识构造与量化知识分析的基础。

学术大师、经典著作与领域主干知识往往相互关联、相互印证。比如柏拉图、康德、维特根斯坦等属于哲学学术大师,其相应著作《理想国》《未来形而上学导论》《逻辑哲学论》等成为哲学经典著作,柏拉图理念论、康德先验哲学、维特根斯坦语言分析哲学等则构成哲学主干知识。又如牛顿、爱因斯坦、狄拉克等属于物理学术大师,《自然哲学的数学原理》《广

义相对论基础》《量子力学原理》等成为物理经典著作,经典力学、相对论、相对论性量子物理等构成物理主干知识。其他学科亦然,如经济学中的学术大师及其经典著作如亚当·斯密《国富论》、马歇尔《经济学原理》、凯恩斯《就业利息和货币通论》等,化学中主干知识如元素周期律、化学结构、化学反应等。这些关联既是对书理学<sup>[37]</sup>的补充,也是支撑知识构造与知识分析的素材。

知识构造与知识分析涉及的理论问题和技术问题还有很多,理论问题涉及知识基因变异及进化,技术问题关联知识工程和人工智能<sup>[38]</sup>。人工智能(AI)技术的运用有变革知识分析的潜力,可能成为知识学的技术枢纽,并有融通知识的语言文字表型、公式符号表型和图表图谱表型之潜能。

通过质性知识构造与量化知识分析连接定性思想和定量方法,可望独具学术意义和应用价值——以图书情报学促进“知识之花”盛开。

## 8 小结:以知识分析学创新图书情报学

质性知识构造与量化知识分析可以成为哲学知识论和科学知识论之间的桥梁,一方面连接哲学认知,另一方面贯通科学测量,知识量值与学科特征相互参证,可望构筑知识分析主干。

整理已有知识固然必要,但开创新知识更加重要。知识表型与知识构型相结合,如同描述植物花朵般描述“知识之花”,就是期望以图书情报学认知促进“知识之花”盛开,同时也通过知识分析学创新图书情报学。

图书情报学守护着人类知识花园,培育着“知识之花”,在注重守护人类已有知识的同时,更应关注知识创新,这正是探讨质性知识构造与量化知识分析的意义之所在,也是由此关注学术评价的原因。而以知识分析学创新图书情报学,不仅能为图书情报学寻求发展机遇和突破方向,也能实现品质主导、量化辅助之学术评价对标法,以期通达美好之未来。

## 参考文献

- [ 1 ] Fichte J G. Science of knowledge (Wissenschaftlehre), with the first and second introductions[M]. Heath P, Lachs J, trans. New York: Meredith Corp., 1970.
- [ 2 ] Popper K R. Objective knowledge: an evolutionary approach[M]. Oxford: Oxford University Press, 1972.
- [ 3 ] 王知津, 李培, 李颖, 等. 知识组织理论与方法[M]. 北京: 知识产权出版社, 2009. (Wang Z J, Li P, Li Y, et al. The theory and methods of knowledge organization[M]. Beijing: Intellectual Property Press, 2009.)
- [ 4 ] 曾民族. 知识技术及其应用[M]. 北京: 科学技术文献出版社, 2007. (Zeng M Z. Knowledge technology with application[M]. Beijing: Scientific and Technical Documentation Press, 2007.)
- [ 5 ] 史忠植. 知识工程[M]. 北京: 清华大学出版社, 1988. (Shi Z Z. Knowledge engineering[M]. Beijing: Tsinghua University Press, 1988.)
- [ 6 ] 柯平. 知识学研究导论[J]. 图书情报工作, 2006, 50(4): 6-10, 34. (Ke P. On research of knowledge science [J]. Library and Information Service, 2006, 50(4): 6-10, 34.)
- [ 7 ] Dawkins R. The selfish gene[M]. 2nd ed. Oxford: Oxford University Press, 1989.
- [ 8 ] Sen S K. A note on the idea gene and its relevance to information science[J]. Annals of Library Science and Documentation, 1981, 28(1-4): 97-102.
- [ 9 ] 刘植惠. 知识基因理论的由来、基本内容及发展[J]. 情报理论与实践, 1998(2): 8-13. (Liu Z H. The origin, basic contents and developments of knowledge gene theory[J]. Information Theory and Practice, 1998(2): 8-13.)
- [ 10 ] 刘植惠. 知识基因理论新进展[J]. 情报科学, 2003, 21(12): 1243-1245. (Liu Z H. About the new development of knowledge gene theory[J]. Information Science, 2003, 21(12): 1243-1245.)
- [ 11 ] 文孝庭, 罗贤春, 刘晓英, 等. 知识单元研究述评[J]. 中国图书馆学报, 2011, 37(5): 75-86. (Wen X T, Luo X C, Liu X Y, et al. Knowledge unit research: review and comment[J]. Journal of Library Science in China, 2011, 37(5): 75-86.)
- [ 12 ] 赵红州, 蒋国华. 在科学交叉处探索科学——从科学学到科学计量学[M]. 北京: 红旗出版社, 2002: 81-86. (Zhao H Z, Jiang G H. Exploring science at the crossroads of scientific disciplines; from the science of science to scientometrics[M]. Beijing: Red Flag Press, 2002: 81-86.)
- [ 13 ] Bergocist J T. A type-oriented approach to knowledge-based system [D]. Helsinki: Helsinki University of Technology, 1987.
- [ 14 ] Alani H, Kim S, Millard D E, et al. Automatic ontology-based knowledge extraction and tailored biography generation from the web[J]. IEEE Intelligent Systems, 2003, 18(1): 14-21.
- [ 15 ] 孙成江, 吴正荆. 知识服务战略: 创建增值联盟[J]. 情报科学, 2002, 20(10): 1028-1030. (Sun C J, Wu Z J. Knowledge service strategy: creating value-added alliance[J]. Information Science, 2002, 20(10): 1028-1030.)
- [ 16 ] 温有奎, 徐国华. 知识元链接理论[J]. 情报学报, 2003, 22(6): 665-670. (Wen Y K, Xu G H. Knowledge element linking theory[J]. Journal of the China Society for Scientific and Technical Information, 2003, 22(6): 665-670.)
- [ 17 ] 文庭孝. 知识单元的演变及其评价研究[J]. 图书情报工作, 2007(10): 72-76. (Wen T X. Research on the evolution of knowledge unit and its evaluation[J]. Library and Information Service, 2007(10): 72-76.)
- [ 18 ] 索传军. 知识转移视角下的学术论文老化与创新研究[J]. 图书情报工作, 2014, 58(5): 5-12. (Suo C J. Study on obsolescence and innovation of academic papers from perspective of knowledge transfer[J]. Library and Information Service, 2014, 58(5): 5-12.)
- [ 19 ] 戎军涛. 学术论文内容知识元语义描述模型研究[J]. 情报科学, 2019, 37(7): 30-35. (Rong J T. Research on the semantics description model of knowledge unit in academic papers[J]. Information Science, 2019, 37(7): 30-35.)
- [ 20 ] 邓方, 陈文颖. 智能计算与信息处理[M]. 北京: 北京理工大学出版社, 2020. (Deng F, Chen W J. Intelligent computing and information processing[M]. Beijing: Beijing Institute of Technology Press, 2020.)

- [21] Vickery B C. Knowledge representation; a brief review[J]. *Journal of Documentation*, 1986, 42(3): 145-159.
- [22] Fensel D, Simsek U, Angele K, et al. Knowledge graphs: methodology, tools and selected use cases[M]. Cham: Springer Nature Switzerland AG, 2020.
- [23] 刘知远, 孙茂松, 林衍凯, 等. 知识表示学习研究进展[J]. *计算机研究与发展*, 2016, 3(2): 247-261. (Liu Z Y, Sun M S, Lin Y K, et al. Knowledge representation learning: a review[J]. *Journal of Computer Research and Development*, 2016, 3(2): 247-261.)
- [24] Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(8): 1798-1828.
- [25] 徐增林, 盛泳潘, 贺丽荣, 等. 知识图谱技术综述[J]. *电子科技大学学报*, 2016, 45(4): 589-606. (Xu Z L, Sheng Y P, He L R, et al. Review on knowledge graph techniques[J]. *Journal of University of Electronic Science and Technology of China*, 2016, 45(4): 589-606.)
- [26] 刘峤, 李杨, 杨段宏, 等. 知识图谱构建技术综述[J]. *计算机研究与发展*, 2016, 5(3): 582-600. (Liu Q, Li Y, Yang D H, et al. Knowledge graph construction techniques[J]. *Journal of Computer Research and Development*, 2016, 53(3): 582-600.)
- [27] 黄晨, 赵星, 卞杨奕, 等. 测量学术贡献的关键词分析法探析[J]. *中国图书馆学报*, 2019, 45(6): 84-99. (Huang C, Zhao X, Bian Y Y, et al. Measuring academic contributions via keyword analytics[J]. *Journal of Library Science in China*, 2019, 45(6): 84-99.)
- [28] 叶鹰. 试论图书情报学的主干知识及有效方法: 兼论双证法和模本法之效用[J]. *中国图书馆学报*, 2021, 47(3): 58-66. (Ye F Y. Identifying backbone knowledge and effective methods in library and information science: useful methods as double-evidence and model-sample[J]. *Journal of Library Science in China*, 2021, 47(3): 58-66.)
- [29] 文献著录总则: GB3792.1-83[S]. 国家标准局, 1983. (General bibliographical description: GB3792.1-83[S]. National Bureau of Standards, 1983.)
- [30] Cassirer E. An essay on man: an introduction to a philosophy of human culture[M]. New York: Doubleday & Co. Inc., 1954.
- [31] 量子学派. 公式之美[M]. 北京: 北京大学出版社, 2020. (Quantum School. The beauty of formula[M]. Beijing: Peking University Press, 2020.)
- [32] Ye F Y. Measuring knowledge: a quantitative approach to knowledge theory[J]. *International Journal of Data Science and Analysis*, 2016, 2(2): 32-35.
- [33] Wittgenstein L. Tractatus logico-philosophicus[M]. Pears D F, McGuinness B F, trans. London: Routledge, 2001.
- [34] 叶鹰. 图书情报学的学术思想与技术方法及其开新[J]. *中国图书馆学报*, 2019, 45(2): 15-25. (Ye F Y. An essay on the academic thoughts and technical methods with their renewing in library and information science[J]. *Journal of Library Science in China*, 2019, 45(2): 15-25.)
- [35] 刘则渊, 陈悦, 侯海燕, 等. 科学知识图谱: 方法与应用[M]. 北京: 人民出版社, 2008. (Liu Z Y, Chen Y, Hou H Y, et al. Mapping knowledge domains: methods and application[M]. Beijing: People Press, 2008.)
- [36] 曹倩, 赵一鸣. 知识图谱的技术实现流程及相关应用[J]. *情报理论与实践(ITA)*, 2015, 12(38): 127-132. (Cao Q, Zhao Y M. The realization process and related applications of knowledge graph[J]. *Information Studies: Theory & Application (ITA)*, 2015, 12(38): 127-132.)
- [37] 叶鹰. 书理学引论[M]. 北京: 中国书籍出版社, 2019. (Ye F Y. An initiatory bookics[M]. Beijing: China Book Press, 2019.)
- [38] Lucci S, Kopec D. Artificial intelligence in the 21st century: a living introduction[M]. Boston: Mercury Learning and Information, 2016.

叶 鹰 南京大学信息管理学院教授, 博士生导师。江苏 南京 210023。

(收稿日期: 2021-05-18)