

机协同及知识与数据双驱动需求。该阶段的主要活动可分为：应用算法模型及系统开发、应用测试及验证、系统部署和系统维护，活动流程及主要角色分工如图7-5和图7-6所示。

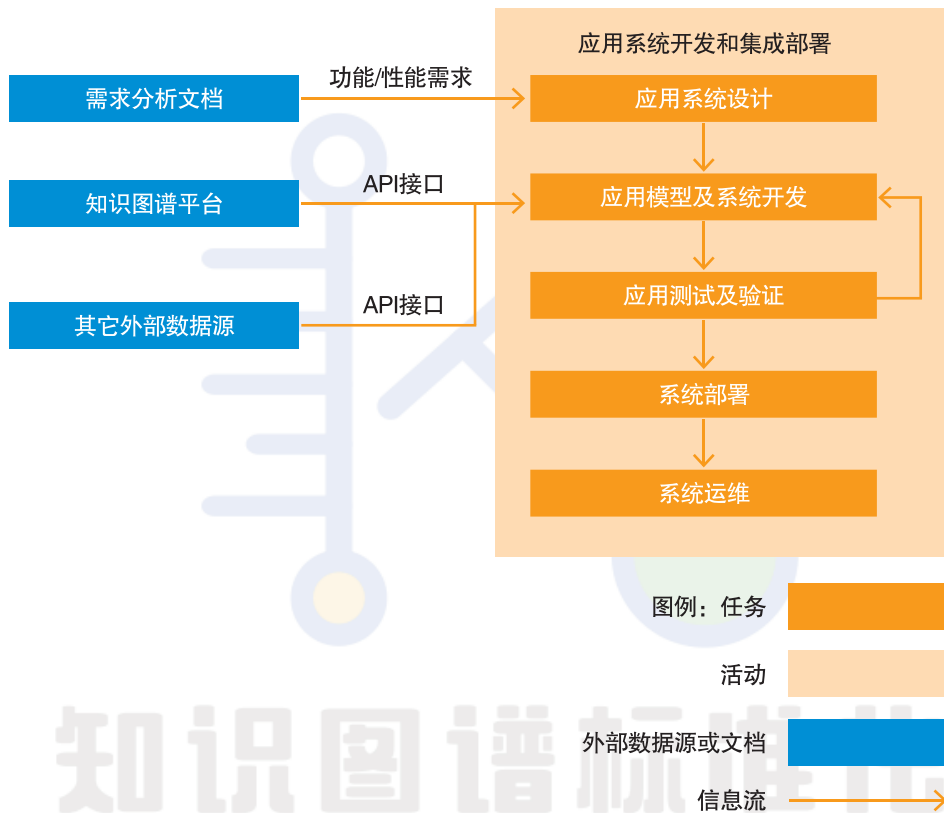


图7-5 应用开发与集成部署阶段活动流程图

应用开发部署活动与角色分工						
	业务需求确认	应用系统设计	应用模型及系统开发	应用测试及验证	系统部署	系统运维
业务分析人员	业务开发细化			系统验证		
开发人员		业务模型设计 应用模块功能设计	算法模型开发 系统功能开发			
知识图谱技术人员		知识计算算法及 接口设计	知识计算开发			
测试人员				系统验证		
系统部署 运维人员					应用部署	应用运维

图7-6 应用开发与集成部署阶段中活动及角色分类

## 1、任务流程构成

### (1) 应用模型及系统开发

基于系统设计文档及要求进行应用模型及系统开发，可包括如下三个方面：

- **应用模型算法开发**：开发满足应用需求的模型算法，并保障业务模型可通过接口的方式对外提供知识计算服务；
- **知识和数据接口开发**：开发对接知识图谱的知识查询以及已有第三方业务系统数据查询的接口，实现知识和数据的接入；
- **知识图谱应用系统开发**：根据设计文档开发应用系统，并通过接口和页面等方式与其它应用业务系统实现集成整合。

### (2) 应用系统测试及验证

知识图谱应用系统作为人工智能应用系统的一类，需同时兼顾软件系统与人工智能系统的测试指标要求。此外，由于应用系统中集成了算法模型，还需测试验证冷启动和模型自我测试能力，保障应用系统中模型的持续迭代和优化。

### (3) 系统部署

本活动通过自动化方法来进行源码编译并部署到目标环境，例如，基于Jenkins和容器化进行持续集成、持续交付和持续部署。该方式可节省运维的人力，也可让系统具备自动水平扩展能力。

### (4) 系统监测

本活动需开发出对知识图谱应用系统各关键状态参数监控的脚本，进而为后续监控其健康程度提供支持，以实现满足知识和数据持续更新要求的应用系统智能运维。

## 2、与其他数据的融合

部分知识图谱应用系统由于在本阶段需完成与外部业务系统间的集成，因此还需关注各活动中与外部数据的融合问题。知识图谱与其他数据的融合可分为两种类型：原始数据的融合和计算数据的融合。其中，原始数据的融合是指从知识图谱和其它数据源中批量获取原始数据，根据数据的内在关联制定融合策略，并实现获取数据的有机整合，如合并为统一的表；计算数据的融合是指在知识图谱上经过知识计算获得衍生知识，并将其附加到其它数据源中作为扩展特征或其他知识图谱的计算结果进行融合。

融合方式可包括接口融合和页面融合。其中，接口融合的方式是将知识图谱开放的知识查询服务、各数据源获取的数据、业务模型的输入输出通过接口的方式进行集成使用，实现知识图谱应用系统与传统业务系统间的集成融合。

#### (四) 角色与分工

本阶段的参与人员角色及其任务简介如表7-9所示：

表7-9 应用开发与集成部署阶段参与人员角色及其任务简介

序号	角色	承担任务简介	交付物名称（如有）
1	业务专家	参与知识图谱应用系统建设及集成效果的评估，并对应用模型开发与迭代、知识服务相关模块功能完善提供建议。	知识图谱应用系统 知识图谱应用系统开发 相关文档
2	需求分析师	针对知识图谱应用系统开发与部署过程中，用户或开发人员提出的需求问题或变更请求等进行沟通，确认真实需求。	需求变更说明书
3	项目经理	对知识图谱应用系统开发与部署过程进行监控，完成风险把控、开发进行控制、工作任务分配及资源调配，对项目的进度、质量、安全等全面负责。	需求变更说明书 技术规格书
4	开发工程师	负责知识图谱应用系统的开发，并参与系统集成。	知识图谱应用系统 知识图谱应用系统开发 相关文档 知识图谱应用系统使用 相关文档
5	测试工程师	完成静态文档审查，根据需求编写测试用例并完成测试，提出测试中发现的问题。	测试用例 测试报告
6	系统架构师	根据开发过程中遇到的问题对系统架构设计、组件规约、功能模块设计进行更新和调整。	知识图谱应用系统功能 清单 知识图谱应用系统设计 说明书
7	系统部署人员	完成有效的代码集成、编译和系统部署，使得整个流程自动化，并且易扩展。此外，可根据应用系统接口获取系统的健康、知识规模等相关信息，对系统状态进行监控。	系统部署相关文档



## 五、系统评估与验收阶段

### （一）阶段建设目标

系统评估和验收阶段需客观和真实地评估知识图谱应用系统的能力水平，保障系统无论在业务层面还是在技术层面都能达到用户预期要求，解决用户问题并顺利上线交付。本阶段的建设目标主要包括：

#### 目标1：完成评估准则的制定。

需要明确定义系统评估涉及的范围和相应的评估指标；完成验收文档清单、项目组织管理制度及风险管控建设，为后续活动提供指导。

#### 目标2：完成评估计划的制定。

需要制定详细的过程评估计划，包括验收计划、实施计划、测试计划、评审计划、风险管理计划等，避免系统建设与验收时出现偏差，导致系统无法按时按质量交付。

#### 目标3：完成系统评估。

依据评估准则和评估计划，进行具体系统评估，从需求分析与设计、系统架构、系统部署、系统实施方案、数据安全等方面，就知识图谱应用系统是否满足相关方的需求进行评审，确保项目执行过程和研发成果与相关方的需求相符合。

#### 目标4：完成系统内部验收。

对待交付上线的知识图谱应用系统开展功能测试、性能测试、安全测试、文档测试等，评估待交付系统符合上线要求，并移交用户正式使用。

### （二）阶段输入输出

#### （1）系统评估与验收阶段输入

本阶段的输入可包括但不限于：

- 待交付的业务系统；
- 业务需求说明书；
- 系统设计原型；
- 系统概要设计文档；
- 系统详细设计文档；
- 系统部署文档；
- 系统集成方案；
- 知识图谱基础产品或服务；
- 业务数据。

### (2) 系统评估与验收阶段输出

本阶段形成的交付物及内容简介如表7-10所示：

表7-10 系统评估与验收阶段交付物清单

序号	交付物名称	简介	重要性 (1-5级)	是否可选
1	系统评估方案	可包括系统评估范围、评估指标等	5	必须交付
2	验收文档清单	验收需要交付的文档列表	5	必须交付
3	项目人员配备情况表	可包括项目组的组织机构、管理制度、人员分工等	3	可选交付
4	内部验收计划	可包括验收时间、验收内容及验收形式等	4	推荐交付
5	评审计划	可包括需评审的内容、评审的方式、评审要求、参与评审等人员等	4	推荐交付
6	测试计划	可包括测试任务及测试质量等		推荐交付
7	算法评测报告	可包括算法评测结果及问题分析说明等	2	可选交付
8	系统质量评估报告	可包括系统功能测试、系统性能测试、系统安全测试等测试结果及整体评估结论等	5	必须交付
9	系统内部验收报告	可包括系统功能验收清单及验收结论等	5	必须交付

### (三) 主要活动

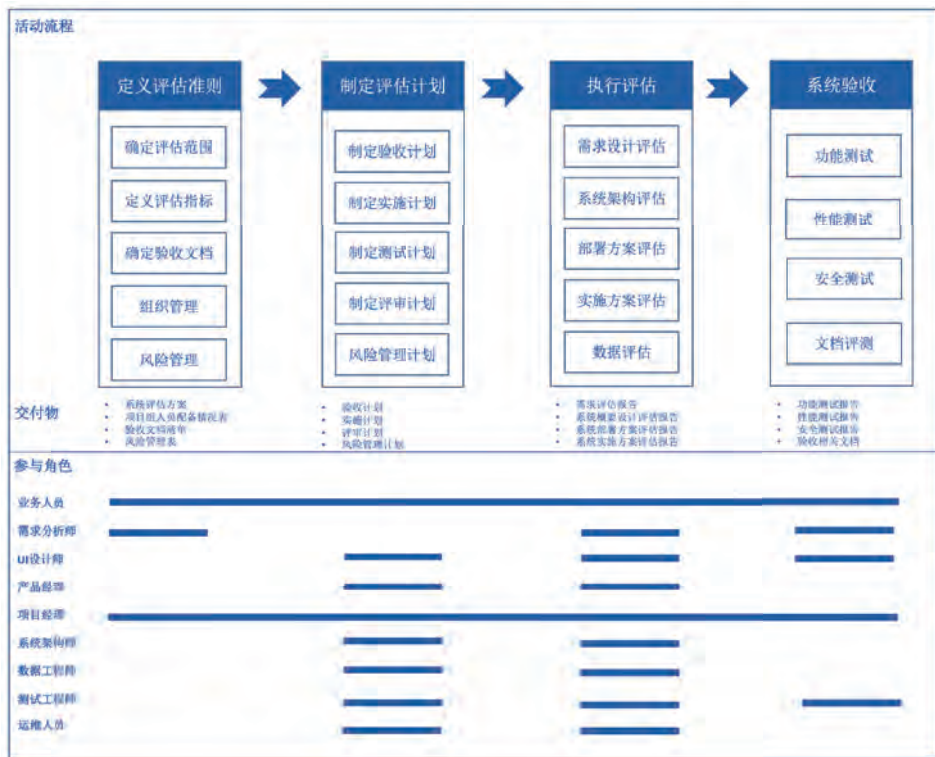


图7-7 系统评估与验收阶段活动流程图

系统评估与验收阶段根据拟实现目标的不同可划分为定义评估准则、制定评估计划、执行评估、系统验收等四项活动，本阶段活动流程及主要参与角色如图7-7所示。各项活动主要内容如下：

#### (1) 定义评估准则

定义评估准则是系统评估与验收阶段的基础活动，主要包括确定评估范围、定义评估指标、确定验收文档、组织管理等四项任务。

- 确定评估范围要求业务人员确定系统评估的维度（如功能、性能、安全、集成、部署等），系统需在确定的评估维度下达到用户的预期。
- 定义评估指标要求明确评估范围的具体量化要求。业务相关人员

需根据不同评估范围给出具体评估指标，如：在性能维度，给出系统吞吐量、最低响应时间、业务UI操作最小步骤等。

- 确定验收文档需定义系统所需的文档清单，清单中的文档是待验收项目的证明材料，也是后续系统维护期的重要输入。清单主要可包括：需求规格说明书、概要设计说明书、详细设计说明书、系统测试报告、验收方案等。

- 组织管理要明确项目组织架构、项目制度、项目管理相关事宜，确保项目顺利进行。

### (2) 制定评估计划

制定评估计划是保证后续评估活动是否能够正确开展的先决条件。制定评估计划可包括制定验收计划、制定测试计划、制定评审计划、风险管理计划等四项任务。

- 制定验收计划要求业务人员、项目经理、测试人员明确验收时间、验收内容及验收形式，后续需按此计划执行系统的验收工作。

- 制定测试计划需根据项目实施计划安排具体的测试任务及测试质量要求。通常，测试计划应按迭代方式进行，优先安排主要系统用例的测试任务。

- 制定评审计划是确保计划能按预期执行的关键任务，精细化的评审可以尽早纠偏，减少不必要的成本浪费。制定评审计划要求详细阐述需评审内容、评审方式、评审要求和参与评审的人员等细节。

- 风险管理计划作为项目管控的重要环节，需尽可能详细定义风险管理的介入阶段、可能出现的风险因素、原因及风险的应对策略，同时考虑何时以何种方式进行阶段性的风险评估。

### (3) 执行评估

执行评估是系统评估与验收的执行环节，是依据已定义评估准则、评估计划等，进行具体系统评估的活动。评估主要从业务、技术、方案等层面展开，对待交付知识图谱应用系统是否满足相关方的需求进行评审。

执行评估活动是一个在项目周期中迭代进行的活动，需在此过程中不断反馈、调整和修订系统建设情况，确保项目执行过程与相关方的需求相符合。执行评估可包括系统功能评估、系统性能评估、部署方案评估、实施方案评估和数据评估等五项任务。

- 系统功能评估要求业务相关方对系统当前的研发成果进行审查，评估系统是否满足用户的真实意图，是否能够解决用户问题。此外，在此过程中还需评估用户UI设计原型及相关视觉设计方案是否易用，并符合用户所处行业的视觉标准。

- 系统性能评估要求相关技术人员审查当前的系统是否符合用户和行业提供的质量要求，是否存在技术上的严重缺陷，所用技术是否存在法律风险等。

- 部署方案评估需审查部署方案是否合理，是否符合实践用户的生产环境要求，是否与现有基础设施相容，是否存在网络风险等。

- 实施方案评估要求项目业务、技术相关方对实施方案的实施过程、实施内容、实施约束进行审查，评估实施方案的合理性和风险点。

- 数据评估需对知识图谱应用系统所需数据的质量、安全性进行审查，确保使用的数据完整和准确，能够支撑业务需要并保护用户的数据隐私。

### (4) 系统验收

系统验收是评估已建设系统是否满足用户要求的重要保障。系统验收需要业务人员和测试人员根据用户功能要求、性能要求和其他质量要求，从全局把控待交付系统的质量。本活动主要包括功能测试、性能测试、安全测试、文档验收等四项任务。

- 功能测试需测试人员、需求人员、UI设计师等从应用角度评估系统功能是否能够反映需求人员的设计，视觉上是否能满足UI人员的设计，总体是否满足用户需求。

- 性能测试要求测试人员从性能角度评估系统的鲁棒性和响应性，确

定被测系统是否能够满足用户的并发及业务操作响应需求。

- 安全测试要求测试人员从系统安全角度对待交付系统进行评估。针对系统安全体系架构的正确性和合理性进行分析，并通过渗透性测试、脆弱性评估、安全配置检测、源代码安全审查等维度对知识图谱应用系统进行全面的安全性评估。

- 文档评测要求业务人员和项目管理人员评估待交付文档是否完整并符合用户要求的文档交付清单，并确定待交付文档内容能否客观准确地反映系统建设的内容及过程。

#### （四）角色与分工

本阶段的参与人员角色及其任务简介如表7-11所示：

表7-11 系统评估与验收阶段参与人员角色及其任务简介

序号	角色	承担任务简介	交付物名称（如有）
1	评估人员	对接需求分析师，充分了解用户需求，负责对待交付系统及材料进行评估。	系统评估方案 验收文档清单 验收计划 需求评估报告 系统实施方案评估报
2	测试工程师	负责系统功能、性能和安全测试，并完成知识质量评估等。	测试计划 系统质量评估报告
3	需求分析师	跟踪评估工作，确保系统满足用户需求。	需求评估报告 系统验收报告
4	项目经理	对项目过程进行监控，对项目的进度和质量负责，并在验收期间进行资源调配。	项目人员配备情况表 风险管理表 验收计划 实施计划 评审计划 风险管理计划 系统验收报告 问题追踪表
5	安全管理人员	负责系统安全方面的评估和验收。	问题追踪表 系统质量评估报告

### （五）需用户方参与的内容

在系统评估与验收阶段，用户方需配合参与的工作可包括：

#### （1）确定评估准则

用户方业务人员需参与评估指标的定义，以降低评估指标的冗余，提高指标体系的完整性、专业性和准确性。此外，用户方还需介入验收文档清单及组织管理的讨论，明确验收文档内容及双方组织人员职责，促进项目沟通协作。

#### （2）制定评估计划

用户方项目管理人员需参与评审计划和验收计划的讨论，明确项目的验收条件、验收时间和验收内容。同时，制定评审的周期、评审的资源与要求，确保项目按各方步调协同进行。

#### （3）方案评审

用户方技术人员要参与对系统架构、系统部署方案及数据的评审，确保当前系统在技术层面符合用户的技术体系要求。

## 六、运营推广阶段

### （一）阶段建设目标

运营推广是知识图谱业务成熟和扩大的主要阶段。本阶段通过在生产环境中验证知识图谱的实际表现并加以运营推广，达到以下目标：

#### 目标1：保障知识图谱的应用效果。

在“系统评估与验收阶段”，开发者与业务方主要针对限定的场景和数据做验证；然而，本阶段则在真实业务环境中对知识图谱的性能表现进



行跟踪，收集用户反馈，并可能增加持续的人工运营服务。

### 目标2：扩大知识图谱应用系统的影响力和知名度。

在持续优化知识图谱应用系统的同时，进行典型案例及应用方法宣传，提升知识图谱应用系统的认知度和影响力。

### 目标3：树立知识图谱创新能力的示范标杆。

梳理知识图谱在方法、技术和应用方面，为行业创新能力树立示范标杆，促进企业创新产品和服务，创造更多的经济效应和社会效益。

### 目标4：提升知识图谱创造的应用价值。

将知识图谱应用创造的效益进一步横向扩展和纵向延伸，提升知识图谱创造效益空间的能力，吸引更多行业组织和服务对象普及知识图谱应用。

### 目标5：依托已有的基础。

面向知识图谱的发展前景与建设要求，充分进行资源优化整合与推广，激发各类人群的兴趣，进行有针对性的推广活动。例如，依据各类人群已有相关专业知识的分情况进行讲解以激发兴趣，并进行分阶段有重点的建设。

## （二）阶段输入输出

运营推广阶段的主要输入是实际生产数据、推广运营材料；主要输出为运营推广文案、推广效果统计报告、竞品分析报告等。

### （1）运营推广阶段输入

本阶段的输入可包括但不限于：

- 实际生产数据
- 知识图谱应用系统
- 系统相关文档，包括系统架构设计、系统使用手册等；



### (2) 运营推广阶段输出

本阶段形成的交付物及内容简介如表7-12所示：

表7-12 运营推广阶段交付物清单

序号	交付物名称	简介	重要性 (1-5级)	是否可选
1	运营推广文案	宣传知识图谱应用系统功能、业务效果、行业价值等信息的文档，形式不限，可通过广告、新闻稿等形式呈现。	3	推荐交付
2	推广效果统计报告	知识图谱应用系统运行指标、性能、效果等的量化统计报告。	4	必须交付
3	竞品分析报告	市场上同类知识图谱应用系统的技术及业务表现对比分析。	3	推荐交付
4	优势分析报告	相比较于同类竞品分析，重点分析本知识图谱应用系统的竞争优势。	2	可选交付
5	销售策略	根据用户需求和同类竞品特点，制定本知识图谱应用系统或其服务的销售方式和渠道等。	4	必须交付
6	报价方案	进行市场行情调研，规划本知识图谱应用系统或其服务的定价方案。	4	可选交付
7	宣传材料	收集系统技术特点、业务推广效果的宣传文章、视频等材料，编制宣传方案材料。	3	推荐交付
8	期刊论文	将系统建设中具备开创性意义的技术或工程经验，形成理论等成果，并以论文等形式发表。	1	可选交付
9	联合专利	业务方与技术方合作突破的技术，可联合对系统开发中的技术申请专利。	2	可选交付
10	运营推广系统对接文档	可包括本知识图谱部署运维环境及与其他运营推广系统对接的规范文档。	3	推荐交付
11	数据对接文档	可包括本知识图谱数据接口及与其他系统对接的数据标准文档。	4	必须交付

序号	交付物名称	简介	重要性 (1-5级)	是否可选
12	新需求记录	系统上线后运维过程中，根据实际系统应用效果和用户体验，记录产生的新需求。	3	推荐交付
13	产品培训资料	为了保证系统使用者规范化操作和系统正常运行，编制使用指南和培训材料。	3	推荐交付
14	运营管理手册	系统上线后，为了保证运营及管理效率，编制系统运维相关人员的工作手册。	3	推荐交付
15	受众调研报告	调研知识图谱应用系统受众的使用效果和体验，编制受众调研报告。	3	推荐交付

### (三) 主要活动

运营推广阶段是在实际生产环境中验证知识图谱应用系统效果，并进行用户运营和市场推广的重要阶段。本阶段根据拟实现目标的不同可划分为系统监控、系统运营、系统推广、使用意见总结、文档沉淀等五项活动，本阶段活动流程及主要参与角色如图7-8所示。各项活动主要内容如下：



图7-8 运营推广阶段活动流程图

### (1) 系统监控

本活动可包括知识图谱应用系统的健康度监控及业务效果监控等任务。从用户和管理员视角，对知识图谱应用系统的线上表现进行监控，关注整体健康度、业务满意度，而非内部技术问题。同时，对系统产生的业务效果做监控，衡量知识图谱是否对预定的业务目标产生正面影响。此外，针对可量化指标项产生的具体数据进行统计和深层次分析。

### (2) 系统运营

系统运营范围可包括与知识图谱应用系统及上下游对接的系统、相关用户组织、相关网络社区等一切运营行为。主要对应为：知识图谱应用系统及上下游系统的运营人员日常工作；相关用户组织的定期活动、培训；网络社区的舆情监控、控评、意见反馈等。

### (3) 系统推广

本活动中可宣传知识图谱应用系统的正面效果，并推广至更多受众。形式可分为组织内推广和社会性推广，如发布宣传材料、广告、地推等。

### (4) 用户意见总结

本活动中需总结用户对于系统可用性和效果等方面意见，用于改进用户体验，帮助系统优化，方式可包括：用户访谈、调查问卷、工作坊等。

### (5) 文档沉淀

本活动针对运营推广阶段工作的问题、经验和待办等，形成各类文档沉淀，主要包括：运营推广文案、推广效果统计报告、竞品分析报告、优势分析报告、管理运营手册等。

## (四) 角色与分工

本阶段的参与人员角色及其任务简介如表7-13所示：

表7-13 运营推广阶段参与人员角色及其任务简介

序号	角色	承担任务简介	交付物名称（如有）
1	市场运营人员	负责知识图谱应用系统、产品、生态、社群运营和管理，与企业外部的集成方和需求用户方合作推广，组织用户参与应用交流和分享活动，汇总推广方案和物料。	运营推广文案 宣传材料 受众调研报告
2	产品经理	负责知识图谱同类竞品调研，用户需求意见总结，产品功能开发和数据规范文档撰写，产品技术特点分析。	竞品分析报告 优势分析报告 运营推广系统对接文档 数据对接文档 新需求记录 产品培训资料
3	销售人员	负责商务侧运营，与知识图谱需求用户沟通，挖掘用户需求内容，推介本产品功能与特色，销售策略制定与调整。	销售策略 报价方案
4	售后服务人员	负责知识图谱售后支持，与技术人员合作，协调解决用户遇到的问题。	/
5	运营人员	负责知识图谱应用系统部署后的产品、生态、社群运营，与企业外部的集成方和需求用户方合作推广，文档及产品侧的统筹。	推广效果统计报告 宣传材料 运营管理手册
6	市场人员	与企业外部的集成方和需求用户方合作推广，组织用户参与应用交流和分享活动，市场推广，用户意见总结。	推广效果统计报告 宣传材料 受众调研报告
7	开发人员	负责知识图谱协调功能开发和技术攻关，总结开发过程中的技术要点和关键技术，系统状态监控。	期刊论文 联合专利
8	商业分析师	负责知识图谱同类竞品商业分析，用户需求分析，本产品部署应用后用户反馈及解决情况监测，及产品投放市场后效果评价监控。	优势分析报告
9	数据分析师	负责收集知识图谱用户需求数据和反馈信息，收集市场需求和用户评价信息，分析本产品指标和效果，对业务数据指标进行监控。	推广效果统计报告

### （五）需用户方参与的内容

在运营推广阶段，用户需作为需求领导方和直接使用方，对真实环境中的知识图谱应用系统表现作出反馈，阐述直观真实的使用体会，并明确对于系统效果的期望。用户方需要参与的主要工作如下：

#### （1）系统试用及反馈

从需求视角出发，试用生产环境中的知识图谱应用系统，用户在业务流程操作应用过程中，对于系统输入、处理、输出等环节，发现业务流程中的功能、数据存在的质量问题，及时反馈并提出改进的建议。

#### （2）协调上下游系统对接

在用户企业中，知识图谱应用系统需要用于对接上下游系统的资源，用户方需对此类资源进行协调。同时，用户方还需指定系统对接相关等的技术负责人等角色。

#### （3）新需求提炼

根据系统使用体验、企业内反馈、市场反响等维度信息，梳理用户与系统交互的可用性，分析企业内人员协作的便捷性，调研市场对知识图谱应用的认可度，总结优缺点并提炼新需求。

#### （4）文档撰写

各类文档的撰写离不开用户企业的配合，例如，企业作为建设主体的过程中，相应专家人员与服务方一同联合专利申请或联合发表期刊文章、论文等，可从技术市场角度为系统进行推广。

#### （5）系统质量纠正

对照标准化要求，如数据规范、数据集规范、专业基础标准、过程标准、质量标准、技术与管理标准、工具与方法标准等，检测知识图谱应用过程中的质量问题，并确定系统是否具有稳定性、安全性、易操作性等性能，及时反馈并提出改进的建议。

### (6) 用户行为

采集使用者的操作行为，通过了解行为数据信息，分析什么情况下使用活跃度高，并采取相应措施，更大程度地提高使用黏性。使用者操作行为体现了他们的诸多特征，可以将使用者分级更加精准化，并为运营策略和运营动作提供了执行依据。此外，收集到的使用者操作行为数据越多，所能做的运营策略也会更加精准有效。

## 七、管理维护阶段

### (一) 阶段建设目标

管理维护是知识图谱落地后可持续输出效益的保障。本阶段通过对知识图谱生产环境的实际使用监控和管理，保障系统正常运行，来持续不断的输出知识图谱价值。管理维护需要达到如下的目标：

**目标1：** 监控知识图谱应用系统的使用，保障系统正常使用，防止对系统的滥用以及系统本身的不完善导致的使用障碍，及时发现导致使用障碍的缺陷，维护系统安全，抵御网络攻击及其他形式的恶意攻击。

**目标2：** 对知识图谱应用系统的使用实施管理，保证既定的使用流程顺畅进行，防止未经许可的访问，保证信息安全和隐私安全。

**目标3：** 维护数据安全，使用多元方式定时对数据进行备份，防止因意外事故丢失数据导致不可挽回的损失。

**目标4：** 维护系统稳定，及时更新过时的组件，检查新版组件间的兼容性，实施最优的系统升级策略，保证系统长期稳定运行。

**目标5：** 实施系统复制和迁徙，使知识图谱应用系统适应新环境，或者在分布式应用场景下引入更为多元的不可篡改拷贝，提升系统可信度和使用价值（需要时执行）。

## (二) 阶段输入输出

### (1) 管理维护阶段输入

本阶段的输入可包括但不限于：

- 需求分析输出的管理端需求；
- 需求分析输出的数据安全指标；
- 系统复制与迁徙需求。

### (2) 管理维护阶段输出

本阶段形成的交付物及内容简介如表7-14所示：

表7-14 管理维护阶段交付物清单

序号	交付物名称	简介	重要性 (1-5级)	是否可选
1	系统使用手册	围绕应用系统各项功能使用注意事项及操作步骤的说明文档	5	必选交付
2	系统运维手册	可包括知识图谱应用系统错误代码表、故障排查方法等	5	必选交付
3	培训手册	为将非专业人士培训成为系统管理者，关于系统整体架构、应用场景、使用说明、注意事项等的指导文件	5	必选交付
4	系统后台管理工具	可包括知识图谱应用系统的后台管理组件和使用方法	5	必选交付
5	监控工具	可视化实时监控系统状态的软件工具	4	推荐交付
6	系统巡检工具	对系统安全性、稳定性进行定时扫描的工具	4	推荐交付
7	巡检报告	对系统安全性、稳定性定时扫描输出的结果报告	5	必选交付
8	问题报告	对系统使用过程中出现的各类问题进行梳理形成的文档	4	推荐交付

序号	交付物名称	简介	重要性 (1-5级)	是否可选
9	日志文件	对系统访问和应答行为进行按时记录形成的文档	3	可选交付
10	补丁文件	针对系统错误的临时修补文件	4	推荐交付
11	修改记录	运维过程中关于系统修改历史的记录文档	4	推荐交付
12	版本管理	维护并记录知识图谱应用系统代码及其集成的知识图谱的版本信息，以便必要时的溯源或回滚	4	推荐交付
13	版本升级或管理说明	在版本更新时形成的更新内容相关的说明文档	4	推荐交付
14	FAQ常见问题清单	就系统可能遇到的常见问题，给出对应解决方法的说明文档，需定期更新	4	推荐交付
15	售后服务记录	记录售后服务中用户方的请求、服务方的应答与完成情况的文档	5	必选交付
16	保修清单	保修中涉及的要素清单	2	可选交付
17	用户满意度调查	对用户满意度的调查报告	2	可选交付
18	授权文件/授权函	关于授权对象、授权期限和范围等的说明文件	5	推荐交付
20	知识维护和管理记录	对知识维护和管理过程及变更情况进行记录形成的文档	5	必选交付
21	风险应对策略	针对可预见的风险形成的应对策略说明文档	1	可选交付



### （三）主要活动

管理维护阶段根据拟实现目标的不同可划分为系统运维管理、知识运维管理、技术支持、使用培训、授权管理等四项活动。各项活动主要内容如下：

#### （1）系统运维管理

保障系统正常运行，围绕流程管理、事件管理、问题管理、变更管理、发布管理、运行管理、知识管理、综合分析管理等类型形成管理机制，全面提高运行维护的快速响应能力；建立自动化分析报表，为业务知识积累和业务考核建立完善的数据模型。

#### （2）知识运维管理

对知识图谱应用系统进行运维管理，可包括：检查、更新、删减、优化、发布等知识单元的管理；流程、事件、问题、运行等系统质量控制的管理。

#### （3）技术支持

帮助用户方解决其在使用知识图谱应用系统过程中出现的具有明显特征且可能由系统导致的技术问题。

#### （4）使用培训

对用户方的使用人员和管理人员进行培训，以便使其能正确有效地使用知识图谱应用系统。

#### （5）授权管理

对使用人员的授权进行管理，确定授权范围、授权依据、授权方式、授权制度，并对授权效果进行及时评估。

### （四）角色与分工

本阶段的参与人员角色及其任务简介如表7-15所示：

表7-15 管理维护阶段参与人员角色及其任务简介

序号	角色	承担任务简介	交付物名称
1	知识运维工程师	负责对知识图谱的维护和管理	知识维护和管理记录
2	系统运维工程师	负责系统部署架构的设计及相关配套资源的识别，并对系统集成方案中涉及的部署活动输出设计方案。	版本管理文档 数据修改记录 保修清单
3	系统安全工程师	负责系统的安全和稳定，保证系统正常运行，防止非授权的访问。	日志文件 监控平台 巡检报告
4	培训人员	提供培训支持	培训手册 系统使用手册 运维手册
5	技术客服	提供关于技术方面的问题解决	FAQ 常见问题清单 售后工单 用户满意度调查
6	开发人员	提供技术支持，完成故障修复，并解决系统缺陷和简易需求优化	补丁文件 版本升级说明 问题报告
7	授权管理人员	对授权进行运行管理和调整	授权函 授权文件

### (五) 需用户方参与的内容

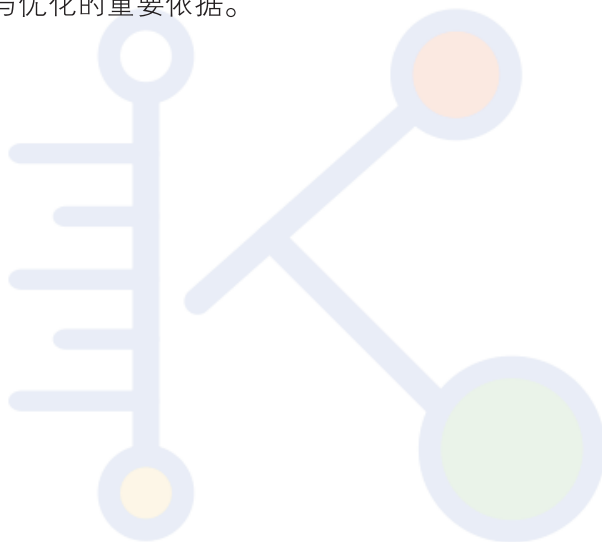
**在管理维护阶段，用户方需配合参与的工作可包括：**

#### (1) 参与培训

用户方需委派特定人员接受关于系统使用的培训。培训内容包括对系统的使用和对系统的运维等。同时，为了更好地使用系统，使之发挥出应有的效能，用户方受培训人员需积极参与熟悉和使用系统，特别是出现问题时的解决方式和流程。

## (2) 参与反馈

为了使系统朝着更为稳定、高效、有价值的方向发展，用户方需配合服务方对系统在使用中出现的问题进行及时、如实、详细的反馈。同时，服务方的技术支持和开发人员也要积极地对用户的反馈作出反应，完成故障排除、效果优化和效能提升，充分地利用用户方的使用体验数据，使其作为系统测试与优化的重要依据。



知识图谱标准化  
Knowledge Graph Standardization

## 第八章 面向知识图谱应用系统建设的 服务方选择

### 一、面向知识图谱应用系统建设的服务定义

面向知识图谱应用系统建设的服务是指为知识图谱应用系统全生命周期提供的服务，包括咨询服务、设计与开发服务、集成实施服务、运维服务等。

当前，知识图谱相关的服务方通常可提供多种服务类型，而且因技术积累方向和定位的差异，各服务方之间擅长的服务类型也可能存在不同。因此，具有知识图谱应用系统建设或维护需求的用户方，可基于自身能力基础评估结果和整体建设规划，确定所需的服务类型，并针对性地筛选单个或多个服务方，以便完整覆盖自身需求。面向知识图谱应用系统建设的服务分类及代码如表8-1所示。

### 二、面向知识图谱应用系统建设的服务分类

表8-1 面向知识图谱应用系统建设的服务分类

标识	类别名称	说明
01	咨询服务	指围绕企业数字化转型战略、发展路径、管理提升、诊断评估、人员培训等方面，向需方提供的咨询服务。
0101	战略咨询	指结合需方战略规划和业务发展需要，基于相关政策、标准，提供实施知识图谱一定时间周期内战略层的咨询建议。
0102	方案咨询	指具有相关资质（能力）的咨询机构，提供相关方案以指导需方进行知识图谱技术产品、应用系统、服务的构建、运营和维护等。

标识	类别名称	说明
0103	测评认证	指具有相关资质（能力）的测评认证机构受需方的委托，依照相关国家标准、行业标准或经确认的技术规范，按照规范程序，对供方提供的知识图谱应用系统产品和服务进行的科学、公正的评价服务。
0104	培训	指围绕需方实施知识图谱所需的能力培训活动，提供包含培训需求研究、培训方案设计、实施、管理等在内的服务。
0105	监理	指具有相关资质（能力）的监理单位受需方的委托，依照相关法律、法规和知识图谱应用系统建设监理合同及其他知识图谱应用系统建设合同，代表需方向知识图谱应用系统建设服务方的建设实施监控的服务。
0199	其他咨询服务	指未列明的咨询服务。
02	设计与开发服务	指受需方委托，以承接外包的方式提供的知识图谱产品及应用系统等的设计和开发服务。
0201	本体模型设计与开发	指构建知识图谱的本体及其形式化表达，如实体类型定义、关系定义及属性定义，以支撑后续应用。
0202	知识获取/数据处理与供给	指从不同来源和结构的输入数据中提取结构化知识并输出，以支撑后续应用。
0203	知识图谱可视化	指将数据通过关系预处理程序处理为图数据库等可以查询的数据，支撑展示与分析。
0204	图算法设计与开发	指聚焦图结构的特殊性，对知识图谱中的图算法进行设计与开发，如路径分析、节点聚类、关联分析等，目的是提升算法效率，支撑对知识图谱的探索与挖掘。
0205	数据标注/元数据管理工具设计与开发	指遵循相关元数据标准，依照元数据模型，提供或基于数据标注工具等对数据进行标注和管理的服方。

标识	类别名称	说明
0206	知识图谱构建设计与开发	指使用数据、知识等构建、开发知识图谱，提供基于知识图谱的基础产品或服务，以满足特定需求的服务；其中，基础产品和服务是指复杂的应用程序或系统的中间件。
0207	系统方案设计与开发	指依据需方的应用需求，设计并开发基于知识图谱的应用模块或核心子系统，包括知识图谱通用应用系统方案设计与开发、知识图谱行业应用系统方案设计与开发。
0299	其他设计与开发服务	指未列明的设计与开发服务。
03	集成实施服务	指依据知识应用需求，将知识图谱、信息系统或服务进行整合，提供知识图谱应用系统及服务。
0301	系统集成实施与部署	指系统集成实施与部署者制定集成计划，按照集成设计方案实施系统集成工作，以持续迭代知识图谱应用系统整合服务，形成功能完整的系统或服务。
0302	知识融合/集成	指整合和集成知识单元（集），形成拥有全局统一知识标识的知识图谱。
0303	知识服务	指将知识图谱的应用形态与领域特征和场景结合，助力业务转型。
0304	环境部署	指配置运行环境，监控和管理知识图谱应用系统的环境部署，交付可正常使用、无故障的软硬件环境。
0399	其他集成实施服务	指未列明的集成实施服务。
04	运维服务	指对 IT 设备（硬件、软件、网络）和知识图谱应用系统的运行环境、业务系统和运营人员等进行配置、监控、更新、维护等的综合管理服务。
0401	知识运维	指根据需方的需求，对故障进行排除及操作给定知识条目，提供知识新增、更新、删除、备份与恢复等维护服务及知识质量管理等服务。

标识	类别名称	说明
0402	系统运维	指根据需方的需求，提供知识图谱应用系统功能和性能的维护、更新、升级的技术支撑方案，进行应用系统的维护、版本检测和优化升级等服务。
0403	图数据库运维	指围绕满足知识图谱技术产品和服务构建、集成和应用所需的图数据库需求和指标，提供远程技术支持、现场服务、系统升级和改造、故障检测和维修等服务。
0404	网络与安全（基础环境）运维	指根据需方的需求，为确保 IT 设备的运营环境、业务平台的高可靠性、高可用性及网络与安全等，提供如基础设施的故障诊断、评估、维修、优化等运维服务。
0499	其他运维服务	指未列明的运维服务。

## 第九章 总结与建议

尽管目前知识图谱应用已取得显著成效，工程化、体系化路径日益明晰，知识图谱产业生态体系骨架也基本显现。然而，在技术发展、部署实施、数字基础设施等方面存在大量挑战，有赖于汇聚用户方、服务方、科研院所、政府管理部门、学会协会等多方力量共同破解瓶颈与障碍，打造我国人工智能领域的知识产业，赋能各行业知识驱动的转型升级。围绕产业生态建设、协同机制构建、先进经验推广、基础短板补齐、国内外合作等需多方共同协作的方面，提出如下建议：

### 一、国家支持层面

#### 1、建设汇聚产学研用多方的产业联盟，打造知识图谱产业生态圈。

建议建立凝聚大中小型企业、科研院所、高校、行业组织共同参与的产业联盟，建设覆盖知识图谱各类利益相关方的生态圈，以产业需求为导向，从业务方向、技术方向统筹推动各方有机协作，并逐步向行业分会等方面体系化建设，不断壮大我国知识图谱供应商、集成商及配套服务商队伍。

#### 2、加大知识图谱相关项目扶持和投入，推动体系化技术攻关与产业化应用。

建议加大知识图谱相关科研课题投入，并在省市区等各级通过专项建设或揭榜挂帅等形式促进应用类知识图谱项目的建设与示范，可鼓励

有能力的区域对企业知识图谱应用系统的引进、部署、升级或出售等行为核实后进行政策补贴。同时，探索建立揭榜项目成功案例的推广机制，例如区、市、省级赛事评审等，强化技术攻关、产业化应用、成果推广间的有机衔接，促进知识图谱作为信息技术领域的重点和急缺领域之一获得更多关注。



### 3、加快推进知识图谱标准化建设，强化知识图谱技术专利追踪。

建议加快推动知识图谱标准体系研制，系统分析和树立迫切标准化需求，聚焦技术固化、产业发展、互联互通等重点方向明确标准推进方向，在知识图谱跨域融合与互操作、产品模块化与规范化、项目建设实施各方权责指引等方向深化标准立项与研制工作。同时，建议强化对国内外专利申请趋势、专利申请、受理分布等跟踪技术动态，监测技术走向。

### 4、加强国家知识基础设施梳理与建设，夯实知识图谱发展基础

建议围绕重点领域，探索建立包括知识图谱算法库、本体模型库、国家级/行业级知识库、算力、开源架构等资源的知识图谱公共服务平台，强化对中小企业、科研院所在大模型构建、先进技术验证等方面的支撑，推动技术创新。同时，推动以知识图谱为载体、以机器可读和可理解为方向的公共资源开放与共享，例如标准、政务数据等可开放信息，并探索建立公共服务平台资源发布、完善、维护与更新的准入机制，形成可持续发展基础设施。此外，也可探索细分领域特殊场景中的共性小众知识资源发布和维护机制，促进知识共享共建。

## 二、行业部署层面

### 1、强化领域专业术语集、语料库建设，增强知识图谱基础数据积累。

建议加大以本行业权威标准、文献等材料专业术语集建设，并围绕非结构化数据知识获取、知识融合等技术方向建立权威语料库，加快对标准化的各类术语集中资源进行整合和推广，形成行业统一的语境，为语义统一奠定基础。

### 2、推进行业标准立项、发布与执行，强化知识图谱与行业需求融合。

建议加快典型领域知识体系标准研制与推广应用，强化知识图谱在典型领域中特殊标准化需求树立与挖掘，如知识图谱与业务系统间集成、知

识图谱与典型业务数据间融合转化、细分领域知识表示和本体模型构建要求等，推动重要标准化需求向行业标准转化，凝聚行业内共识，固化先进实践经验。

### 3、遴选和树立典型领域应用企业标杆，推动一批示范性知识图谱项目的建设与实施。

建议综合考虑知识图谱优秀实践案例对知识图谱推广落地的带动作用，筛选一批数据基础好、知识应用需求强、知识图谱技术积累扎实或已取得显著成效的应用企业，推动示范性知识图谱项目的建设、实施与宣传，形成细分领域应用场景中的典型样例、使用方式及成功经验，深入发掘知识图谱应用经济效益和社会价值。同时，鼓励对先行者提供必要的政策支持，进一步巩固和发挥先行先试企业或机构的探索热情。

## 三、产学研协同层面

### 1、强化知识图谱测试评估工具和平台建设，提升知识图谱质量和价值保障。

知识图谱的应用与推广离不开知识图谱的测试评估。为规范知识图谱的应用并制定科学的评判依据，建议建立健全知识价值评估与评测体系和知识图谱与知识图谱应用系统测试评估体系，界定测试范围和评估标准，以明确知识图谱相关系统的建设基础能力评估方式和最终验收边界，管控用户企业预期和系统验收质量。同时，建议形成标准符合性测试规范与工具，并搭建知识图谱标准的符合性测试平台，提高测试执行的准确率和效率。

### 2、加强专业化和专职化的标准符合性测试、认证、监理机构建设，提升第三方服务供给水平。

建议加强专业化、专职化的标准符合性测试机构建设，鼓励适应知识

图谱技术和产业发展且具有领域影响力和公信力的第三方检测认证服务机构发展，重点支持1-2家检测机构成为国家级标准符合性检测单位，提供知识图谱标准化符合性评测服务和技术指导，提升标准符合性检测服务供给能力和专业化水平。同时，强化知识图谱产品认证体系以及项目实施过程监理体系建设，打牢产品供应方质量意识，提升产品成熟性和项目工程化实施水平，降低用户企业审核验收难度。

### 3、强化科研院所与企业间协作攻关，引导企业需求与科研力量精准对接。

建议加强技术研发与行业应用的结合，筛选知识图谱相关前瞻性、示范性技术攻关方向，以解决企业重大技术难题为目标，强化科研院所、高校与企业间的协作攻关，并引导先进成果的转化和示范应用。同时，依托产业联盟等组织，促进企业研制需求与科研力量间的精准对接，也可鼓励多方共同搭建具备领域特点的联合实验室或工程中心，形成常态化的协作渠道和平台。

### 4、加大知识图谱人才培养，探索产教融合新模式。

建议加强知识图谱领域人才培养力度，畅通产业端人才需求与科研端、教育端人才供给匹配渠道，探索建设产教融合模式，形成人才合作培养机制，为优秀人才提供实际项目参与和学习机会。此外，结合人员能力认证体系建设，推动知识图谱相关知识体系和核心能力的掌握水平。

### 5、促进先进经验推广与研究成果转移转化。

建议加强技术或产业峰会、刊物等多种形式的经验实践分享途径，积极加强国际国内先进经验、优秀研究成果对接，汇聚更多行业案例，推动先进技术的互学互鉴，分析和预判知识图谱领域重点技术、重大问题和发展趋势，加快知识图谱相关技术的突破效率和速度，为知识图谱产业化应用奠定创新基础。

# 案例一：基于知识图谱的大数据风控系统建设与实施

## 一、案例背景

商业银行传统风控体系，普遍存在信息不对称、成本高、时效性差、效率低等问题。一方面客户经理现场踩点、自主汇总各方面数据，人力成本较高且效率低下，另一方面由于金融领域数据典型的“4V”特征（数量海量Volume、多结构多维度Variety、价值巨大Value、及时性要求Velocity），使得数据获取难度大，隐藏在数据中的深层次关系难以挖掘。因此，解决银行风控难点痛点迫在眉睫。

为解决金融领域风险管控中的数据孤岛、数据不对等、数据关联关系难以挖掘等问题，综合应用知识图谱、大数据等技术，广泛引入外部数据，整合内部系统和流程，构建覆盖全用户、涵盖全流程的大数据风控系统，提供从用户准入、关联关系、用户画像、预警管理等功能的风险管理能力。

## 二、案例所处阶段及成效

（1）所处阶段：运营推广阶段

（2）案例成效：

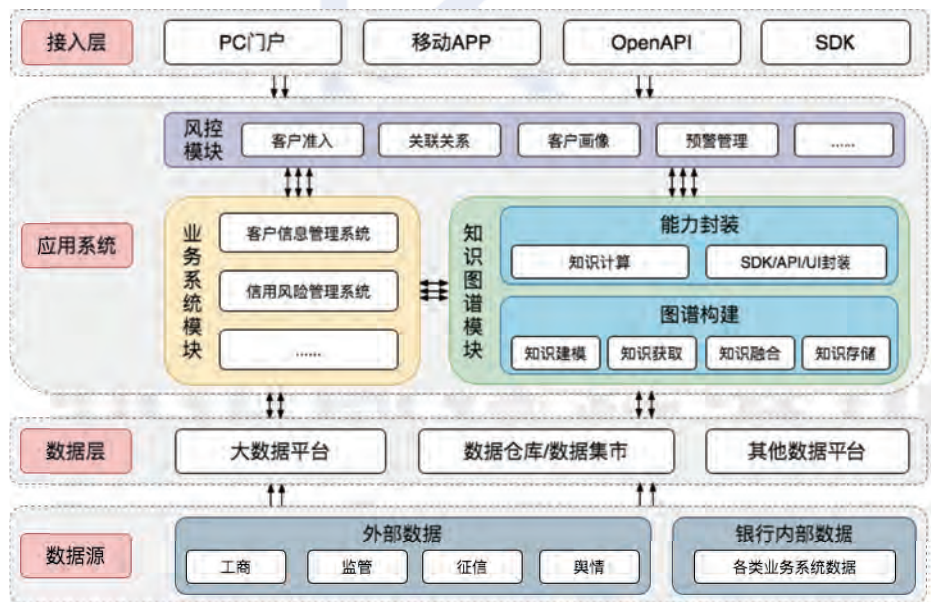
基于知识图谱的大数据风控系统上线后多次完成系统的迭代升级，目前处于运营推广阶段，且良好的经济和社会效益：

在用户质量方面，累计通过该系统准入的公司授信用户超过5万家，低风险/次低风险用户占比为16.7%，有效提升了用户质量。

在反欺诈能力方面，通过反欺诈模块建立团伙等级、特殊名单分等反欺诈模型，累计完成超过500万用户风险识别，拦截43.58万疑似欺诈用户。

在预警能力方面，系统累计产生超过450万条预警信号，自动定级信号占比94.34%，实现92.14%的预警率，系统平均预警时间为224.33天，大幅提升了预警能力。

### 三、案例系统架构



图A1-1 基于知识图谱的大数据风控系统架构

如图A1-1，基于知识图谱的大数据风控系统的数据来源包括外部数据和银行内部数据，外部数据引入了工商、监管、征信、舆情等权威渠道的数据，银行内部数据为各类业务系统的数据。

### 系统分为数据层、应用系统层和接入层：

(1) **数据层**：由银行内部的大数据平台、数据仓库/数据集市和其他数据平台组成，进行各类数据的处理、加工。

(2) **应用系统层**：由业务系统模块、知识图谱模块和风控模块组成，业务系统模块集成用户信息管理系统、信用风险管理系统等业务系统，知识图谱模块包括图谱构建模块和能力封装模块，风控业务模块提供用户准入、关联关系、用户画像、预警管理等业务功能。

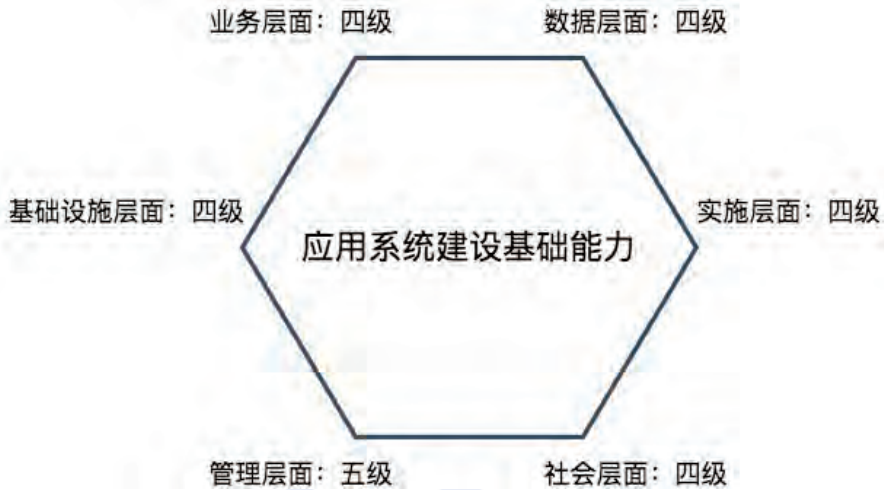
(3) **接入层**：提供PC端门户展示、移动端APP以及OpenAPI、SDK能力。

## 四、案例开发前基础能力评估情况

### (1) 基础能力评估结果

目前基于知识图谱的大数据风控系统在业务层面处于L4级别，在数据层面处于L4级别，在基础设施层面处于L4级别，在实施层面处于L4级别，在管理层面处于L5级别，在社会层面处于L4级别，综合得到该系统建设基础能力为L4就绪级别。

知识图谱标准化  
Knowledge Graph Standardization



图A1-2 基础能力评估

## (2) 短板提升措施

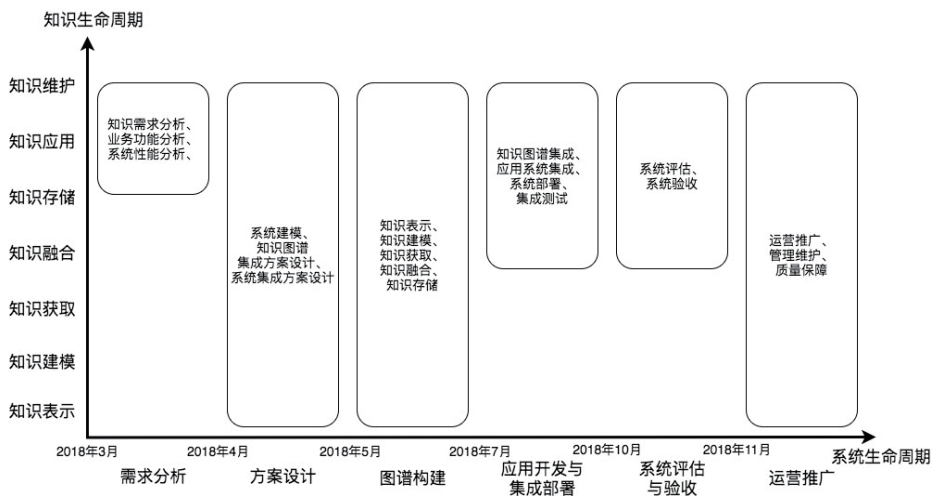
**基础设施层面：**人员储备上缺乏知识图谱专业化人才。目前缺乏知识图谱领域的知识专家、研发人员储备，需持续加强知识图谱相关技术人才储备和人才培养，强化知识图谱研究、应用能力，以满足可持续开发建设知识图谱需求。

**业务层面：**知识图谱建设预算投入不充分。需从战略层面重视人工智能技术的研究应用，实现公司业务的智能化建设，进一步提升人工智能、知识图谱建设的预算投入。

**实施层面：**需不断加强知识图谱应用系统的实施能力。基于公司已有的知识图谱应用案例，持续拓展应用场景，不断进行业务应用试验验证，提升知识图谱应用的系统实施能力。



## 五、案例开发与实施历程及难点



图A1-3 案例开发与实施历程

### （1）需求分析阶段（2018年3月-4月）

在该阶段业务人员、数据人员、项目经理、产品经理、架构师等人员，主要完成行业调研、可行性分析、业务需求梳理和分析、数据收集和分析、相关数据和技术标准约定、开发技术选型等工作。

该阶段的难点主要在于对业务目标的准确把握，大型的复杂系统很难通过一次或多次分析得到完整清晰的功能、性能等要求。通过对同业产品的深度调研，与业务专家和技术专家反复沟通，澄清业务需求，对需求目标分类定级，实现业务目标的拆分和明确。

### （2）方案设计阶段（2018年4月-5月）

在该阶段业务人员、产品经理、项目经理、系统架构师、研发工程师、算法工程师等人员，主要完成系统架构设计、功能详细设计、知识图谱设计、数据库设计、模块集成方案设计等工作。该阶段的主要难点在于



知识图谱的设计和系统集成中各应用模块的互联、互操作设计，这主要通过知识图谱建设方案的不断试验验证、对各应用模块约定统一的数据、接口、通讯标准解决。

### **(3) 图谱构建阶段 (2018年5月-7月)**

在该阶段业务人员、产品经理、项目经理、知识图谱工程师、研发人员等人员，主要完成知识建模、知识获取、知识融合、知识存储等工作。该阶段的主要难点在于知识建模不够准确、通用和非结构化知识抽取难，这主要通过业务专家和知识图谱工程师协同采用人工自顶向下的建模方式完善模型，以及采用自然语言处理技术抽取非结构数据解决。

### **(4) 应用开发和部署阶段 (2018年7月-10月)**

在该阶段业务人员、项目经理、知识图谱工程师、研发人员、测试人员、运维人员等人员，主要完成功能开发、功能测试、系统部署等工作。该阶段的主要难点在于不同数据来源的数据融合，这主要通过相关知识融合技术解决。

### **(5) 系统评估与验收阶段 (2018年10月-11月)**

在该阶段业务人员、产品经理、项目经理、研发人员、测试人员、运维人员等人员，主要负责功能评估验收、系统评估验收、系统上线等工作。

### **(6) 运营推广阶段 (2018年11月至今)**

在该阶段业务人员、产品经理、研发人员、运维人员、运营人员等人员，主要负责系统监控、系统运营、系统迭代、系统推广等工作。该阶段主要难点在于知识图谱应用系统的知识维护和管理，主要通过定期组织专门的培训来提升运维能力。

## 六、涉及的关键技术

### 关键技术1：基于OCR的非结构化数据处理技术

针对银行内部包括财务报表、权证、票据等凭证的非结构化数据，采用基于机器学习、基于caffe深度学习框架等的OCR引擎进行识别和提取。

### 关键技术2：基于HAWQ的大数据分析引擎

针对来自工商、监管、征信等外部数据源以及银行内部的海量数据，采用基于HAWQ的大数据分析引擎，完成结构化数据的统一存储、加工、计算。

### 关键技术3：基于NLP的知识抽取技术

金融领域数据具有规模大、维度多、价值难以挖掘的特点，给知识抽取带来很大困难。采用LTP、BERT等NLP技术对数据进行去重、分类、词性标注、命名实体识别等处理，实现实体、关系、属性的抽取。

### 关键技术4：基于关系数据库、图数据库、ES的混合存储架构

数据管理上，系统采用基于关系数据库、图数据库、ES的混合存储，实现数据的高效插入、更新和查询。目前图谱规模已达十亿级节点和关系，支持20+图算法，具备亿级别规模的图计算能力。

## 七、当前面临的重要挑战

公司致力于应用知识图谱、人工智能等创新技术赋能银行传统业务的智能化、数字化建设，后续将持续开展知识图谱的研究和应用。对于后续持续开展知识图谱的研究和场景应用面临着许多挑战：

挑战1：在数据层面，金融领域海量多源异构数据造成的数据歧义、关联性难以挖掘的问题，使得图谱的构建难度大。

挑战2：在技术层面，在包括知识获取、知识融合、知识存储、知识

计算等图谱构建关键技术的技术研究上有待加强。

挑战3：在应用层面，针对各类复杂业务场景、业务需求转换为知识图谱模型的建模能力需持续提升。

挑战4：在人才方面，需要通过引进知识图谱专业人才和开展技术培训，增强人才储备。



知识图谱标准化  
Knowledge Graph Standardization

## 案例二：测绘地理信息领域知识图谱系统建设与实施

### 一、案例背景

地理实体作为新型基础测绘的标准化成果，是自然资源领域为适应新时代高质量基础测绘发展，进行的基础测绘改革，旨在提升基础地理时空数据在自然资源部门和社会公众管理部门的服务能力和水平，促进自然资源管理和治理能力现代化，服务经济社会各领域高质量发展。

当前，地理实体打破了传统基础地理信息数据面向制图的要素化管理模式，转而面向地理空间对象，从多模态，时空化，语义化等方面形成的新型地理时空数据。但管理部门在对地理实体数据进行管理，提供社会化服务的过程重，面临了几个核心需求：

核心需求1：地理实体数据内容种类繁多，关系庞杂，实体概念交叉重叠，包括建筑，水系，交通，植被，管线等多领域的空间数据。目前多以某一领域的线性分类方法，分级分类构建地理实体数据模型结构，这种方式存在难以兼顾多领域实体，难以扩展领域知识和关系等问题。因此，通过提取地理实体抽象类别和关系，构建地理实体语义网络，可以有效解决上述问题。

核心需求2：地理实体数据最显著的特征是具有几何图形结构，将现实世界中地理对象的表现形态抽象表示，如河流可以抽象为一条多节点的折线，也可以抽象成一个河流边界线围绕而成的多边形。同时，地理实体数据还具有丰富的属性信息。

传统方法是借助关系型数据库和几何图形数据库分别存储基础地理信息的要素属性和几何图形结构。这种方式容易丢失数据间的关系，而且随着数据量的增加，查询复杂性也大大增加。利用图技术管理地理时空数据

间的关系，同时结合GIS空间表达能力，可以更加直观地表达地理实体数据间的关系，降低查询复杂度。

核心需求3：地理实体数据难以提供社会化应用，是因为它只是测绘管理部门获取的地理空间上某一时刻的现状数据，虽然在很多应用中可以提供一些空间位置信息，但是在大多数政府问题、商业问题中还需要提供物体、人、空间、时间、事件、位置、行为之间的关系。目前图技术可以借助各要素之间的复杂关联和数以百万计的因果路径，表达海量数据间的各种复杂关系，进而扩展业务规则，提供更多的应用情景。

核心需求4：地理时空数据管理部门除了掌握地理实体数据，还掌握了大量经过空间化的专题数据，如地理国情普查数据，不动产房产测绘数据等。这些数据保留了丰富的社会化应用涉及的元数据，但目前这些尚未得到有效利用。利用算法语义规则，从现有的数据中获取新的推理知识，提供更深层次的查询应用。如通过模拟人的思考，实现自然语言查询功能，计算城市体检指标，提升城市居民满意度。

## 二、案例所处阶段及成效

### (1) 所处阶段：图谱构建阶段

### (2) 案例成效：

测绘地理信息领域知识图谱系统将地理时空数据与不同领域的专题数据进行知识融合，示范知识图谱在基础测绘数据领域的应用前景，增强了测绘管理部门的地理时空数据管理水平，扩展了基础测绘数据应用广度，提升了自然资源领域的服务能力和水平。此外，结合地理时空数据和专题数据构建的知识图谱，支持计算城市体检指标，有效提升城市体检工作效率，进而提升城市居民满意度。

### 三、案例系统架构



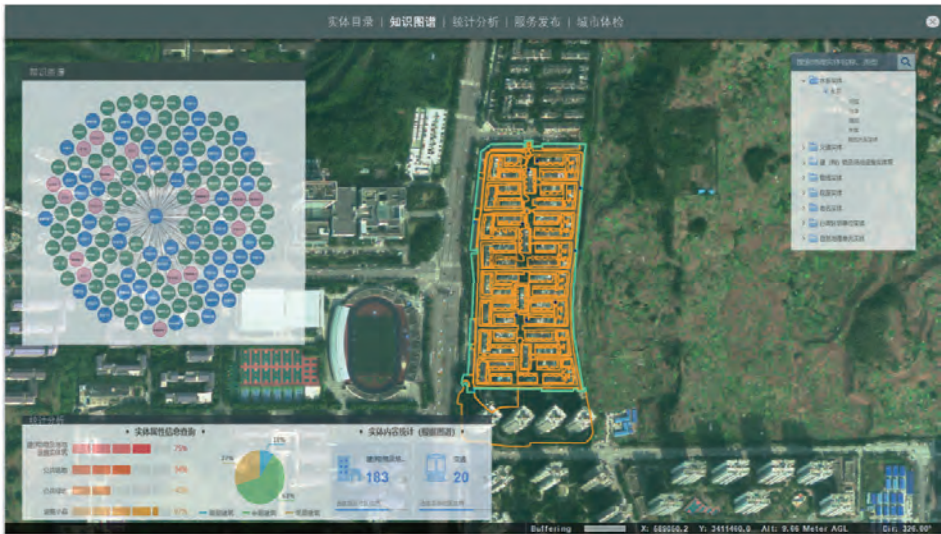
图A2-1 案例系统架构图

#### (1) 数据接入与预处理

将地理实体数据和专题数据分开管理，降低数据的耦合度，以便后续更新维护语义关系库。以某城市某区域内200平方公里范围内的地理实体数据为例，从三个方面接入和预处理数据：一是基于统一的数据标准与技术规范，生产、处理数据，挖掘数据间的语义关系，将图谱实例化；二是提取该范围内专题数据的本体类，构建专题领域的知识图谱；三是通过地理实体空间身份编码关联专题数据和地理实体。

#### (2) 知识图谱和地理实体数据的联动可视化

基于梳理的各地理实体间的语义关系，构建测绘地理信息领域知识图谱，实现知识图谱和地理实体数据的联动可视化。



图A2-2 地理实体属性展示页面

### 1) 智能查询

从地理实体数据出发，面向不同的应用场景，构建与之对应的地理实体查询规则数据集，训练学习地理实体查询的规则特征，以支持地理实体的查询，实现“所得即所想”，做到“人机交互”。

### 2) 城市体检指标的计算及展示

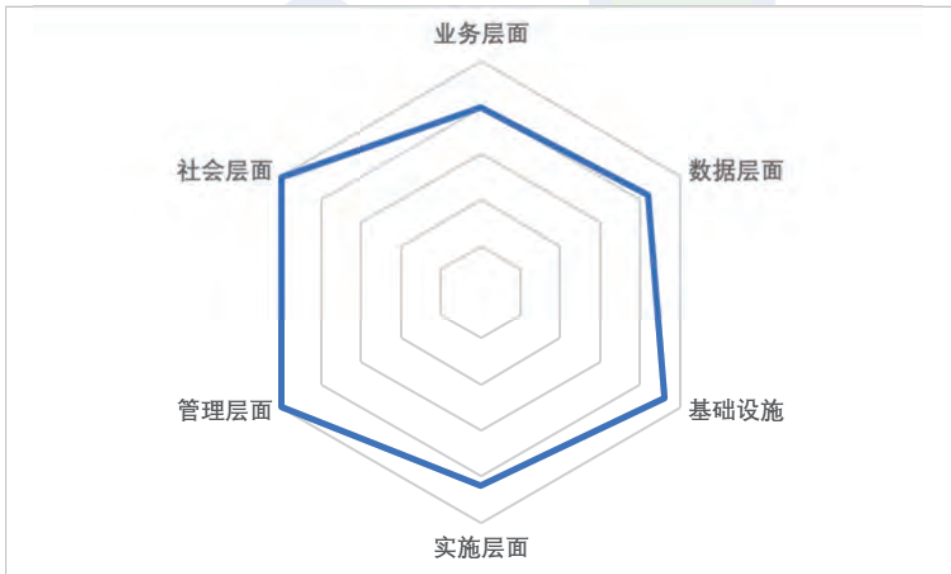
系统与多行业专题相结合，可以满足大多数应用服务需求。以城市体检为例，基于地理实体数据的属性信息及语义关系，计算分析城市体检的各项指标。同时，综合考虑城市体检规程设计和客户意见两个因素，从“宜居”、“宜业”、“宜乐”、“宜游”四个专题挑选典型指标，并融入知识图谱，以区分细化图谱中的数据，实现图谱的专题查询展示。





图A2-3 城市体检指标计算页面

#### 四、案例开发前基础能力评估情况



图A2-4 基础能力评估



目前测绘地理信息领域知识图谱系统在业务层面处于L4级别，在数据层面处于L4级别，在基础设施层面处于L5级别，在实施层面处于L4级别，在管理层面处于L5级别，在社会层面处于L5级别，计算得到其综合得分为93分，系统处于L5完备级别。

## (2) 短板提升措施

### 短板1：缺少具体的应用情景。

用户需求以管理需求为主，在战略层面重点规划应用情景，但在实际应用中缺少具体的应用场景。

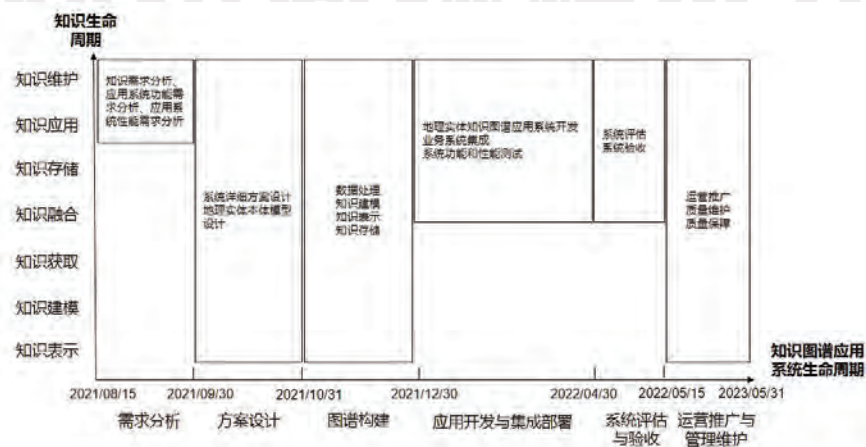
提升措施：基于本项目构建地理时空知识图谱典型应用案例，拓展具体的应用情景。

### 短板2：地理实体关系构建复杂

地理实体数据的天然多领域复杂性，导致数据内容过大，需要一系列的关系计算和一致性检查工具来处理实体关系。

提升措施：结合数据特点，开发关系计算和语义一致性检查工具。

## 五、案例开发与实施历程及难点



图A2-5 案例开发与实施历程

### (1) 需求分析阶段（2021年8月15日至2021年9月30日）

该阶段涉及角色包括测绘地理信息领域专家、知识图谱领域专家、项目经理、需求分析师、数据治理负责人、业务人员等，主要完成行业调研、业务需求梳理和分析、数据现状梳理和分析、可行性分析等工作。

该阶段的难点主要在于对精确把握业务需求、明确工作目标。知识图谱技术在测绘地理信息领域的应用尚处于探索阶段，缺乏具体应用场景，需要通过业务专家、技术专家等多方人员深入交流和探讨，并借鉴其他行业知识图谱应用案例，以实际业务应用场景为基础，完成对需求的分类定级，明确业务目标。

### (2) 方案设计阶段（2021年10月1日至2021年10月31日）

该阶段涉及角色包括项目经理、业务专家、系统架构师、研发工程师、数据工程师、算法工程师、UI设计师等，主要完成本体设计方案、系统详细设计方案、系统设计原型等工作。

该阶段的难点在于多源异构的地理信息数据语义化处理的技术路线设计。测绘地理信息数据来源、数据结构、数据模态多样，如何以较高的自动化程度进行语义化处理和语义融合，当前行业内并没有非常好的解决方案。在本项目中，泰瑞数创结合多年的语义化技术研发经验和研发成果，设计了一整套地理信息数据语义化处理和语义化地理实体生产的工艺流程和标准规范体系，能够快速构建的地理实体语义化数据。

### (3) 图谱构建阶段（2021年11月1日至2021年12月30日）

该阶段涉及角色包括项目经理、数据工程师、知识图谱工程师、研发人员、业务人员等，主要完成数据处理、知识建模、知识融合、知识存储等工作。

该阶段的主要难点在于设计合适的、满足业务应用的知识表示模型和本

体模型，提取应用情景强关联的实体关系。这需要项目组采用“自顶向下”的构建方式构建模型，并通过质量评测对构建的知识图谱进行质量管控。

#### (4) 应用开发与集成部署阶段（2022年1月1日至2022年4月30日）

该阶段涉及角色包括项目经理、知识图谱工程师、研发人员、测试人员、业务人员等，主要完成系统功能开发、系统测试和系统部署等工作。

该阶段的主要难点在于保障系统功能和性能满足业务需求，以及业务流程习惯。这需要在实施开发前，业务人员、研发人员进行充分沟通和确认开发需求，同时开发过程中进行多次确认和测试，以保证系统开发满足需求。

#### (5) 系统评估与验收阶段（2022年5月1日至2022年5月15日）

该阶段涉及角色包括项目经理、研发人员、测试人员、运维人员、业务人员等，主要完成系统评估、系统验收、系统上线等工作。

该阶段的主要难点是制定有效的系统评估指标，标准化的系统测试，验证系统稳定性，业务逻辑是否满足应用需求。

#### (6) 运营推广与管理维护阶段（2022年5月16日至2023年5月31日）

该阶段涉及角色包括业务人员、运维人员、销售人员、市场人员等，主要完成系统运营、系统推广、用户意见总结等工作。

该阶段的主要难点一是建立系统运营推广体系。二是如何指导用户自主扩展系统知识，尤其是在图数据库中新增实体和关系，从而扩展系统应用情景。

## 六、涉及的关键技术

### 关键技术1：地理实体语义化技术

地理实体成果包括结构化、非结构化及半结构化的数据，属性和关联

关系通常采用非结构化的自然语言进行描述，这种描述方式不利于计算机处理与理解。目前计算机难以理解地理实体数据中的三维数据。

语义化尤其是三维语义化是解决上述问题的关键，三维语义化把实体对象的几何信息与语义信息结合起来，将实体转化成计算机可以理解的语言模型，实体对象变得可查询、可分类、可统计、可渲染、可计算、可分析等。泰瑞数创提出地理实体语义化技术，通过人工智能技术研发三维数据的语义信息提取和表达工具、地理实体语义关系的构建工具，有效解决了知识图谱在新型基础测绘领域的应用难点。

### 关键技术2：开源共享的语义化地理实体识别解译样本库

针对不同类别、不同区域的地理实体特点，建立涵盖不同数据类型（如地面照片、遥感影像等）、不同波段（如可见光波段、微光波段等）的地理实体识别解译样本，关联必要的地理实体语义信息，覆盖地理实体类别、纹理、光谱、形状和范围等特征，同时建立样本标准化存储、索引与管理规则，形成开源共享的语义化地理实体识别解译样本库，为地理实体的自动化、智能化提取奠定基础。

## 七、当前面临的重要挑战

### 挑战：基于情境提供知识服务

传统的地理信息服务通过硬编码的方式，实现若干事先设定情况下的信息分析与服务，无法根据实际情况灵活、智能地提供服务。随着传感网、物联网技术和智能感知技术的发展，信息获取的实时化正在逐步成为现实。为此，需要发展融合智能分析的实时感知情境化服务。

对时空信息而言，要发展时空情境觉知的能力，需要融合认知化的智能分析，以提供情境化服务。虽然基于位置的服务(location based services, LBS)

也属于一种时空情境化服务,但时空情境的觉知远远不限于此。情境化服务的关键不仅在于信息的感知,更在于如何把感知到的位置、时间、空间场景、地理行为习惯等时空信息有效融合,实现时空情境的深度觉知,并基于此驱动相关应用场景的服务。



知识图谱标准化  
Knowledge Graph Standardization

## 案例三：通信认知管控系统 建设与实施

### 一、案例背景

随着5G无线通信网络的快速部署，用户终端数不断增加，业务类型种类也日益增多，这给通信运营商带来了挑战。一方面，随着用户规模的持续性增长，通信运营商需要提供大量的咨询服务，另一方面，如何在满足大量用户通信需求的情况下高效、低成本地维护通信网络，提升运营效率变得尤为重要。

为解决通信领域频谱管控中的数据孤岛、数据关联性差等问题，综合利用深度学习、知识图谱等技术，引入频谱数据，构建覆盖全国频谱使用规则的大规模频谱管控系统，实现客服智能问答，鉴别恶意用频或非法用户、伪基站、黑飞无人机，高效调度和分配网络资源等，提升用户满意度，降低运营成本。

### 二、案例所处阶段及成效

(1) 所处阶段：应用开发与集成部署阶段

(2) 案例成效：

从无线通信的结构化数据、半结构化数据和非结构化数据中，挖掘用户、通信业务、无线资源等相关的管理关系和通信知识，构建无线通信知识图谱，为更好地服务用户、实现非法用频预警、诈骗电话识别、通信资源的高效智能管理提供重要的技术手段。