

信息时代的智慧——兼谈科普

心蛛

2013-02-22

目录

1 从神话到科学：世界观的变迁	1
2 个人的世界与真实的世界——浅谈人的认知	2
3 什么是知识？	4
4 知识的载体	6
5 信息时代的数据与知识——浅谈“知识图谱”	7
6 信息时代的智慧：构建你自己的“知识图谱”	10
7 浅论信息时代的教育与科普	12
8 后记	15

有幸于 2012 年 7 月 14 日参加了 科学公园 组织的第一届“科学与无神论者聚会”，讨论的主题是“无神的未来”。与会中认识了一些朋友，听了几个讲座，有颇多感受。当时，我正在思考一个问题：如何组织自己的知识，使之使用更合理高效？近来得闲，把一些个人认识及与会中关于科普的思考写出来，希望能够抛砖引玉，也算是聚会的个人总结报告。

1 从神话到科学：世界观的变迁

自古以来，无数先贤凡人都问过同样的问题：“我是谁？生命是什么？我们要到哪里去？”仰望星空并开始审视世界，发出类似这样的疑问，这表明了自我意识的一种觉醒。意识的第一个标志，就是将自己与周围的世界区别开来，就如儿童第一次意识到：他是他，爸爸是爸爸，妈妈是妈妈。意识独立之后，人类就开始了对外部环境的探索，进而开始思考更深的问题，比如前面那些，而这些思考的结果则不尽相同。

在古罗马，人们认为混沌神造出了日与夜俩兄妹，他们结合后生出了十二泰坦，然后泰坦间的彼此结合又生出了天庭诸神，之后就是天神创造了人界。在中国汉族，我们有盘古开天、女娲造人，其他各民族也都有类似的兄妹结合并成为民族起源的传说。这些神话无论多么荒谬，数千年

来，人们仍口口相传并信以为真。人们将自己的日常生活与这些神话结合起来，并于其上建立了诸多的习俗。然后又不断加入新的想象，使得那个神话的世界也越来越完整，越来越像人类的世界，比如汉族的玉皇大帝就活脱脱是人间皇帝的翻版。

以神话为基础的对世界的认知，最基本的思维模式就是模仿与想象，这种认知的进一步发展就是宗教以及类似于宗教的各种学说，比如三大宗教、中国的阴阳五行与“周易”玄学。这类认知的特点是以与现实世界的相似性为基础的，而不是真实性，它们也并不追求真实性，结果就与真实世界相去甚远。这些认识对于简单事物的描述则没什么问题，但对稍微复杂一些的，则往往荒谬不堪，对于那些行不通的说法，则变成了禁忌。这种认知模式最常见的思考方法有猜测、顿悟、感受等，然后就把这样得来的结果当成真理，顶多在其中再多加一点似是而非的逻辑。比如以阴阳五行为理论基础的中国传统医学，在其中就有许多食物相克之类的禁忌；在旧医入药上，因为人参长得奇怪，有点像人形，就变成了大补之物，非常奇怪。

好在人类的思维模式不止是模仿与想象，还有归纳、综合、逻辑推导，等等。自古希腊始，在实证与逻辑的基础上，人类发现了科学，人类的认知自此不同。

在神话、宗教及一些宗教似的学说里，关于对世界的认知，你只需要相信结果，而不管它是否与周围的世界相符。如果有牧师说是上帝安排犹太出卖了耶稣，你想质疑“那么，就是上帝摆了犹太一道喽？”这样则是万万不可的。但在科学中，你可以质疑一切，只要你有证据与逻辑，即可反驳一切现有的认识。而在科学中，通过这样不断地质疑与实证，我们对世界的认识也越来越接近真实。

什么是世界观？就是人对周围世界的认识。这个“人”可以是个体，也可以是一群人，乃至整个人类。这样，就有了两种类型的世界观：个体的及群体的。群体的世界观则指的是某个群体对周围世界的共识，这儿只考虑数量上占优的那些共识。

脱胎于神话与想象的宗教世界观是人类认知的初级阶段，因为它们只利用了人类一部分思维能力，即猜测、想象与类比，玄学与哲学比它们又多加了一点点逻辑。因此，宗教的、玄学的及哲学的世界观，彼此是类似的，它们都缺乏实证，不妨把这些统称为“宗教世界观”。历经实证与逻辑检验的科学世界观则是人类智慧的最优美的结晶。就整个人类而言，全球七十亿人中大部分仍然被形形色色的宗教控制着，宗教势力仍然占据主流，因此，目前人类的世界观并非统一的，大致上分成两类：宗教的（包括玄学的、哲学的）与科学的。

这个世界最奇妙的就是，人类可以发现科学并且借助其正确地认知世界。

2 个人的世界与真实的世界——浅谈人的认知

现实生活中，道德是道美丽的篱笆，而真正起到防护作用的是那堵法律的围墙，但无论道德还是法律，都建立在人们对世界的共识之上，或者说，建立在人们的共同世界观之上。世界观的建立过程也就是人类的认知过程。而人类的认知是一个科学的专业领域，其研究内容与方法都很复杂。人工智能 (AI) 是认知科学与计算机科学的交叉领域，它希望通过计算机来模拟人类的智能，这个要建立在人们对自己的智能有充分认识的基础之上，其中涉及到的内容非常广泛，如基于神经元的生理构造、神经元的工作模式、思维本身的特点及方式，等等。人们曾经对认知科学持非常乐观的态度，早在 1958 年有人工智能学者认为“十年之内，数字计算机将成为国际象棋世界冠军”。而这一目标直到 1997 年 5 月 11 日才实现，比预期推迟了约三十年。尽管认知科学非常复杂，进展没有预期的神速，但奇妙的是，人类可以认知到自己的思维，不是吗？

这儿仅就人的认知做一点粗浅的展开介绍。

人的认知过程就是一个信息处理过程，这其中记忆尤为关键。现代的信息处理理论将人的记忆分成三类：感官记忆、短期记忆和长期记忆，而长期记忆则还可进一步划分为显性的与隐性的，等等。借用一幅来自网站 The Human Memory 的图片（我加入了汉语译文）如下：

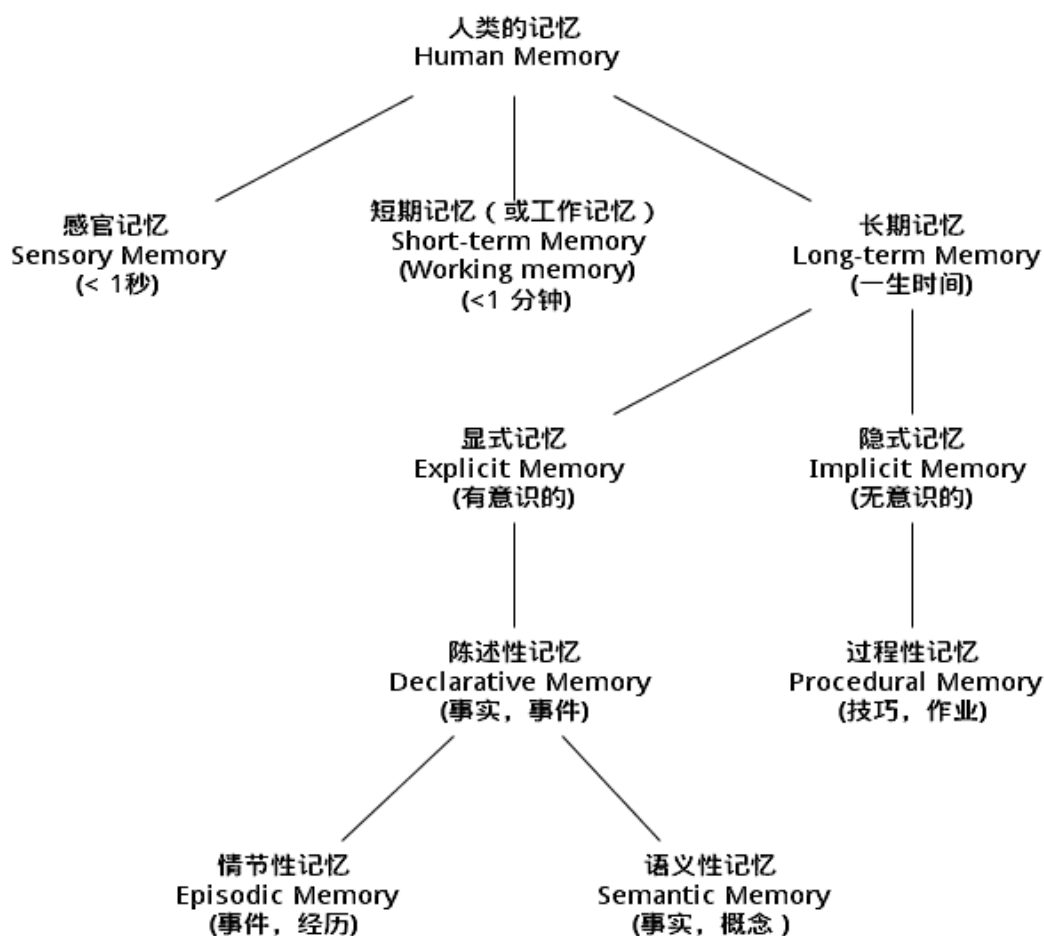


图 1: 记忆的类型

个人对世界的认知过程，就是通过自身的感官构造出一个世界的记忆副本的过程。每个人所感觉到的、接受到的、构造成的不尽相同，则个体的认知世界也不一样。因此，可以说，每个人的大脑有一个独立的认知世界，不妨把这个独立的个体认知世界称为“意识世界” (mental world)。

个体的意识世界存在极大的差异，但相互之间存在重叠，这种重叠就是大家的“共识”。如果了解一点数学中的集合论，你不妨将共识想象成许许多多的个体意识世界的交集。共识反映在记忆上通常是“语义性的记忆”，最基本的就是概念与重大的史实。

科学的认知过程则是许多人采用相同的实证与逻辑的方法与世界反复交互的过程，参见科学公园中的 证据、理论与真理 一文。

在前一节中谈到两种类型的世界观，宗教的与科学的。基于神话、猜想或胡乱的类比所构造的意识世界，包括那些基于玄学或某些哲学假说所构造的意识世界，这些都是宗教的精神世界，亦即宗教世界观。而基于科学得来的共识就是科学世界观，这也就是人类通过科学所获得的在头脑中的世界映象，这个映象也是人类各种各样的意识世界中与真实最为接近的。

3 什么是知识？

什么是知识？汉语中收录词汇最全当属《辞海》，它对“知识”给出的解释如下：

人类的认识成果。来自社会实践。其初级形态是经验知识，高级形态是系统科学理论。按其获得方式可区分为直接知识和间接知识。按其内容可分为自然科学知识、社会科学知识和思维科学知识。哲学知识是关于自然、社会和思维知识的概括和总结。知识的总体在社会实践的世代延续中不断积累和发展。有关学术文化的：知识界 | 知识分子。相识；朋友：朝廷大臣多有知识的。

上述解释中的第一条与本文所谈的内容最为接近，但过于简略。实际上，我们很难给知识下一个准确的定义。维基百科的英文词条 Knowledge 及中文词条 知识 中也只给出了描述性解释，大致如下：

知识是对个体或事物的认识，它包括事实、信息、描述，或者通过经验与教育获得的技巧。

不过，如果利用对计算机的认识，再结合一些相关的概念一起类比地理解，可能会稍微清楚些。这儿引用一下 PpcSoft 公司在 个人知识管理 的介绍中对“知识”的解释：

- 数据：未经组织与处理的事实。
- 信息：处理后的数据，包含有特定的含义（什么，谁，何时，何地）。
- 知识：人们通过学习与思考而获得的对信息的理解（怎么样）。
- 智慧：历经实践考验的知识（为什么）。

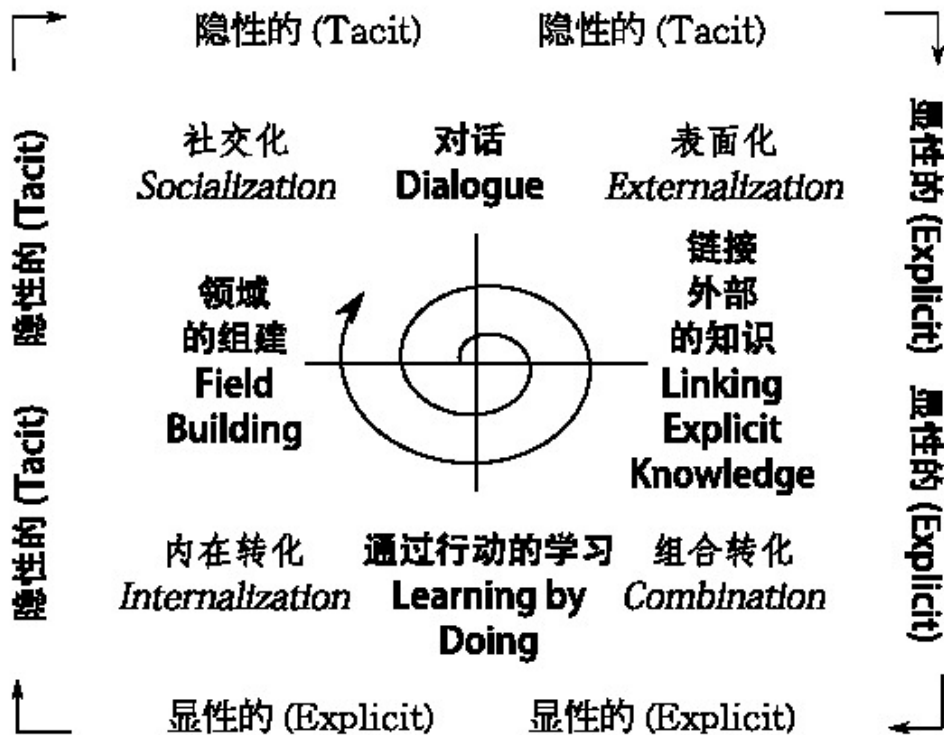
同“科学”的概念一样，尽管“知识”的定义很难下，但可以罗列出它的一些特性。Sveiby¹给出的一些特性如下：

- 知识是源源不断的——你永远有用不完的原始材料；
- 知识通过分享可以壮大（经常传授知识的人通过这个过程会变得更加有知识）；
- 交流与人际关系在知识获取过程中很关键。

Nonaka 和 Takeuchi²认为有两种类型的知识：隐式的（主观的）和显式的（客观的）。隐式的知识 (tacit knowledge) 建立在经验基础之上。它包括洞察力、直觉、预感等，这些都不太容易看见与表达。这些都非常个人化的，没什么特定形式，也不容易分享。显式的知识 (explicit knowledge) 则是规范化与系统化的，可以用词汇与数字轻松地表达出来。它是理性的知识，可以很容易通过数据、公式、固定的程序或通用的原则来交流与分享。下图是由这二位所提出来的“知识的螺旋”。

¹Karl Erik Sveiby, The New organizational wealth: managing & measuring knowledge-based assets, San Francisco: Berrett-Koehler Publishers, Inc., 1997, p.29

²Nonaka and Takeuchi, The Knowledge-Creating Company, 1995, p.71



In Nonaka and Takeuchi, *The Knowledge-Creating Company* 1995, page71

图 2: Nonaka 和 Takeuchi 所定义的“知识的螺旋”

人们认知活动的目的是获得“知识”与“智慧”。通过上述的解释，人类认知的过程，实际上就是不断地处理信息的过程。处理信息，实际上就是不停地把接触到的事实贴上 5W³ 的标签：是什么 (What)，谁 (Who)，什么时候 (When)，哪儿 (Where)，怎样做的 (How)。我们想得到的是最后一个 W，就是 Why，也就是“智慧”。

所以，当孩子不停地问你“为什么”时，千万不要烦，帮助他们认识世界吧！

正如 Sveiby 所罗列的那样，知识的原料是无穷无尽的。事实上，你每天接触的任何一样东西、任何一个人人都可以挖掘出无尽的知识出来。比如影响极大的韩寒代笔案件，得出他有代笔这个结论需要很丰富的知识，只不过，这些知识对于方舟子等人是直觉反应，对于一些韩粉，则可能由于认知缺陷而永远也无法理解。诸多网民参与的分析求证过程中，挖掘出越来越多素材，也涉及了越来越多的知识，如文本分析、语音分析、笔迹分析、经验逻辑分析，等等，由此可见知识的无穷尽。当然，韩仁均等人将一个有学习障碍的人推成“公知”、“天才作家”，这也需要丰富的知识。不过，绝大部分知识对于个人或群体则未必有用，或者在当下未必有用。这就引出了一个问题：我们需要

³4W 很好理解，就是：What, Who, When, Where；而 How 可以取首字母 H，也可以取尾字母 W，因此有 5W 或 4W1H 的说法。相应地，把 Why 加进来，则有 5W1H 或 6W 的说法。但用 Five Ws 的说法多一些。5W 恰当地反映了人类对世界的认识过程。比如拉雅德·吉普林的小诗：

I keep six honest serving-men:
 (They taught me all I knew);
 Their names are What and Why and When
 And How and Where and Who.

什么样的知识？

当然，我们需要有用的知识。屠龙之技华丽，修仙之法艰深，人鬼交流之道神秘，这些也是知识，只不过，它们象把韩寒推为“天才”的知识一样，只有在存心行骗时才有用，而这类用途对群体的发展是有害的。我们需要的有用的知识是指对绝大部分以及全人类都有益的那些。而对全人类有用的知识也只有科学知识了。科学的有用性显尔易见，即便是与科学格格不入的教皇也需要乘坐安装有防弹玻璃的高级汽车（因为它比上帝更实用），街头的算命师甚至也搞起了电脑算命。

由于知识的范围太广，本文默认所指的“知识”只限于科学知识。

4 知识的载体

语言是知识的天然载体，也是人类共同意识世界的载体。实际上，一个人的语言能力往往也反映了他的认知能力。人类语言的发展过程，也是人类认知逐渐丰富的过程。原始部落及发展落后的人群，词汇往往很简单，比如东巴文中“万”以上的数字是没有的，而“万”就表示了“无穷”的意思。方舟子写过一篇很有意思的科普文章《从鹦鹉学舌和猩猩学电脑看智力开发》，其中谈到，即便是动物在掌握了少许语言之后，也能够表现出一些逻辑思维能力。

语言承载知识的最直接方式就是概念。概念通常表现为新名词或新词汇，或者是旧词新解。当然，并非所有的新词都有用，这与并非所有的知识都有用一样。不过，一种语言的词汇量还是可以较直观地反映出它承载的知识量。比如，《汉语大词典》是中国大型的汉语词典，最新的 2.0 光盘版共收入汉字 20902 个，复词 343307 条，成语 23649 条，释义 515524 项。一直引为权威的牛津英语词典在线版目前收录单词超过六十万个，这个可与汉语词典中的三十四万多个复词相比较，可见，英文中的概念还是比汉语相对多一些。当然，这样的比较非常粗糙。

语言的演化，最直接的表现就是词汇量越来越大，语法也越来越规范。词汇量的扩大表明了人类认知范围的扩大，而语法越来越规范是因为人们需要语言工具本身更加方便高效。比如德语从 1876 年至今就进行过多次正字运动；百年以来，汉语的白话运动、罗马化及简化运动也持续进行着；其它流行的语言也都经历过类似的正字或规范化运动。为了促进全人类都能够高效地进行交流，从早期的沃拉普克语 (Volapük)、世界语 (Esperanto)，再到后来的伊多语 (Ido)、国际语 (Interlingua)，以及 2002 年发布的大同语 (Mondlingo) 等，有几十种较为流行的人工设计语⁴正在努力成为国际辅助语，国际语运动方兴未艾。为了研究需要，其它人造语言也频频现身，并且广为传播，比如大家熟知的《指环王》中的精灵语、《星际迷航》中的克林岗语等。为了证明萨丕尔-沃夫假说，詹姆士·库克·布朗还依照谓词逻辑设计出了逻辑语⁵，它的目标是让计算机能够直接听懂人的话。至于未来哪一种语言会成为人类的普通话，就由这些语言去相互竞争吧。但我相信，就象民族语言的正字运动一样，人类需要一种语言来承担共同交流的任务，因此，人类需要一种规范的、高效的国际辅助语或通用语，不妨把通用语的形成与确定当成全人类的“正字运动”。

不同的语言会影响人们的思维方式吗？同样的知识用不同语言表示会完全一样吗？会不会有一种超级语言，简单、强大、承载全部的人类知识，并且能够适用于所有的人？这些问题很有意思，

⁴人工设计语言，英文中叫 Constructed Language，简称为 conlang，是指其发音体系、语法、词汇等均由个人或群体从头开始设计的语言。许多人工设计语言的最初目标是成为国际辅助语 (International Auxiliary Language，简称为 IAL，或 auxlang)，但也包括一些研究性的语言。国际辅助语的目标不是取代人类现有的民族语言，而是成为不同民族、不同文化之间交流的桥梁，它与民族语言之间的关系就象汉语的普通话与方言的关系一样。

⁵逻辑语最初于 1955 年发表在《科学美国人》上，当时叫 Logban。自 1987 年始，由逻辑语言组织 (Logical Language Group, LLG) 接管其开发，其名称规范化为 lojban [lo ban]。其特点是：言文一致、形意一致、文化中立。它的语法源自谓词逻辑，它没有标点符号，其文本甚至可以直接用 YACC 类似的程序进行解析。逻辑语的全部语法都在 The Lojban Reference Grammar 这本书中，该书用 DocBook 格式开发并托管在 GitHub 上。逻辑语的主页参见 - <http://www.lojban.org/tiki/Lojban>

不是吗？不妨把这些问题留给语言学家与科学家吧。

语言承载知识，除了口头交流、演讲外，最规范的形式就是正式文本与图书。并且，除了可以发声能够口头交流的语言外，还有其它形式的语言，如身势语、旗语、信号灯、盲文、哑语等等，不妨将它们称为广义的语言，这些都是重要的辅助。

除了语言之外，人类的知识还以其它的形式承载，比如艺术品、建筑、设计图纸、工业产品等。随着计算机技术的发展，知识的载体进一步扩大，影像、声音能够得以长期保存，这又使得知识的传承形式更加丰富，数量也呈现出爆炸式增加。而计算机程序本身不仅是知识的载体，也变成了一种新的知识。

为减少讨论内容，本文将知识的载体局限在图书、正式的文本及电子化的资源上。

5 信息时代的数据与知识——浅谈“知识图谱”

如前所述，知识是无限的。而人类的记忆却是有限的。譬如，你的英文词汇量达到三万已经非常了不起了，但《牛津英语词典》中有六十多万呢。当科学的发展程度越来越高时，所获得的有用知识也越来越多，即便抛开宗教的、习俗的知识，只记忆那些有用的，你的大脑也不够用。而且，尽管大部分知识现在你可能用不上，但你无法知道未来的某个时候是否会用到。你不可能学习全部的科学知识，又不知道哪些会在什么时候用到，怎么办呢？

四万年前的澳大利亚土著会用歌声来记忆地图，一旦记忆出错，地图也就错了，把原来有蜂巢的危险树洞当成安全的，这就很糟糕了。公元前 2300 年巴比伦人用泥板绘制广场地图，这就安全和实用多了。如果不是托尔·海尔达尔，波利尼西亚群岛上的康·提基神也不会被留意，就没有文字的记录，也不会有后人知道那一千多年前群筏过重洋的壮举。你看，人们的记忆是有限的、易出错的，而只有借助了外在的工具才能够让思考完整。并且，随着时代的发展，绝大部分人必须借助外部工具才能进行高效的思考。

我们需要的思考工具，除了词典，还有各种数据库、图表，等等。听朋友说，当初电报很流行时，他的一位亲属作为发报员硬是把几千个汉字的数字编码背了下来，可以迅速地在摩尔斯电码与汉字之间形成转换。这样的技能作为职业可能需要，但绝大部分人一生都没必要去记一行编码的，顶多在高考填卡时去查一下自己的姓名所对应的数字。若非职业需要，记忆编码表这样的任务，不仅低效，而且极易出错；但借助外物查询则容易得多了。

信息时代的特点是：信息的获取与分享，随时随地，方便、高效。越来越多的信息都开始电子化，各种各样的信息分散在世界各地各式各样的设备中，如智能手机、平板电脑、磁盘阵列、各类服务器、笔记本、车载智能终端、电视机顶盒、网络存储设备、GPS，等等。网络连接变成了信息的高速公路，信息的发布变得无比快捷，信息检索甚至比你从记忆中搜索还要方便、详实与准确。我们也越来越习惯于借助这些信息终端。信息高速公路也越来越变成一项公共事业，互联网公司的关键是服务，搜索引擎等公司则相当于公路的服务站一样普遍。

这儿借用一下传统微型计算机的架构做个类比：如果把大脑当作中央处理器，可以把感官记忆与短期记忆当作中央处理器的一级缓存，长期记忆与潜意识当成二级与三级缓存，则你手头可用的词典、笔记、包括电脑内的资料就相当于内存，位于图书馆、档案室、书店、互联网或其他人的私人藏书中的资料则相当于外部存储了，不妨把这些大脑之外的介质统称为人脑的“外存”。

那么，我们看看互联网“外存”中有哪些内容可用：

- 物理学基础知识，推荐 HyperPhysics. 如果你对核物理与粒子物理感兴趣，推荐在线的 核素

图及 Particle Data Group (PDG)。

- 化学的基础概念，推荐 IUPAC 的“金书” – <http://goldbook.iupac.org/>
- 数学，推荐 Wolfram 的 MathWorld – <http://mathworld.wolfram.com/>
- 实际上，每个专业都有自己的专业网站可用，比如生物医学的 PubMed. 相关领域的人士自行搜索即可。
- IMSLP, Petrucci 音乐图书馆, 目前提供了二十多万份乐谱可供下载。
- 对于一般性了解的，则推荐一些百科类的网站，比如 维基百科, HowStuffWorks, British Encyclopaedia, WikiHow, 世界数字图书馆, 等等。
- 语言资料更不必说了。汉语推荐 汉典、中国哲学书电子化计划 及 汉字字源网站。英语推荐 Merriam-Webster 及普林斯顿大学的 WordNet.

只要有时间，通过 Google 你可以找到许许多多类似这样的资料性网站。如果把这些内容都保存下来，其数据量会超乎你的想象，并且这些内容每天都在增加着。有许多有意思的研究也在产生越来越多的数据，比如已经完结的人类基因组计划 (ENCODE), 比如奥巴马所批准的投资 30 亿美元的人脑地图计划, 等等。这些科研不仅会产生海量数据，实际上更酷的是，任何人均可通过互联网分享这些成果。整个互联网都是你大脑的外存，你不需要专门记忆，只需在用到时简单搜索即可，很欣喜吗？你是否感觉到自己抱着一个宝藏呢？是的，这也是全人类的宝藏。这样说来，互联网隔绝就相当于把人的记忆硬生生割掉一部分一样，是很不人道的。

然而，新的问题来了，面对这些近乎无穷尽的资料，你有没有感到不知所措？甚至，你都不知道自己的外存有多大。很多时候，这些数据库并不能直接被你利用，你仍然需要各种帮助才能够找到对自己有用的东西，因为你可能不会使用，或者对检索到的信息无法理解。拥有宝藏却无法使用，这是为什么呢？

信息不等于知识。在“什么是知识”一节中把几个概念一块介绍了，就是数据、信息、知识与智慧。数据转化成信息需要附加一些 5W³ 的内容，信息转换成知识则需要进一步整理出相互的联系，而最终把知识形成一整套的体系结构，找出其中的因果，那就是智慧了。无论信息还是知识，都是一种特殊类型的数据。我们的目标是知识与智慧。

保存在大脑外存中的那些数据与信息，可能是仔细整理过的，比如《辞海》和《大英百科全书》；更多的则可能仅仅是信息的罗列，比如一个城市的电话号码簿。仔细整理后的信息，已经是知识了，但在你理解之前，那还只是某个群体的共识，是他们的意识世界的交集。只有理解了这些知识，在你自己的意识世界中形成了一个映像，那才变成了你的。

这样看来，互联网这个大脑“外存”拥有三种类型的数据：

1. 未经分析与整理的原始数据；
2. 整理后的信息；
3. 以及整理后的知识。

尽管互联网数据量非常之大，但这个大脑外存中的知识量却相对较小。即便如此，仍然可观，比如前面所罗列的那些网址资料，这些知识个人仍然不可能完全掌握，知识量仍然非常庞大。

那么，如何高效地利用这些数量庞大的知识呢？

事实上，有许多人正在朝这个方向努力着，并于二十世纪九十年代中期开始形成一个概念——知识管理 (Knowledge Management)。不过，这个概念最初局限于商业领域。一般企业的知识管理是从认知科学的基本思想出发，比如“知识螺旋”、记忆结构等，开发出一系列技术与工具，使得企业的知识获取、记忆、组织与积累更加合理高效。而本文则希望能够将其中的基本思想推广到全人类所有的知识当中。

知识虽然量非常庞大，但其本身是有规律的。知识管理要处理的主要内容包括：

- 知识的 **结构**，这是相对静态的；
- 知识的 **获取与演化**，参见前面知识的螺旋，这是一个动态的过程；
- 知识的存储、检索、表示与分享等。

知识的获取与演化是需要人参与的一个动态过程，而知识的结构、存储、检索、展示、分享等则是对这个过程的重要帮助，这些是可以寄存在大脑“外存”之中的。知识管理的最基础内容就是如何去组织知识，而知识的组织又与后续的存储、检索和展示形式密不可分，通常把知识展示与组织当成一个整体来讨论，这就是知识表示 (Knowledge Representation, KR)。

知识表示是认知科学与人工智能两个领域的交叉点，通常借助本体语言 (Ontology Language) 对知识点进行描述，从而构造出庞大的知识库。这儿的本体 (Ontology) 不同于哲学中的“本体” (或“实在”)，是指信息科学中对要共享的概念体系一种形式化的、明确而又详细的说明。信息论中的“本体”实质上就是概念之间的关联体系。下图取自维基百科，这是一个简单的使用本体表示的示例：http://upload.wikimedia.org/wikipedia/commons/1/17/Semantic_Net-zh_cn.svg

知识表示的底层技术相对复杂，但展现在人们面前的通常是相对简单的关系图，比如 TheBrain 的脑图，FreeMind 或 Xmind 中的层级结构图，Visual Understanding Environment (VUE) 及 CmapTools 所展示的概念图，等等。不管具体的形式有多大差异，这些工具都表明这个领域正在蓬勃发展之中。

知识表示的努力结果之一就是各种各样的知识库，与传统的百科全书、维基百科类网站均不相同，它们更专注于知识点 (对应于传统的词条) 之间的联系。大脑的互联网外存中有许多知识库，其中比较知名的有：

- FreeBase (FB), 于 2007 年建站。FB 的基本单位为“条目” (entity), 可以是人、地点或任何事物，而 FB 则专注于这些条目之间的联接。FB 旨在提供对人类全部知识的图示。它采用专门的查询语言 MQL, 你可以通过公开的 API 构造自己的应用程序，从而只专注于某一个特定领域的知识，例如，你可以利用 FreeBase API 构建一个类似于 IMDB 的网站。
- Knowledge Atlas (KA), 一个基于网页的学习工具，提供知识点之间的关系，有多种展示形式，包括树状图与叙事图 (Storyline). 其中很有意思的是它的叙事图，你可以将一些关联的概念按时间线组织成一篇研究文章、一篇博客、甚至一本书。
- Exploring Data, 这个网站主要是展示一些数据挖掘的技术，其展示效果令人印象深刻，比如这个人类疾病网络图 – Human Disease Network. 该网站的程序源代码都托管在 GitHub 上，感兴趣的人可自行研究。

知识管理是在知识表示的基础上，综合知识存储、检索、分享等技术，对知识获取与演化过程提供重要辅助，以达到获取有用知识与智慧的目的。有人提出的“数据挖掘”概念，其实也属于知识管理的一部分。

网站 KA 提出一个知识管理工具的概念，即 Knowledge Atlas，我翻译成“知识图谱”，这是他们对自己知识管理方法的一个统称。特定领域的知识，概念之间的关系很清晰，不妨将这些概念及它们之间的联系所构成的一张概念之网称之为这个领域的知识地图。例如物理学，HyperPhysics 给出了一个很好的尝试，不妨将之看成物理学这个领域的知识地图。当多个领域的知识累积起来时，会形成一系列的知识地图，而这些知识地图之间又存在紧密的联系：比如元素周期表不仅关联着化学，也关联着物理学；医学更是与分子生物学、病毒学、化学等等领域交叉频繁、密不可分。那么，不妨将网站 KA 的 **知识图谱** 概念从工具本身做一个推广，将之视为全人类知识结构的一个综合体，从而能够将特定领域的知识地图、各个领域知识地图之间的联系都包含在内。**知识图谱，可以视为人类全部认识的骨架**。为辅助理解，你不妨简单将之视为在词典与百科全书中添加上条目（或概念）之间的关系结构。

这样，在信息时代，我们大脑的互联网外存就形成了这样的结构：

- 各种各样的原始数据与信息，如监控录像、照片、短消息、论坛讨论帖子，等等；
- 结构化的知识体系，如专业性很强的 核素图 和通用型的知识库 FreeBase 等；
- 全人类知识的整体结构，那就是“知识图谱”。这个结构目前还没有完整可用的工具，网站 KA 提供了一个雏形，至于它能否完全承担起这样的任务，则视其发展而定。不过有一点可以肯定，这样的“知识图谱”型的工具必然会与维基百科类似，是集体合作的结晶。

知识图谱包含着科学、社会、历史、甚至宗教等每一个领域的知识地图；沿着特定领域的知识地图可寻找到特定的知识点；再将知识点关联到特定的信息与数据上。这样，知识与数据就构成了我们对外在世界的一张完整的 **认知之网**，这张网不仅包括科学的世界观，也包括宗教的世界观，它是人类认识的意识世界的合集。人类各式各样的认知活动，最终都是在编织着这张认知之网，并试图用它来覆盖住整个世界。

知识图谱中科学认知的框架部分不妨称之为 **科学知识图谱**。以这个科学知识图谱为骨架，填充进具体的科学知识与信息，这就形成了人类对世界的一张 **科学认知之网**，它是人类对世界认知的最真实的合集。

6 信息时代的智慧：构建你自己的“知识图谱”

人们需要知识与智慧的原因如下：

- 满足对世界本身的好奇；
- 利用这些知识和智慧去预测未来可能发生的事情，从而避开危害，给自己的生活带来便利，从而更好地享受生活。

前一节中谈到了信息时代的特点。不能够充分利用信息时代特点的人，不妨称之为信息时代的低能儿，他们大致有这样的表现：

1. 不会检索信息，这个不必多言。
2. 不会获得正确的信息。尽管获得信息很方便，但你有非常大的机会获得虚假的、错误的、充满欺骗的信息。因此，甄别信息的真伪非常关键。多种信息摆在面前时，不会分析与分辨真伪，这是一种低能的表现。

3. 更低能的，就是试图以欺骗自己的方式来欺骗他人。信息的公开、发布、检索的便捷则意味着试图通过隐瞒、篡改信息达到欺骗目的变得越来越困难，这种情况下仍然把他人当作瞎子来愚弄的人，自己会很容易变成被愚弄的对象。

如何适应信息时代，摆脱低能，拥有信息时代的智慧呢？个人认为，你需要充分利用大脑外存中的海量信息，建立一个自己的**个人知识图谱**，将你自己的知识与全人类的知识之间建立起联系，从而拥有自己的知识与智慧。

什么是**个人知识图谱**呢？这是我对前一节“知识图谱”概念的一个推广。它是一个抽象的概念，指的是个人对世界认知的骨架结构。英文可为 Personal Knowledge Atlas, 为后文叙述方便，简称为 PKA. 实质上, PKA 也可理解为个人的知识结构、知识框架等。之所以采用这个名词，因为我觉得，它的内涵又不同于单纯的知识结构，它应该成为一个个人的内在的认知工具。

全人类知识图谱内容甚巨，甚至彼此矛盾，比如同时包含有科学与宗教的知识，其中的大部分内容对个人可能无用。比如，除非兴趣或研究需要，你没有必要知道二千年前消亡的苏美尔语与阿卡德语之间相互影响的历史。那么，对个人而言，有哪些知识有用呢？至少应该包括这些内容：

- 知识图谱本身。比如一些信息学中基础的“本体”概念，如人、实物、一些抽象概念等。
- 科学的框架结构。你至少应该了解有哪些基本的科学门类，如物理、化学、数学、生物学、医学，等等；对于常见的科学，你应该了解它的基本结构，比如物理学是研究物质结构与运动的学科，可分为力学、热学、光学、原子物理、粒子物理与核物理等。
- 与你自身环境密切相关的知识，比如你所处的国家、地理位置、社会阶层、民族习性等。
- 与你工作密切相关的知识。

根据上述的有用知识，个人知识图谱 (PKA) 可以用科学知识地图为蓝本，结合与个人密切相关的环境，进行构造与完善。因此，每个人的 PKA 会彼此不同。

PKA 是你头脑中一个认知的框架，它越合理与完善，你能够恰当、准确处理的信息也越丰富、准确。继续借用传统微机架构进行类比，若把全世界的信息看成一个庞大的网络，把每个人当作是一个信息处理终端，则个人的大脑相当于信息处理终端的主要运算芯片，而 PKA 则相当于这款芯片的“指令体系”。基于宗教的 PKA 与基于科学的 PKA 相比，同样的信息很难得出一样的结果，因为彼此的“指令体系”不同。目前，最准确、高效的“指令体系”只有基于科学的 PKA.

信息时代，最快捷的是信息，而信息仍然需要一个过程才能够转化成个人的知识，进而转化成智慧。PKA 可以帮助你接触到的信息迅速定位，从而参与到个人决策中去。比如一条微博消息称《武汉协和医院发生严重放射事故造成多名女教授集体患癌》，由于我的核物理背景，第一反应就是仪器设备屏蔽不合格或操作不遵守程序，接着，大脑就自动把电离辐射导致生物体损伤相关的内容与之关联起来，将消息定位为技术使用不当的一类内容；而有的人的第一反应，可能是“天哪，当医生怎么这么不安全”，或者是“我们禁核吧”，然后把这条消息定位为一起丑闻了事；这就是个人知识图谱对信息定位的一个典型示例。

除了能够将信息快速定位，PKA 应该是你个人那条知识螺旋的旋转轴线，你的个人知识将沿着它进行累积、演化与壮大，进而获得更多有用的知识，从而将个人的学习与经验转化成智慧。

信息时代的智慧，主要体现在如何利用时代的特点进行信息的管理与知识的获取。对于这个问题的思考，目前有两方面的内容，即个人信息管理 (Personal Information Manage, PIM) 与个人知识管理 (Personal Knowledge Manage, PKM), 彼此间有诸多交叉。那么，把 PIM 与 PKM 的方法与你的 PKA 结合起来，你在信息时代的智慧则体现在：

- 具备快捷、高效、准确的信息检索技巧；
- 具备识别大部分信息真伪的能力；
- 具备将信息迅速分类、组织，并进一步推理出新内容的能力；
- 具备在知识间建立因果关联从而进行决策的能力。

实际上，上述这些智慧中最主要最关键的是最后一条：高效、准确的决策能力。

如何获得在信息时代的高效、准确的决策能力？或者说，如何获得信息时代的智慧？个人认为，你需要经历这样的过程：

1. 构建自己的 PKA.
2. 通过学习拥有基本的信息获取、检索、发布的技能。
3. 通过练习，将 PKA 与大脑外存中的海量信息建立联系，尤其是对你的工作与生活密切相关的信息，你可以迅速、准确地进行定位、归类、分析与判断。这个练习的过程可能需要一定的时间。

PKA 的构建，需要一个正规的、系统的学习过程。如果你系统地完成了中学教育，然后又认真修习完一套专业课程，实际上你的 PKA 已经基本建立。如果你再能够完成硕士或者博士的教育过程，则你的 PKA 相对完善了。如果你没有完成基本的中学教育，不妨寻找帮助，借助互联网这个大脑外存完成一下系统的教育（这个现在真的可以做到吗？）。而 PKA 的进一步完善则可以耗费你全部的余生，因为人类的认知是无止境的。而如何在完成系统教育之后继续完善你的 PKA，这且留给我们共同去思考。

在远古时期，为存储一点信息，我们需要编制歌曲、制作泥板或刀刻木片，思考本身反而不如制作工具麻烦。就在不久之前，我们的信息处理仍然以纸与笔为主，有人戏称为 PPP，即 Paper + Pencil + Patience，因为记录与检索信息仍然需要很大的耐心。在前信息时代，尤其是在古代，拥有信息都可被视为是智能。而在我们这个时代，信息就象深空中的星云一样浩瀚，信息的存储、检索等已经成为最简单的事情，大部分人不需要去记忆编码、不需要去背地址或邮编、不需要记住地形，大部分人已经从记录、传播信息的繁琐劳动中解放出来，大脑已经不再是信息的主要存储场所，甚至不再是大部分知识的存储场所。当信息的存储、检索已经主要依靠大脑的外存时，大脑就获得了解放。除了享受之外，解放了的大脑的主要任务就变成了获取知识与智慧，从而更加专注于智能本身。那么，我们不妨把这个时代的信息处理称为 III，即 Internet + Information + Intelligence.

如果将全部世界的信息与知识类比成银河系，我们人类的大脑是一艘星际飞船，则 PKA 就是这艘飞船的星图与导航系统，我们智慧就是能够在这个浩瀚的星系间给自己的心灵寻找一片舒适的空间。

让我们的大脑在信息高速公路上好好地展现其智能吧！

7 浅论信息时代的教育与科普

如果把所有的科学认知构造成一张 **科学的认知之网**，那么科学家就是那些辛勤织网的人。他们用自己的智慧构造这张网的骨架，那就是一个个科学概念与一条条概念之间的联络。然后，这张网覆盖到整个世界之上时，我们每个人都在检查这张网是否与现实吻合、是否存在漏洞、是否留有空

白，然后用我们自己的信息与知识进一步去完善这张网。实际上，任何人只要有能力去修建与弥补这张网的骨架，你就是科学家。原来局限在象牙塔中的科学研究活动实际上已经成为日常生活中很普通的活动，只要有能力，每个人都能成为织网的人。如果把给网中填充信息也当作织网的话，每个人都是织网的人。

这张科学的认知之网如何变成每个人的智慧？这就需要教育与科学普及。

首先谈谈教育。

个人认为，在信息时代，教育概念的内涵与外延已经与前大大不同。中国古代的教育，是为了普及“圣人”的智慧，或者说“圣人”的认知，其目的是训化良民，让百姓们知道哪些是“圣人”想让他们做的，哪些是禁忌。近代，教育的目的已经多样化了，如宗教的教育与中国古代类似是为了驯化教民，职业的教育是为了培养技工，等等。各类教育中，应该只有科学的教育目的最为特别，为了培养人们获得智慧的能力。个人认为，信息时代的科学教育，应该以获得信息时代的智慧为目的而进行。因此，这样的教育应该包括：

1. 系统的、基础的认知能力训练，其目的是让受教育者的知识螺旋旋转起来。这些认知能力包括感知、记忆、分类、排序、比较，等等，参见皮亚杰的《发生认识论》。
2. 具体知识的传授，在受教育者大脑内形成一套科学的个人知识图谱 (PKA)。
3. 特定的信息处理能力训练，包括信息获取、分类、分析与综合等。

由于信息时代的特点，需要单独谈谈基础认知能力训练中的记忆力训练。当然，记忆力训练仍然重要，但目的已经与以前大不相同。记忆力已经不再是单纯为了记住更多的信息，因为你有更庞大的外存，甚至，在知识螺旋的旋转过程中获得的新知识也没有必要专门记忆，你完全可以保存在外部介质中。若把记忆力类比成中央处理器的缓存，缓存越大，能够一次性处理的信息量也越大，这样，记忆力的训练，它的目的就变成了提高你的大脑一次性处理信息的量。

再次，谈谈科普。

什么是科学普及，或者科普呢？科普有什么作用？

个人以为，科普的本质，就是将人类的科学共识（包括科学方法与知识）普及到更多的人，让他们的意识世界与真实的世界更加接近。

为什么要科普？

个人以为有三方面的原因：

1. 为了文明的传承。科学认知是人类最伟大的成果，也是人类文明最精华所在。科普，是为了传承这个文明精华。
2. 科学的认知之网是与真实最接近的认知，也是最实用的认知。好东西的分享不带来损失，反而有好处，为什么不分享给更多的人呢？
3. 尽管科学的认知之网已经几乎覆盖了人类周围全部的世界，但全球大部分人仍然被宗教的认知之网控制着，这张科学的认知之网还没有被覆盖到足够人群的意识世界。尽管科学认知已经取得非常惊人的成就，但外部真实的世界仍然有许多未知的东西，科学的认知之网仍然没有覆盖住全部的世界，这张网仍然存在漏洞或破绽。科学的认知之网需要许多的织网人，除了人力，织网还需要许多的资源。只有这张科学的知识之网铺展到足够的人群，才会有足够的织网人，才会获得足够的织网资源。因此，科普也是科学与宗教进行人力与资源之争的主要手段之一。

怎样科普？这个问题至少涉及到两个方面：一是科普的内容，二是科普的方法与手段。

实际上，如果把传播知识、培养技能当成广义的教育，科普则属于系统教育的扩展，它是对系统教育的补充与更新。科普与系统的科学教育在内容上有重叠，同样包括科学方法、思想与具体的科学知识。但与科学教育不同的是，科普尽可能把最新的内容补充进来，以弥补教育中的不足。对于未能得到科学系统训练的人，科普可以吸引他们对科学的关注，从而激发兴趣，让他们自行去弥补知识结构的不足，完善其个人知识图谱。这样，科普在内容上就有如下的特点：

- 科普内容会显得相对零散，不一定如科学教育一样系统。比如物理学的科普作品，它不可能如朗道十卷本的理论物理系列教程那样完备。但作为科普作家，则必须具备完整的个人知识图谱，然后，他才能够从这个图谱中摘取一点形成作品，摆放到公众面前。
- 面向的对象不一定具备完整的个人知识图谱，因此要尽可能兼顾知识层次较低的人群，内容的组织上应该有“就低不就高”的原则。
- 内容与新闻一样，具有一定的时效性。当日本福岛核电站被地震撕毁时，公共陷入恐慌之中，此时传播核电站与放射性危害的科学知识比平时会有更佳的效果。那么，所谓的时效性，就是人们在不同的时间对具体科学知识的需求不同，科普应该及时去满足公众的这种需求。

科普的方法与手段指的是针对其内容特点所采用的技术，以达到受众更广、接受程度更高的目的。方玄昌在第一届“科学与无神论者大会”上的讲话中曾提到用做新闻的方法来做科普，由于科普与新闻在时效性与传播受众上有相似性，这个想法是非常好的，也是很自然的，值得推广。

科普的手段以前以科普文章为主，但在信息时代，科普的手段则愈加丰富：

- 综合文字、图片、视频的“富媒体”，比如浏览器和网络世界 20 大须知，比如 YouTube 上有许多一分钟物理学的视频，等等。
- 游戏形式的科普。比如将力学规律的认知历史做成角色扮演型游戏，你愿意扮演牛顿，还是第谷呢？这个是个个人借鉴自 RPG 游戏的一个构思，我不知道有没有具体的作品。
- 计算机程序形式的科普，尤其是面向 iPad 等平板设备的教育型的程序。“每个孩子一台笔记本” (OLPC) 的计划导致了“糖果操作系统” (Sugar OS) 的诞生，个人以为，在这样的教育平台上开发科普性质的应用程序应该很不错。
- 在微博、论坛等互动媒体上的科普。
- 更多新奇的想法等待着你来展示。

如果把科学认知之网当成风景优美、趣味横生的迪斯尼乐园，科技论文则相当于乐园中具体设施的技术资料，科学领域的系统教材（如朗道十卷本的理论物理系列教程）和综述性文章则相关于具体设施的详细使用手册，而科普作品，则相当于乐园的宣传手册。科普，是引领游客抵达快乐的向导。如果你是名科学认知之网的织网人，那么，写一篇宣传手册吧，把你的编织乐趣与所织就的成果传达给更多的人；如果你见识了很漂亮的风景，那么，写一篇宣传手册吧，把美丽传达给更多的人。

8 后记

本文中有一些自创的概念或者对概念的引申，如 **知识图谱**，**个人知识图谱**，**知识地图**，**认知之网**。之所以自创或引申概念，是因为觉得这样更能准确表达我的思想。但这些概念的由来，除了文中的许多网址链接外，主要受如下内容的启发：

- Frand, J. & Hixon, C. (1999). Personal Knowledge Management : Who, What, Why, When, Where, How?. Los Angeles: UCLA. Retrieved July 28, 2011 from <http://www.anderson.ucla.edu/faculty/jason.frand/researcher/speeches/PKM.htm>
- Knowledge Atlas – <http://knowledgeatlas.coop/>
- 知识制图学 (Knowledge Cartography) – <http://knowledgecartography.org/>

读了不少作品，把自己关于知识图谱、知识管理与科普的一点粗浅想法写出来之后，发觉其实仍然比较凌乱。但传播想法是知识壮大的有效途径，因为觉得这点想法是有点意思的，敝帚自珍吧，希望传播出来后能够得到更多的回馈，令它更完善。