

## 动态社区下基于用户应用环境的开放描述

李宇

(中国科学院文献情报中心 100080)

文摘: 本文探讨了在点对点的协作模式下使用 FOAF 和 XTM 对用户的应用环境进行描述, 以支持动态聚合和组建动态社区。

关键词: 用户应用环境, 描述, FOAF, XTM, 动态社区

Organize User Context to Support Person in Dynamic Community

Li Yu

(Library of Chinese Academy of Sciences 100080)

Abstract: The essay described an approach which uses FOAF and XTM to describe user context information, and it is supposed to support dynamic community.

Keyword: user context, describe, FOAF, XTM, dynamic community

互联网十年的发展已经深入人心, 信息铺天盖地, 用户面对 Google 成千上万的检索结果只能期望自己能够得到需要的信息。这里的问题在于, 当用户向搜索引擎提交一个检索式的时候, 双方都忽略了应用环境。搜索引擎面对的是一个抽象的人, 一个概念化了的抽象类, 而不是实实在在输入检索式获取信息的用户。应用环境被作为一种双方都假定对方已经了解的背景知识而被隐藏, 重新显性化的表达这种被隐藏的知识, 找到其合适的方面并将其模型化, 正确的定义应用环境及其相应的关系, 从而能够被系统化的应用到推理和决策支持上。减少信息过载并提高用户所使用的信息的质量。

### 1 用户应用环境 (User Context) 的表达

“某个事物的应用环境包含了主意、情境、事件或者与之相关的信息, 这些信息使得该事物能够被完整的理解”<sup>[1]</sup>。应用环境的表达难点在于, 如何侦测出众多的应用环境的元素中适合应用并并能起作用的元素。现有的个性化服务的主要问题, 在于缺乏对用户及其信息需求的理解。传统的用户个性化建模通常将焦点放在用户兴趣、偏好等认知模式上。而要真正理解用户及其信息需求, 就要求对用户的应用环境进行清晰的理解和表达。

在 AmbieSense 中将用户应用环境分为 5 个部分<sup>[2]</sup>:

(1) 环境上下文 (Environment context) 包括用户周围的实体。其实例可以是东西、服务、温度、光线、湿度、噪音和人。用户在当前用户应用环境中访问的信息 (例如文本、图像、影像、声音) 都属于环境上下文。用户周围不同的计算机网络也可以被描述为环境上下文。

(2) 个人上下文 (Personal context) 包含生理上下文和心理上下文。生理上下文包括脉搏、血压、体重、视网膜模式和头发颜色。心理上下文包括专家知识 (相对静止)、情绪、生气、压力 (相对变化较大) 等。

(3) 任务上下文 (Task context) 描述用户在用户应用环境中做什么。可以被显性的描述为目标、任务、行动、活动、事件等。可能包含在同一处境下其他人的任务。

(4) 社会上下文 (Social context) 描述社会方面的问题。包括朋友、中立者、敌人、邻居、合作者以及其他相关的实例。社会上下文的一个重要方面是用户在其中所扮演的角色。

(5) 时空上下文 (Spatio-temporal context) 描述用户应用环境中和时间空间相关的部分。可能包含: 时间、地点、方向、速度、形状、路线、住所、用户的衣物等环境空间的扩展。

CYC 系统<sup>[3]</sup>将应用环境划分为 12 个维度, 其中包括时间、时间类型、位置、位置类型、

文化、复杂化/安全、主题、粒度、形式/部署/认识论、意见-偏好、理由、领域假设。在 CYC 中，应用环境被定义为关于能进行共享的事物的命题假定的一致集，这种集合形成了关于某个主题的理论。公式不仅仅是为真或假，而是公式在特定的应用环境中为真或假。这使得人们能够利用关于应用环境的陈述，尽管应用环境并不能完全地被陈述。

目前对应用环境的研究主要表现在两个方面：一个方面是对应用环境敏感（Context-aware）的智能设备的研究，他们对于应用环境的研究主要侧重于当前的地点、时间、人及其使用的设备等因素。例如智能设备使用的应用环境信息；另一个方面则是更多考虑了通用性的应用环境模型。例如在<sup>[4]</sup>中使用扩展的分面结构将应用环境分为四维：认知模式、关系、任务和环境。

由于应用环境的复杂性和应用的广泛性，因此，对应用环境的进行分面探讨无疑是比较理想和实际的。

## 2 基于用户应用环境的动态社区架构

P2P（Peer to Peer）技术，社会化计算技术，网格技术等技术的一个共同特点，就是让以前单纯的信息或资源的获取者，同时成为信息或资源的提供者。这些技术的出现，使得信息的提供者、组织者和获取者之间的界限变得模糊。通过动态的组织用户与用户之间的联系，形成一个支持用户应用环境的动态社区。

HP 的研究结果表明<sup>[5]</sup>，知识工作者在利用 Web 信息的时候，具有如下特点：

- (1) 使用熟悉的资源。质量、标准或者内容得到保证。即使对于新发现的资源和工具也如此。
- (2) 项目的本性。用户的任务和其已经拥有的资源在很大的程度上确定了信息是否有用。
- (3) 熟悉的领域。知识工作者不仅习惯于访问已知的资源，同时他们也习惯于待在他们自己的专家领域中。
- (4) 在实践中学习。例如在知识工作者的搜索实践中，他们倾向于在使用中学习及提高搜索技巧。
- (5) 个人计算机只是工作台。计算机中的文档常常只是一个摆放位置，只看一个文档是没有办法了解该文档的创建目的以及它与其它文档或者工作的关系。用户大脑中的应用环境通常不会放到计算机里，除非有良好的机制协助用户表达其应用环境。
- (6) 使信息可共享的工作。在共享一个文档之前，知识工作者通常会对文档进行一些修改，使之能够被使用者所理解。通常需要描述文档所针对的任务以及文档与其它文档的关系，从而使隐藏的应用环境显性的表达出来。
- (7) 对共享者的反馈。接收者对信息的反馈提高了信息共享的动机。共享者可以根据接收者的情况修改和组织内容使之更符合接收者的需要，同时支持该共享的工具也应当使共享者能够清楚掌控和明确哪个文档共享给了谁。

从以上的研究结果可以看出，要想协助用户更加有效的进行动态社区环境下的合作推荐及信息过滤<sup>[6][7]</sup>，除了利用社会化计算<sup>[8][17]</sup>的弱链接特点<sup>[9][18]</sup>以外，还需要对个人的信息及知识架构以及与之相应的应用环境进行清晰的描述和揭示以同时支持人能够浏览及机器可处理。

## 3 用户个人信息及应用环境信息描述

Dan Brickley 和 Libby Miller 发明的 FOAF<sup>[10]</sup>词汇表为社区成员提供了一种基本表达：描述人及其基本特性，如姓名、地址、电子邮件、兴趣、参与的项目等等。

FOAF 是一个 RDF（Resource Description Framework）<sup>[11]</sup> 词汇表。它的常规用法类似于 RSS<sup>[12]</sup>：在 Web 服务器上创建一个或多个 FOAF 文件并共享 URL，这样，使得软件可以使用该文件内的信息。和 RSS 一样，当聚集数据然后探究和交叉链接这些数据时，FOAF 的作用

就凸现出来：找到某一方面相似的人（例如找到具有相似任务的人或者具有相似兴趣的人）。与 vCard<sup>[13][14]</sup>不同的是，FOAF 的词汇表<sup>[15]</sup>规定了 FOAF 从 5 个方面进行描述：FOAF 基础、个人信息、即时通讯帐号、项目和组、文档和图像。

FOAF 的设计使其能够成为管理社区的重要工具。例如：为社区新成员提供帮助、查找兴趣相投的人、查找参与过某个项目的人等等。下面的表 1 是使用 foaf-a-matic<sup>[16]</sup>创建并修改的 FOAF 说明：

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:admin="http://webns.net/mvcb/"
<foaf:PersonalProfileDocument rdf:about="">
  <foaf:maker rdf:nodeID="me"/>
  <foaf:primaryTopic rdf:nodeID="me"/>
  <admin:generatorAgent rdf:resource="http://www.ldodds.com/foaf/foaf-a-matic"/>
  <admin:errorReportsTo rdf:resource="mailto:leigh@ldodds.com"/>
</foaf:PersonalProfileDocument>
<foaf:Person rdf:nodeID="me">
<foaf:name>Yu Li</foaf:name>
<foaf:title>Mr</foaf:title>
<foaf:givenname>Yu</foaf:givenname>
<foaf:family_name>Li</foaf:family_name>
<foaf:nick>carp</foaf:nick>
<foaf:mbox>liy@mail.las.ac.cn</foaf:mbox>
<foaf:workplaceHomepage rdf:resource="www.csdl.ac.cn"/>
<foaf:schoolHomepage rdf:resource="www.las.ac.cn"/>
<foaf:interest rdf:resource="www.w3c.org"/>
<foaf:knows>
  <foaf:Person>
    <foaf:mbox rdf:resource="mailto:yangxh@mail.las.ac.cn" />
    <foaf:name>xiaohang yang</foaf:name>
  </foaf:Person>
</foaf:knows>
</foaf:Person>
</rdf:RDF>
```

表 1

表 1 说明了 FOAF 的词汇表如何描述姓名，头衔，工作地主页，毕业学校的主页，感兴趣的资源，以及电子邮件地址。同时说明 Yu Li 认识邮箱地址为 yangxh@mail.las.ac.cn 的用户 xiaohang yang。通过聚集和合并多个 FOAF 文件，可以获得同操作集中式目录服务相同的效果，而没有单点故障或单点控制的任何问题。对于许多需要分散或转移控制的动态社区而言，FOAF 提供了一个很好的选择。

#### 4 用户知识架构描述

用户的知识架构是存在于用户的大脑当中并和许多应用环境相关的。他们具有在相对时

间内稳定但会随着用户的知识结构变化而变化的特点。知识可以被分为三个部分：知识的内容、知识的结构、知识的表达。支持用户应用环境的知识结构及其表达应当具备足够的灵活性、可扩展性、设备平台无关性。XTM<sup>[19]</sup>提供了成熟的选择。

ISO/IEC 13250 标准<sup>[20]</sup>定义了“描述知识结构并将其与信息资源联合”的国际标准——主题图 (Topic Maps)<sup>[21][22]</sup>，对信息资源结构的可互相转换的信息描述提供了使用定义主题和主题之间关系的方法。

通过主题图，可以实现：

(1) 限定包含在作为主题的信息对象中的内容或数据并可以形成导航工具，如：索引、相互参照、引用系统、术语表。

(2) 通过主题之间的关系和交流，将主题联结在一起。

(3) 过滤信息资源集合为特定的目的或用户创建个性化门户。包括管理多种语言的资源，和依赖安全标准的访问模式的管理等。

(4) 组织非结构化信息对象，或者促进创建面向主题的用户界面以提供将非结构化信息合并到结构化信息效果。

主题图以一种独立实施的方式捕获并描述资源中所论及的主题以及这些主题之间的关系。XTM 是基于 ISO13250 标准定义的用 XML 描述和标记主题图的方式<sup>[21]</sup>。XTM 的基本组成单元是主题 (Topics) 及其关系 (Associations)。每个主题代表具体实体或概念对象，可出现在信息资源中 (被描述、讨论或提及)。在 XTM 中，主题索引主要由三部分构成：一系列 topic 主题、出现、主题之间的关系。

主题图通过范围 (scope) 的概念来表达应用环境有效性 (contextual validity) 的问题 [23]。通过简单的给定相关断言的范围而使得一个名称，出现和关系角色被认为是有效的应用环境声明。范围的作用与命名空间类似：基本名称在某个范围内应该是惟一的，如果两个主题具有相同的基本名称，而且在同一个范围内，它们两者就可以合并。

通过表 2<sup>[24]</sup>可以看出 XTM 是如何工作的：

莎士比亚及其著作的 Topic Maps 节选

```
<!-- A topic representing the Elizabethan playwright
      William Shakespeare.  No occurrences because you cannot download
      a person -->
<topic id="shakespeare">
  <baseName>
    <baseNameString>William Shakespeare</baseNameString>
  </baseName>
</topic>
<!-- A topic representing the play "Hamlet" -->
<topic id="hamlet">
  <instanceOf><topicRef xlink:href="#play"/></instanceOf>
  <baseName>
    <baseNameString>Hamlet, Prince of Denmark</baseNameString>
  </baseName>
<!-- An occurrence given by Project Gutenberg's plain text download
      of the Hamlet -->
<occurrence>
  <instanceOf>
    <topicRef xlink:href="#plain-text-format"/>
  </instanceOf>
</occurrence>
```

```

    </instanceOf>
    <resourceRef
xlink:href="ftp://www.gutenberg.org/pub/gutenberg/etext97/lws2610.txt"/>
    </occurrence>
</topic>
<!-- An association representing an authorship relationship -->
<topic id="written-by">
    <baseName>
        <baseNameString>written by</baseNameString>
    </baseName>
</topic>
<!-- Used here to associate Shakespeare and the play Hamlet -->
<association>
    <instanceOf><topicRef xlink:href="#written-by"/></instanceOf>
    <member>
        <roleSpec><topicRef xlink:href="#author"/></roleSpec>
        <topicRef xlink:href="#shakespeare"/>
    </member>
    <member>
        <roleSpec><topicRef xlink:href="#work"/></roleSpec>
        <topicRef xlink:href="#hamlet"/>
    </member>
</association>

```

表 2

主题图为语义网带来了高度的精确性。高度的精确性导致规范中定义的名词和细节变化多端，因此模型<sup>[25]</sup>理解起来比较困难。但主题图为合并、出现、关联等等定义了非常详细的规则，因此非常适合一个总是在变化的动态的社区。方便易用的集成主题图功能的工具或者 Web 服务的出现，将为每个用户创建其自己的概念架构提供直观、有效、简洁易用的知识结构表达方式。

HP 的原型测试程序 ePerson 片断管家<sup>[26]</sup>是为了在特定的节点组中组织并共享小片断的信息资源（例如新闻文章、论文、图片、书签、注释等）而设计的。每个用户通过片断管家自行管理自己的知识库。ePerson 的基础架构包括：对代理服务器日志进行分析从而推理用户兴趣的网关、将 URI 映射到 DMOZ<sup>[27]</sup>的主题类别下、一套具有图形界面和命令行界面的支持开发者访问和测试 ePerson 网络的工具、一个支持绑定本地 ePerson 网络服务的发现服务器。这套程序试图表达用户的个人信息，并通过适当的共享机制支持社区用户的共享发现、组织、过滤信息的工作。共享的个人信息包括用户的 profile，共享的内容以及共享的元数据（例如注释、评论、评级以及分类）。与之相似的项目还有 MIT 的 haystack<sup>[28]</sup>。

片断管家作为对个人信息表达及社区共享的探索是非常成功的，例如其 RDF 查询机制 QBE (query by example)、显性的本体设计与管理等等。但也要看到它还存在的不足：对对安全和信任的支持，以及用户应用环境、社区应用环境的支持需要加强。

## 5 结论及展望

个性化的发展已经不能依靠对用户的认知模式和兴趣偏好的模型来满足了。新技术的发展要求新的应用及设计。除了支持可互操作的标准之外，还必须考虑复杂多变的应用环境。

应用环境本身的多面性和复杂性使得现在的应用环境基本上是手工定义的，其模型的建立以及相应的应用方法还需要进一步的探讨和研究。动态社区的构建除了用户应用环境之外，还需要社区应用环境的支持，进一步的，还需要考虑数字对象本身的应用环境。

#### 参考文献

- [1]Akman, V., Surav, M. Steps toward formalizing context[J]. AI Magazine, 1996, 17:55-72.
- [2]Goker, A., Myrhaug, H. User context and personalization[A]. Proceedings of the European Conference on Case Based Reasoning (ECCBR 2002) - Workshop on Personalized Case-Based Reasoning[C], Aberdeen, Scotland: Springer-Verlag, 2002.4-7
- [3]Lenat, D.B. The dimensions of context space [EB/OL]. <http://courses.csail.mit.edu/6.803/lenat2.pdf>, 2005-04-01
- [4]Nieder é, C.J., Stewart, A., Mehta, B., Hemmje, M. A multi-dimensional, unified user model for cross system personalization[A]. Proceedings of Advanced Visual Interfaces International Working Conference (AVI 2004) - Workshop on Environments for Personalized Information Access[C], Gallipoli, Italy: ACM, 2004. 26-39
- [5]Gathering and Sharing Web-based Information: Implications for “ePerson” Concepts [EB/OL]. <http://www.hpl.hp.com/techreports/2003/HPL-2003-19.pdf>, 2005-04-01
- [6]Thomas, J., Kellogg, W.A. and Erickson, T. The Knowledge Management Puzzle: Human and Social Factors in Knowledge Management[J]. The IBM Systems Journal, 2001, 40(4): 863-884
- [7]Erickson, T. and Kellogg, W. A. Knowledge Communities: Online Environments for Supporting Knowledge Management and its Social Context[EB/OL]. <http://www.research.ibm.com/SocialComputing/Papers/KnowCommunities.pdf>, 2005-04-01
- [8]Social Computing Group[EB/OL]. <http://research.microsoft.com/scg/>, 2005-04-01
- [9]The Eureka Search[EB/OL]. <http://eureka.com/>, 2005-04-01
- [10]the foaf project[EB/OL]. <http://www.foaf-project.org>, 2005-04-01
- [11]Resource Description Framework (RDF) [EB/OL]. <http://www.w3c.org/RDF/>, 2005-04-01
- [12]RSS 定义[EB/OL]. <http://www.google.com/search?hl=zh-CN&lr=&newwindow=1&oi=defmore&q=define:RSS>, 2005-04-01
- [13]RFC 2426 (rfc2426) - vCard MIME Directory Profile[EB/OL]. <http://www.faqs.org/rfcs/rfc2426.html>, 2005-04-01
- [14]Renato Iannella, Representing vCard Objects in RDF/XML, W3C Note 22 February 2001[EB/OL]. <http://www.w3.org/TR/vcard-rdf>, 2005-04-01
- [15]FOAF Vocabulary Specification[EB/OL]. <http://xmlns.com/foaf/0.1/>, 2005-04-01
- [16]FOAF-a-matic -- 使用 RDF 描述你自己[EB/OL]. <http://www.ldodds.com/foaf/foaf-a-matic>, 2005-04-01
- [17]Microsoft Research Social Computing Symposium 2004[EB/OL]. <http://research.microsoft.com/workshops/SCS2004/>, 2005-04-01
- [18]THE STRENGTH OF WEAK TIES: A NETWORK THEORY REVISITED Mark Granovetter[EB/OL]. [http://www-personal.si.umich.edu/~rfrost/courses/SI110/readings/In\\_Out\\_and\\_Beyond/Granovetter.pdf](http://www-personal.si.umich.edu/~rfrost/courses/SI110/readings/In_Out_and_Beyond/Granovetter.pdf), 2005-04-01
- [19]XML Topic Maps (XTM) 1.0[EB/OL]. <http://www.topicmaps.org/xtm/1.0/>, 2005-04-01
- [20]ISO/IEC 13250 Topic Maps[EB/OL]. <http://www.y12.doe.gov/sgml/sc34/document/0129.pdf>, 2005-04-01
- [21]What Are Topic Maps? [EB/OL]. <http://www.xml.com/pub/a/2002/09/11/topicmaps.html>, 2005-04-01
- [22]The TAO of Topic Maps: finding the way in the age of infoglut[EB/OL]. <http://www.ontopia.net/topicmaps/materials/tao.html>, 2005-04-01
- [23]Ten Theses on Topic Maps and RDF[EB/OL]. <http://www.ontopia.net/topicmaps/materials/rdf.html>, 2005-04-01
- [24]Think XML: 从书本学习 XML Topic Maps[EB/OL].

<http://www-900.ibm.com/developerWorks/cn/xml/rdf/part19/index.shtml> , 2005-04-01

[25]XML Topic Maps (XTM) Processing Model 1.0[EB/OL].<http://www.topicmaps.org/xtm/1.0/xtmp1.html> , 2005-04-01

[26]The ePerson Snippet Manager: a Semantic Web Application[EB/OL].

<http://www.hpl.hp.com/techreports/2002/HPL-2002-328.pdf> , 2005-04-01

[27]ODP - Open Directory Project[EB/OL].<http://dmoz.org/> , 2005-04-01

[28]David Huynh, David Karger, and Dennis Quan. Haystack: A Platform for Creating, Organizing and Visualizing Information Using RDF. Draft submission to journal 2002[EB/OL].

<http://haystack.lcs.mit.edu/papers/computer-network-2002.pdf> , 2005-04-01

作者简介：**李宇**，男 1977 年生，中国科学院文献情报中心博士研究生。联系地址：北京北四环西路 33 号中国科学院文献情报中心 619 室，邮编：100080 电话：010-82629002。E-mail：[liy@mail.las.ac.cn](mailto:liy@mail.las.ac.cn)