

### 推荐国家科技进步奖项目公示

|  |                     |
|--|---------------------|
| 项目名称   | 基于知件的知识获取、管理和知识服务平台 |
| 推荐单位   | 教育部                 |
| 推荐单位意见： <p>我单位认真审阅了该项目推荐书及附件材料，确认全部材料真实有效，相关栏目均符合国家科学技术奖励工作办公室的填写要求。</p> <p>该项目在国际上率先提出“知件”这一全新的概念，作为知识系统的基础元素，并以知件为核心，知识中间件为支撑，知件获取为突破口，基于知件的知识服务为目标，在知识内容模块化、知件获取在线化、知件推理目的化和知识服务个性化等方面取得突破性进展。</p> <p>在 2013 年教育部组织的科技成果鉴定中，鉴定委员会一致认为：“该成果在基于知件的知识获取、管理和服务等方面取得了突破和创新，在知识工程的若干关键技术研究上与国际前沿同步，研究成果具有原创性，在基于知件的按需知识模块动态创建方面达到国际领先水平。”</p> <p>该项目出版英文专著 2 部，发表论文 8 余篇，做特邀报告 40 余人次，获 8 项授权国家发明专利，7 项软件著作权，并获教育部高等学校优秀科技成果一等奖。</p> <p>该项目成果在东软熙康健康管理等不同行业知识系统构建中得到集成和应用，累计新增销售额约 2.5 亿元，新增利润约 1488 万元。目前，基于该成果构建的东软熙康健康管理服务平台，创建和管理了 30000 多个领域知件，服务覆盖国内 30 个省（含直辖市、自治区）、20000 多家基层医疗服务机构，服务人群近 3000 万人，取得巨大社会效益。</p> <p>对照国家科学技术进步奖授奖条件，推荐该项目申报 2017 年度国家科学技术进步奖二等奖。</p> |                     |

## 项目简介:

基于知识的系统（知识系统）指利用知识库并提供推理能力以支持问题求解，是软件智能化的重要途径。传统知识系统的典型代表是专家系统，由于知识源匮乏易导致知识获取瓶颈，不同知识系统间互操作困难易形成知识孤岛，其规模化和个性化都受到限制。互联网海量信息提供了丰富的知识来源，但如何有序地组织知识、有目的地获取和适时更新知识、以及有效地利用知识仍然是目前面临的挑战。

该成果完成人陆汝钤院士在多年知识系统研究经验基础上，提出“知件”概念和将“知件”与“硬件”、“软件”作为信息技术三大基石的理念，旨在刻画智能化软件的知识特征。该成果以知件为核心，知件获取为突破口，知件中间件为支撑，基于知件的问题求解为目标，在知识内容模块化、知件获取在线化、知件推理目的化和知识服务个性化等方面取得突破性进展，为知识服务平台和智能化软件的构造提供关键技术支撑和实践指南，为知识产业奠定基础。主要科技创新：

- 在国际上率先提出“知件”概念，揭示知识系统中知识的有目的构建性，为海量信息环境下知识的有序组织、有效利用和商品化奠定基础。创造性地提出了基于知件中间件的知识系统框架，明确了知识系统中知件、知件中间件和软件的互操作关系，建立了知识系统的开发规范和实践指南。
- 系统地发明了一系列离线和在线知件获取方法，多种方法的技术指标位于国际前列，发明了基于知识规约和知识编排的知件融合技术，支持知件的按需获取和融合，提出了基于本体比对和一致性度量和维护的知件质量保障方法，解决了知件的有效获取、适时更和质量保障的问题。
- 发明了多种基于知件的知识服务模式，包括交互式知识浏览，个性化定制式知识综合，基于知件的问题求解，知件制导的智能软件架构等。研制了基于知件的知识获取、管理和知识服务平台，和面向领域的东软SaCaDeepCogni知识服务平台，解决了按需知识服务及其平台构建问题。

该成果形成英文专著2部，授权发明专利8项，在IJSEKE、IEEE IS、IJAR、Science in China等知名刊物，和AAAI、ACL、KSEM等著名国际会议上发表论文80余篇，成果完成人受邀在国内外主要论坛和学术会议上做特邀报告40余次。

该成果在东软熙康健康管理平台构建等项目中得到全面实施，有效提高了知识获取的效率和数量，规范了知识系统的构建。目前熙康健康管理平台已累计创建和管理30000多个领域知件，服务覆盖国内30个省（含直辖市、自治区）和20000多家基层医疗服务机构，服务人群近3000万人，取得巨大社会效益，受到国家领导人的高度关注。该成果整体直接应用于平安银行等20多家单位的知识系统构建和智能系统知识库构建项目中，近三年直接新增销售额5033万元，新增利润517.2万元。用于支持航天光达、江苏蓝安、上海汇和等企业完成行业知识系统的构建，并使其智能化程问题求解结果的准确性等指标达到行业实际应用要求，自2013年来累计新增销售额14312.8万元，新增利润743.6万元。

周巢尘院士和吕建院士等专家组成的鉴定委员会认为：该成果在基于知件的知识获取、管理和服务等方面取得了突破和创新，在知识工程的若干关键技术研究上与国际前沿同步，研究成果具有原创性，在基于知件的按需知识模块动态创建方面达到国际领先水平。该成果获2013年度教育部高等学校优秀科技成果科技进步一等奖。

## 客观评价：

- AIWATCH杂志发表评论指出这是他们所见到的“以领域知识库支持主流软件开发的第一例”。杨芙清院士在为专著《Domain Modeling based Software Engineering》作的序中，称该研究成果“不仅具有重大的理论创新和突破，而且是对软件工程、CASE 和软件自动化的重要贡献”
- 意大利Polimettrica 出版社的“Best Innovation”丛书评价专著《Knowware, The Third Star after Hardware and Software》：“The concept of knowware is original, innovative and topical. I think that many researchers, CEOs (and also the general public) will be interested in this book”
- 所提出的本体匹配方法被华盛顿大学等采用，作为医学本体知识库的每周自动检测工具，并发现了其中的多处错误。本体匹配系统FCA-Map在国际本体匹配比赛（the Ontology Alignment Evaluation Initiative 2016 Campaign）中获得疾病和基因类本体匹配奖（共三个系统获奖）。国际上多名研究者引用了关于本体匹配的论文，并称论文是“首先提出把大型解剖学本体转换为OWL语言的工作”，提到将使用了论文“提出的规则进行框架系统的变换”，将根据“提出的方法做进一步的工作”。
- 针对所提出的基于树卷积神经网络的自然语言推理模型，NAACL2016的论文《Dependency Based Embedding for Sentence Classification Tasks》进行了引用和评价称该项目提出的“树卷积神经网络模型是对问题语句进行分类的最新模型；取得了当前的最好结果。”《MGNC-CNN: A Simple Approach to Exploiting Multiple Word Embeddings for Sentence Classification》进行了引用并给出评价“树卷积神经网络模型以96%的准确性，取得了当前的最好结果。”
- 针对所提出的基于LSTM的关系分析模型，论文评审人明确指出，“相比于其他方法，该方法取得了当前最好的分析结果”，ICLR2016论文《Towards Universal Paraphrastic Sentence Embeddings》进行了引用指出“该方法所使用的的LSTM模型是对序列信息分析最有效的模型”；The Journal of Biological Databases and Curation上发表的论文《Exploiting syntactic and semantics information for chemical-disease relation extraction》引用该成果并指出“该方法能够有效获取自然语言语句中的语义信息。”
- 针对所提出的基于树卷积神经网络的语义推断模型，论文评审人指出，“这篇论文显示了基于树型结构的分析方法在自然语言语义分析方面所具有潜力”，并指出，“该方法可以用来有效识别具有相似语义的自然语言语句”。在ACL2016上，来自斯坦福大学的学者Sam Bowman在论文交流中明确指出，“我非常喜欢这个模型，在我看来，这个模型是目前基于对语句模型进行向量求取所取得的最好结果。这也是一项我正想要进行的研究工作”。
- 针对所提出的基于稀疏词向量表示的压缩自然语言模型，2016年09月22日，在由“阿里技术”发表的博文《深度学习即将攻陷的下一个领域：NLP—ACL2016优秀论文解读》中，重点应用了该成果的论文，并指出：该方法“针对现在神经网络模型过大带来的时间开销以及内存占用的问题，提出了一种压缩算法使得模型的参数伴随着词典的膨胀缓慢增加，并在Perplexity上的评估性能也有所改善。最终的实验结果从性能（Perplexity）与内存开销减少两方面说明了价值。”
- 中国科学院院士、著名计算机科学家周巢尘研究员，中国科学院院士、著名计算机软件专家吕建教授，著名计算机软件专家王怀民教授等组成的鉴定专家委员会一致认为：“……该成果在基于知件的知识获取、管理和服务等方面取得了突破和创新，在知识工程的若干关键技术研究上与国际前沿同步，研究成果具有原创性，在基于知件的按需知识模块动态创建方面达到国际领先水平。”

## 推广应用情况：

通过与经典知识系统开发方法和技术的对比分析，该成果对知识系统构建和知识工程在海量信息环境下的发展，具有极大的推动和促进作用。具体体现在：

- 提出了一种新的大规模知识库的构建方法，揭示知识模块化（知件）和知件封装的机理，提出从知识的内容、形式和目的三个维度确定知件边界；发明了基于模型（和模型驱动）的知件开发方法，极大提高知识系统的开发效率，降低知识系统开发的门槛，推动知识系统的普及。
- 提出知件工程和知件/软件协工程开发方法和原理，依照知识构建理论，提出了知识开发的熔炉模式和知识开发的结晶模式，并通过发明不同知识获取技术诠释了知件工程化开发模式，推动了知识工程向知件工程的发展，孕育新型信息技术产业——知识产业，并为知识产业的建立提供了一套可行的技术路线。

该成果已经在中国最大的基层健康医疗服务平台“熙康健康管理服务平台”的建设中得到直接应用，“熙康健康管理服务平台”利用该成果共创建30000多个领域知件，知件库及相关智能推荐服务系统、健康风险预测系统成为“熙康健康管理服务平台”相对于其他互联网医疗服务商的差异化竞争优势，使得东软熙康的健康管理服务业务持续保持在国内互联网医疗健康市场的领先地位。目前该平台服务已覆盖国内30个省（含直辖市、自治区）、20000多家基层医疗服务机构，服务人群总数近3000万人。熙康健康管理服务平台促进了东软熙康公司医疗健康服务的发展，2015年业务收入1.08亿，2016年预计业务收入2.2亿，实现100%的增长，获得良好的经济效益和社会效益。习近平主席、李克强总理都曾现场考察过东软熙康的“健康城市”、“云医院”等项目，对东软熙康所构建的智慧健康云平台给予了高度关注和肯定。

该成果自2014年以来还整体直接应用于平安银行、中国日报、央视新闻、台达电子、中国大百科全书、中国文联等20多家单位的知识系统构建和智能系统知识库构建项目中，近三年直接新增销售额5033万元，新增利润517.2万元。

该成果研制了一组支持知件工程的关键技术及其支撑工具，这些工具已经直接得到实际应用，授权国家技术发明专利8项。所提供的方法、平台工具和知识处理技术用于支撑航天光达重大事件预警、江苏蓝安士兵心理评测、上海汇和教育软件等不同行业和类型的知识系统开发和知识服务平台建设中，提高了知识系统的开发效率，降低了知识系统的开发门槛，同时使系统的智能化程度和问题求解的准确性和针对性都达到了实际应用的水平，取得显著经济和社会效益。自2013年以来，累计间接新增销售额14312.8万元，新增利润743.6万元。

## 主要知识产权证明目录:

- 授权发明专利, 分类知识获取方法和装置 (ZL201310219725.5); 发明人: 李戈, 金芝, 牟力立, 邢雪源; 授权日: 2016年5月18日
- 授权发明专利, 创建物联网控制系统的方法及装置 (ZL201310218328.6); 发明人: 金芝、李戈, 鲁扬扬、邢雪源、郝逸洋; 授权日: 2016年8月10日
- 授权发明专利, 应用建模工具展现本体的方法和装置 (ZL201310218408.1); 发明人: 金芝, 李戈, 邢雪源, 施心悦; 授权日: 2016年4月20日
- 授权发明专利, 基于受力情况的元素排布方法 (ZL201310219421.9); 发明人: 李戈、金芝、郝逸洋、邢雪源; 授权日: 2016年7月13日
- 授权发明专利, 语义Web应用中检测DoS攻击的方法与系统 (ZL201310751402.); 发明人: 陈德彦、张霞、赵宏、赵立军、平安、蒋理成; 授权日: 2016年11月2日
- 授权发明专利, 基于混合主题的文本标注方法及系统 (ZL20131069189.9); 发明人: 王勇、赵立军; 授权日: 2016年6月2日
- 授权发明专利, 程序注释方法和装置 (ZL20131028270.5); 发明人: 李戈; 金芝; 郝逸洋; 张路; 授权日: 2016年8月10日
- 授权发明专利, 物联网设备的控制方法、装置和物联网 (ZL201210110369.9); 发明人: 李戈、金芝、鲁扬扬、邢雪源、郝逸洋; 授权日: 2014年6月4日
- 计算机软件著作权, 北京大学知件库系统 (2013SR070705); 权利人: 北京大学; 登记时间: 2013年7月22日
- 计算机软件著作权, 东软SaCa Deep Cogni知识服务平台软件 (2016SR072253); 权利人: 东软集团股份有限公司; 登记时间: 2016年3月20日
- Ruqian Lu and Zhi Jin, Domain Modeling based Software Engineering, Kluwer Academic Publishers, 2000
- Ruqian Lu, From Hardware to Software to Knowware: IT's Third Liberation? IEEE Intelligent Systems, Vol.20, No.2, 82-85, 2005
- Ruqian Lu, Knowware: The Third Star after Software and Hardware, Polimatrix Publishers, 2007
- Ruqian Lu and Zhi Jin, From knowledge based software engineering to knowware based software engineering. Science in China Series F: Information Sciences 51(6): 638-660, 2008

### 主要完成人情况：

1. 金芝，排名1，教授，工作单位：北京大学，完成单位：北京大学，是该项目主要负责人，对创新点一、二、三均有重要贡献。对技术创新点一中的基于知识的软件工程有创造性贡献，对技术创新点二知件载体和知件工程模型研究，技术创新点二中的知件获取方法研究，技术创新点三中的知识服务模式的研究等有重要贡献。合作撰写专著1部，是48篇代表性论文的作者，负责或参与获得发明专利授权6项。
2. 陆汝钤，排名2，院士，研究员，工作单位：中国科学院数学与系统科学研究院，完成单位：中国科学院数学与系统科学研究院，对创新点一、二、三均有重要贡献。对技术创新点一的基于知识的软件工程研究、知件概念的提出、以及基于知件的知识产业的规划，技术创新点二中知件工程模型研究，技术创新点三中的知件服务模式的研究等有重要创造性贡献。撰写和合作撰写专著2部，在30多个国际国内大会上做特邀报告，是18篇代表性论文的作者。
3. 张霞，排名3，高级副总裁，教授，工作单位：东北大学，东软集团股份有限公司，完成单位：东北大学，东软集团股份有限公司，对技术创新点一、三均有重要贡献，对技术创新点一中的基于知件的知识服务产业规划，技术创新点三中的知识服务需求调研和知识服务模式研究，以及知件技术的推广应用部署有重要实质性贡献。
4. 李戈，排名4，副教授，工作单位：北京大学，完成单位：北京大学，对创新点二、三均有重要贡献。对技术创新点二中知件载体的研究、知识获取方法和技术研究、知件库管理方法研究，以及技术创新点三中知件服务模式的实现技术研究，平台研制和开发均有重要贡献，是13篇代表性论文的第一或者通讯作者，负责或参与获得发明专利授权6项，负责开发软件工具10个，登记软件著作权6项。
5. 张松懋，排名5，研究员，工作单位：中国科学院数学与系统科学研究院，完成单位：中国科学院数学与系统科学研究院。对创新点一、二均有重要贡献，对技术创新点一中的智能化软件开发模式的研究，以及领域本体的研究，技术创新点二中的知件载体表示和本体匹配算法的研究等有重要贡献。12篇代表性论文的作者。
6. 赵立军，排名6，工程师，工作单位：东软集团股份有限公司，完成单位：东软集团股份有限公司，对技术创新点三有贡献。对知识服务需求调研和知识服务模式研究，以及协调资源在健康管理平台和企业知识社区等领域的应用部署。
7. 崔朝辉，排名7，工程师，工作单位：东软集团股份有限公司，完成单位：东软集团股份有限公司，对技术创新点三有贡献。参与完成特定领域知识服务技术的需求提炼，参与成果在企业知识管理和健康服务平台等领域中的应用。
8. 牟克典，排名8，副教授，工作单位：北京大学，完成单位：北京大学，对创新点二有贡献。对技术创新点二中的不一致知识的处理理论和方法的研究有创造性贡献，是6篇代表性论文的第一作者。
9. 平安，排名9，副教授，工作单位：东北大学，完成单位：东北大学，对创新点三有贡献。参与特定领域知识服务技术需求提炼，参与成果在企业知识管理和健康服务平台等领域中应用。

## 主要完成单位及创新推广贡献:

### (一) 北京大学

作为本成果的第一完成单位,对研究实施所需的人力、物力、实验条件和工作实践等提供保障,积极推动研究的深入和实际应用,促进科技成果转化;承担成果的鉴定、专利申请、软件登记等相关工作,主要贡献如下:

1. 负责本成果研究方案的制定和组织实施;负责组织相关单位定期进行围绕本成果技术方案的专项研讨和责任分工,保证本成果研究的持续深入;
2. 作为成果第一完成单位,完成项目基础原理和方法的研究,负责项目核心技术的需求调研和技术规划,负责并组织实施项目核心技术的研发及应用部署;
3. 联合项目合作单位一起推广项目的理论创新和技术应用成果,取得了良好的经济和社会效益;
4. 对技术创新点一中的基于知识的软件工程、特定领域知识模型的研究,技术创新点二中知件载体和知件工程模型研究、以及不一致知识的处理理论和方法的研究,技术创新点三中的知件获取方法和知件库管理方法的研究,知识服务模式的研究、平台研制和开发等均有重要贡献;
5. 合作撰写专著1部,22篇代表性论文的作者单位,6项授权发明专利的申请单位,负责开发软件工具10个,登记软件著作权6项。

### (二) 中国科学院数学与系统科学研究院

中国科学院数学与系统科学研究院计算机科学研究室长期从事知识工程的理论和方法研究。在本成果中,对研究实施所需的人力、物力、实验条件和工作实践等提供保障。主要贡献如下:

1. 参与组织本成果研究方案的制定和实施,完成项目基础原理和方法的研究,参与实施项目核心技术研发;
2. 对技术创新点一中的基于知识的软件工程研究、知件概念的提出、智能化软件开发模式的研究和领域本体的研究以及基于知件的知识服务产业规划,技术创新点二中知件工程模型研究、知件载体表示和本体匹配算法研究,知识服务模式的研究等有重要贡献;
3. 开展本成果的宣传、推广和技术应用,在30余个国际国内大会上做特邀报告;
4. 撰写或合作撰写专著2部,是54篇代表性论文的作者单位。

### (三) 东北大学

东北大学课题组具有丰富的特定领域软件开发的经验,以及企业管理软件开发的经验。在本成果中,对研究实施所需的人力、物力、实验条件和工作实践等提供保障。主要贡献如下:

1. 参与组织本成果研究方案的制定和实施,为实施项目核心技术研发提供需求;
2. 对技术创新点一中的基于知件的知识服务产业规划,技术创新点三中的知识服务需求调研和知识服务模式研究,以及知件技术在企业知识管理和健康管理平台等领域的应用推广和部署有重要贡献。

### (四) 东软集团股份有限公司

东软集团股份有限公司课题组具有丰富的特定领域软件开发的经验,以及企业管理软件开发的经验。在本成果中,对研究实施所需的人力、物力、实验条件和工作实践等提供保障。主要贡献如下:

1. 参与组织本成果研究方案的制定和实施,为实施项目核心技术研发提供需求;
1. 对技术创新点一中的基于知件的知识服务产业规划,技术创新点四中的知识服务需求调研和知识服务模式研究,以及知件技术在企业知识管理和健康管理平台等领域的应用部署有重要贡献。

## 完成人合作关系说明：

该成果第一完成人金芝教授自1992年起开始与成果第二完成人陆汝钤院士及第五完成人张松懋研究员开展知识工程和非规范知识处理方面的合作，共同承担国家攻关项目，基金委重大和重点项目，以及863项目。第一完成人金芝教授与第二完成人陆汝钤院士共同撰写英文专著，第一完成人金芝教授与第二完成人陆汝钤院士、第五完成人张松懋研究员共同发表论文等。

自2005年起，该成果第一完成人金芝教授与第二完成人陆汝钤院士和第九完成人牟克典副教授开展知识不一致性检测和决策消解方面的合作，共同完成多篇论文，并在国际期刊和会议上发表。

自2008年起，该成果第一完成人金芝教授与第四完成人李戈副教授开展知识获取、知件库开发和隐式知识模型求取等方面的合作，共同获得多项国内授权专利，发表多篇论文，并开展知件库技术和基于知件的知识获取和管理技术的应用推广工作，完成多个领域应用。

自2011年起，该成果第一完成人金芝教授、第二完成人陆汝钤院士与第三完成人张霞教授开展合作，共同进行基于知件的知识获取、管理和知识服务平台的顶层设计和部署规划。第三完成人张霞教授和第六、七、九完成人赵立军、崔朝辉和平安开展知件技术的推广，并完成多个不同行业知识系统构建和健康管理平台等的领域应用。

该成果第一、二、三、四、五、八、九完成人合作完成“基于知件的知识获取、管理和知识服务平台”的研制，并获得2013年度教育部高等学校科学研究优秀成果奖科学技术进步奖一等奖。