http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

基本形式化本体 BFO 及中文版介绍

朱彦[1]，ZHENG Jie[2]，李晓瑛[3]，杨啸林[4*]，HE Yongqun[5*]

（1. 中国中医科学院中医药信息研究所，北京 100700；

2. University of Pennsylvania Perelman School of Medicine, PA, 19104 USA

3. 中国医学科学院医学信息研究所，北京 100005；

4. 中国医学科学院基础医学研究所，北京 100005；

5.University of Michigan Medical School, Ann Arbor, MI, 48109, USA）

［**摘要**］本体在人类可解释和计算机可解释的数据和知识标准化、共享、集成和高级分析中发挥着重要作用。基本形式化本体（BFO）作为一种应用广泛的顶层本体，在生物、生物医学等领域已经被 300 多个本体所采用。BFO 也已被开放生物/生物医学本体工场（OBO Foundry）作为官方的顶层本体。为了支持 BFO 中文使用与推广，我们作为 OntoChina 的成员，在本体研究社区的支持努力下，将 BFO 翻译成中文并予以共享。本文主要介绍了 BFO 的主要术语及其中文翻译情况。

［**关键词**］本体；顶层本体；基本形式化本体；中文翻译；本体中国；开放生物/生物医学本体工场

Introduction to Basic Formal Ontology and its Chinese Translation

Yan Zhu[1], ZHENG Jie[2], Xiaoying Li[3], Xiaolin Yang[4*], HE Yongqun[5*]

（ 1. Institute of Information on Traditional Chinese Medicine, Medicine, China Academy of Chinese Medical Sciences, Beijing, 100700

2. University of Pennsylvania Perelman School of Medicine, PA, 19104 USA
3. Institute of Medical Information (IMI) & Library, Chinese Academy of Medical Sciences and Peking Union Medical College (CAMS & PUMC)
4. Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences (CAMS) & School of Basic Medicine, Peking Union Medical College (PUMC)
5. University of Michigan Medical School, Ann Arbor, MI, 48109, USA)

[Abstract] Ontology has emerged to play important roles in human- and computer-interpretable data and knowledge standardization, sharing, integration, and advanced analysis. The Basic Formal Ontology (BFO) is a widely used top-level ontology that has been adopted by over 300 ontologies in the biological and biomedical domains and other areas. BFO has been adopted as the official top-level ontology of the Open Biological/Biomedical Ontology (OBO) Foundry as well. To support Chinese usage of BFO, we as members of the OntoChina have translated BFO to Chinese and shared it by a community effort. Key BFO terms and their Chinese translated are introduced.

[Keywords] Ontology; Top-level ontology; Basic Formal Ontology, Chinese translation, OntoChina, OBO Foundry

［**基金项目**］ 1. 国家重点研发计划资助（2019YFC1710400；2019YFC1710401）2. 中央级公益性科研院所基本科研业务费专项资金资助（ZZ13-YQ-126）

［**通讯作者**］ 杨啸林，女，副教授，中国医学科学院基础医学研究所副教授，研究方向：本体学与生物信息学 E-mail: yangxl@pumc.edu.cn

He Yongqun 何勇群，男，密西根大学医学院副教授，研究方向：本体学与生物信息学，E-mail: *yongqunh@med.umich.edu*

［**作者简介**］ *朱彦 副研究员，研究方向：中医药信息学，Tel：（010）64089623，E-mail：zhuyan166@126.com

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

1 引言

随着大数据时代的到来，海量数据管理和集成为知识发现带来了大量的机会。本体是用人和计算机都可以理解的术语（terms）及关系（relations）来描述实体（entity）以及实体之间的相互关系，提供对事物本质的统一认识，能有力支持多源异构数据的集成并最终实现知识发现。基因本体(Gene Ontology, GO)[1]的巨大成功，激励了越来越多的研究者和社区投身到生物医学领域的本体研究和开发中。OBO Foundry 是其中的典型代表[2]，成为目前生物与生物医学本体领域最有影响力的社区。中国的本体学者也开始借鉴其先进理念和成功模式，成立了本体中国（OntoChina）社区，致力于国内生物医学信息本体的研究与推广[3]。完成了 OBO Foundry 中的部分核心本体的本地化，其中就包括 OBO Foundry 中多数本体都采用的顶层本体基本形式化本体(Basic Formal Ontology, BFO)[4]。本文将系统介绍 BFO 的来源、特点以及中文版 BFO（BFO-zh）的翻译和推广工作，并简要梳理其层次结构，希望能为国内本体研究同行提供有益参考。

2 BFO 基本情况介绍

BFO 来源于 2002 年德国大众基金会(Volkswagen Foundation)资助的项目，最初由 Barry Smith 和 Pierre Grenon 发起，后由大量的专家参与持续改进更新，目前最新版本为 2019 年 8 月发布[5]。BFO 已提交国际标准化组织（ISO）的数据管理和交换技术委员会投票审核，于 2020 年 3 月份正式发布[6]。

顶层本体（Top ontology），也称上层本体（Upper ontology）或形式化本体（Formal ontology），只对最抽象、领域中立的的知识进行表征。。与其他顶层本体相比，BFO 有以下特点：

（1） 在生物医学领域取得了巨大成功

依托于 OBO Foundry 的发展，目前，以 BFO 作为顶层框架的本体已经超过 300 个[7]，尤其在生命医学领域取得了广泛的应用。

错误!未找到引用源。展示了部分重要的基于 BFO 开发的顶层领域本体和领域本体。

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

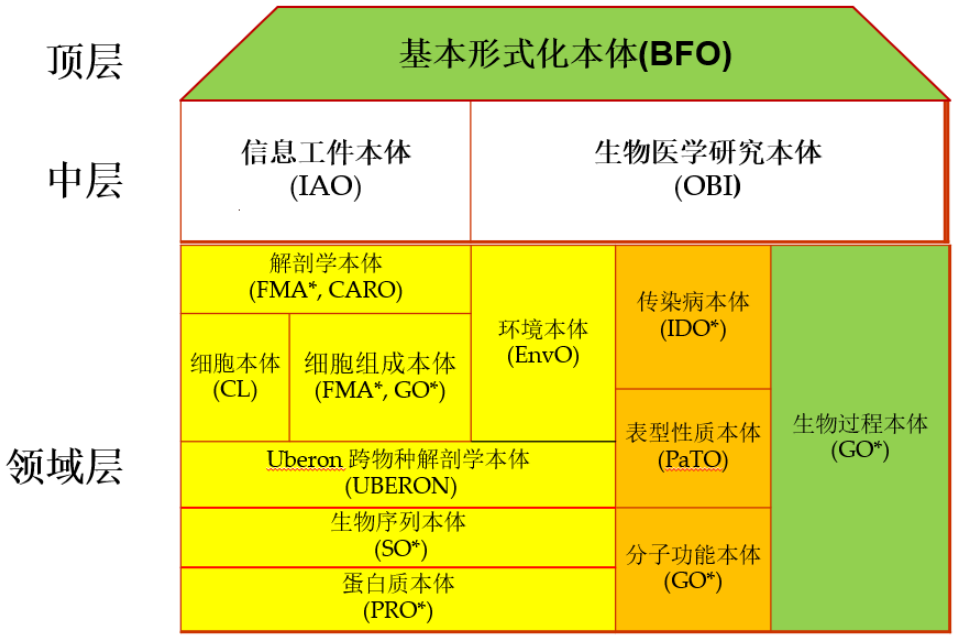

图 1：以 BFO 为顶层本体的中层本体及领域本体

中国的研究者也尝试基于 BFO 来构建领域本体。如传统中药本体(Traditional Chinese Drug Ontology，TCDO)[8]和中国植物物种多样性领域本体[9]。

（2） 体量小而精

BFO 最初构建时借鉴了部分 SUMO 和 DOLCE 的核心概念，但与之不同的是，BFO 的设计初衷是为支持集成、分析、整合科研数据而开发的顶层本体，所以设计得非常简明，方便不同的领域和不同的粒度级别科学家复用。

（3） 基于本体实在论的哲学基础

BFO 是以本体实在论（Ontological Realism）为哲学基础的，即认为本体所表征的是现实本身的普遍特征。这一点与现有很多学者，尤其是信息学研究领域一些专家所持的概念取向（concept orientation）不同，他们认为本体表征的是科学家关于现实的看法。在表述上最明显的区别是就是 BFO 及相关论著中，尽量避免了使用“概念（concept）”这一术语，取而代之的是用哲学意义的“实体（entity）”、“类（class）”和“共相（universal）”等来直接指称现实的普遍特征。

3 中文版 BFO 的翻译与推广

3.1 背景

为了推进国内生物医学本体的研究与应用，2017 年在国家人口与健康科学数据共享服务平台（现为国家人口与健康科学数据中心）的支持下， OntoChina 社区系统引入并翻译了 OBO Foundry 的一些重要本体，其中包括（Basic Formal Ontology，BFO）[10]，Ontology for Biomedical Investigations (OBI)[11]，Relation Ontology (RO)[12]，Ontology for General Medical Science (OGMS)，Human Phenotype Ontology （HPO）[13]， 和 Cell Line Ontology （CLO）[14]等本体，并开发了 MedPortal 本体资源库[15]来存储和共享。

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

受 Barry Smith 教授委托，OntoChina 于 2017 年底正式开始 BFO 的中文翻译工作，并同步翻译了 MIT 出版社出版的 Barry Smith 等撰写的 BFO 本体著作“Building Ontologies with Basic Formal Ontology”[10]，该书即将由人民卫生出版社在 2020 年出版[16]，将是第一本向国内读者系统介绍 BFO 及基于 BFO 构建本体的理论、方法和技术的书籍。

3.2 翻译的基本原则

根据 BFO 特点及工作组的实际情况，制定了翻译工作的几点基本原则，除了力求做到信、达、雅以外，要尽可能了解作者在原语言环境及学科背景下力图所达到的效果，然后试图在新的语言环境和文化背景下重建这种效果。同时多从中国传统文化，尤其是传统哲学中寻找灵感。

具体从术语选择来说，遵循大道至简的原则，因为 BFO 中本身的实体就是高度抽象和概括的，翻译成中文也务必简洁直白，尽量采用最常见的字和词。一方面尊重术语的哲学根源，同时也需要考虑具体用户主要是具体领域的专业人员背景，不能太艰涩难懂。对于平级关系的术语，尤其是一分为二的情形，尽量使用对偶形式的词；同时也要兼顾父子关系的术语，体现出从属关系。

3.3 翻译工作的流程

翻译工作分以下步骤进行：

（1） 人员筹备：针对预计面临的挑战，我们组织了来自不同背景的专业人员，以志愿者的形式加入到工作组中，并邀请了 Barry Smith 教授担任顾问。

（2） 平台搭建：由于参与者来自全球各地，我们使用了 WebProtege 作为协同编辑和讨论平台，并进行问题的跟踪。

（3） 初步翻译：工作组在花了近半年的时间，对 BFO 中的每一个术语逐个溯源、讨论和翻译。

（4） 征求意见：通过 OntoChina 和微信公众号等平台广泛发布公开征求意见，基于意见再进行讨论调整，加入相应的中文同义词。

（5） 定稿公开：定稿后将 BFO-zh 上传到 MedPortal 平台（http://medportal.bmicc.cn/ontologies/BFO-ZH）和 GitHub（https://github.com/zhuyan166/BFO-zh）公开共享。

4 BFO 的类层次结构介绍

本节内容将根据 BFO 中类的层次结构（如图 2）从上至下地简要介绍 BFO 的每一个术语。更多详细内容可以参考 BFO 网站和相关书籍[10]。

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

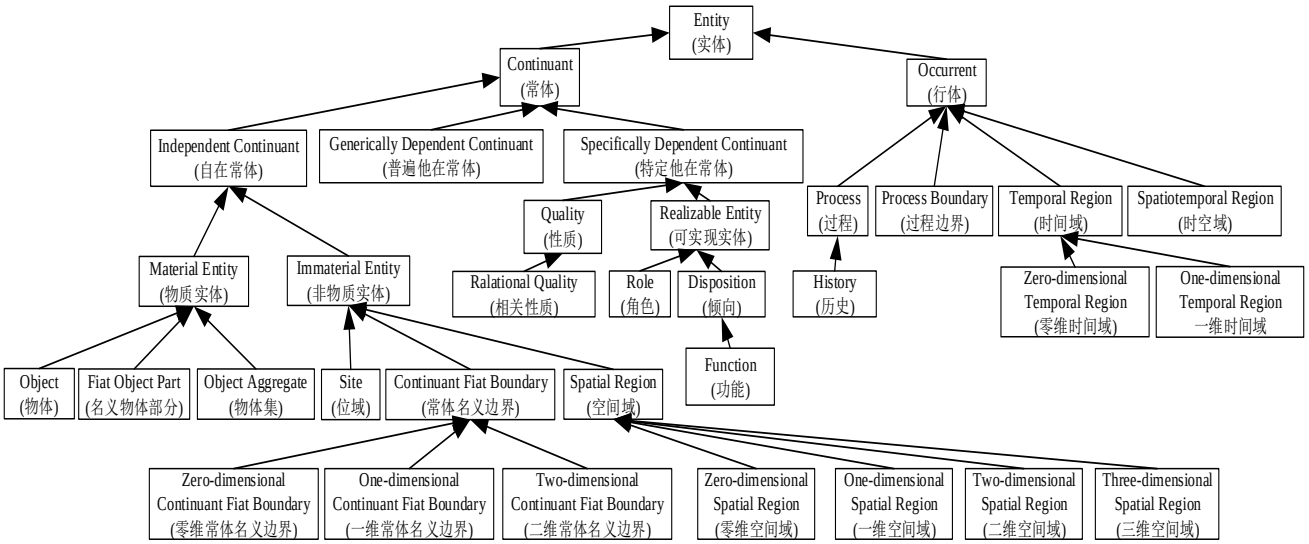

图 2：BFO 的层次结构图

4.1 实体及其基本类型

BFO 首先从现实的运行方式出发，认为现实是由实体（entity）组成的。哲学家和科学研究者常用“实体（或存在体）”一词，来指称以任何方式存在的任何事物。

基于不同的两种世界观，BFO 将实体分为为两种基本类型：

- 常体（continuant ）：历经时间而能持续或保持不变的实体。
- 行体（occurrent）：指发生或出现的实体。

两者最直观地区分就是，常体没有时间部分（temporal parts），而行体有。

我们在翻译时没有找到现成的参照，故采用了汉字中常见的、起源很早的并且是对偶的两个字“常”和“行”。灵感来自于儒家典籍《荀子·天论》“天**行**有**常**，不为尧存，不为桀亡。”和佛家典籍《宗镜录-卷第四十二》中的“诸**行**无**常**，是生灭法”。“常”字能很好地概括历经时间不变的特征；而“行”也能直观地体现四维时空中伸展穿行的时空蠕虫（spacetime-worms）的意象[17]。

4.2 常体（continuant）

常体的特征： (1) 能历经时间持续（persist）、持久（endure）或连续（continue）存在，并能保持其同一性，以及(2)没有时间部分。根据存在性是否依赖于其他事物以及依赖的特征，BFO 中常体又分为三个子类型：

4.2.1 自在常体（independent continuant）

自在常体，简单说就是不依赖于其他实体自己能独立存在的实体。它是作为性质（quality）的持有者（bearer）的一类常体。

（1） **物质实体**（material entity)

指有某一份物质作为其组成部分的自在常体。因此，它可以在三维空间中延展，并且可以持续存在一段（无论多短的）时间。进一步可以分为：

a) 物体（object）：现实存在的基本单元或颗粒。其特性包括：在三维空间中延展；自成一体(causally unified)；各部分最大的程度地自连接。例如原子、细胞、器官、

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

生物体、恒星等。

b) 物体集（object aggregate）：由物体的集合所组成的物质实体。例如一堆石头、你血液中的一团细菌、一群鹅、某家医院的所有患者。

c) 名义物体部分（fiat object part）：指某个更大物体的真子部分（proper part），但没有任何物理间断将其与该物体的其余部分分割开（因此它本身不是物体）。“名义(fiat)”一词表示边界的规范性并非物理间断，其存在性也不依赖于认知主体。例子包括人的上躯干，实心金属勺的手柄，地球的西半球等。

（2） 非物质实体(immaterial entity)

指不包含任何物质实体作为其部分的自在常体。包括：

a) 常体名义边界 （continuant fiat boundary）：是某种物质实体的边界，它精确地存在于物体与其周边的相遇的地方。例如地球表面，或细胞膜表面。BFO 识别了三种不同子类型，即二维、一维和零维常体名义边界。

b) 位域（site）：直观来说是由某物质实体所包含的孔，例子包括你的鼻腔、你的静脉（血液流过的空腔）、苏伊士运河（沟渠）、你的胃肠道的内腔等。

c) 空间域 （spatial region）：是空间的某部分，独立于物质实体而存在。BFO 识别了四种不同的空间域，即三维、二维、一维和零维空间域（又称为空间体、面、线和点）。

需要注意的是：常体名义边界和位域可随其宿主的移动或形状或大小的变化而改变其位置、形状和大小。而空间域相对于参照系永远不会发生移动。

4.2.2 特定他在常体（specifically dependent continuant）

特定他在常体是依赖于一个或多个特定的自在常体而存在的常体。特定他在常体又两种子类型：性质（quality）和可实现实体（realizable entity）。性质与可实现实体的不同之处在于，性质在其完全依附的实体中得到充分展现（exhibited）、表现（manifested）或实现（realized ）。而可实现实体可以依附而不被实现，或者不同的程度（包括不同的可能性程度）地实现。

（1） 性质（quality）。例子包括这个肾脏的质量、这一份血液的颜色、这只手的形状等。在本体构建过程中制定定义时，通常将物体及其部分的性质称为种差（differentiae）。

（2） 可实现实体（realizable entity）。例子包括医生的角色，金属导电的倾向、生殖器官的功能等。对应其更具体的类型：角色、倾向和功能。角色是一种外部决定性的（externally grounded）可实现实体，它存在是因为持有者处于特定的物理、社会或制度环境中，且这些环境对于持有者来说是非必需的。而倾向则不同，是一种内部决定性（ internally grounded）的，也就是说，如果一个实体在物理上以某种方式存在，那么它就会拥有某特定倾向，如果这种方式不

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

存在了，那么它就会失去这种倾向。功能是基于持有者的物理构造而存在的一种倾向，该特定物理构造是因为自然选择（在生物学实体的情况下）或者有意的设计（在人为的情况下）而产生的。

4.2.3 普遍他在常体（generically dependent continuant）

特定他在常体受到所谓的非迁移公理（axiom of non-migration）的约束：它们不能从一个持有者迁移到另一个持有者。然而，信息实体（如 pdf 文件）一方面依赖于某持有者，它必须保存到某种物理存储设备上；另一方面，数值上等同的信息实体又可以存在多个副本，即可以从一台电脑复制到另外一个电脑。为了处理这种特殊的客观存在，BFO 引入了普遍他在常体（generically dependent continuant）这个范畴，定义为：依赖于一个或多个自在常体作为其持有者的一种常体。

4.3 行体（occurrent）

行体：简单来说是指发生或出现的实体，如“事件”、“过程”或“发生的事情”等。

BFO 的观点是将整个时空视为四维的整体，过程在其中开始、持续和结束。每个过程可看作是一个在时间上延展的连续体，像时空蠕虫一样在一个统一的容器（即整个时空）中伸展。

行体的子类型包括：过程（process）、过程边界（process boundary）、时间域（temporal region）和时空域（spatio-temporal region）。

4.3.1 过程（process）

是因为在时间上发生或出现而存在的一种行体，它具有时间部分，并依赖于一个或多个物质实体。例子包括：这个有机体的生命过程、这个疾病的病程、那只鸟的那次飞翔、这个细胞分裂过程等。其中历史(history)是过程的一个重要的子类型，是物质实体或位域所占据的时空域内发生的所有过程的总和。

4.3.2 过程边界（process boundary）

指过程的瞬时时间边界。过程边界是其所属过程的开始和结束。例子包括突触的形成、快速的眼球运动睡眠的开始等。

4.3.3 时间域（temporal region）

行体的子类型，是（整个）时间的一部分。时间域与时空域的不同之处是它只在时间维度上延展或用作时间边界。时间域是将时空域投影到该时间维度上的结果。又分为：

- 零维时间域（zero-dimensional temporal region）：也称为时刻（或瞬间）——是没有长度的时间域。
- 一维时间域（one-dimensional temporal region）：也称为时间间隔——是在时间上延展的时间域。

4.3.4 时空域（spatio-temporal region）

指行体之所在的时间和空间范围。时空域是时空的一部分(即整个时空的一部分)，并且每个时空域都是相对于某四维坐标系统参照系所定义的。例子包括某个人生命所占据的时空

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

域、一个癌症肿瘤的发展所占据的时空域、一个细胞减数分裂的过程所占据的时空域、一场战争所占据的时空域等。

5 总结与展望

在本体理论研究及技术应用上，我国与国际先进水平尚有不小的差距。可喜的是，近几年来，国内越来越多的研究者开始关注和应用先进的本体构建理念。本文介绍的 BFO 的本地化工作正是本体理论在中国的推广和应用的一个基础性的探索。

未来我们也将遵循可误论（Fallibilism）原则，随着 BFO 的更新，BFO-zh 将随之更新修正；同时，在使用过程中广泛听取中文用户意见建议，修正中文翻译的共识。我们还将致力于将 BFO 应用到实际的领域本体开发中，争取为整个生物医学本体事业添砖加瓦，做出贡献。

致谢

非常感谢 Barry Smith 教授、王其冰老师在 BFO 翻译中提供的殷切指导。也感谢关健、周伟、潘虹洁、余红、万灵、刘萌等老师提出的宝贵意见。

bmr.202007.00009V2

http://www.biomedrxiv.org.cn/article/doi/bmr.202007.00009
此预印本（未经同行评审）的作者拥有该文稿的版权，biomedRxiv拥有永久保存权。任何人未经允许不得重复使用。

参考文献

[1] Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium [J]. Nature genetics, 2000, 25(1): 25-29.

[2] Smith B, Ashburner M, Rosse C, et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration [J]. Nature Biotechnology, 2007, 25(11): 1251-1255.

[3] Pan H, Zhu Y, Yang S, et al. Biomedical ontologies and their development, management, and applications in and beyond China [J]. Journal of Bio-X Research, 2019, Latest Articles.

[4] Smith B, Grenon P. Basic Formal Ontology (BFO) Home 2019-3 2007. http://basic-formal-ontology.org/.

[5] OBO library - Basic Formal Ontology. http://purl.obolibrary.org/obo/bfo/2019-08-26/bfo.owl.

[6] Information technology — Top-level ontologies (TLO) — Part 2: Basic Formal Ontology (BFO) 2020. https://www.iso.org/standard/74572.html?browse=tc.

[7] BFO Users. http://basic-formal-ontology.org/users.html.

[8] Zhu Y, Liu L, Gao B, et al. TCDO: a community-based Traditional Chinese Drug Ontology; proceedings of the The 11th International Biocuration Conference (Biocuration-2018), Shanghai F,2018 [C].

[9] 段宇锋, 黄思思. 基于 BFO 构建中文植物物种多样性领域本体的研究 [J]. 现代图书情报技术, 2015, 265(12): 80-87.

[10] Arp R, Smith B, Spear AD. Building Ontologies with Basic Formal Ontology [M]. MIT Press: Cambridge, MA, USA, 2015.

[11] Bandrowski A, Brinkman R, Brochhausen M, et al. The Ontology for Biomedical Investigations [J]. Plos One, 2016, 11(4).

[12] Smith B, Ceusters W, Klagges B, et al. Relations in biomedical ontologies [J]. Genome biology, 2005, 6(5): R46.

[13] Groza T, Kohler S, Moldenhauer D, et al. The Human Phenotype Ontology: Semantic Unification of Common and Rare Disease [J]. American journal of human genetics, 2015, 97(1): 111-124.

[14] Sarntivijai S, Lin Y, Xiang Z, et al. CLO: The Cell Line Ontology [J]. Journal of biomedical semantics, 2014, 5: 37.

[15] 郭靖文, 杨晟, 史涪仁, et al. MedPortal: 面向精准医学的生物医学本体资源存储和应用平台 [J]. 中国生物医学工程学报, 2017, 05: 557-564.

[16] 朱彦, 何勇群, 郑捷, et al. 基于基本形式化本体的本体构建（待出版） [M]. 北京: 人民卫生出版社, 2020.

[17] Smith, Barry. CLASSIFYING PROCESSES: AN ESSAY IN APPLIED ONTOLOGY [J]. Ratio, 2012, 25(4): 463-488.

bmr.202007.00009V2