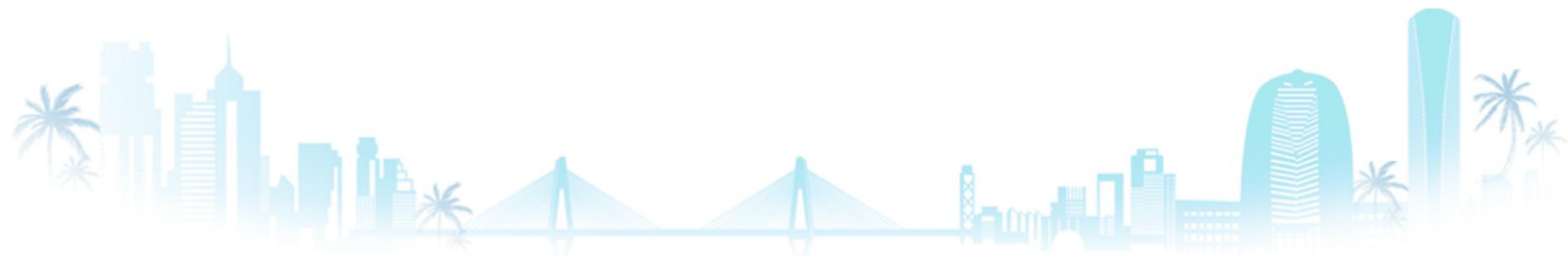


鼎复数据
2017.12.16

FinTech 在智能投研的应用



鼎新革故，复往利亨——鼎复是一家Fintech公司

金融产业链

技术支持

资金端

各类金融机构

- 公募/私募/资管
- 券商/保险
- 会计师事务所
- PE/VC
- 商业银行
- 投资银行
- ...

资产端

传统



网页搜索

Title	Title	Title
o	o	o
o	o	o
o	o	o
o	o	o
o	x	o

数据库



文档/表格处理

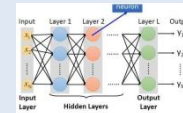


报告



交易

鼎复



深度学习 知识图谱 智能建模

df_x

关键字: Fin-

-tech

鼎复

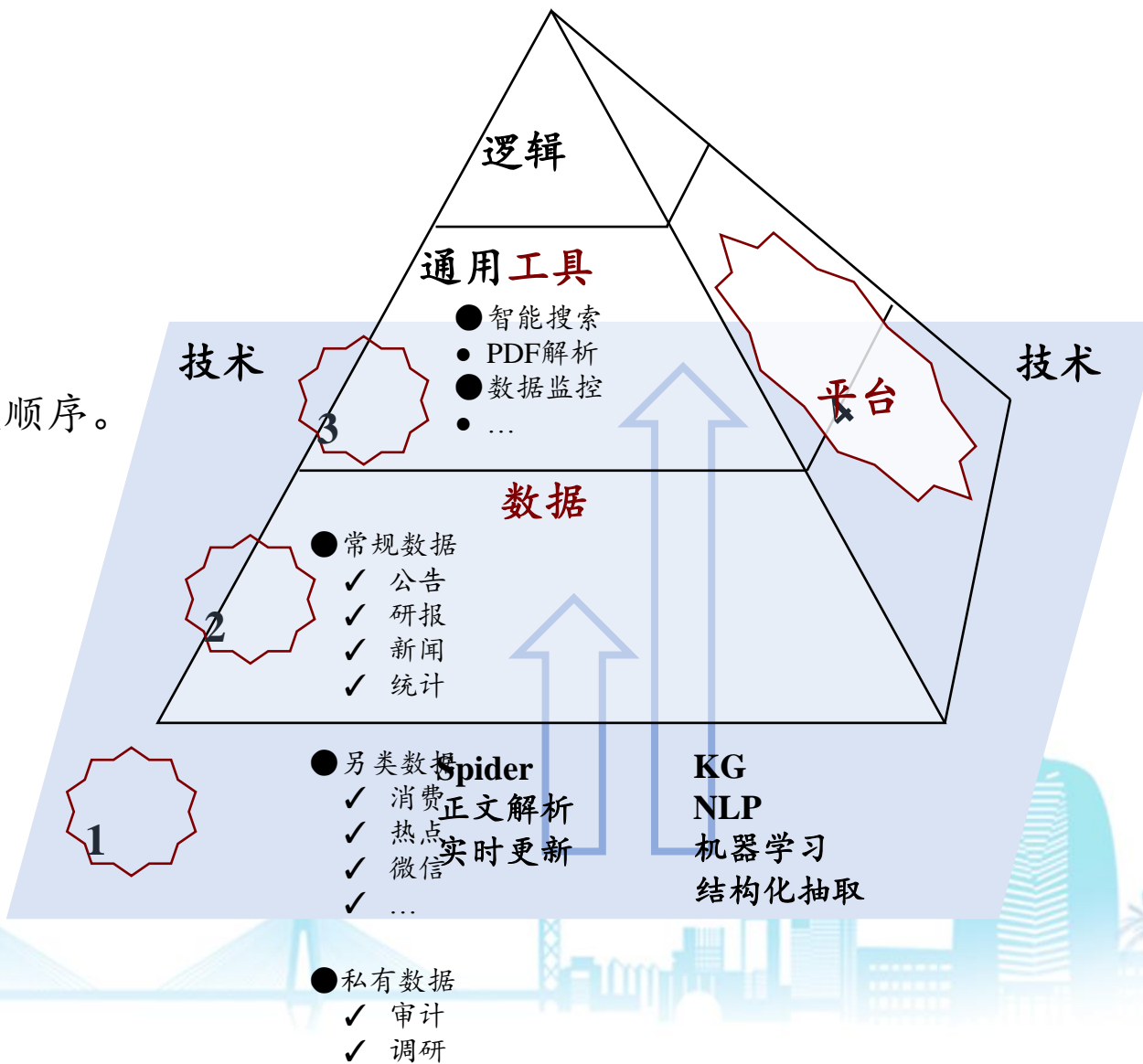
经过2年发展，形成成熟方法论和商业模式

从下至上，依次对应

- 底层技术体系搭建
- 各类数据挖掘积累
- 金融通用工具开发
- 配合场景定制平台

是我们的实现路径与发展顺序。

——亦是从通用到定制的过程，以及技术颠覆人工处理的难度加大过程



团队方面：已形成资深、活力、靠谱的梯队

● 创始人，70s

郭昞

- 百度七剑客
- 工学博士
- 斯坦福Sloan Fellow

张良华

- 中金、黑石高管
- 金融从业经验27年
- 中国人民银行研究生部硕士

吴雪军

- 阿里资深总监
- 百度科学家
- 天涯副总裁

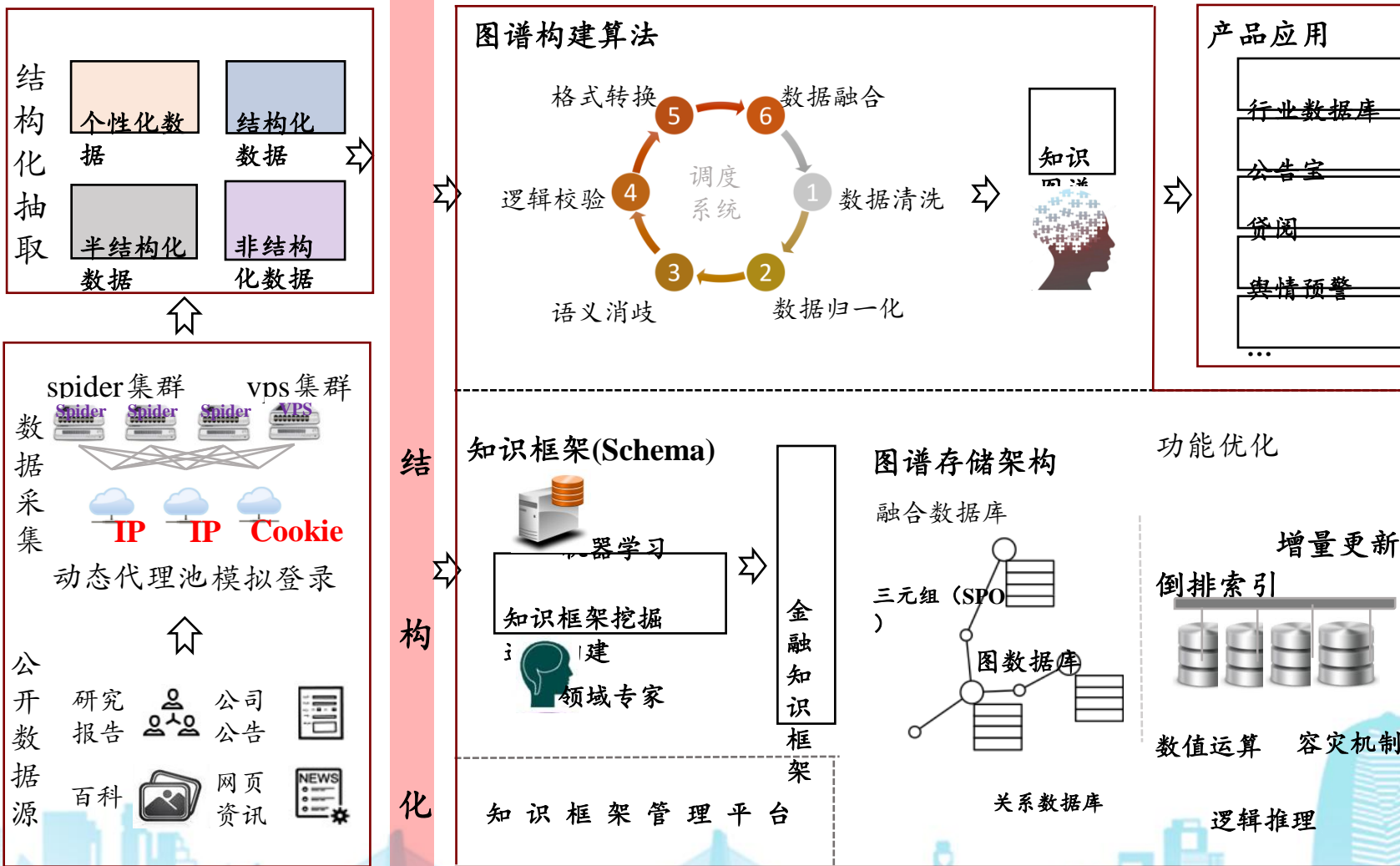
● 执行层，80s

- 来自高盛、中金的专业人士担任金融联席总监
- 来自百度自然语言理解人工智能专家担任技术总监

● 核心团队，80+90s

- 核心研发人员来自BAT、搜狗、小米、微博等主流互联网技术公司
- 自建百人金融标注团队，注重金融专业与标注经验

技术架构：已拥有全面的技术实力与积累



如上模式+团队+构架，已得到广泛认可

资本市场认可

- 2015年8月，
公司创立，红杉
领投2,700万
- 2017年3月，
A轮融资，领沅
领投8,000万

客户认可

- 主动管理类规模第
一梯队的公募基金公
司
- 全国性大型综合券
商
- 百亿以上知名私募

市场认可

- 入选
2017年
毕马威
中国领
先金融
科技50

在数据采集方面

- 目标
 - 获取原始数据并进行初步处理
- 关键技术
 - 爬虫
 - 数据解析
 - 私有数据接入

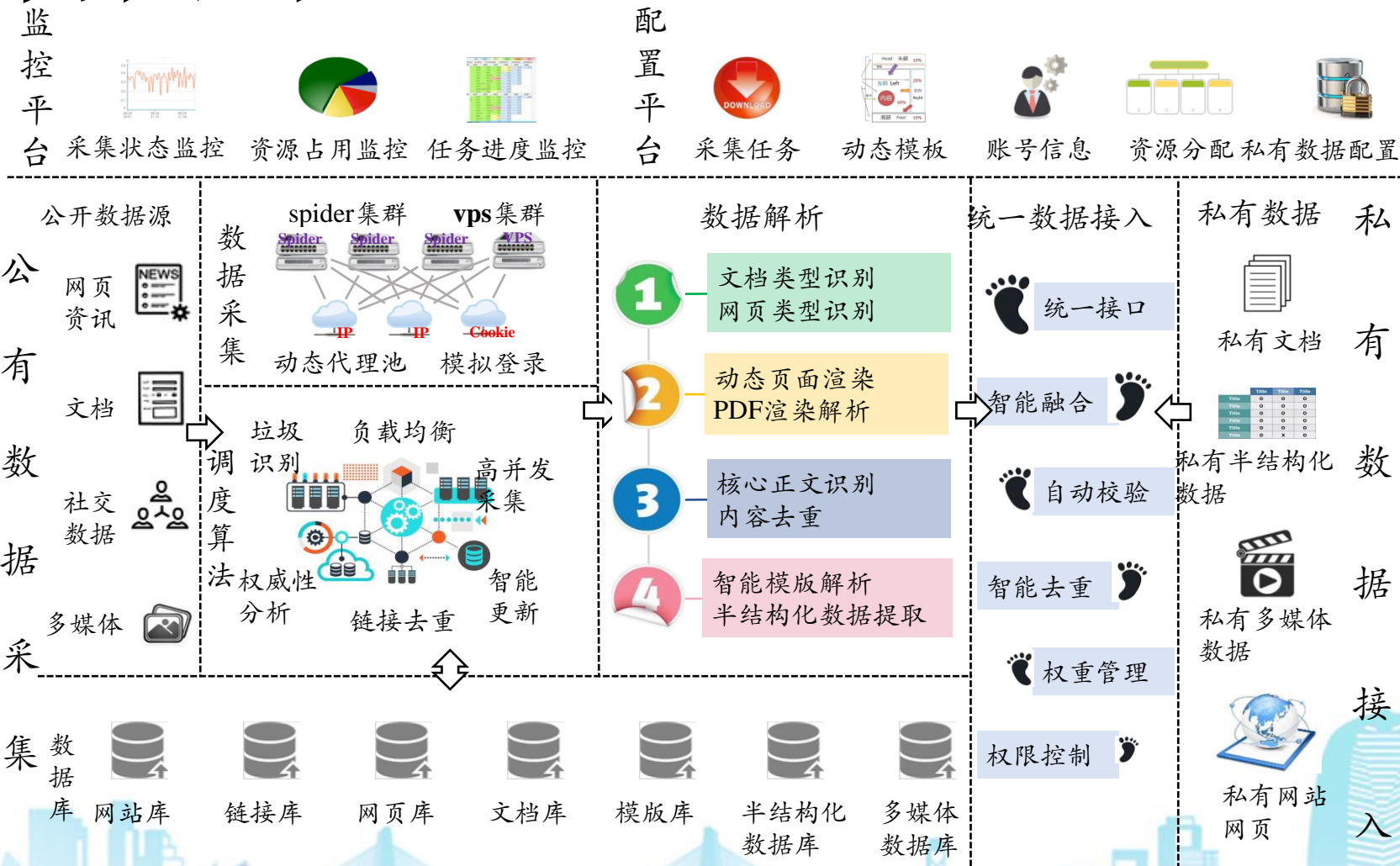
● 关注指标

- 准确率、召回率
- 实时性
- 稳定性
- 可维护、可扩展

● 网页库

- 规模：十亿+
- 每日更新1,500万
- 重要数据秒级更新

数据采集流程

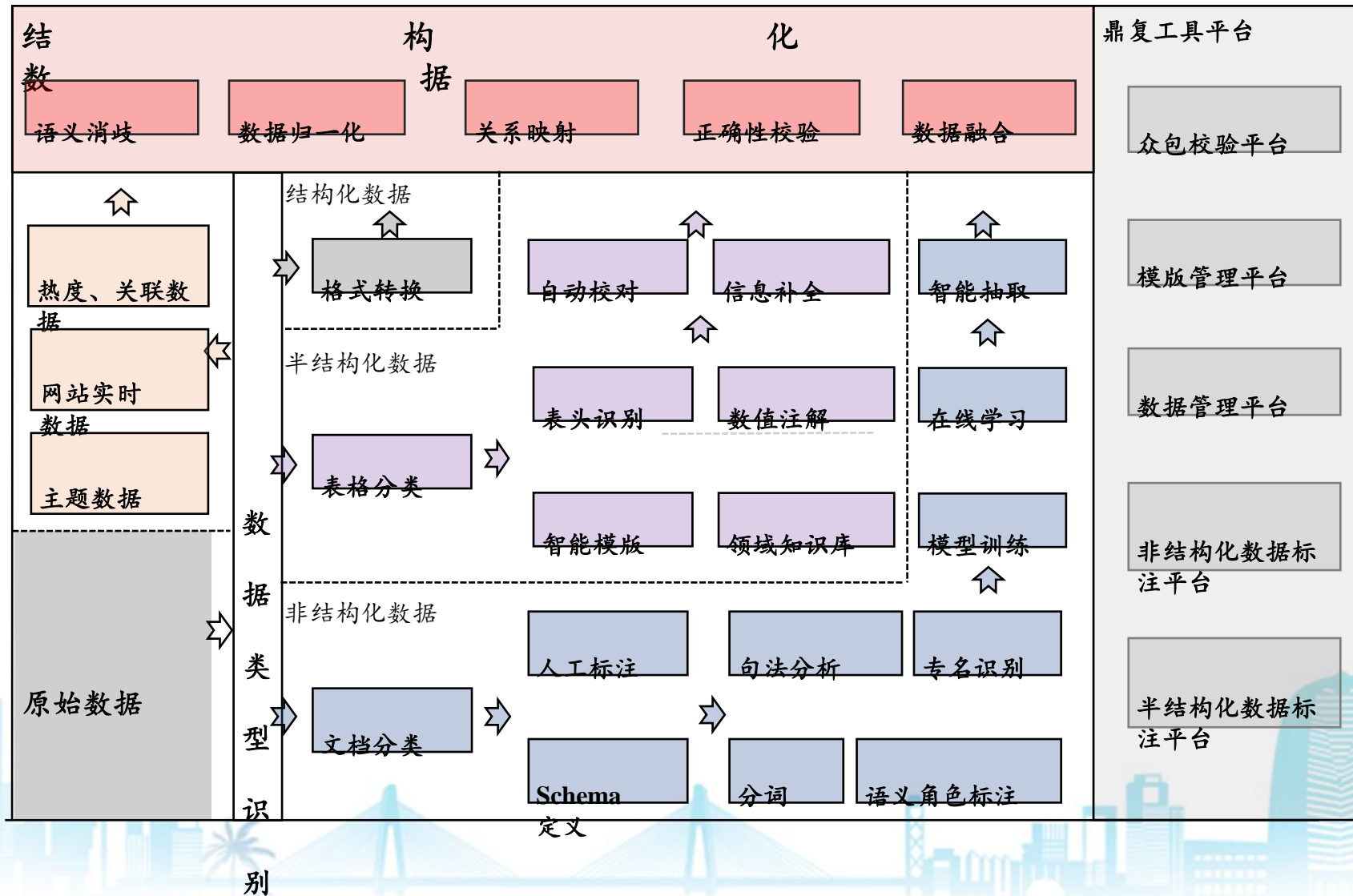


在数据抽取方面

- 目标
 - 从原始数据抽取结构化数据
- 关键技术
 - 结构化抽取
 - 基于模板
 - 基于机器学习模型
 - 数据融合
- 重要指标
 - 准确率
 - 召回率

算法分类	介绍	优点	缺点
基于模板	当数据匹配模板时进行知识抽取	准率高； 效果相对可控；	扩展性不好； 召回低；
基于机器学习模型	通过机器学习算法学习知识各因素关系	扩展性好； 召回较高；	准确率略低； 可控性略低；

数据抽取流程



数据抽取产品举例-公告宝a半结构化数据

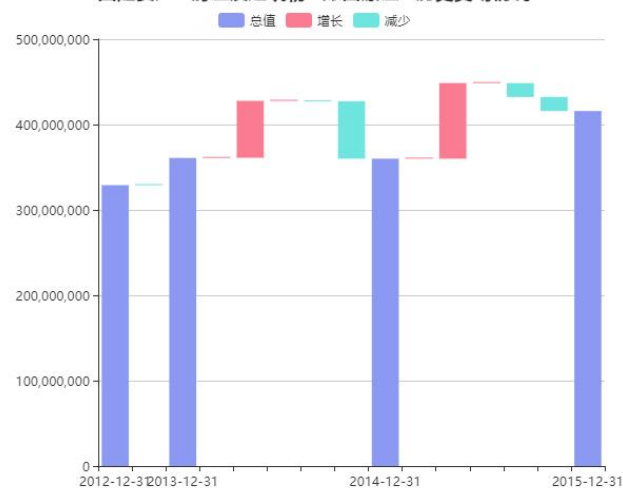
WALVAX 沃森生物 三药沃森生物技术股份有限公司2015 年年度报告全文

划分为持有待售的资产		
一年内到期的非流动资产		
其他流动资产	29,657,963.61	97,326.00
流动资产合计	2,366,545,005.07	2,220,714,287.27
非流动资产:		
发放贷款及垫款		
可供出售金融资产	118,900,000.00	118,500,000.00
持有至到期投资		
长期应收款		
长期股权投资	602,956,568.10	476,323,314.47
投资性房地产	6,068,091.05	
固定资产	769,812,416.17	518,791,188.45
在建工程	625,933,847.02	767,900,319.32
工程物资		
固定资产清理	159,153.61	
生产性生物资产		
油气资产		
无形资产	248,885,791.16	165,569,985.91
开发支出	923,309,337.82	888,158,295.95
商誉	469,575,364.46	689,493,481.27
长期待摊费用	42,742,953.75	51,580,455.72
递延所得税资产	55,477,671.15	70,684,832.79
其他非流动资产	27,474,425.07	
非流动资产合计	3,891,295,619.36	3,747,001,873.88
资产总计	6,257,840,624.43	5,967,716,161.15
流动负债:		
短期借款	865,000,000.00	1,036,000,000.00
向中央银行借款		
吸收存款及同业存放		

300142 固定资产

	房屋及建筑物	机器设备	电子设备	运输设备	其他	合计
账面原值	415823795.5	597515067.1	6503457.53	16651798.67	18804267.38	1055298386.18
累计折旧	90665704.54	168492006.99	4782570.39	9512121.5	6919359.45	280371762.87
减值准备	-	5114207.14	-	-	-	5114207.14
账面价值	325158090.96	423908852.97	1720887.14	7139677.17	11884907.93	769812416.17

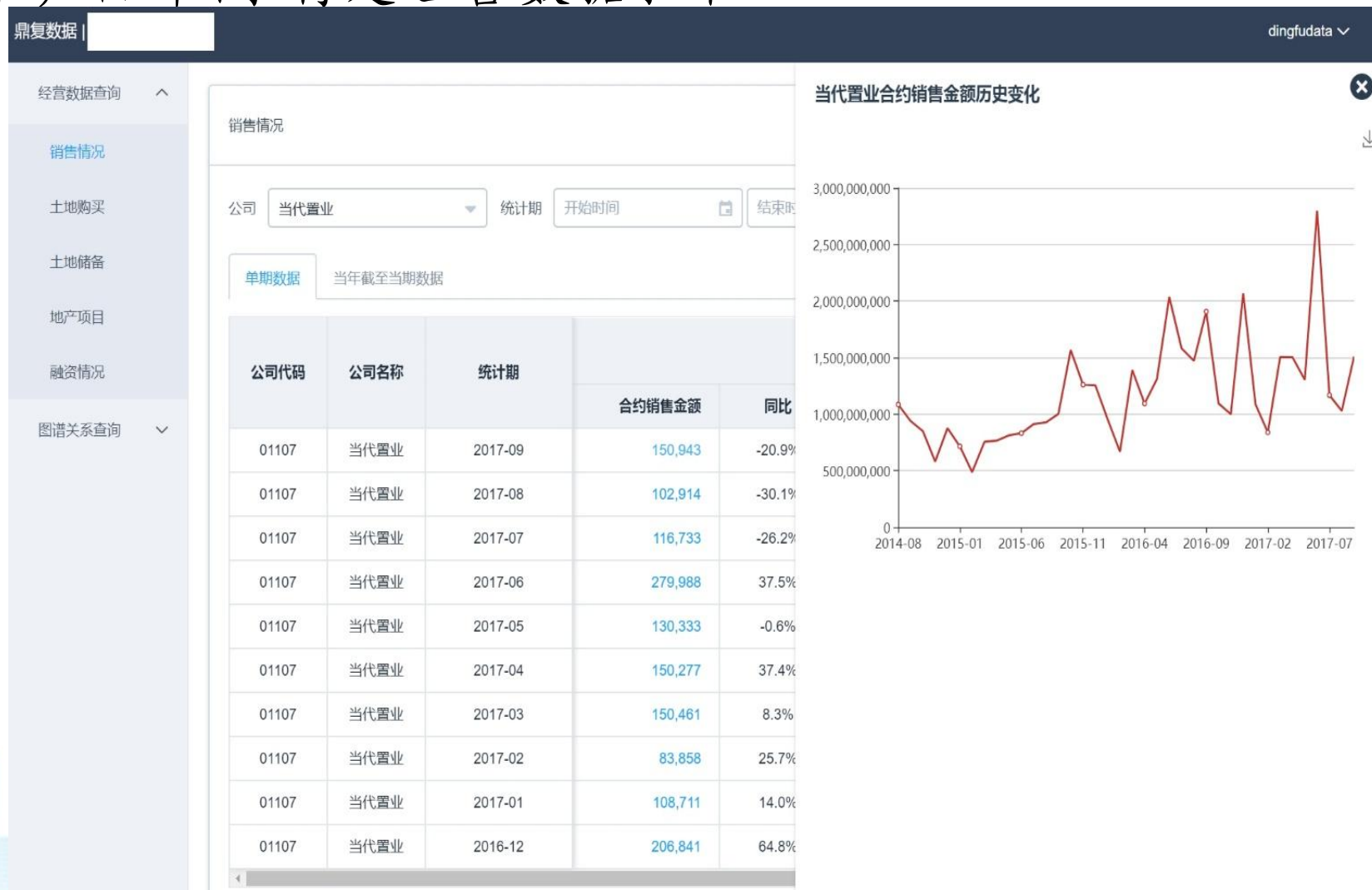
固定资产--房屋及建筑物--账面原值--历史变动情况



数据抽取产品举例-公告宝b非结构化数据

公告时间	公司简称	高管姓名	性别	离职岗位	离职时间	离职原因	继任岗位	继任者	继任者简历	代理人	来源
2017年07月06日	振华科技	筛选		副总经理	近日	工作调动	-	-	-	-	来源
2017年07月06日	岷江水电	搜索		监事会监事 监事会主席	近日	因工作原因	-	-	-	-	来源
2017年07月06日	武汉中商	清除筛选		董事	2017年7月4日	工作变动	-	-	-	-	来源
2017年07月06日	绿盟科技	黎宏	女	董事会秘书	-	-	董事 首席...	赵晓凡	-	-	来源
2017年07月06日	深赛格	曹翔	男	董事 董事会发展战略委...	2017年7月4日	因工作变动原因	-	-	-	-	来源
2017年07月06日	欣天科技	薛枫	男	总经理	近日	工作原因	董事长	林艳金	-	-	来源
2017年07月06日	*ST 中绒	杨晓军	男	副总经理	2017年7月5日	个人原因	-	-	-	-	来源
2017年07月06日	武汉中商	聂冰	女	监事	2017年7月4日	工作变动	-	-	-	-	来源
2017年07月06日	北方股份	杨芙蓉	女	副总经理	近日	因工作变动原因	-	-	-	-	来源
2017年07月06日	汇嘉时代	罗仁全	男	监事	近日	个人原因	-	-	-	-	来源
2017年07月06日	振华科技	李建辉	-	副总经理	近日	工作调动	-	-	-	-	来源
2017年07月06日	跃岭股份	陈圳均	男	监事会监事	近日	因工作调整原因	其他职务	汤重庆 (...)	-	-	来源
2017年07月06日	天奇股份	吴晓锋	女	独立董事 薪酬与考核委...	近日	连续担任公司独立董事...	-	-	-	-	来源
2017年07月06日	欣天科技	林艳金	男	执行副总经理 市场总监	2017年7月5日	-	-	-	-	-	来源
2017年07月06日	深赛格	徐宁	男	监事会监事 监事会主席	2017年7月4日	因工作变动原因	-	-	-	-	来源
2017年07月06日	纽威股份	陆献忠	男	职工代表监事	2017年07月05日	身体原因	-	毛广武	-	-	来源
2017年07月05日	中毅达	徐清波	-	上海中毅达股份有限公...	-	-	-	-	-	-	来源
2017年07月05日	中毅达	杨士军	-	上海中毅达股份有限公...	-	-	-	-	-	-	来源
2017年07月05日	龙建股份	付百彦	男	董事	2017年7月4日	工作变动	-	-	-	-	来源
2017年07月05日	海普瑞	哈继铭	男	独立董事 董事会各委员...	近日	个人原因	-	-	-	-	来源

数据抽取产品举例-特定经营数据抽取



数据抽取产品举例-行业新闻非结构化数据抽取

钛白粉_产量_总产量_年度

数据来源: 所有来源

指标

值(万吨)

2014

同比(%)

14.99

10

5

0

-5

-10

筛选数据源

时间	值	来源	总结	选择
2015	280	www.chinairm.com	通过市场调研,我们认为我国钛白粉行业产量未来仍将保持一定的增速,2013年产量必将突破200万吨,到“十二五”末产量达到250万吨将是大概率事件,如果全球经济在2014年起步入景气周期,到2015年我国产量有望达到280万吨左右	选择
2015	350	stock1.cf8.com.cn	若到2015年我国能达到人均2千克的钛白粉消费量,则我国市场需求约为290万吨,考虑发达国家产业转移等因素,我国市场生产量有可能达到350万吨左右	选择
2015	232	研报	业)的数据统计,2015年,我国钛白粉实际产量达到232万吨,约占全球总量的三分之一	选择
2015	27	生意社	而如今在“去产能”的大背景下,全国的钛白粉产量反而出现了增加的现象,2016年的259.7万吨相比2015年的产量增加了27万吨,增幅11.8%创下历史新高	选择
2015	350	southmoney.com	若到2015年我国能达到人均2千克的钛白粉消费量,则我国市场需求约为290万吨,考虑发达国家产业转移等因素,我国市场生产量有可能达到350万吨左右	选择
2015	320.1	www.wjh.org.cn	加速提升产业集中度中国钛白粉行业规模化发展起始于1998年,17年来无论是产能或产量,一路高速上扬,到2015年,行业产能达到320.1万吨/年	选择
2015	350	www.ctia.com.cn	若到2015年我国能达到人均2千克的钛白粉消费量,则我国市场需求约为290万吨,考虑发达国家产业转移等因素,我国市场生产量有可能达到350万吨左右	选择
2015	232	news.cnstock.com	2015年,国内钛白粉产量仅232万吨,产能利用率为63.91%,创2009年以来新低	选择
2015	232	www.zhicheng.com	2015年,国内钛白粉产量仅232万吨,产能利用率为63.91%,创2009年以来新低	选择
2015	232	guba.eastmoney.com	2015年,国内钛白粉产量仅232万吨,产能利用率为63.91%,创2009年以来新低	选择

数据库

请输入要查找的内容

- 硝酸钾
- 硫酸钾
- 金红石
- 钛白粉
 - 进口
 - 出口
 - 产能
 - 总产量
 - 年度
 - 分企业
- 市场价
- 出厂价
- 经销价
- 成交价

数据抽取产品举例-PDFfree

The screenshot displays the PDFfree web application interface. The main content area is divided into several sections:

- Table 1 (Top Left):** A summary table with 4 columns and 6 rows of financial data.
- Table 2 (Middle Left):** A table titled "2. 合并利润表" (Consolidated Profit Statement) with columns for "项目" (Item), "2017年1-6月" (2017 Jan-Jun), "2016年度" (2016 Annual), "2015年度" (2015 Annual), and "2014年度" (2014 Annual). It lists items like "营业收入" (Operating Income) and "营业总成本" (Total Operating Costs).
- Table 3 (Bottom Left):** A table titled "3-1-2017" with columns for "项目" (Item), "2017年1-6月" (2017 Jan-Jun), "2016年度" (2016 Annual), "2015年度" (2015 Annual), and "2014年度" (2014 Annual). It lists items like "营业成本" (Operating Costs) and "营业税金及附加" (Operating Taxes and Surcharges).
- Table 4 (Top Right):** A table titled "第200页有2个表格" (Page 200 has 2 tables) with 4 columns and 2 rows of data.
- Table 5 (Middle Right):** A table with 4 columns and 15 rows of data, likely a detailed financial statement.
- Table 6 (Bottom Right):** A table with 4 columns and 15 rows of data, likely a detailed financial statement.

数据抽取产品举例-PDFfree



	2017年3月末		2016年末		2015年末		2014年末	
项目	金额	增长率	金额	增长率	金额	增长率	金额	
资产合计	36,224.21	12.06%	32,326.66	25.83%	25,690.25	-0.07%	25,708.01	
负债合计	12,970.44	20.13%	10,797.19	7.24%	10,068.14	5.87%	9,510.04	
股东权益合计	23,253.77	8.01%	21,529.47	37.81%	15,622.10	-3.56%	16,197.97	
归属于母公司股东权益合计	23,222.31	8.02%	21,497.52	38.02%	15,575.34	-3.84%	16,198.00	
资产负债率	35.81%		33.40%		39.19%		36.99%	
智能合并表格								
	2017年3月末		2016年末		2015年末		2014年末	
项目	金额	占比	金额	占比	金额	占比	金额	占比
货币资金	7,350.80	20.29%	6,585.31	20.37%	2,146.78	8.36%	1,832.49	7.13%
应收票据	1,038.00	2.87%	24	0.07%	-	0.00%	-	0.00%
应收账款	7,456.96	20.59%	5,364.10	16.59%	6,824.37	26.56%	5,855.15	22.78%
预付款项	650	1.79%	1,083.09	3.35%	281.84	1.10%	482.6	1.88%
其他应收款	25.75	0.07%	69.88	0.22%	25.07	0.10%	180.01	0.70%
存货	9,174.52	25.33%	8,143.12	25.19%	5,384.06	20.96%	6,250.97	24.32%
一年内到期的非流动资产	-	0.00%	-	0.00%	-	0.00%	-	0.00%
其他流动资产	1,659.23	4.58%	2,713.72	8.39%	2,791.97	10.87%	2,404.81	9.35%
流动资产合计	27,355.26	75.52%	23,983.22	74.19%	17,454.08	67.94%	17,006.04	66.15%
固定资产	5,499.00	15.18%	4,136.23	12.80%	4,443.22	17.30%	5,173.59	20.12%
在建工程	47.81	0.13%	850.76	2.63%	612.6	2.38%	-	0.00%
无形资产	1,717.17	4.74%	1,726.43	5.34%	1,763.46	6.86%	1,791.26	6.97%
长期待摊费用	1,117.04	3.08%	1,167.05	3.61%	1,367.10	5.32%	1,567.16	6.10%
递延所得税资产	91.02	0.25%	52.4	0.16%	34.65	0.13%	36.64	0.14%
其他非流动资产	396.91	1.10%	410.58	1.27%	15.13	0.06%	133.33	0.52%
非流动资产合计	8,868.95	24.48%	8,343.44	25.81%	8,236.17	32.06%	8,701.97	33.85%
资产总计	36,224.21	100.00%	32,326.66	100.00%	25,690.25	100.00%	25,708.01	100.00%

Made by PDFFree(鼎复数据出品)

第203页2个表

第186页1个表 第187页1个表 第198页1个表 第199页3个表 第200页3个表 第201页2个表 第202页1个表 第203页2个表 第204页2个表

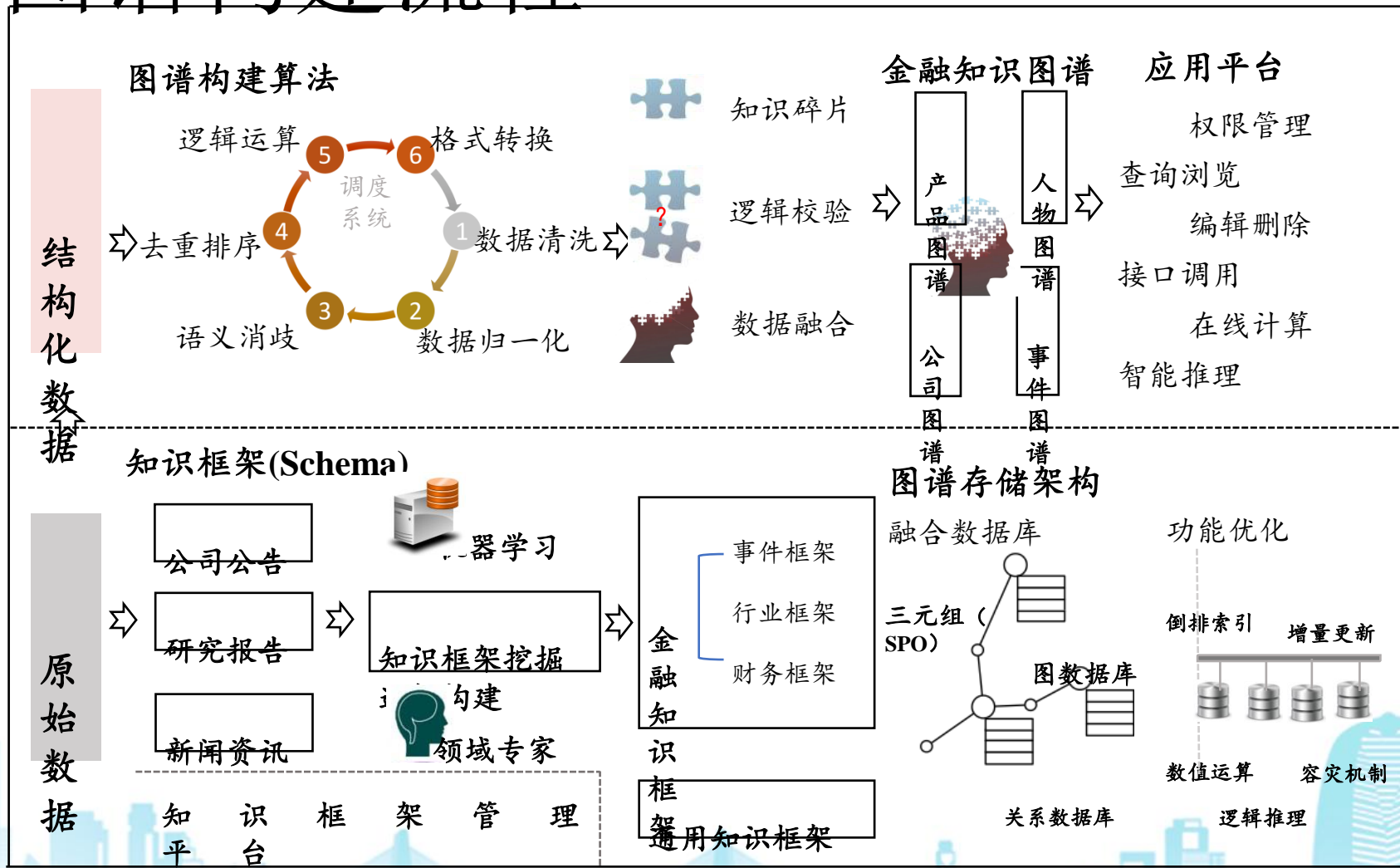
在知识图谱方面

- 目标
 - 基于结构化数据构建金融知识图谱
- 关键技术
 - 知识框架 (Schema)
 - 高性能知识图谱架构
 - 图谱构建算法

● 金融知识图谱特点

- 准确率高
- 覆盖广
- 实时性强
- 逻辑性强

知识图谱构建流程



知识图谱产品举例-贷阅智能摘要



审阅概览 审阅工作台 审计经理 1 v

信贷0208 贷前 [查看底稿](#) [原始doc](#) [关闭原文](#) [全部展开](#)

BQ_A汇总信息 [编辑摘要](#)

张十八公司人民币5亿元
土地储备项目贷款授信调查报告

BQ_B企业背景 [编辑摘要](#)

(一) 借款人概况 业务范围包括：受市政府的委托收购储备开发和经营管理土地 指导本市区、县土地收购储备土地 具有较强的政策优势和垄断地位
土地收储资金主要来源于银行借款、财政拨款和关联方资金支持
(四) 股东及关联企业概况 张十八公司与张十八四公司（简称：集团）实际为两块牌子、两个独立法人、一套班子的管理运作模式

BQ_C经营情况 [编辑摘要](#)

1.借款人经营情况 各级土地储备机构是经市、区（县）人民政府批准成立、具有独立的法人资格
(四)已办理农用地转用、土地征收批准手续的土地
2.借款人土地运作流程 A. 收购储备环节：
报土地管理领导小组审核

BQ_D项目情况 [编辑摘要](#)

(一) 项目合法合规性 2010年5月12日获批由某某市规划和土地资源管理局签发的关于同意某某区土地储备规划（2010-
(二) 项目背景 总投资估算暂定为74亿元（某某市发改委批文）
(四) 授信用途及还款来源 土地储备中心的还款来源包括：1、财政拨款（无需等到土地出让

BQ_E财务分析 [编辑摘要](#)

- > BQ_A汇总信息
- > BQ_B企业背景
- > BQ_F资信情况
- > BQ_C经营情况
- > BQ_E财务分析
- > BQ_I担保分析
- > BQ_D项目情况
- > BQ_H风险分析
- > BQ_G风险分析
- > BQ_J担保分析
- > BQ_K风险分析
- > BQ_L分行意见
- > BQ_M分行意见

知识图谱产品举例-贷阅智能摘要

审阅概览 审阅工作台

审计经理1

信贷218 贷前 [查看底稿](#) [原始doc](#) [关闭原文](#) [全部折叠](#)

BQ_A汇总信息 [编辑摘要](#)

张三建设集团有限公司
15000万元综合授信调查报告

BQ_B企业背景 [编辑摘要](#)

(一)借款人概况 张三建设集团有限公司地处张三一某某301号 注册资本5818万元 由张三二等33位自然人股东共同出资组建 市政公用工程施工总承包二级和建筑幕墙工程专业承包二级资质 并适时进军房地产配套的设计、木门窗、组合家具和大理石制品、智能化、物业管理乃至投资等新领域 2005年增资为31378万元 并变更为集团公司

BQ_C经营情况 [编辑摘要](#)

占据较大市场份额的是具备技术、管理、装备优势和拥有特级资质的大型建筑企业
(1)房屋建筑工程业务基本情况 具有很强的竞争力
3、上下游客户情况: 对本公司的利润率造成一定的影响 期限则视本公司所承担项目的工期和性质而定 本公司就要承担原材料价格波动的风险

BQ_E财务分析 [编辑摘要](#)

(1)偿债能力分析 短期偿债指标 流动比率 1.63 1.62 1.77 1.49 速动比率 1.06 0.88 1.01 0.84

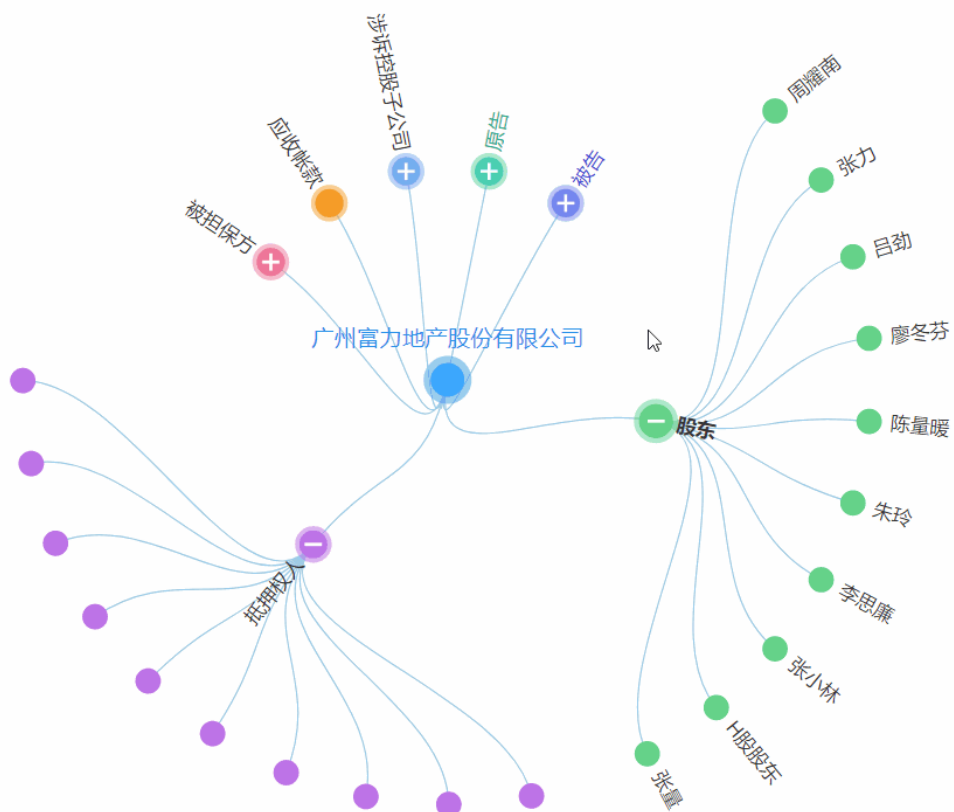
▼ BQ_A汇总信息

张三建设集团有限公司
15000万元综合授信调查报告

▼ BQ_B企业背景

一、授信客户基本概况
(一)借款人概况
张三建设集团有限公司地处张三一某某301号 前身张三一建筑安装工程公司 创建于1958年 2004企业改制变更为有限公司 注册资本5818万元2005年增资为31378万元并变更为集团公司
由张三二等33位自然人股东共同出资组建 法人代表张三二 公司具有房屋建筑工程施工总承包特级、地基与基础、消防设施、装饰装修、起重设备安装、机电设备安装五个专业承包一级资质
市政公用工程施工总承包二级和建筑幕墙工程专业承包二级资质 近年来 张三二公司按照“突出主业、多元经营”的发展思路 扩大了装饰装潢、市政基础、房地产等领域
并适时进军房地产配套的设计、木门窗、组合家具和大理石制品、智能化、物业管理乃至投资等新领域
同时将多元开拓的触角延伸到国外 完成了同单一土施工向多元化方向发展的转型 张三建设集团有限公司为房屋建筑工程施工总承包特级资质企业 拥有对外签约权 集团下设22个控股子公司 拥有一个工程类甲级设计院和一个建筑智能化系统集成专项工程甲级设计院 公司跻身张三五八省建筑业综合实力第8位 施工企业第15位 公司十分注意人才引进和培养 拥有一定数量的中高级职称人才和建造师 并多次参与国家级推荐规范的编写工作 其中公司的“中国住宅工程质量”和“全装修住宅分户验收导则”等被列为行业标准
公司组织架构:
多年来 张三二集团注重科技创新 先后成立了省级技术研发中心 国家级住宅性能认定基地 国家住宅产业化基地 公司先后主编、参编国家标准3项 地方标准8项 获国家级科技进步奖2项 省级科技进步奖3项 获国家级工法10项、省级工法45项 拥有发明专利12项 实用新型专利29项
公司获得荣誉:公司发扬“诚信守法 励精图治 以人为本 追求高效”的企业精神 艰苦创业 不断进取 强化管理 以得了优异的成绩 多年来 张三二集团始终坚持“诚信为本、质量兴业”的方针 先后创“鲁班奖”8项 国家优质工程7项 全国用户满意住宅工程6项 白玉兰杯、扬子杯、长城杯、泰山杯等省(市)级优质工程400多项 荣获中国企业500强、中国民营企业500强、中国建筑竞争力百强企业、中国建筑工程承包商60强、全国优秀

知识图谱产品举例-公司图谱违约预警系统



知识图谱产品举例-公司图谱违约预警系统

被执行人

诉讼 供应链情况 重大权益事件

被执行人

文号	立案时间	执行法院	执行标的(元)	操作
(2017)粤0106执8681号	2017年07月27日	广州市天河区人民法院	2,882,052	查看
(2017)粤0104执2705号	2017年03月20日	广州市越秀区人民法院	526,525	查看
(2017)粤0106执2870号	2017年03月15日	广州市天河区人民法院	192,805	查看
(2017)粤0106执2602号	2017年03月10日	广州市天河区人民法院	438,678	查看
(2017)粤0106执1648号	2017年02月13日	广州市天河区人民法院	82,980	查看
(2016)粤0106执10579号	2016年11月10日	广州市天河区人民法院	81,146	查看
(2016)粤0106执9208号	2016年10月10日	广州市天河区人民法院	274,042	查看
(2016)粤0106执9210号	2016年10月10日	广州市天河区人民法院	195,785	查看
(2016)粤0106执9209号	2016年10月10日	广州市天河区人民法院	179,999	查看
(2016)粤0106执7892号	2016年08月17日	广州市天河区人民法院	391,172	查看

裁判文书

开庭公告

开庭时间	案由	原告/上诉人	被告/被上诉人
12-05 2:30	肖像权纠纷	苏炳添	广州富力地产股份有限公司
17-04-25 9:00	合同纠纷	佛山国际旅行社有限公司	广州富力地产股份有限公司 海南三林旅业开发有限公司

裁判文书

发布时间	文号	文书名称	案件类型	当事人地位	涉案金额	案件相关人及身份	法院	操作
2017-09-09	(2013)万民初字第412号	王明华与敬勇、广州天力建筑工程...	民事	被告		原告王明华,男,1972年7月...	太原市万柏林区人民法院	查看
2017-09-08	(2017)渝0106民初10221号	王德梅与重庆富力房地产开发有限...	民事	被告		原告:王德梅,女,1962年5...	重庆市沙坪坝区人民法院	查看
2017-09-08	(2016)黑0102民初8685号	史济裕与北京恒富物业服务有限公...	民事	被告	1,420	原告:史济裕,男,1994年3...	哈尔滨市道里区人民法院	查看
2017-09-05	(2016)粤0104民初9929号	广州天力物业发展有限公司与张国...	民事	原告	25,378	原告:广州天力物业发展有限公司...	广州市越秀区人民法院	查看
2017-09-05	(2017)粤01民终8243号	邓永锋、广州富力恒盛置业发展有...	民事	被告	1,830	上诉人(原审被告):邓永锋,<...	广州市中级人民法院	查看
2017-09-04	(2017)川0105民初7015号	成都袁房宝互动科技有限公司与成...	民事	被告		原告成都袁房宝互动科技有限公...	成都市青羊区人民法院	查看
2017-09-01	(2017)渝0106民初10971号	重庆鑫立宸建筑装饰工程有限公司与...	民事	被告		原告重庆鑫立宸建筑装饰工程有限...	重庆市沙坪坝区人民法院	查看
2017-09-01	(2017)晋01民终342号	段德丰与太原富力城房地产开发有...	民事	被告		上诉人(原审被告):段德丰,男...	太原市中级人民法院	查看
2017-09-01	(2016)晋0107民初3369号	张志秀与广州天力物业发展有限公...	民事	被告	7,851	原告张志秀, 委托代理...	太原市杏花岭区人民法院	查看
2017-09-01	(2017)晋0107民初1410号	李政与太原富力城房地产开发有限...	民事	被告		原告:李政, 委托诉讼...	太原市杏花岭区人民法院	查看

总结，鼎复在产品端形成的BPC三条业务线



✓ 有影响力的机构认可
✓ 验证价值

相互渗透与转化

✓ 用户数
✓ 影响力

非常感谢!

欢迎提问