

# 知识图谱驱动的智能标准化转资方法

沈银萱<sup>1</sup> 程之曼<sup>2</sup> 刘锦屏<sup>2</sup> 谢振平<sup>1\*</sup> 何谐<sup>3</sup> 刘志<sup>2</sup>

<sup>1</sup>(江南大学人工智能与计算机学院 江苏 无锡 214122)

<sup>2</sup>(四川长宁天然气开发有限责任公司 四川 成都 610051)

<sup>3</sup>(中国石油天然气股份有限公司西南油气田川东北作业分公司 四川 成都 610051)

**摘要** 针对油气田企业资产管理需求与国家高质量发展趋势,该文以长宁公司业务为依据,提出智能标准化转资方法。通过系统智能化技术支持,引入知识建模和知识图谱驱动的数据处理算法,重点解决了人工操作繁琐的转资过程,实现了资产流程标准化管理。这一研究成果优化了资产信息整合与转资管理,显著提升了油气资产管理的准确性、可靠性和时效性。该成果对推动油气田企业资产管理智能化具有重要意义,为业内人员提供了有益的参考和启示。

**关键词** 固定资产管理 智能化转资 流程标准化 知识图谱

中图分类号 TP391.1

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.10.007

## INTELLIGENT STANDARDIZED ASSETS TRANSFER METHOD DRIVEN BY KNOWLEDGE GRAPH

Shen Yinxuan<sup>1</sup> Cheng Zhiman<sup>2</sup> Liu Jinping<sup>2</sup> Xie Zhenping<sup>1\*</sup> He Xie<sup>3</sup> Liu Zhi<sup>2</sup>

<sup>1</sup>(School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, Jiangsu, China)

<sup>2</sup>(Sichuan Changning Gas Development Company, Chengdu 610051, Sichuan, China)

<sup>3</sup>(Northeast Sichuan Operating Branch of Southwest Oil and Gas Field, CNPC, Chengdu 610051, Sichuan, China)

**Abstract** In view of the asset management needs of oil and gas field enterprises and the national high-quality development trend, this paper proposes an intelligent standardized assets transfer method based on the business of Changning Company. Through the intelligent technical support of the system, the introduction of knowledge modeling and knowledge graph driven data processing algorithm, the manual operation of the cumbersome assets transfer process was solved mainly, and the standardized management of the asset process was realized. This research result optimizes asset information integration and assets transfer management, and significantly improves the accuracy, reliability and timeliness of oil and gas asset management. The results are of great significance to promote the intelligent asset management of oil and gas field enterprises, and provide useful reference and enlightenment for the industry personnel.

**Keywords** Fixed assets management Intelligent assets transfer Process standardization Knowledge graph

## 0 引言

油气田企业的固定资产往往具有规模巨大、种类繁多、用途特殊、地域分布广、管理难度大等特征。一般会存在转资不及时、资产管理标准不统一、重视工程建设而轻视资产管理等问题<sup>[1]</sup>。根据事业单位固定资产管理原则,首先要做到管理效率优先;其次是以实物

管理为主,保证物尽其用<sup>[2]</sup>。油气资产管理主要围绕转资、调拨、报废、处置、盘点这五个环节,分别在实物和账务两个方面开展资产的管理工作<sup>[3]</sup>。转资是指所购建资产达到预定可使用状态,经决算审计完成后,按照审定的成本,由在建工程转入固定资产核算及管理的过程<sup>[4]</sup>。转资作为整个管理流程的起点,需要项目建设部门参与结算和提供转资材料,并由财务资产管理部门验收后将相关信息录入系统,再进行固定资产

后续的全生命周期管理。然而转资时间往往长达几个月,还存在转资材料不准确、转资信息不完备、实物资产归集不规范等诸多问题,进而影响其他资产管理环节,最终导致资产管理的精度和质量低下<sup>[4]</sup>。因此,如何因地制宜地对固定资产进行标准化分类,提高转资环节效率和准确率,并进行统一的流程标准化管理,是油气田企业亟须解决的问题。

近年来,国内外在标准化数据分类和知识图谱计算与应用等方面开展了大量相关研究。Shanavas 等<sup>[5]</sup>根据术语与文本分类的相关性引入了加权共现图来表示文本文档,并利用词相似度矩阵形式的语义知识自动丰富加权图,确保了术语和模式的精确匹配。Yuan 等<sup>[6]</sup>基于元数据注册(MDR)概念元模型,对语义关系类型和表示进行扩展,从而构建了一个标准、可扩展、通用的知识图谱语义关系元模型,实现传统数据语义结构向知识语义结构的迁移图。马冬雪等<sup>[7]</sup>提出了一种基于本体的网页自动解析新方法,较好地实现了自适应地对网页中语义信息的结构化解析抽取。Zhou 等<sup>[8]</sup>提出了将知识图谱应用于中石油上游领域的一种整合多源异构数据的方法。陈思等<sup>[9]</sup>提出基于“一致

性检验”与“语料评价”的骨架法应用策略,提高了知识模型构建效率。

基于上述文献在标准化数据分类、知识图谱构建、图谱语义解析匹配等方面的研究思路,本文结合四川长宁天然气开发有限责任公司(以下简称“长宁公司”)<sup>[10]</sup>在油气资产管理方面的现存问题,包括资产分类标准不统一、资产管理流程各环节操作及数据格式不规范、转资工作效率低下等,提出了一种基于知识图谱的标准化智能转资方法,旨在综合多种技术手段,实现转资环节中业务操作的标准化和数据处理的智能化,提升资产管理流程中各环节的规范化和准确性。

## 1 方法过程

### 1.1 方法框架

针对长宁公司人工操作非常繁琐的转资过程,本文提出了基于知识图谱的标准化智能转资方法,主要的体现结构框架如图 1 所示。包括固定资产标准化模板构建、模板图谱构建、资产项目图谱构建、知识计算、文档导出和文档整合,共 6 个模块。

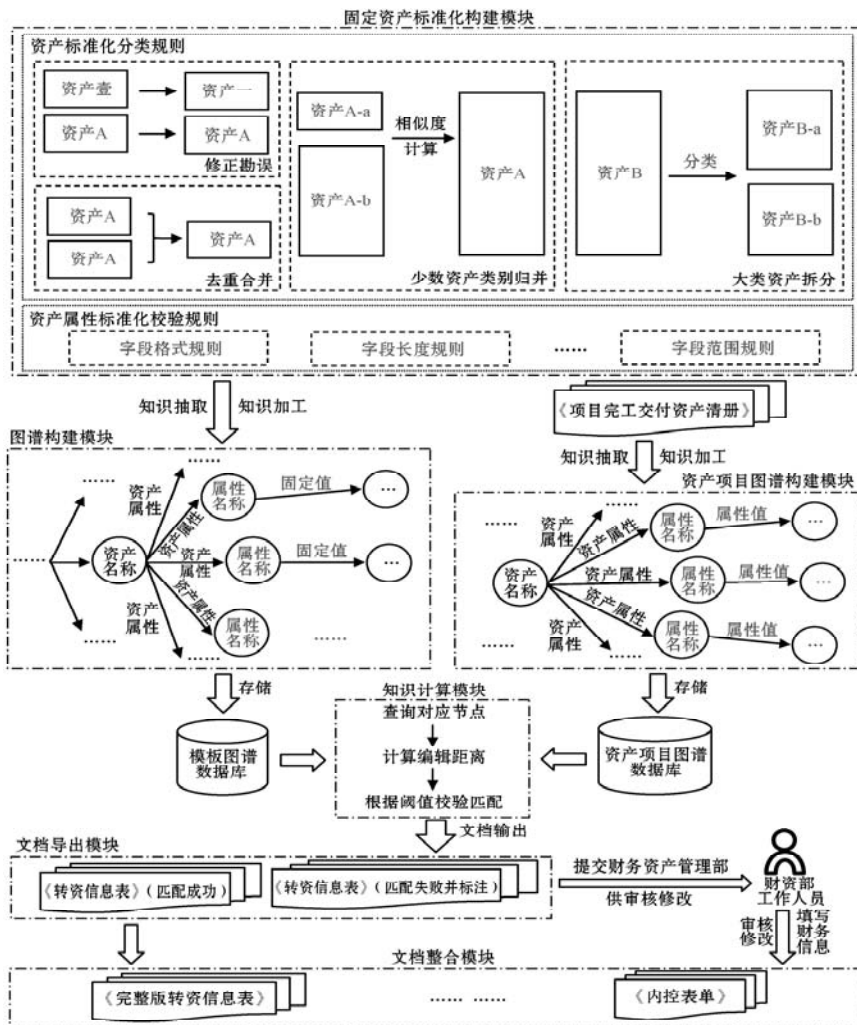


图 1 知识图谱驱动的智能标准化转资方法框架

首先,通过文本处理、相似度计算、层次聚类等技术手段,将长宁公司现有的固定资产进行标准化分类,得到一套完整的固定资产分类标准,并作为准则嵌入转资模板中;其次,利用信息自动化抽取、知识加工,构建资产项目数据图谱和模板图谱并存储到数据库中;然后利用编辑距离计算等方式进行知识计算和匹配校验输出;最后,财务资产管理部门工作人员结合导出的文档进行修改和财务信息填写,形成最终的《完整版转资信息表》以及其他内控表单。

本文提出的技术思路,实现了转资环节从纯手工填表的线下繁琐操作到智能化快速简便操作的质的飞跃,加速跨信息系统的数据库转换过程,并可以依托这些基础数据进一步实现后续资产管理环节的标准化和智能化。

## 1.2 标准化模板构建

本模块分为资产标准化分类规则和资产属性标准化校验两个部分。第一部分包括相同资产的去重合并,修正勘误等处理,少数资产利用相似度匹配和层次聚类的方式归并到大类资产,以及大类资产根据“规格型号”等属性进行类别拆分等。第二部分包括资产的各项属性中所蕴含的校验规则,即字段的格式、长度、范围等规则。比如,“交付日期”属性中所填写的内容必须为“年/月/日”的格式等。

### 1.2.1 数据预处理

本文以编程语言 Python 作为数据预处理工具。

(1) 修正勘误:如图 1 所示,分为空格消除和同音字归并两部分。

空格消除:用正则表达式匹配替换空格为空字符串。调用 re (Regular Expression) 库中的 sub (r'',", text) 函数,用空字符串"替换了“text”文本中的所有空格。

同音字归并:首先,将需要统一的同音字对写入字典,用于同音字映射。比如,将“撬”统一为“橇”的字典示例:dict = {'撬': '橇'}。其次,调用 pypinyin 库中的 lazy\_pinyin ('中文字符串') 函数,将中文字符串转换为拼音列表,定义了一个新函数 unify (),接收拼音列表作为输入,遍历拼音列表中的每个拼音,并根据预先定义的 dict 字典查找相应的同音字映射,将拼音中的同音字统一为一个字;最后,用 joint () 函数将处理后的拼音列表连接为一个字符串并返回。

(2) 去重归并:直接调用 pandas 库中的 duplicates () 函数,将上述处理好的字段 (DataFrame 格式) 进行去重,最终得到预处理结果。

### 1.2.2 语义相似度计算

(1) 计算词对 (A, B) 的语素相似度,如式(1)所示。

$$chasim(A, B) = \frac{2 \times samec(A, B)}{len(A) + len(B)} \quad (1)$$

式中:chasim(A, B) 为词对 (A, B) 的语素相似度,samec(A, B) 为词对 (A, B) 的共同语素个数,len(A) 表示词语 A 的长度。

(2) 计算词对 (A, B) 的语序相似度,如式(2)所示。

$$ordsim(A, B) = \begin{cases} 1 - \frac{revord(A, B)}{|oncec(A, B)| - 1} & |oncec(A, B)| > 1 \\ 1 & |oncec(A, B)| = 1 \\ 0 & |oncec(A, B)| = 0 \end{cases} \quad (2)$$

式中:ordsim(A, B) 为词对 (A, B) 的语序相似度,oncec(A, B) 表示在词 (A, B) 中出现且只出现一次的语素的集合,|oncec(A, B)| 表示该集合中元素的个数。

(3) 计算词对 (A, B) 的词长相似度,如式(3)所示。

$$lensim(A, B) = 1 - \left| \frac{len(A) - len(B)}{len(A) + len(B)} \right| \quad (3)$$

式中:lensim(A, B) 为词对 (A, B) 的词长相似度,len(A) 表示词语 A 的长度。

(4) 对三种相似度加权计算词对 (A, B) 的词语相似度,如式(4)所示。

$$wordsim(A, B) = \alpha \times chasim(A, B) + \beta \times ordsim(A, B) + \gamma \times lensim(A, B) \quad (4)$$

式中:wordsim(A, B) 为词对 A, B 的词语相似度,取  $\alpha = 0.7, \beta = 0.29, \gamma = 0.01$ 。

### 1.2.3 层次聚类

层次聚类作为一种聚类算法,通过逐步合并或分裂数据点来构建聚类层次结构<sup>[11]</sup>。层次聚类分为凝聚的(自下而上)或分裂的(自上而下)。本文采用凝聚层次聚类算法,从单个数据点开始,逐渐将相似的数据点合并为越来越大的类。步骤如下:首先,获取预处理结果中每项资产的“资产名称”、“资产类别”和“资产组名称”三个属性内容;其次,两两计算所有资产项(词对)之间的名称相似度,记为 namesim(A, B),类别相似度,记为 catsim(A, B) 以及组别相似度,记为 groupsim(A, B),计算过程见 1.2.2 节。对三种相似度加权计算词对 (A, B) 的总相似度,如式(5)所示。

$$totalsim(A, B) = \alpha \times namesim(A, B) + \beta \times catsim(A, B) + \gamma \times groupsim(A, B) \quad (5)$$

式中:totalsim(A, B) 为词对 (A, B) 的总相似度,取  $\alpha = 0.6, \beta = 0.3, \gamma = 0.1$ 。

然后,用总相似度作为矩阵元素构造资产项之间的相似度矩阵,设置相似度阈值,进行后续的分类合并。相关概念及过程详解如下:

(1) 相似度矩阵: $x_{ij} \in \mathbf{R}^{n \times n} (1 \leq i < n, 1 \leq j < n)$  其

中,  $R$  为  $n \times n$  的矩阵,  $n$  表示样本的数量, 矩阵中的每个元素  $x_{ij}$  表示第  $i$  个样本和第  $j$  个样本之间的相似度。通常, 相似度矩阵是一个对称矩阵 (即  $x_{ij} = x_{ji}$ ), 对角线上的值为每个样本与自身的相似度。

(2) 合并规则: 一般包含单链接、完全链接、平均链接三种规则<sup>[13]</sup>。本文选择平均链接 (Average Linkage), 即在层次聚类时, 采用两个类别中所有样本的平均相似度作为合并依据。

(3) 合并过程: 首先, 初始化每项资产为一个单独的类别; 其次, 根据阈值判断, 循环合并相似度最高的两个类别, 并且将合并的类别视为一个整体, 重新计算与其他类别的相似度, 同时更新相似度矩阵, 直到所有样本被聚类完毕。如图 2 所示。

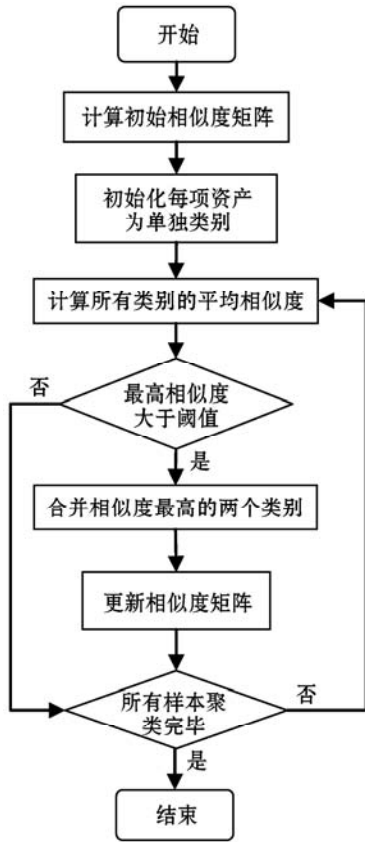


图2 类别合并过程

#### 1.2.4 类别拆分

将大类资产 (即类别中包含超过 100 个资产) 根据“规格型号”等属性的差异性, 进行类别拆分, 以确保每个类别中资产数量在 3 ~ 100 之间。

### 1.3 图谱构建

知识图谱本质上是一种语义网络<sup>[14]</sup>, 用于揭示各类实物之间的关系, 旨在从多种类型的复杂数据中抽取概念、实体和关系, 是事物关系的可计算模型<sup>[15]</sup>。

知识图谱的一种通用表示形式是三元组形式, 即  $G = (Entity_{head}, Relation, Entity_{tail})$ ,  $Entity_{head}$  为三元组  $G$

中的头实体,  $Entity_{tail}$  为尾实体,  $Relation$  为 2 个实体之间的关系, 其中,  $Entity = [Entity_1, Entity_2, \dots, Entity_n]$  表示实体的集合, 其包含了  $n$  种实体的概念,  $Relation = [Relation_1, Relation_2, \dots, Relation_n]$  表述实体之间的关系集合, 其包含了  $n$  种不同的关系<sup>[16]</sup>。本文将知识图谱的形式定义为  $KG = \{E, R, A\}$ , 其中  $E, R, A$  分别为实体、关系和属性的集合<sup>[17]</sup>, 采用 Neo4j 图数据库<sup>[18]</sup> 作为图模型的存储方式, 以节点 (node) 和边 (edge) 来组织数据。node 代表知识图谱中的实体, edge 代表实体间的关系, 关系可以有方向, 两端对应开始节点和结束节点。

构建知识图谱的主要分为知识抽取、知识融合、知识存储与可视化三个部分。下面介绍模板图谱和资产项目关系图的构建过程。

#### 1.3.1 模板图谱

(1) 实体抽取。从资产模板中抽取资产名称作为唯一实体, 其他资产属性作为与资产名称相关联的实体。假设模板中包含  $m$  个资产名称,  $n$  个其他资产属性, 资产名称抽取可以表示为集合  $E$ , 其他资产属性抽取可以表示为集合  $P$ 。

$$E = \{e_1, e_2, \dots, e_m\}$$

$$P = \{p_1, p_2, \dots, p_n\}$$

其中,  $e_i \in E (1 \leq i \leq m)$  为资产名称,  $p_j \in P (1 \leq j \leq n)$  为其他资产属性。

(2) 关系抽取。将资产名称与其他属性建立关系, 同时对具有校验规则的属性也建立相应关系, 表明属性与资产名称的关联。假设属性集合包括  $n$  个属性, 其中每个属性  $p$  可以分为固定内容属性和具有校验规则的属性, 关系抽取可以表示为集合  $R$ 。

$$R = \{r_1, r_2, r_3 \mid r_1 = \text{“属性”}, r_2 = \text{“固定值”}\}$$

(3) 属性抽取。从资产模板中抽取固定内容属性内容, 假设模板中包含  $t (t \leq n)$  个固定属性值, 则可以表示为集合  $A$ 。

$$A = \{a_1, a_2, \dots, a_t\}$$

其中,  $a_q \in A (1 \leq q \leq t)$  为固定值属性内容。

(4) 知识表示。用三元组表示资产名称与属性之间的关系。三元组可表示为集合  $KG$ 。  $KG = \{E, R, A\}$ , 其中  $E, R, A$  分别为实体、关系和属性的集合。

$$KG = \{(e_i, r_1, p_j) \mid e_i \in E, p_j \in P, 1 \leq i \leq m, 1 \leq j \leq n\} \cup \{(p_k, r_2, a_q) \mid p_k \in P, a_q \in A, 1 \leq k \leq t, 1 \leq q \leq t\}$$

其中,  $e_i \in E (1 \leq i \leq m)$  为资产名称,  $p_j \in P (1 \leq j \leq n)$  为其他资产属性,  $r_1 = \text{“属性”}$ 。  $p_k \in P (1 \leq k \leq t)$  为固定值属性,  $a_q \in A (1 \leq q \leq t)$  为固定值属性内容,  $r_2 = \text{“固定值”}$ 。

(5) 知识存储与可视化。根据上述构建步骤,即实体对应资产名称及其他属性名称,关系对应资产属性、属性内容、校验规则,属性对应属性内容/校验格式要求,可将模板中所有资产及相关属性、准则准确映射到图谱中。最后采用 Neo4j 图数据库进行模板图谱的可视化呈现。

### 1.3.2 资产项目关系图

根据长宁公司的业务需求,转资一般以项目为单位,每个项目均会提供相应的《完工交付资产清册》,其中包含该项目中所有资产的信息。资产项目关系图的构建过程与模板图谱大致相似,同样通过知识抽取、融合、存储与可视化等步骤,因此,不过多赘述。最终同样用三元组集合  $KG$  表示资产与属性间的关系。 $KG = \{E, R, A\}$ , 其中  $E, R, A$  分别为实体、关系和属性的集合。

$$E = \{e_1, e_2, \dots, e_m\}$$

$$P = \{p_1, p_2, \dots, p_n\}$$

$$R = \{r_1, r_2, r_3 \mid r_1 = \text{"属性"}, r_2 = \text{"固定值"}\}$$

$$A = \{a_1, a_2, \dots, a_n\}$$

$$KG = \{(e_i, r_1, p_j) \mid e_i \in E, p_j \in P, 1 \leq i \leq m, 1 \leq j \leq n\} \cup \{(p_j, r_2, a_l) \mid p_j \in P, a_l \in A, 1 \leq j \leq n, 1 \leq l \leq n\}$$

其中,  $e_i \in E (1 \leq i \leq m)$  为资产名称,  $p_j \in P (1 \leq j \leq n)$  为其他资产属性,  $r_1 = \text{"属性"}$ 。  $a_l \in A (1 \leq l \leq n)$  为属性内容,  $r_2 = \text{"固定值"}$ 。

最终通过 Neo4j 图数据库存储并可视化资产项目关系图中资产与资产属性之间的映射关系。

## 1.4 数据项智能匹配

### 1.4.1 计算编辑距离

将模板图谱记作三元组集合  $KG_T = \{\{E_1, R_1, A_1\}, \{E_2, R_2, A_2\}, \dots, \{E_n, R_n, A_n\}\}$ , 资产项目图谱记作  $KG_P = \{\{E_1, R_1, A_1\}, \{E_2, R_2, A_2\}, \dots, \{E_m, R_m, A_m\}\}$ 。现将两个三元组集合均转换为字符串集合, 记作  $S_i = \{S_1, S_2, \dots, S_i, \dots, S_n\}$  和  $C_p = \{C_1, C_2, \dots, C_j, \dots, C_n\}$ 。其中,  $1 \leq i \leq n, 1 \leq j \leq m, S_i$  和  $C_j$  的形式均为“实体-关系-属性”。

计算字符串  $S_i (i = 1, 2, \dots, n)$  与  $C_j (j = 1, 2, \dots, m)$  之间的相似度, 如式(5)所示。

$$Sim(S_i, C_j) = 1 - \frac{LD(S_i, C_j)}{N + M} \quad (5)$$

式中:  $LD(S_i, C_j)$  表示编辑距离,  $N$  和  $M$  代表  $S_i$  和  $C_j$  的长度。计算编辑距离如式(6)所示。

$$D_{ij} = \begin{cases} 0 & i=0, j=0 \\ \min(D_{i-1, j-1} + \omega_a, D_{i-1, j} + \omega_b, D_{i, j-1} + \omega_c) & \text{其他} \end{cases} \quad (6)$$

式中: 矩阵  $D_{ij} = D(s_0, s_1, \dots, s_i, c_0, c_1, \dots, c_j) (0 \leq i \leq n, 0 \leq j \leq m)$ ,  $s_0, s_1, \dots, s_i$  表示字符串  $S_i$ ,  $c_0, c_1, \dots, c_j$  表示

字符串  $C_j$ , 通过  $(n+1) \times (m+1)$  阶矩阵  $D_{ij}$  计算得到  $LD(S_i, C_j)$ ;  $D_{i-1, j-1}$  为删除操作,  $D_{i-1, j}$  为插入操作,  $D_{i, j-1}$  为替换操作, 设置删除的权重  $\omega_a$  和插入的权重  $\omega_b$  为 1, 替换的权重  $\omega_c$  为 2, 即  $\omega_a = 1, \omega_b = 1, \omega_c = 2$ 。从字符串  $S_i$  的第一个位置  $s_0$  和字符串  $C_j$  的第一个位置  $c_0$  进行比较, 对已经比较过的编辑距离, 继续计算下一个字符位置的编辑距离。矩阵  $D_{ij}$  从  $D_{00}$  逐行逐列计算, 最终得到  $D_{nm}$ , 即编辑距离  $LD(S_i, C_j)$ 。

### 1.4.2 阈值匹配校验

经过计算得到  $S_i$  和  $C_j$  之间的相似度  $Sim$ ,  $Sim$  值越大表示相似度越高。设置相似度  $Sim$  的阈值: 小于阈值继续进行迭代; 大于等于阈值则保存经过匹配得到的  $S_i$  和  $C_j$  信息, 将匹配成功和失败的  $C_j$  内容按照格式要求分别输出至指定 Excel 文档。随后判断字符串集合  $C_p = \{C_1, C_2, \dots, C_j, \dots, C_n\}$  是否遍历结束, 遍历结束则输出两个文档, 供财资部作为修改参考。具体算法设计流程如图 3 所示。

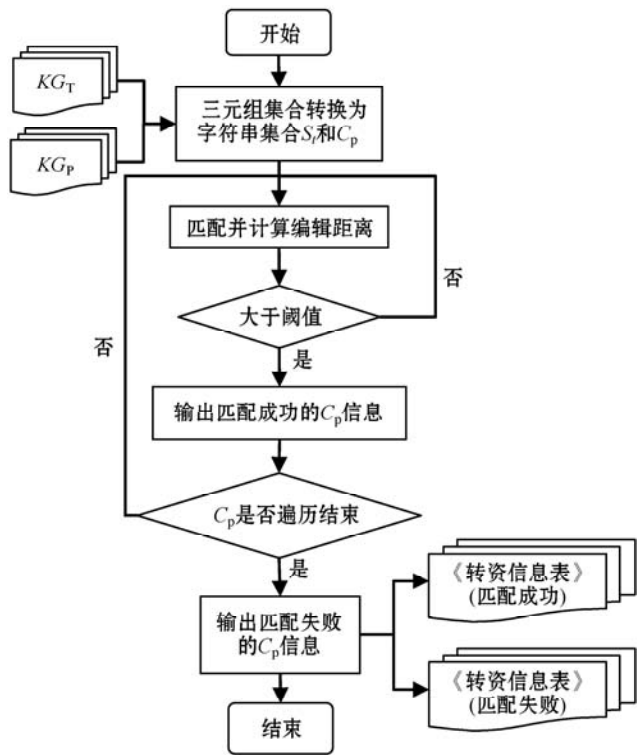


图 3 数据项智能匹配流程

## 2 系统实现与结果分析

本文设计的智能化转资系统流程以标准化模板为基础, 通过图谱的构建和系统的自动化匹配校验, 辅以财务资产管理部门工作人员的交互配合, 将长宁公司繁杂的转资流程以及数据进行标准化、智能化统一处理, 最终准确高效地形成油气资产管理的基础数据和内控表单。

## 2.1 系统实现

图 4 展示了智能化转资流程的系统框架。框架中所实现的转资流程包括转资申请、转资审核、转资验收和转资确认。其中,转资验收包括线上和线下两个部分。线上验收步骤包括系统中实现的资产属性标准化校验(如图 1 所示)、资产清册核对等,系统中对应的线上验收界面如图 5 所示。线下验收步骤包括:① 建设单位对资产进行贴标签、定位、拍照,并在资产航拍图的对应位置进行标注;② 财资部借助 RFID<sup>[19]</sup> (Radio Frequency Identification, 射频识别技术),通过 PAD 对完工现场资产的标签进行远程扫描定位验收。系统中对应的航拍图验收界面如图 6 所示。

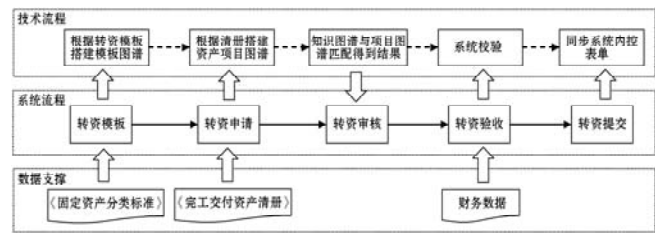


图 4 智能化转资流程框架



图 5 线上验收-资产交付清单界面



图 6 航拍图验收界面

## 2.2 结果分析

智能化与标准化在本文提出的转资流程中发挥了关键作用。智能化的引入大幅提高了转资的效率,同时标准化的实施则极大地提升了转资信息的准确性,从而为油气资产管理全流程的标准化奠定了基础。

### 2.2.1 转资智能化

如图 4 所示,本文提出的智能化转资流程通过知

识图谱等技术的应用,大大减少了人力物力消耗。如表 1 所示,转资操作时长显著缩短,新增了资产标签、航拍图标注等功能使得资产定位准确,并且支持数据痕迹溯源。这意味着系统的自动化和智能化程度使得传统繁琐流程变得更为高效迅捷。如图 5 - 图 6 所示,系统页面直观简洁,操作便捷。

表 1 转资性能提升对比

性能参数对比		传统转资	智能标准化转资
操作耗时		36 天	12 天
信息采集	资产信息数据	有(账实不对应)	有(账实对应)
	资产财务数据	有(账实不对应)	有(账实对应)
	资产标签数据	有(账实不对应)	有(账实对应)
	航拍图	无	有
其他	资产定位	无	有(账实对应)
其他	数据痕迹	无	支持
	准确率/%	70.1	98.4

### 2.2.2 转资标准化

(1) 标准化论证:本文的转资流程设计以《中石油集团公司固定资产转资工作应用指南-油气田企业篇》(以下简称《应用指南》)为参考标准。经长宁公司的内部审查与验证,本文提出的转资方法与参考标准间的对比审查结果如下:

资产归集:《应用指南》中规定,油气井资产与其配套设施均有相应的列示原则。本文通过资产标准化分类得到的模板,经核定,分类结果符合标准。

数据填写规范:《应用指南》中对“资产名称”“规格型号”“计量单位”等属性内容均有明确规范要求。本文设计的转资方法在“资产属性标准化校验规则”模块有对应设计,并在系统中进行了校验规则的嵌入的实现,经核定,完全符合标准要求。

转资流程:《应用指南》中仅规定了转资条件与表单提交等必要流程。本文根据规定,并结合长宁公司的具体业务及表单,设计实现了如图 4 所示的转资流程。经审查核定,符合标准规定。

(2) 标准化优势:提升信息准确率。准确率指转资信息的正确性,即信息处理后与实际情况的一致性程度。计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (7)$$

式中: $T_p$  为正确匹配的样本数, $F_p$  为错误匹配的样本数, $P$  为信息准确率,衡量正确匹配的样本占全部匹配样本的比例。

传统的转资方法主要通过反复地人工填写表格、

交接资料,以及手动录入系统来完成所需的转资信息。往往会因为填写和手工录入的勘误等问题,加上后续没有设置智能化校验环节,且人工校验存在疏漏,从而导致准确率较低。

本文设置的准确率统计实验基于 Windows11 系统, CPU 3 GHz,内存 16 GB,实验工具为 Python。实验统计了长宁公司使用传统方法和本文方法对公司所有完工资产进行转资录入的信息准确率,对比结果如表 2 所示。据显示,通过标准化方法的引入,转资信息的准确率得到了显著提高,从原本的 70.1% 提升至 98.4%,由此可见,本文方法大幅提升了转资过程中信息的可靠性。

表 2 准确率对比统计结果

方法	正确样本数	错误样本数	准确率/%
传统方法	4 515	1 925	70.1
本文方法	6 339	101	98.4

### 3 结 语

资产标准化管理和智能化转资是当前油气田资产管理领域的热点。本文就相关问题进行了探讨研究,重点分析了长宁公司的资产管理难点,并通过标准化模板的构建、信息自动化抽取、知识图谱构建和相似度匹配等一系列技术手段,为长宁公司实现了资产标准化管理和转资智能化流程,减少了大量人工操作内容,极大地提高了资产信息的管理效率,且能为后续资产管理环节提供精确的数据支撑。

经研究发现,制定统一的资产管理标准和流程,以及加强系统的智能化设计、减少人工操作,是实施资产智能化管理的关键点。长远来看,油田气企业可以根据自身情况建立企业内部行业知识体系,以图谱形式描述多个资产之间的复杂关系,对于不同的部门或业务流程,提供一套标准的资产信息管理及校验体系,从而串联起多个模块的功能,提升资产管理效率。本文研究内容可为其他企业的资产智能化管理提供极有价值的建设参考。

### 参 考 文 献

[ 1 ] Xia T W, Dai Z X, Huang Z, et al. Establishment of technical standard database for surface engineering construction of oil and gas field[J]. Processes, 2023, 11(10): 2831.

[ 2 ] 李涛. 事业单位固定资产管理标准化研究[J]. 现代商业, 2016(20): 160 - 161.

[ 3 ] Ghazali M, Anuar H. Value management: implementation of asset life cycle in one of oil and gas service company[J]. International Journal of Advanced Engineering Research and

Science, 2017, 4(10): 67 - 72.

- [ 4 ] Bukvič V. How to manage and monitor tangible fixed assets effectively[J]. Journal of Accounting and Finance, 2021, 21(5): 187 - 199.
- [ 5 ] Shanavas N, Wang H, Lin Z W, et al. Knowledge-driven graph similarity for text classification[J]. International Journal of Machine Learning and Cybernetics, 2021(12): 1067 - 1081.
- [ 6 ] Yuan M, Liu M Q, Mu M N. Research on semantic relationship and representation standardization model of knowledge graph based on MDR[J]. Journal of the China Society for Scientific and Technical Information, 2023, 42(7): 832 - 841.
- [ 7 ] 马冬雪,宋设,谢振平,等. 领域本体驱动的招投标网页解析方法[J]. 计算机应用,2020,40(6):1574 - 1579.
- [ 8 ] Zhou X G, Gong R B, Shi F G, et al. PetroKG: Construction and application of knowledge graph in upstream area of PetroChina[J]. Journal of Computer Science and Technology, 2020, 35(2): 368 - 378.
- [ 9 ] 陈思,王亚平,胡灿灿,等. 单兵装备人机工效知识图谱构建技术[J]. 兵工自动化,2023,42(2):6 - 12.
- [ 10 ] 何骁. 页岩气规模效益开发模式创新:长宁页岩气钻井“日费制”探索实践[J]. 北京石油管理干部学院学报, 2021,28(5):55 - 60.
- [ 11 ] 张景越,肖小玲,王鹏飞,等. 基于多信息融合的层次聚类测测井曲线自动分层方法[J]. 断块油气田,2024,31(1): 42 - 49.
- [ 12 ] 王胜,苏艳苹. 一种矩阵分解和相似度矩阵学习的多视图聚类算法[J]. 郑州航空工业管理学院学报,2023,41(5): 64 - 69.
- [ 13 ] 韩鑫. 基于核的层次聚类算法研究[D]. 西安:西安石油大学,2021.
- [ 14 ] Zhu L, Liu X T, He S, et al. Keywords co-occurrence mapping knowledge domain research base on the theory of Big Data in oil and gas industry [J]. Scientometrics, 2015, 105: 249 - 260.
- [ 15 ] Yun W, Zhang X, Li Z D, et al. Knowledge modeling: A survey of processes and techniques[J]. International Journal of Intelligent Systems, 2021, 36(4): 1686 - 1720.
- [ 16 ] 张吉祥,张祥森,武长旭,等. 知识图谱构建技术综述[J]. 计算机工程,2022,48(3):23 - 37.
- [ 17 ] Aswadi F, Chan H, Gan K. Automatic ontology construction from text: A review from shallow to deep learning trend[J]. Artificial Intelligence Review, 2020, 53: 3901 - 3928.
- [ 18 ] Zhao H X, Pan Y L, Yang F. Research on information extraction of technical documents and construction of domain knowledge graph[J]. IEEE Access, 2020(8): 168087 - 168098.
- [ 19 ] Xia Y. Study on fixed assets management system based on bar code and RFID technology[J]. Electronic Technology, 2021, 50(9): 230 - 231.