

科技知识组织体系 (STKOS) 相关工具集成服务系统的设计与实现*

方安¹ 吴思竹¹ 洪娜¹ 钱力² 王颖² 胡佳慧¹

¹ (中国医学科学院医学信息研究所 北京 100020)

² (中国科学院文献情报中心 北京 100190)

摘要: 【目的】科技知识组织体系课题产出了大量知识组织相关工具, 为了避免工具重复建设, 实现对这些工具的复用和共享。【应用背景】“面向外文科技文献信息的知识组织体系建设和应用示范”项目是国家“十二五”科技支撑计划项目, 构建了包含超级科技叙词表、本体、范畴在内的, 覆盖理、工、农、医的多层次知识组织体系。在科技知识组织体系建设过程中, 各课题组开发了大量知识组织相关工具。【方法】基于 OSGi 的 Equinox 实现框架搭建了插件型知识组织工具集成服务系统。设计及实现了插件自动封装流程, 利用文件和数据库相结合的方式存储工具和插件, 并提出基于 jBPM 工作流的插件组配机制。【结果】完成了 STKOS 相关工具的仓储及对外发布系统建设, 基于插件的知识组织工具集成框架和组件建设以及基于工作流的知识组织工具集成三方面的建设内容。【结论】实现了对 STKOS 相关工具的集成、规范化管理和共享共用。

关键词: STKOS 工作流 OSGi 插件 工具集成

分类号: G350

Design and Implementation of Integrated Service System for STKOS

Related Tools

Fang An¹ Wu Sizhu¹ Hong Na¹ Qian Li² Wang Ying² Hu Jiahui¹

¹(Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing 100020, China)

² (National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: [Objective] Avoiding the repeat of tools construction, this paper aims to realize the efficient reusing and sharing of tools related to knowledge organization, which are outputs of Science and Technology Knowledge Organizing System (STKOS). [Context] *The Construction and Application Demonstration of Knowledge Organizing System Oriented Foreign Scientific and Technical Literature Information* is one of the National Science & Technology Pillar Program projects during the Twelfth Five-year Plan Period, which aims to construct a multilayer system of science, engineering, agriculture and medicine involving thesaurus, Ontology and category. During the construction of STKOS, a large number of tools are developed. [Methods] The Equinox OSGi core framework implementation is used to build a plug-in integrated service system for knowledge organization tools. Plug-in automatic packaging process is designed and implemented, tools and plug-ins are stored in files and databases, and a plug-ins allocation mechanism based on jBPM workflow is proposed in this paper. [Results] The construction of

*本文系国家“十二五”科技支撑计划项目课题“科技知识组织体系共享服务平台建设” (项目编号: 2011BAH10B03) 和其子课题 3 “科技知识组织体系 (STKOS) 相关工具集成服务系统建设” (项目编号: 2011BAH10B03-1) 的研究成果之一。

STKOS is completed, including the related tools storage and release system, the tools integration frame and components and the knowledge organization tools integration based on workflow. [Conclusions] The related tools integration, standardized management and sharing of STKOS are realized.

Keywords: STKOS Workflow OSGi plug-in Tools integration

1 引言

“面向外文科技文献信息的知识组织体系建设和应用示范”项目是国家“十二五”科技支撑计划项目，构建了包括超级科技叙词表、本体、范畴在内的，覆盖理、工、农、医的多层次知识组织体系。在科技知识组织体系（Science and Technology Knowledge Organizing System, STKOS）的建设过程中，各课题组开发了大量知识组织相关工具，为了避免工具重复建设，实现在 STKOS 课题组下对相关工具进行规范管理、共用共享和有效利用，课题组提出构建基于 OSGi 的插件型 STKOS 相关工具集成服务系统，定义了工具集成框架，通过标准插件开发规范，将 STKOS 项目的各课题中构建的一些重要知识组织工具封装为可控、可管理的插件，集成到系统之中，形成知识组织工具插件库，用户可以按需求选择各种知识组织工具插件，将这些插件有机地组合起来，满足某项知识组织体系建设的需要。

OSGi（Open Service Gateway initiative）是面向 Java 的动态模型系统。作为 Java 模块化发展的代表，其具有面向服务的组件模型（Service-oriented Component Model）的核心设计思想，具有高层次的模块管理、良好的封装性、较低的模块耦合度、易扩展等优点。OSGi 以 bundle 为单位实现模块化，其提供了一套完善的机制用于管理和控制 bundle、服务（Service）的生命周期，以及 bundle 和服务在其生命周期内的交互。由于已存在的项目和若干来自不同 STKOS 课题开发的知识组织相关工具，均经历了长时间的开发，不能抛弃以前所有的成果重新编写，根据 OSGi 规范可对各课题组的核心工具进行改造、封装形成灵活的插件，工具插件化能够极大地促进 STKOS 相关工具的集成、复用和共享。因此，本文提出基于 OSGi 规范的插件型 STKOS 相关工具集成服务系统的构建思路，设计了系统框架，论述了系统的数据存储方式、服务模式和功能模块，并介绍了当前系统的实现效果。

2 相关研究

知识组织系统是传统信息检索语言的发展，传统信息检索语言的主要功能是为标引者和检索者之间的信息沟通建立共同表达的桥梁。随着计算机技术的发展和人们对知识需求的日益迫切，知识组织系统的功能不断扩展，不仅能帮助用户查找和感知信息，帮助用户描述问题和给出答案，也能够提供有意义的概念结构作为资源组织和描述的形式化基础，还能够提供计算机程序理解、分析和知识推理。当前知识组织系统建设关注多种知识组织工具和方法的整合和集成，包括多词表、多语种的综合知识组织系统构建。如美国国立医学图书馆的医学一体化语言系统（Unified Medical Language System, UMLS）^[1]、欧洲多语言叙词表 EuroVoc^[2]、联合国粮农组织的多语言农业叙词表 Agrovoc^[3]等。当前，也有许多

项目目标在于多词表的整合集成,如 FOS 项目,整合了 4 个词表系统,包括 FIGS 门户的参考目录、ASFA 叙词表、Agrovoc 叙词表的渔业部分、OneFish 主题树^[4]。OntoFusion 项目是面向基因和临床医学数据库的本地集成系统,采用半自动化的映射 (Mapping) 和全自动化的统一 (Unification) 两种方式将大规模生物医学数据库资源关联起来^[5]。在知识组织系统构建过程中,产生了大量知识组织系统编辑、构建等相关工具,这些工具或以功能模块的形式集成在平台上,如 AGROVOC 概念服务器工作台 (AGROVOC Concept Server Workbench)^[6]提供概念编辑、加工等功能;或以独立、集成工具的形式发布在网站上,提供用户下载,如美国国立医学图书馆将知识组织工具 MetaMap 等^[7]提供给用户下载和使用,并没有形成关于知识组织工具的统一、规范的管理和应用。

我国正在建设的“面向外文文献的科技知识组织体系建设”项目,集成了来自理、工、农、医 4 个领域的上百部词表资源构建超级科技叙词表 (STKOS),用于提供更深入的知识服务。在 STKOS 的建设和示范方面,中国科学院国家科学图书馆的王兴兰等^[8]研究了图书馆领域知识组织体系的文本自动分类方法,分析得出现有方法需要结合基于机器学习的算法对分类模式进行改善,指出完善的知识组织体系支撑的重要性,并介绍了 STKOS 作为新的自动分类体系,拥有强大的数据支撑,可以在极大程度上消除分类体系本身的不足。NSTL 智能检索平台也是“面向外文科技文献信息知识组织体系建设和示范应用”项目的建设内容之一,其目标是为科技创新主体提供高效便捷的知识服务和科研文献信息支撑环境,中国医学科学院医学信息研究所参与构建的 STKOS 协同工作平台为该检索平台构建关联网络提供其所需的基础数据^[9]。中国科学院文献情报中心的相关研究人员基于 STKOS 引擎 API 设计了开放查询和推理接口模块化方案和规范^[10],目前已完成子课题“科技知识组织体系 (STKOS) 的开放查询和推理接口建设”的主体建设。通过标准接口协议,所设计的 STKOS 接口为第三方系统实现知识组织体系的深度开发和利用提供帮助。综上所述,以实现我国海量外文科技文献信息知识组织和利用为目标^[11],STKOS 的研究工作目前已取得了一定的进展,在 STKOS 的建设过程中,产出了大量知识组织建设、编辑工具。但是各课题组之间所建设的知识组织相关工具存在交集,不可避免地出现大量工具重复建设的情况,因此有必要对这些工具进行集成、规范管理,提高知识组织工具的共享、共用程度。

20 世纪 90 年代后期,软件开发从面向对象的开发阶段转向基于构件的开发阶段。构件是可以被复用的软件构成成分,由构件规约和构件实现两部分组成。基于构件的软件工具具有灵活的使用性和可扩展性。为了便于软件工具的集成、复用,各种构件库建设兴起,如欧盟信息技术计划 ESPRIT 中开发的 REBOOT (Reuse Based on Object Oriented Techniques) 系统^[12]、北京大学开发的青鸟构件库管理系统^[13]和美国 Colorado 大学开发的 CodeBroker 构件库^[14],支持大量构件的描述、存储、检索和管理。构件插件化是当前工具软件开发的新兴发展趋势之一。插件是构件的组成部分,粒度更细、更灵活并易于管理、组配和扩展。Eclipse^[15]平台就是一个典型的开放式插件体系结构,以微内核扩展形式构成,通过插件扩展开发功能。尹铨^[16]利用 Eclipse 插件技术、数据集成技术,并采用 GEF 和 JFACE 配置 workflow 实现软件的集成,用于软件生产线的复用。2008 年,欧盟第七研究框架 (the 7th Framework Programme for Research, FP7) 启动了 LarKC (Large Knowledge Collider) 项目^[17],研发基于插件的 LarKC 语义数据处理平台,集成了包括识别插件、转换插件、选择插件、推理插件、决策插件等几大类插件,通

过 workflow 可将一些插件贯穿在一起, 实现特定任务。该平台主要用于处理语义数据, 实现语义推理。这些项目和研究在不同应用领域对工具集成进行了实践, 并且有的仅作为工具或项目的托管平台, 支持工具的下载而不能支持工具的在线使用和灵活组配。一些平台虽然能够支持 workflow 组配, 但其组配的工具多由平台本身创建, 而非由多方提交、集成和封装, 并且现有研究中缺乏在知识组织建设的工具集成和管理方面的实践。基于当前不足, 本文借鉴已有研究对 OSGi 的实践作为工具集成系统建设的经验, 探索基于 OSGi 的插件型知识组织相关工具集成建设, 设计可集成的系统框架, 对各种知识组织工具的核心模块安插模式进行封装, 规范这些插件的类型、输入/输出接口, 使这些核心模块能够被集成到知识组织工具的集成框架中, 形成可被管理的组件对象。本文将详细介绍 STKOS 相关工具集成服务系统的设计方案, 并结合目前的研究进展给出系统的实现效果。

3 系统设计

3.1 总体框架

图 1 为 STKOS 相关工具集成服务系统的总体框架。STKOS 相关工具集成服务系统采用基于 OSGi 的 Equinox 框架作为运行环境, Equinox 是 OSGi 最早和应用最为成熟的实现框架, 具有规范的代码和清晰的结构, 并且其与 Eclipse 的结合, 能提供便利的开发环境配置^[18]。各课题提交知识组织相关工具, 如文本抽词工具、选词评价工具、编辑工具等, 知识组织体系相关工具集成服务系统对这些工具进行登记、注册, 存储到工具库, 遵循 OSGi 规范对工具进行规范接口封装, 并对插件描述信息进行记录, 将插件及其相关信息存储到插件库中。按插件功能对其进行组织分类, 如抽取类、词法分析类和分类去重类等, 这些类集成了来自不同课题、不同单位、不同用户开发的实现相同功能的插件。集成服务系统通过动态加载插件, 将插件发布到 Web 页面, 提供用户服务。系统通过构建 workflow 引擎调度实现对不同功能的若干插件的组配, 如在一个流程中实现对一篇文本的词汇抽取、词形规范、规范后的词频次统计等功能, 形成为满足特定应用需求的工作流, 并支持对外发布和服务。

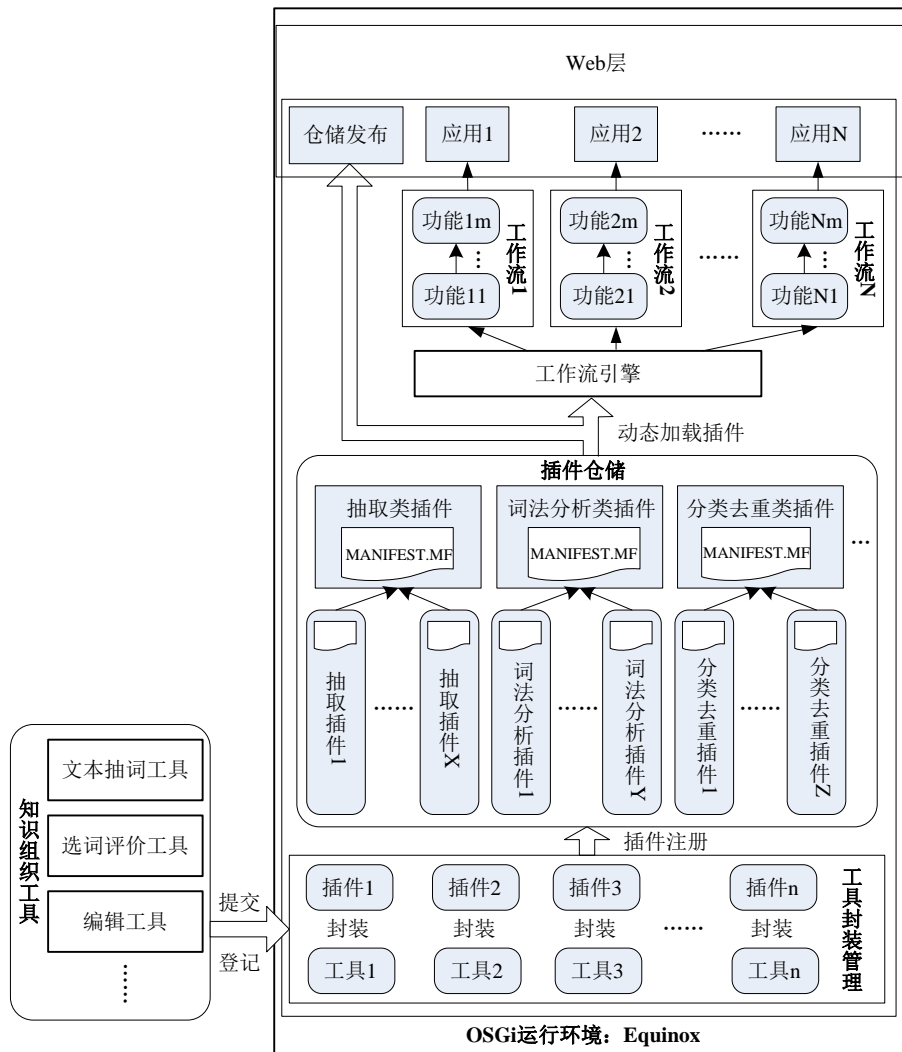


图 1 STKOS 相关工具集成服务系统总体框架

3.2 存储方式

STKOS 相关工具集成服务系统中包含多种存储对象，主要有待集成工具、封装后的插件以及工具和插件的相关描述信息。工具，主要是课题各成员单位提交的工具源码压缩包或 jar 包、工具运行所需的第三方 jar 包或运行需调用的资源等。封装后的插件 bundle 也是 jar 包，与工具 jar 包不同的是其在 jar 包中的 MANIFEST.MF 文件中增加了插件描述信息。为了降低系统运行负荷及便于工具、插件的存储、调用和关联，知识组织体系相关工具集成服务系统采用文件系统和数据库相结合的存储方式，利用文件系统存储工具源码压缩包或 jar 包、工具运行需依赖的第三方 jar 包、插件 bundle 等，利用数据库存储工具和插件的相关描述信息，包括插件编号、安装运行环境、说明文档，插件状态信息、用户注册信息、系统日志信息、系统资源信息、用户信息、统计信息等。此外，工作流是插件组配形成的流程，集成服务系统并不对构成工作流的多个插件进行再次存储，而是记录工作流中插件的组配信息，包括插件编码、组配顺序、工作流描述说明、工作流执行状态及中间结果等。

3.3 服务模式

插件服务化是 STKOS 相关工具集成服务系统实现的重要目标之一，系统除提供工具、插件的检索、浏览、分类、统计等服务外，对于工具及插件的使用主要提供如下三类服务模式，其中包括离线和在线方式，用以满足用户的不同使用需求。

(1) 插件在线服务：用户可准备和上传符合系统要求的规范格式数据，在线调用插件完成特定的数据处理需求，系统提供插件运行结果的在线显示和下载服务。

(2) 基于工作流的插件在线服务：根据插件功能，并提供按照用户特定需求，进行插件的工作流程组配、建模，形成有效的工作流。实现用户所需的功能，建立规范的数据输入、输出接口，用户通过输入和配置规范数据格式，完成所需的数据处理。系统提供 workflow 功能节点的插件执行结果的显示、修改和下载。

(3) 工具及插件的下载服务：向授权用户提供工具和插件的下载，并提供工具及插件使用的相关运行环境、配置信息、第三方 jar 包以及资源等用于离线运行。

4 功能模块

STKOS 相关工具集成服务系统的功能框架如图 2 所示，主要包括工具管理模块、插件封装模块、插件管理模块、插件运行模块、workflow 模块以及基础模块。

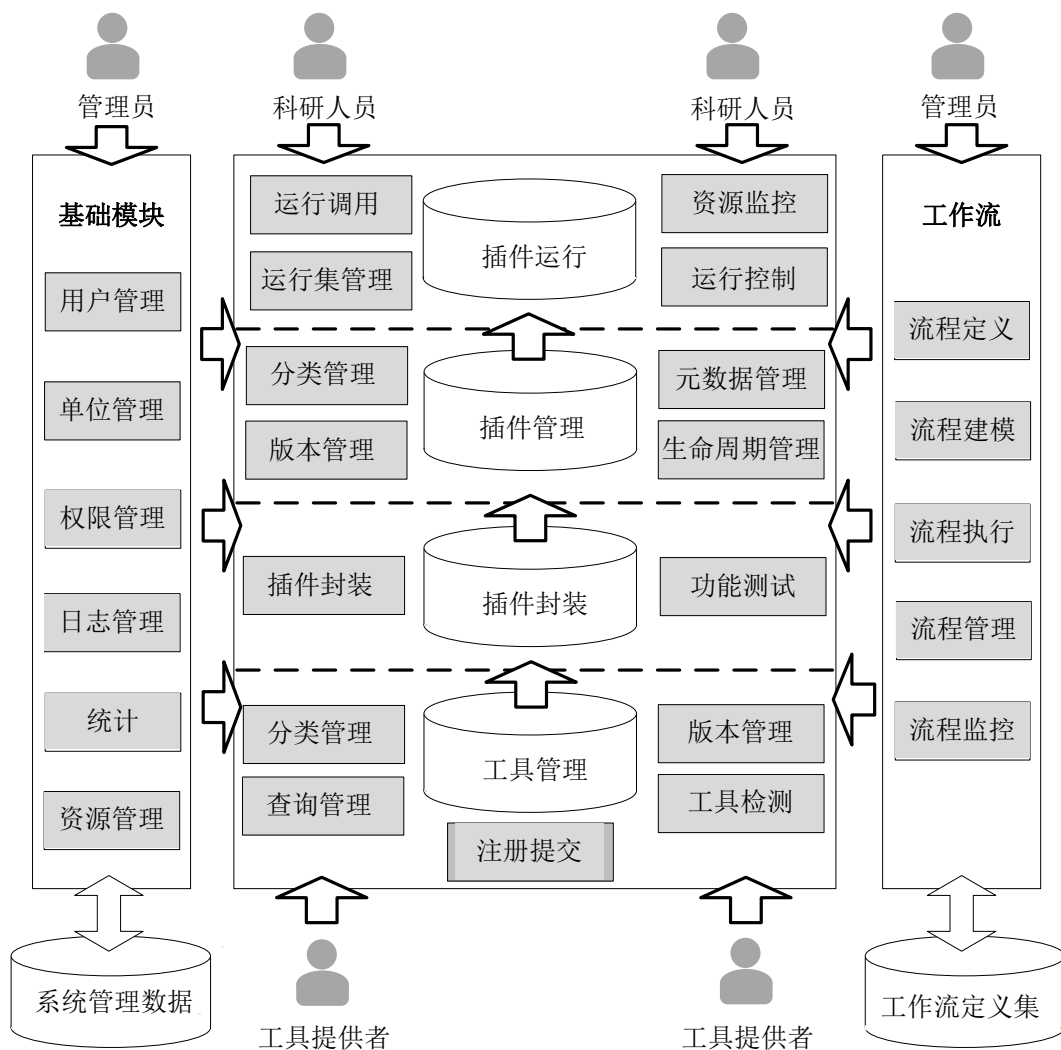


图2 系统功能框架

(1) 工具管理模块

工具管理模块主要实现包括知识组织工具的注册提交、查询管理、工具检测、分类管理、版本管理等功能。

①注册提交，STKOS 相关工具集成服务系统为各工具提交者提供工具存储的入口，提交内容包括工具描述和工具包。工具描述的元数据项，包括工具名称、功能描述、所属机构、运行环境、开发环境、运行参数、工具版本号等，工具有需要上传的工具源码压缩包或 jar 包、工具运行依赖的第三方 jar 包或所需资源。

②查询管理，用于对已提交的工具进行查询。

③工具检测，STKOS 相关工具集成服务系统提供包括关键词、工具分类及工具提交单位等不同检索点及组合检索点的插件检索功能。

④分类管理，按实现功能进行工具分类。

⑤版本管理，管理工具的历史版本。

(2) 插件封装模块

考虑到工具的复杂性，STKOS 相关工具集成服务系统中的插件封装功能分为自动封装和手动封装。对于粒度小、功能独立的工具进行自动封装，而对于需要开放特定接口或复杂的工具则需要提交管理员进行手动封装。工具封装需要解析工具包结构，对工具接口进行规范，将插件依赖信息写入元数据并重新编译打

包。封装后对插件提供在线功能测试，确保插件可用，最后将封装后的插件存储到插件库中，插件封装流程如图 3 所示：

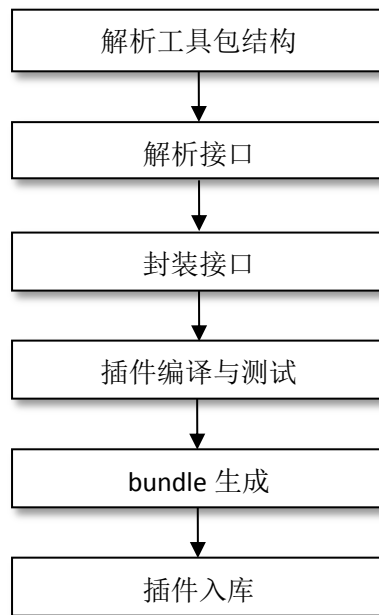


图 3 系统插件封装流程

①解析工具包结构：工具提供者提供的工具以 jar 包形式上传到系统中，STKOS 相关工具集成服务系统实现 jar 向 bundle 的转换，首先需要对原有 Jar 包解压，这是 jar 的接口解析与封装的准备条件。

②解析接口：利用 Java 反射技术，对 jar 的类、函数进行解析，并与上传 jar 包时所填写的执行类、函数信息进行验证。

③封装接口：把解析后的 jar 集成到 bundle 模板中(bundle 模板是符合 bundle 所要求的最简洁封装形式，本身并不具有任何的功能函数)，并已经解析后的接口在 bundle 模板重新声明，并以直接调用 jar 包相应的函数，生成新的 MANIFEST.MF 文件，从而实现原有 jar 函数在 bundle 中的封装。

④编译与测试：对生成的集成了 jar 的 bundle 源码进行编译，并重写包中的 MANIFEST.MF 文件中的信息，如 Bundle-Name 等信息检测。

⑤bundle 生成：实现 bundle 源码的打包，重新生成 jar 包，即 bundle。

⑥插件入库：生成后的 bundle 入插件库进行管理，通过 bundle 容器进行启动、停止等控制，从而实现 bundle 的在线运行服务。

(3) 插件管理模块

插件管理模块实现插件的分类管理、版本管理、元数据管理和生命周期管理功能。

①分类管理，插件封装后存入插件库，STKOS 相关工具集成服务系统提供按插件功能的分类管理。

②版本管理，管理插件的历史版本。

③元数据管理，管理插件描述的元数据信息，包括插件名称、版本号、插件类型、运行环境、调用的类名、方法名、传递参数、封装时间等用于插件表示和展示。

④生命周期管理，系统还提供插件的在线发布、插件的生命周期管理，包括插件的启动、解析、运行、卸载、停止的状态管理。

(4) 插件运行模块

STKOS 相关工具集成服务系统框架监控插件运行中的状态、服务的注册、发现和绑定情况，如图 4 所示，能够对运行中的插件实现数据加载和调用，控制插件的停止、启动，并监控插件运行对系统资源的消耗。

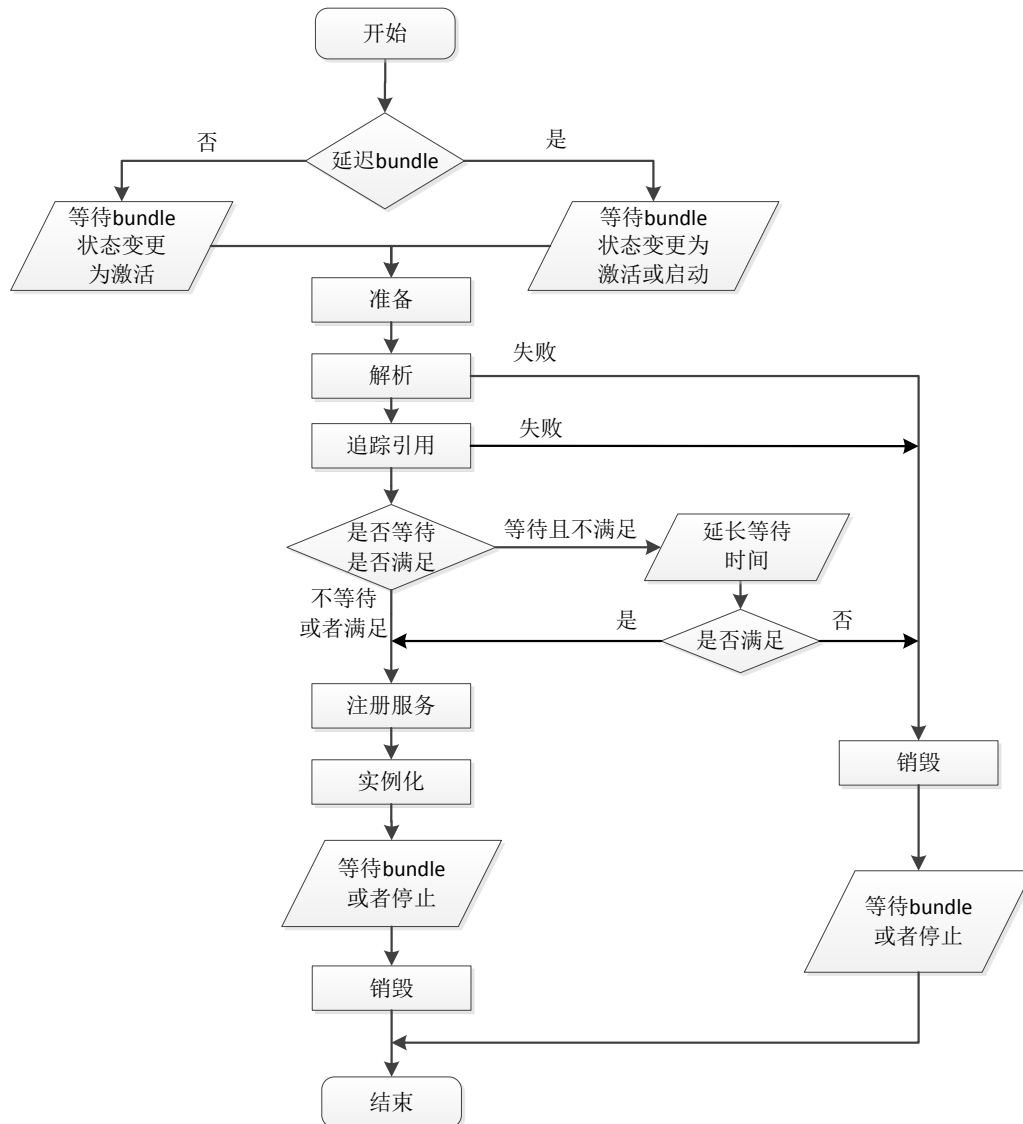


图 4 bundle 运行模块机制

在 Java 虚拟机环境中，类的生命周期开始于虚拟机的加载，终止于被提前卸载或虚拟机停止。STKOS 相关工具集成服务系统中，系统通过实现接口，并在 bundle 的 MANIFEST.MF 文件中获取 bundle 在 OSGi 环境中运行的上下文。当 bundle 被启动或停止时，OSGi 框架将该 bundle 的运行上下文注入到该实现类实例中，系统可以在该 bundle 的运行过程中使用该上下文访问 OSGi 环境信息，访问处于 OSGi 环境中的其他 bundle，注册事件监听，发布或引用服务等。并且，系统实时监控插件运行中的系统资源消耗情况，包括 CPU 和内存的占用情况。

(5) workflow 模块

workflow 模块包括流程定义、流程建模、流程执行、流程管理和流程监控。

①流程定义，主要是对 workflow 进行命名和描述流程功能。

②流程建模，基于开源软件 jBPM (Java Business Process Management) [19]

构建基于可视化的 workflow 模板，对 jBPM 进行本地化改进，在插件面板中引入由本系统封装、分类的插件，并附有每个插件的输入、输出的详细说明信息。可通过在流程建模面板内的插件拖拽实现多个插件的组配，构建 workflow 模板，系统提供模板的保存和修改。

③流程执行，支持启动执行 workflow，包括全过程一键执行和按节点分步执行。系统可以显示 workflow 中的每个功能节点的运行结果，在分步执行时，可对作为本功能节点输入的上一功能节点的结果进行调整和修正。

④流程管理，支持对 workflow 进行发布、删除、编辑、修改。

⑤流程监控，监督每个 workflow 的功能节点的插件生命周期状态。

流程执行过程如图 5 所示，用户执行 workflow 时，系统从 workflow 库中读取 workflow 定义及配置信息，创建工作流引擎实例，启动执行 workflow 引擎，按节点执行，返回每一步或最终的结果。

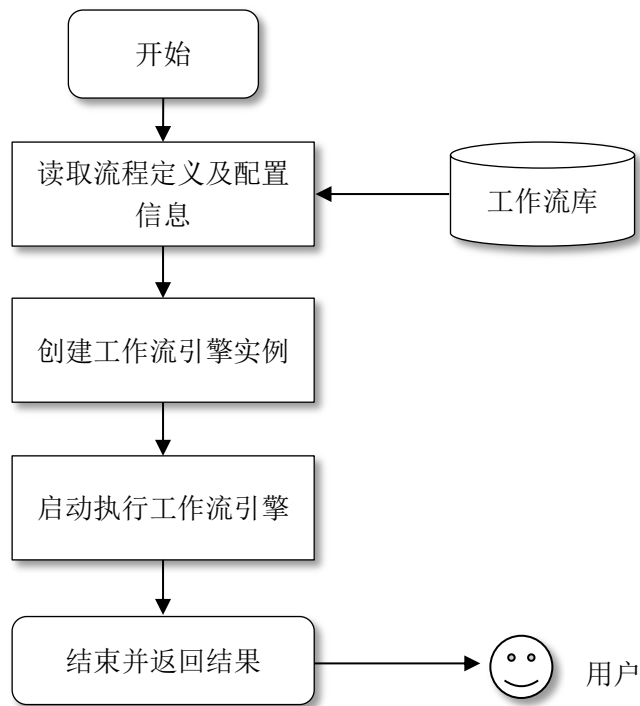


图 5 流程执行过程

(6) 基础模块

实现 STKOS 相关工具集成服务系统的基础业务功能，包括用户管理、单位管理、权限管理、日志管理、统计及资源管理。

①用户管理主要涉及对用户的用户名、密码和权限的管理。

②单位管理对使用系统资源的单位提供单位名称的添加、编辑、修改、删除。

③权限管理对使用系统资源的所有用户进行权限分配和管理。

④日志管理实现对系统操作情况的记载，便于对发生问题的回溯和问题跟进。

⑤统计模块实现系统对工具或插件的收录存储、使用、下载情况的统计分析。

⑥资源监控，对用户创建和使用插件所占用的资源（包括 CPU 使用率、内存占有率及数据库的连接数等。）进行监控，对资源占用过大的节点进行调整，减少过度占用资源。

5 实现效果

本文基于前文提出的系统总体框架及其功能模块的设计思路，实现了 STKOS 相关工具集成服务系统，主界面如图 6 所示。该系统实现了工具存储、插件封装、插件仓储以及 workflow 组配等一系列功能。用户可以进行工具、插件和工作流的检索和浏览。



图 6 STKOS 相关工具集成服务系统界面

进入插件浏览界面（如图 7 所示），可浏览、检索系统中封装的插件。左侧面板可按功能分类或来源工具提交单位浏览插件，右侧为浏览插件的运行环境、版本号、封装时间等信息。点击相应的插件，可看到插件的详细描述信息。图 8 给出了名称为“trans”的英译汉插件的展示页面，用户可在线使用该插件。在“参数”栏输入需要翻译的英文文字，启动 trans 插件后即可得到相应的中文译文。获得授权的用户也可下载该插件或原始工具。

NS TL 科技知识组织体系相关工具集成服务系统
Integrated Service System for STKOS Related Tools

首页 工具 插件 workflow 统计 关于

分类统计

- 分词(2)
- 词形规范(2)
- 词频统计(1)
- 词性标注(0)
- 词表校验(0)
- 翻译(3)
- 聚类(1)

单位统计

- 中国科学院文献情报中心(0)
- 中国农业科学院农业信息研究所(0)
- 中国科学技术信息研究所(1)
- 中国医学科学院医学信息研究所(8)
- 其他单位(0)

搜索

数据列表

--插件类型-- --单位查询--

工具名称	分类	运行环境	提交机构	版本号	封装时间
<input type="checkbox"/> splitword	分词		中国医学科学院医学信息研究所	1.0	2014-05-16 18:13:07.0
<input type="checkbox"/> lovinnew	词形规范		中国医学科学院医学信息研究所	2.0	2014-05-16 18:44:01.0
<input type="checkbox"/> microtranslation	翻译		中国医学科学院医学信息研究所	1.0	2014-05-16 18:49:14.0
<input type="checkbox"/> counthum	词频统计		中国医学科学院医学信息研究所	1.0	2014-06-16 20:58:48.0
<input type="checkbox"/> wekaclustering1	聚类		中国医学科学院医学信息研究所	4.0	2014-06-17 09:08:27.0
<input type="checkbox"/> trans	翻译		中国医学科学院医学信息研究所	5.0	2014-06-17 11:11:20.0
<input type="checkbox"/> splitword2	分词	winidows	中国医学科学院医学信息研究所	1.0	2014-05-28 01:57:27.0

第 1 页 / 共 2 页

首页 上一页 下一页 末页

Copyright(C)2011 国家科技图书文献中心 版权所有 开发单位：中国医学科学院医学信息研究所 中国科学院文献情报中心

图 7 插件浏览界面

NS TL 科技知识组织体系相关工具集成服务系统
Integrated Service System for STKOS Related Tools

首页 工具 插件 workflow 统计 关于

名称: trans 测试

插件描述: 翻译工具

版本号: 5.0

所属机构: 中国医学科学院医学信息研究所 上传时间: 2014-06-17 11:11:20.0

插件类型: 翻译

运行环境: 插件包: micro_fat_5.0.jar

测试: 显示 隐藏

类名称: com.memetix.test.MicroTranslate 方法名: translation

参数:

* 输入字符串

启动

下载: 下载插件 下载原始工具

最新插件

- microtranslation
- lovin
- splitword2
- trans
- wekaclustering1
- counthum
- microtranslation

常用插件

- splitword
- lovinnew
- microtranslation
- counthum
- wekaclustering1
- trans
- splitword2

图 8 trans 插件界面

图 9 给出了系统的工作流建模示例。定义 workflow 名称为 trans_count，功能描述为实现英文翻译为中文，对中文进行分词，对分词后的文本进行词频统计。在系统提供的可视化图形库中分别选择英汉翻译插件 trans 1、中文分词插件 chinese_lexcion 1 以及词频统计插件 count 1，从左侧面板的插件分类中拖拽到右侧面板中实现流程的配置，按插件执行顺序组配完成 trans_count workflow 模板的创建和保存。STKOS 相关工具集成服务系统对给出每个插件的输入、输出、实现功能都给予提示，便于插件的前后组配。

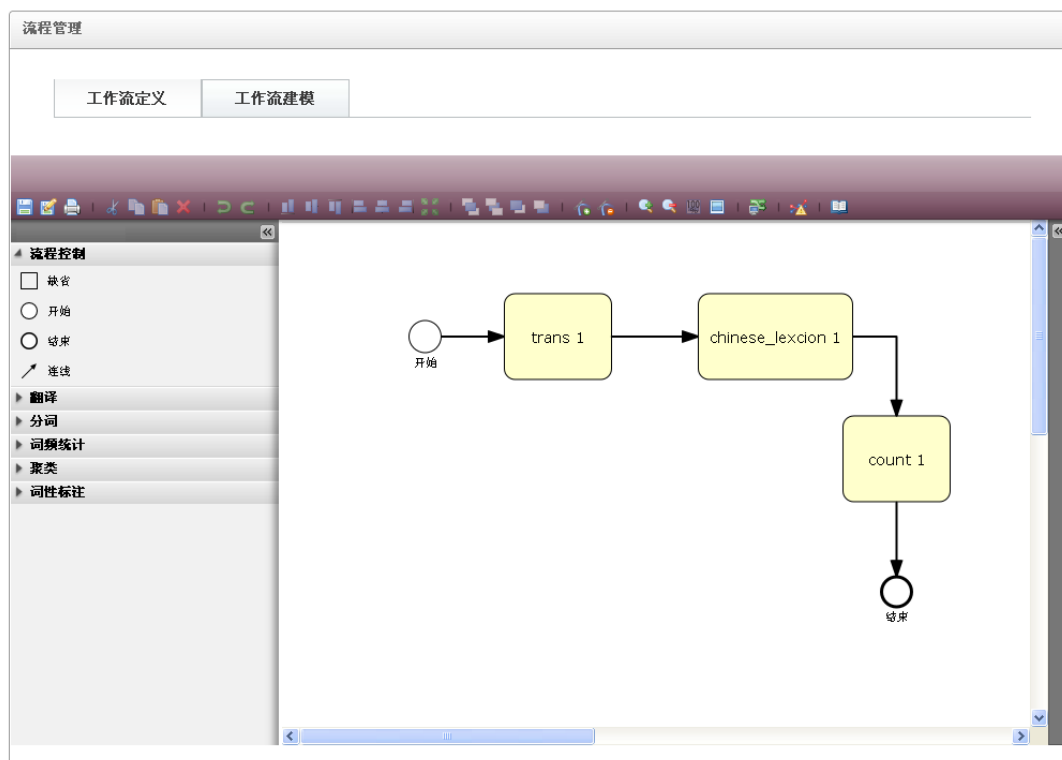


图 9 workflow 建模示例

图 10 给出了翻译分词统计的 trans_count workflow 运行过程。在文本框中输入一段英文文字，经过翻译英译汉插件 trans 1，输出相应的中文翻译文本，该文本作为插件 2 的输入。通过中文分词 chinese_lexcion 1 插件，插件 2 对插件 trans 1 的输出文本进行分词，该分词结果作为插件 3 词频统计 count 1 的输入。从该功能显示的结果可以看出，在 STKOS 相关工具集成服务系统中，前一个插件的输出可以作为下一个插件的输入，系统提供对每一步执行结果的修改和调整。系统支持对 workflow 运行结果的保存。



图 10 trans_count workflow 运行过程

6 结 语

综上所述，本文实现了 STKOS 工具集成服务系统的主体功能，完成了 STKOS 相关工具的仓储及对外发布系统建设，基于插件的知识组织工具集成框架和组件建设，基于工作流的知识组织工具集成三方面的建设内容，形成了能对来自不同课题、实现不同功能的知识组织相关工具进行组织、集成和发布的系统。下一步，课题组将进一步完善和细化系统功能，进行系统性能测试，加强系统的友好性建设，为用户提供细致、友好的使用文档，协调项目相关单位进行工具存缴。另外，STKOS 相关工具集成服务系统并非独立存在，而是科技知识组织体系共享平台的一个重要组成部分，因此，需要考虑与大平台的对接，包括用户管理、权限管理、角色分配等，为用户体提供更好的使用体验。

参考文献

- [1] <http://www.nlm.nih.gov/research/umls/>
- [2] <http://eurovoc.europa.eu/>
- [3] <http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>
- [4] Building an Integrated Formal Ontology for Semantic Interoperability in the Fishery Domain[EB/OL].[2014-11-24].<ftp://ftp.fao.org/docrep/fao/008/af242e/af242e00.pdf>.
- [5] Aldo Gangemi, Domenico M. Pisanelli, Geri Steve. An Overview of the ONINS Project: Applying Ontologies to the Integrating of Medical Terminologies [J]. Data & Knowledge Engineering - Special issue on formal ontology and conceptual modeling archive, 1999, 31(2): 183-220.
- [6] <http://naist.cpe.ku.ac.th/agrovoc>

- [7] <http://metamap.nlm.nih.gov/>
- [8]王兴兰, 宋文. 基于知识组织体系的自动分类研究[J].图书馆论坛,2013,33(6):8-13. (Wang Xinglan, Song Wen. Automatic Classification Based on Knowledge Organization System[J]. Library Tribune, 2013,33 (6): 8-13.)
- [9]王莉, 梁冰, 白海燕.以数据空间理念建立关系发现应用——NSTL 智能检索平台的实践[J].数字图书馆论坛, 2014,(6): 48-54.(Wang Li, Liang Bing, Bai Haiyan. Relationship Discovery Based on the Concept of Dataspace - A Case Study of NSTL Intelligent Retrieval Platform[J]. Digital Library Forum, 2014,(6): 48-54.)
- [10]刘毅, 汤怡洁, 周子健, 等. 科技知识组织体系共享服务平台服务接口建设研究[J]. 现代图书情报技术, 2014,(7-8):9-16. (Liu Yi, Tang Yijie, Zhou Zijian, et al. Research and Construct of the Service Interface in STKOS Sharing Infrastructure[J]. New Technology of Library and Information Service, 2014,(7-8): 9-16.)
- [11]孙坦, 刘峥. 面向外文科技文献信息的知识组织体系建设思路[J]. 图书与情报,2013(1):2-7. (Sun Tan, Liu Zheng. Methodology Framework of Knowledge Organization System for Scientific & Technological Literature [J]. Library & Information, 2013,(1): 2-7.)
- [12]Morel J M, Faget J. The REBOOT Environment[C]. In: Proceedings of the 2nd International Workshop on Software Reusability, 1993: 80-88.
- [13]杨芙清, 梅宏, 李克勤, 等. 支持构件复用的青鸟III型系统概述[J]. 计算机科学,1999,26(5):50-55. (Yang Fuqing, Mei Hong, Li Keqing, et al. An Introduction to JB3 System Supporting Component Reuses[J]. Computer Science, 1999,26(5):50-55.)
- [14]Ye Y, Fischer G. Supporting Reuse by Delivering Task-Relevant and Personalized Information[C]. In: Proceedings of the 2002 International Conference on Software Engineering, Orlando, FL, USA. 2002:513-523.
- [15]Eclipse[OL],[2014-09-01].<http://www.eclipse.org>.
- [16]尹铨. 基于插件体系结构的软件生产线集成框架设计与实现[D]. 北京: 国防科学技术大学,2008. (Yin Quan, Design and Implementation of Integrated Framework of Software Production Line Based on Plugin Architecture[D]. Beijing: National University of Defense Technology, 2008.)
- [17]Larck: The Large Knowledge Collider[OL],[2014-09-01].<http://www.larkc.eu>.
- [18]McAffer J, VanderLei P, Archer S. OSGi and Equinox: Creating Highly Modular Java Systems[M]. Addison-Wesley Educational Publishers, 2010.
- [19]jbpm[OL],[2014-09-01].<http://www.jbpm.org/>.

(通讯作者: 方安 E-mail:fa.an@imicams.ac.cn)
ORCID:0000-0002-9526-9306