Lenovo Data Intelligence

Lenovo Enterprise Analytics Platform LeapAI and Case Studies

Jun Luo

Machine Intelligence Center
Lenovo Data Intelligence Business Group

Lenovo Digital Transformation

big data continues to drive Lenovo's global business

Lenovo Data Intelligence

9 Data centers

3000+ Servers

Total storage **16PB**

Data Volume **>12PB**

Data processed **4.3PB/day**

New data **>30TB/day**

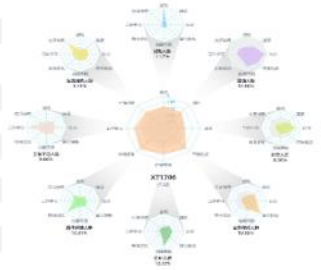
Task processed **>15K/day**

Availability **>99.99%**

Lenovo Digital Transformation

from 2011, big data continues to drive Lenovo's full business optimization

Lenovo Data Intelligence



Precise insight

Improve supply chain management

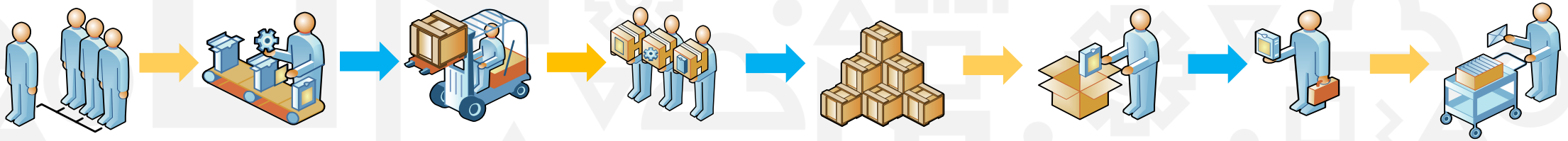
Improve product marketing ROI

Establish global insight & analysis capabilities

Increase customer satisfaction

.....

Data-driven through whole value chain



Product design | Manufacturing | Logistics | Product marketing | Inventory & storage | Sales channels | User experience | Customer service

Lenovo Way to Data Intelligence

Lenovo Data Intelligence

2011 — Lenovo Big Data team and platform founded

2012 — Data of mobile phone and PC on Bigdata platform was over Peta Byte

2013 — Lenovo digital transformation based on Data Intelligence. Support over one billion end users

2014

2015 — Lenovo global data lake support over 100 internal apps.

Machine Intelligence Center founded

2016 — Promotion Lenovo data culture internally, 95% data was integrated on LUDP

2017

2018 — Data Intelligence platform iterate and improve continuously…

2019

Lenovo BigData business team founded

Lenovo Big Data → Lenovo 联想 数据智能 Data Intelligence

2016 — **Leap** product portfolio announced and enter into market

2017 — Enable 30+ enterprises on digital transformation. Revenue Y2Y tripled

2018 — Enter into more segments. Revenue Y2Y double increase

2019 — **DIBG**

We Offer End-to-End Data Intelligence Solution

Lenovo Data Intelligence

DI Application

AI Platform

Big Data Platform

IoT & Data Collection

Industry DI App

Data Insight

R & D: 智能表检 工艺分析

Manufacture: 产线预警 能效提升

Device: 故障预测 备件优化

Supply: 路径优化 仓库选址

Marketing: 销量预测 商机洞察

Service: 智能巡检 产品舆情

3rd Party App

AI Platform

LeapAI

BI | Analytics | Data Mining | Search | API

Prediction Lib | Optimization Lib | Knowledge Lib

Deep Learning DeepNEX | Machine Learning HyperMiner | Knowledge Graph HyperGraph

Data Governance

MetaData	Data Quality
Standard	Sequrity
Architecture	MasterData

BD Platform

LeapHD

DW | MPP

Parallel Filesystem | HDFS

Parallel Computing

Resource Management

No SQL

Cluster Management

Memory DB | Graph DB | Temporal DB

Data Trusted

Authentication	Application
BIOS	Platform
Network	Server

IoT Data & Computing - LeapIOT

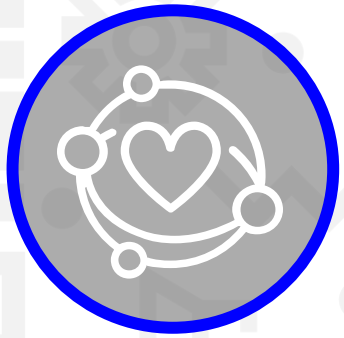
IT System Data Integration - DataHub

CRM SCM ERP MES PLM ...

DMS CAE MRP MDC OA ...

Leap Product Portfolio

Lenovo Data Intelligence



LeapIOT

IoT Platform, both centralized and edge platform to configure, connect, manage and implement devices.

LeapHD

Big Data Platform, a mature and easy to use platform for data cleaning, data integration, data governance etc.

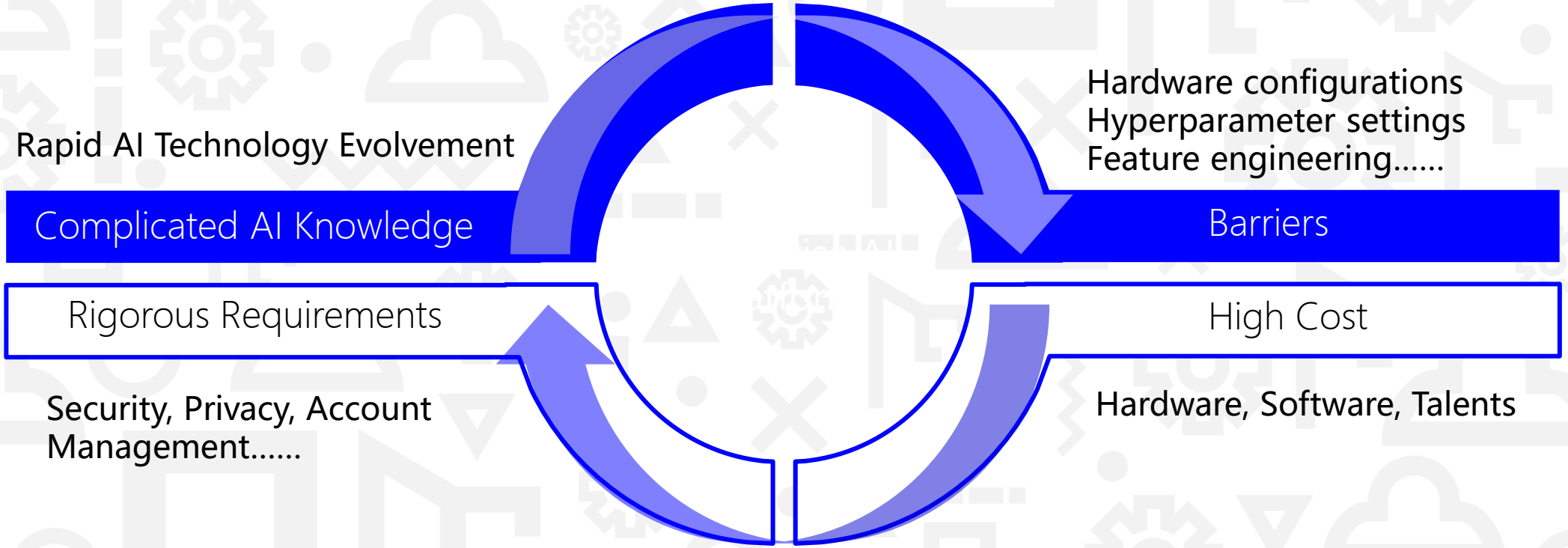
LeapAI

AI Platform, an enterprise AI platform to provide overall capability on AI training and inference.

Make A Great Leap Forward!

Four Challenges for AI Landing in Enterprise

Lenovo Data Intelligence



Goals

Lenovo Data Intelligence

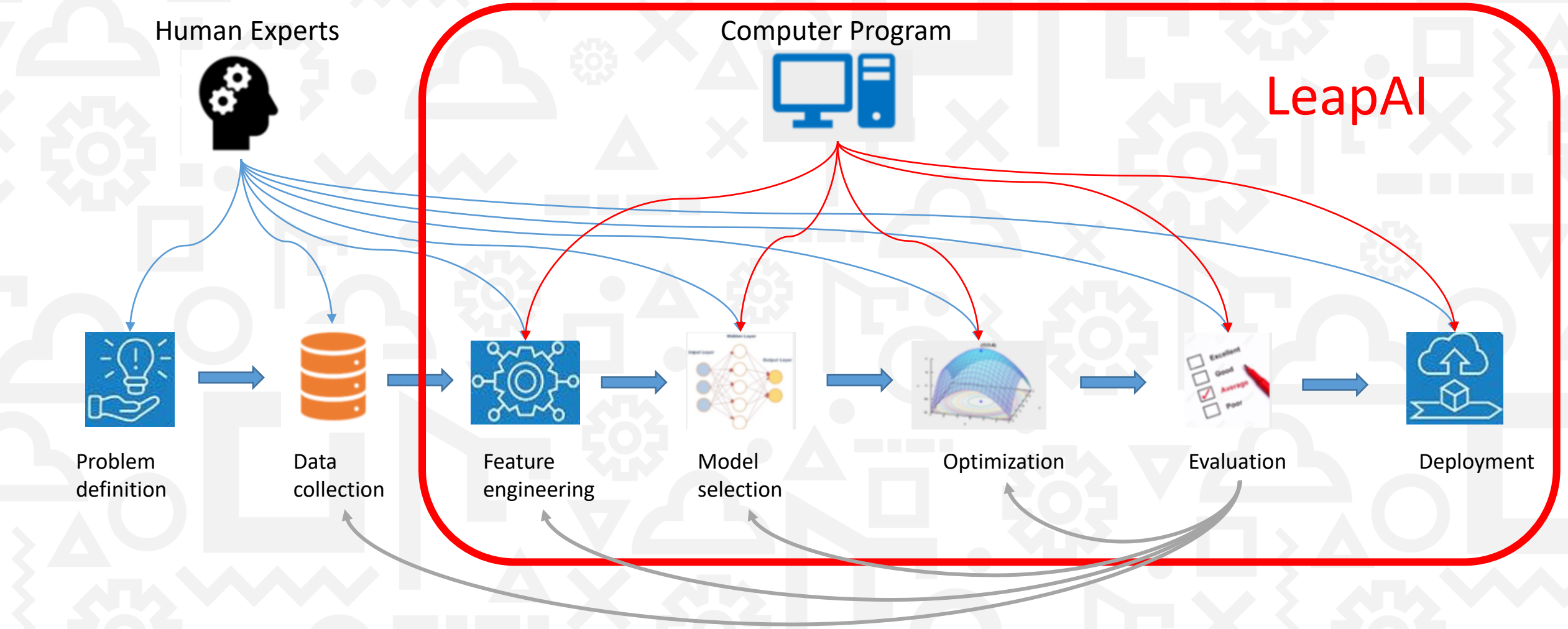
Fish ----------> **LeapAI** Net

Lower barriers for landing AI cases

Enable customers to build their own AI abilities

Machine Learning Workflow

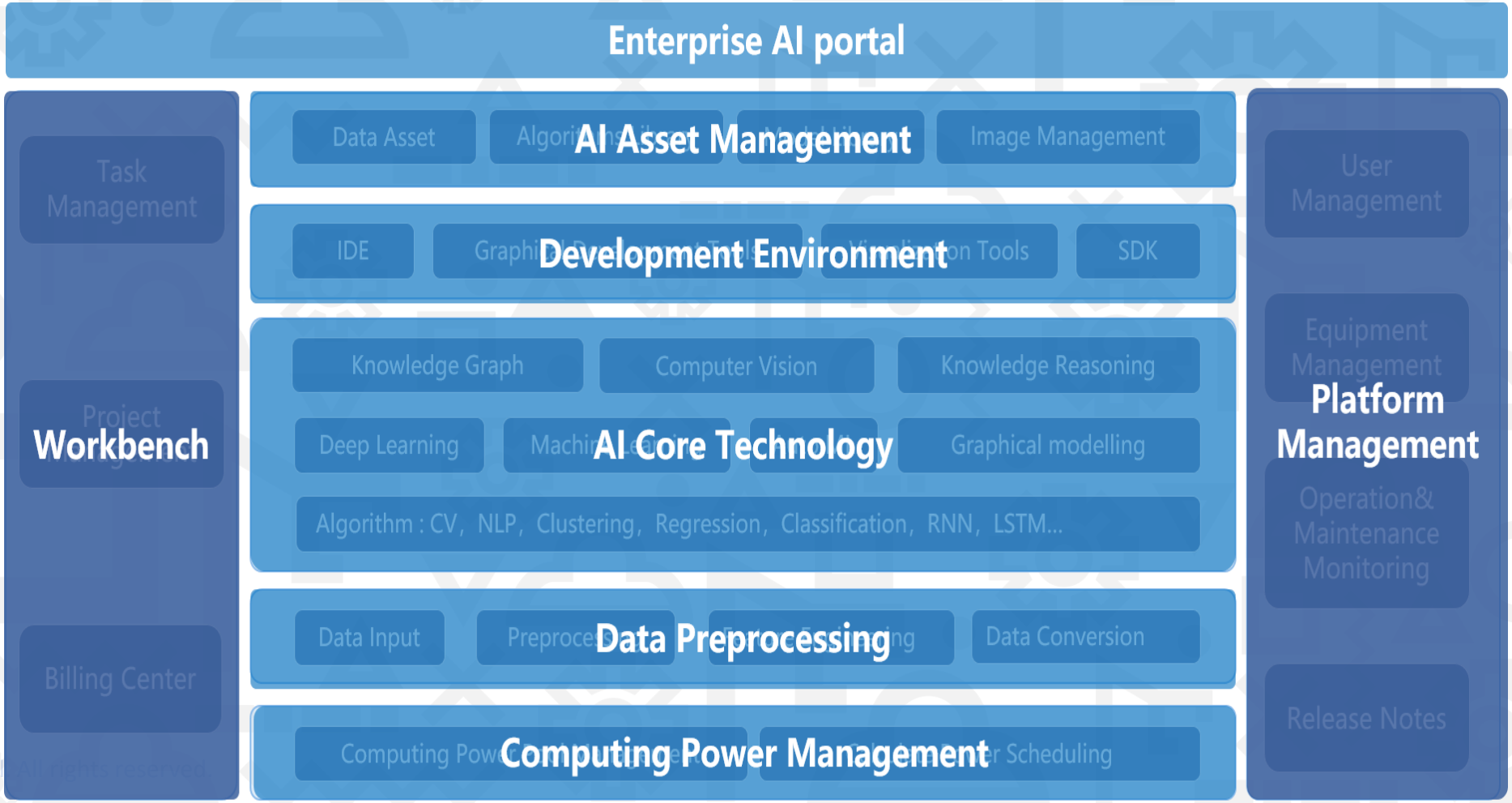
Lenovo Data Intelligence



LeapAI Platform

Lenovo Data Intelligence

LeapAI provides full-scale and end-to-end AI empowerment with data, algorithm, and computing power management; with easy-to-use AI development tools and interactive user interface, it lowers the barrier of enterprise AI application



Lenovo Internal AI Cases: LUDP AI

Lenovo Data Intelligence

Challenges

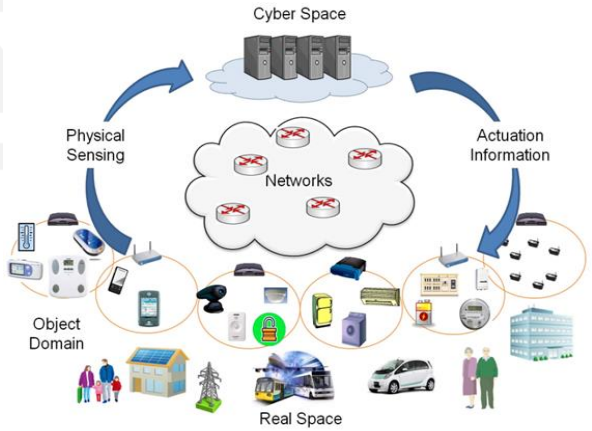
Laptop sales prediction

Quality predciton

SMT product line optimization

AI

Smart investment



Data Analysis

Data

Device Data
- Mobile
- PC
- Smart IOT

Internal Application Data
- App Store
- Preinstall App
- Internal Web

Enterprise Data
- Sales
- Customer Service
- Supply Chain

Externall Data
- Social Media
- Product BBS

Data Lake

Extract Transform Load

Device Structured Data

Apps Structured Data

Users Structured Data

Enterprise Operation Structured Data

Sentimental Structured Data

Analysis

Algorithm Library

Product Quality Control

Product Line Optimization

Data Mining and Modeling

Industry Template

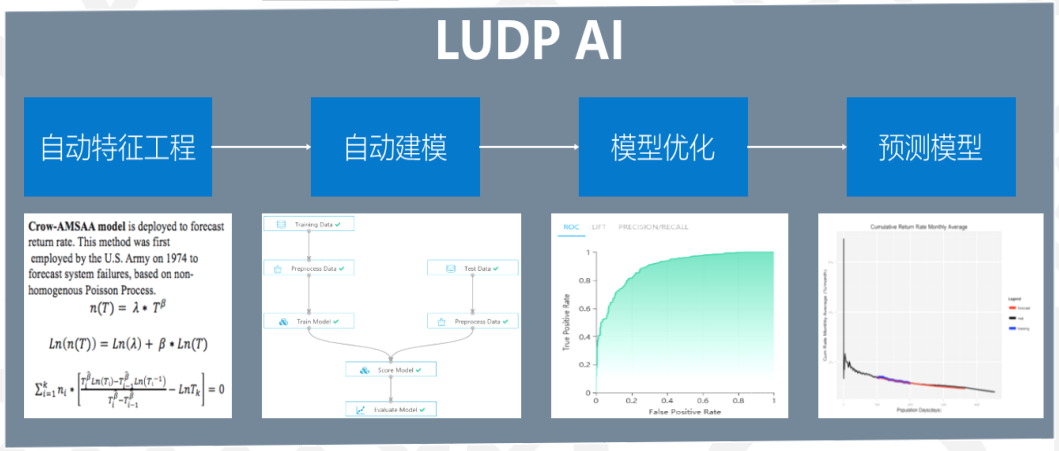
Sales Prediction

Supply Chain Optimization

Product Analysis and Optimization

Business Operation Analysis and Predction

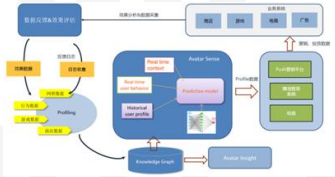
Smart Supply Chain



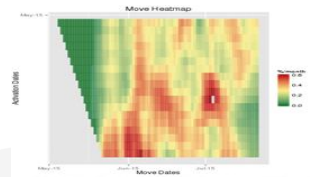
Applications

Results

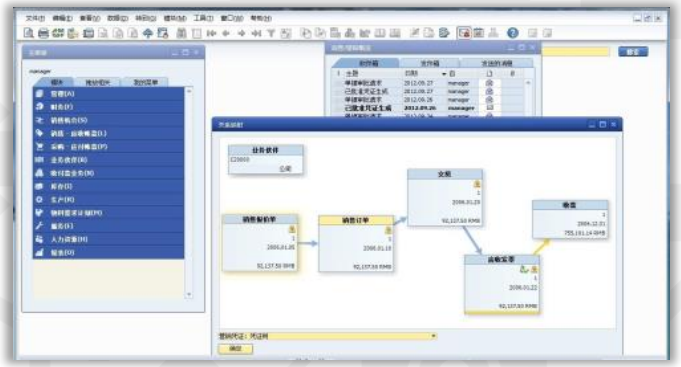
- Dozens of business domain
- Application scenarios **100+**
- Users **1000+**



Supply chain prediction



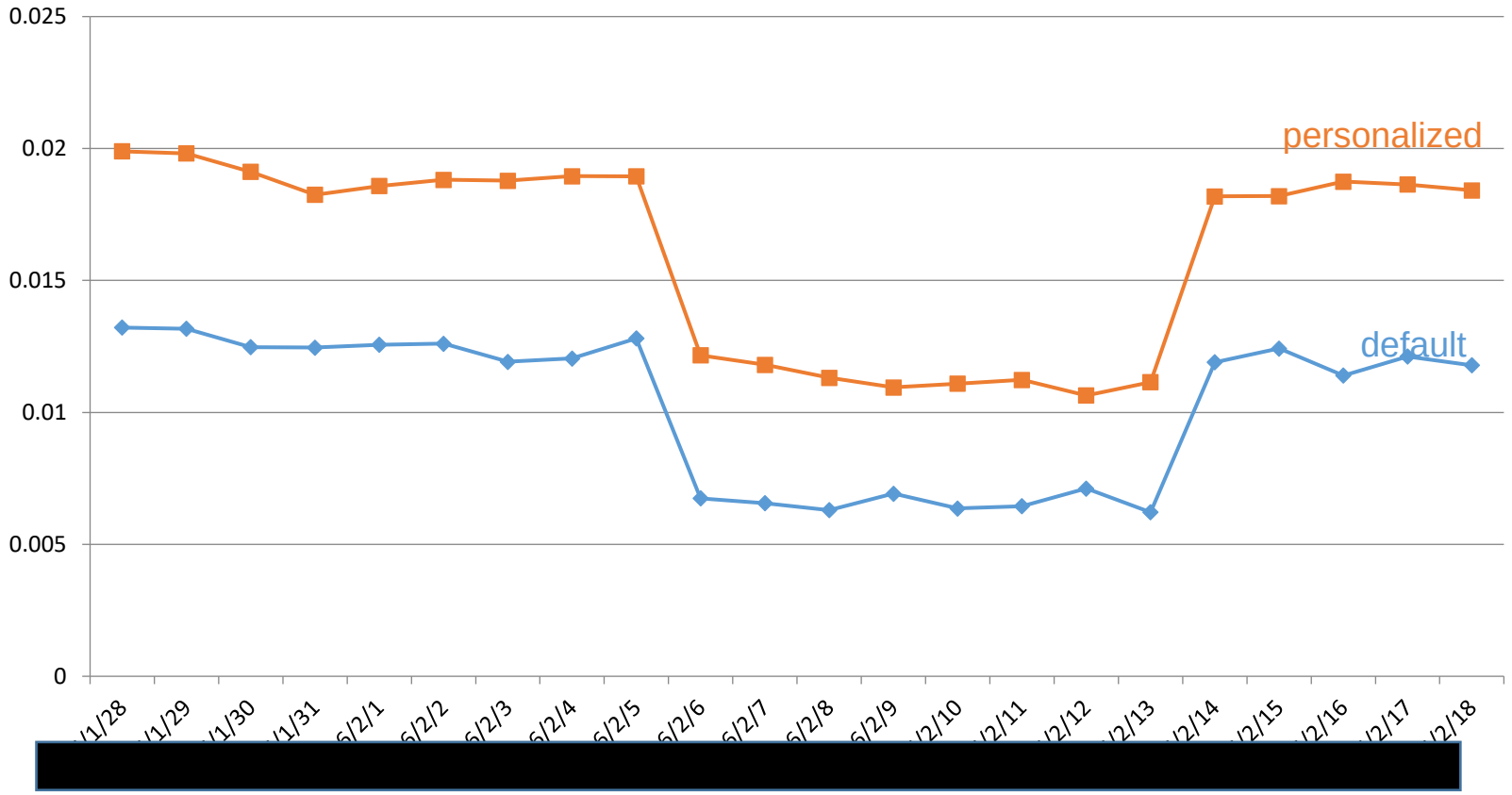
Quality prediction



Product line optimization

Lenovo Mobile Big Data Project-Personalized Game App Recommendation

- **Background：** Lenovo app store has thousands of game apps. Originally, all users have the same ranking order of game apps.
- **Goal：** Provide different game apps ranking order for different users in order to increase the click through rate.
- **Difficulties**: millions of features and users, sparse matrix.
- **Solution**: ensemble methods: FFM + LR
- **Results**: CTR increases 50% over default ranking.

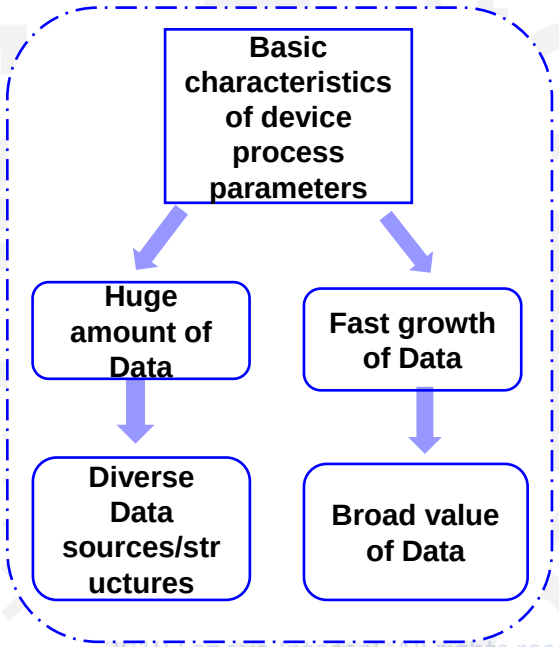


Oil-Refinery Process Optimization

Lenovo Data Intelligence

Challenges

Refining and chemical plants are large in scale, have many processes, have high concentration, and have complicated management systems.



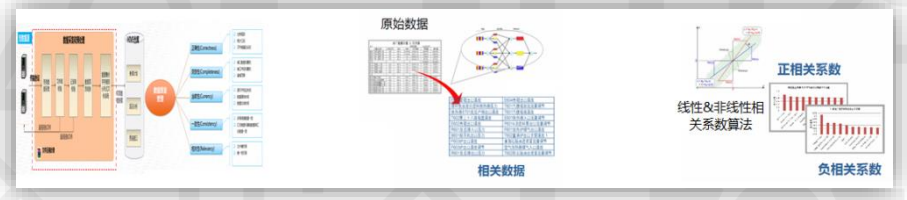
Data layer

- Operational data
- Quality data
- Corrosion data
- Cost data
- Material balance
- Energy supply data
- Energy consumption
- Catalytic device data
- Process parameters
- Process run data
- Mechanical perform

...

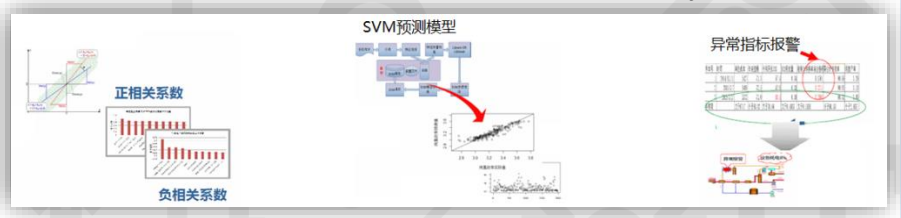
Big Data analysis

Process parameter correlation analysis



Data Acquisition | Cleaning & standardization | Correlation algorithm analysis

Indicator anomaly monitoring



Correlation analysis and feature selection | Prediction model construction | Indicator anomaly judgment

Index parameter optimization

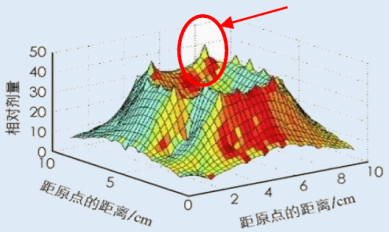
Raw material clustering analysis | Material classification model | Sample database of operations | Parameter optimization

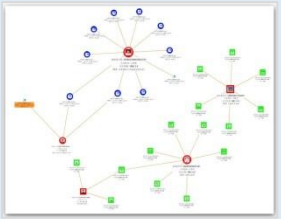
Warning of key points

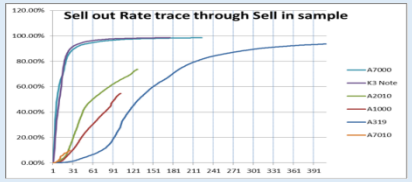
Data Acquisition | Model training | Model prediction

Analysis results

- Gasoline yield rate improves 0.5% ~ 1%
- Abnormal detection rate increases 26%







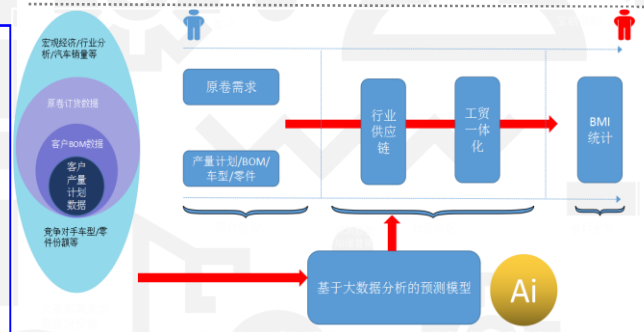
Steel Demand Prediction

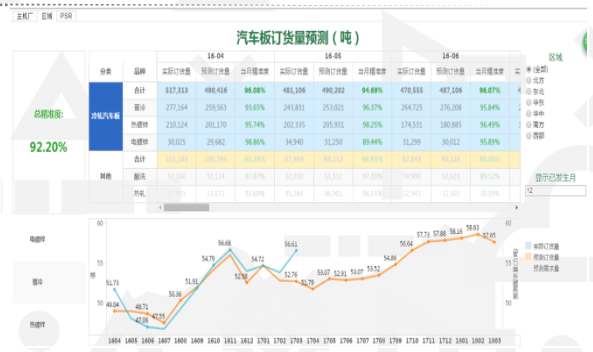
Lenovo Data Intelligence

Challenges

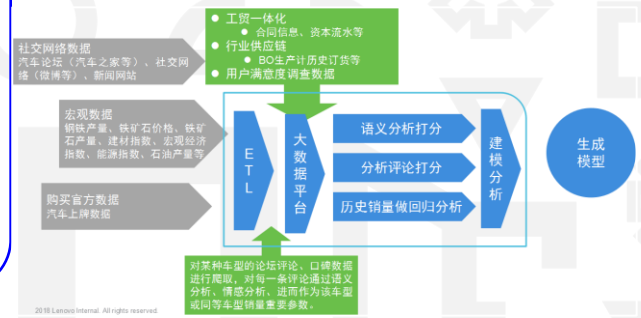
- Internal data are not enough
- Demands change abruptly

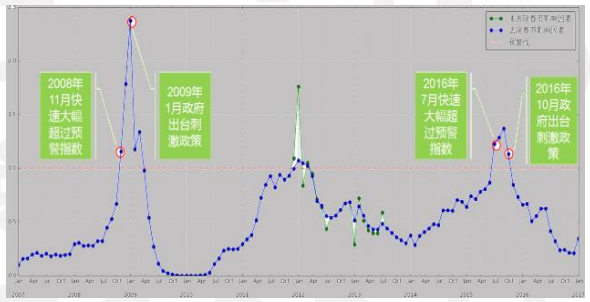
Solutions





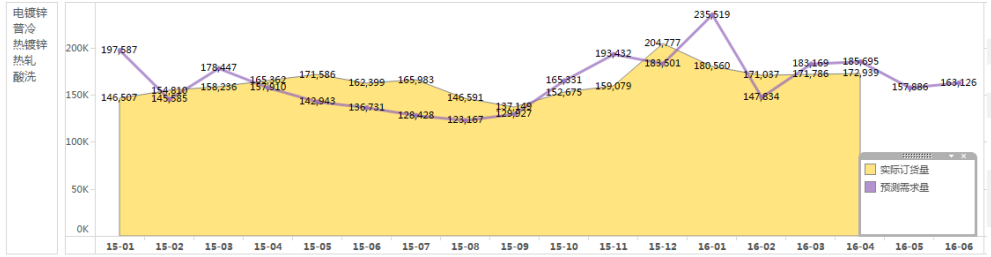
下游行业综合预警指数追踪图





Steel demand prediction of east china region (ton)

PSR	2015-01			2015-02			2015-03			2015-04			2015-05		
	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准
总计	197,587	146,507	92.91%	145,585	154,810	89.33%	178,447	158,236	92.86%	157,910	165,362	89.11%	142,943	171,586	8
普冷	84,089	62,350	95.08%	66,481	70,694	91.68%	79,154	70,189	97.80%	72,498	75,919	91.37%	60,804	72,988	9
热镀锌	17,607	13,055	97.36%	12,365	13,148	89.10%	16,405	14,547	97.62%	12,930	13,540	88.17%	13,260	15,917	8
电镀锌	66,976	49,662	95.45%	48,437	51,506	99.62%	58,903	52,231	94.94%	50,545	52,930	84.97%	45,989	55,204	9
酸洗	27,217	20,181	98.10%	16,819	17,885	73.88%	21,898	19,418	90.00%	20,301	21,259	83.65%	16,453	19,750	9
热轧	1,698	1,259	78.59%	1,483	1,577	92.36%	2,087	1,851	83.95%	1,636	1,713	97.39%	6,436	7,726	2

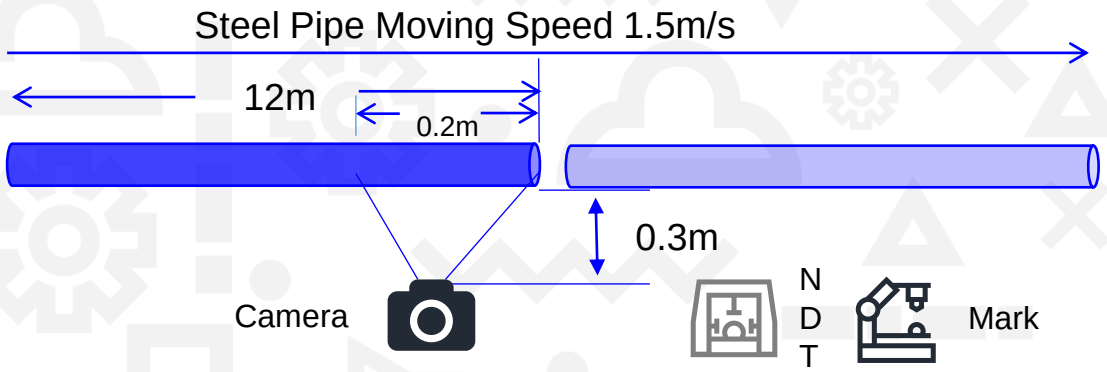


Results

- Prediction Accuracy
  - By regions and steel types **90%.**

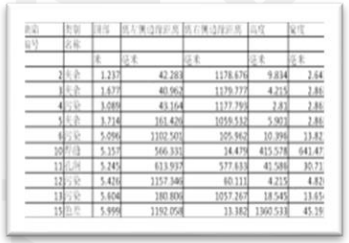
Steel Pipe Defects Detection

Lenovo Data Intelligence

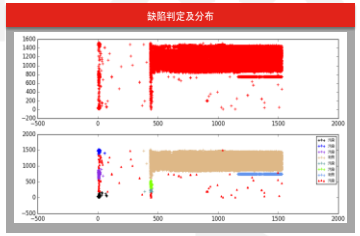


System Functions

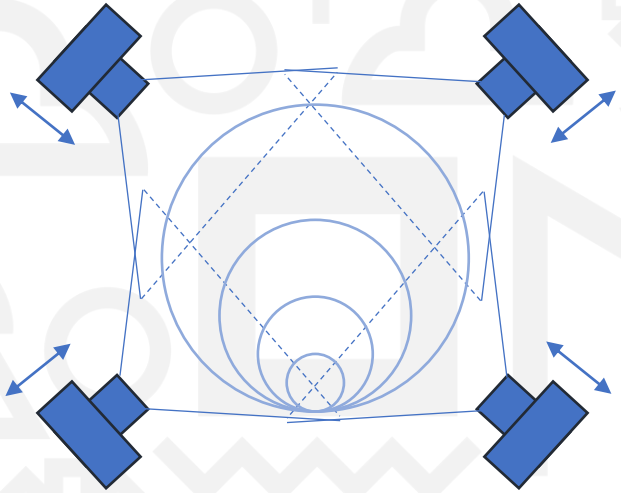
Process Data Collection



Predictive Modeling



Real-Time Monitoring & Analysis



Basler/acA2440-20gm*4
Basler/FL-CC1214A-2M*4

Business value

Automatic Quality Grading:

Type：A（Serious）、B、C （Minimal）

Automatic Defect Detection:

Holes, Rolling concave, Outer fold, Rolling mark, External scar, Mechanical injury, Pock surface

Type A Defect Detection Rate 100%

Type B & C Defect Detection Rate >95%

Lenovo Data Intelligence

Inspecting Power Lines Defects with Drones

Power line defects category

Lenovo Data Intelligence

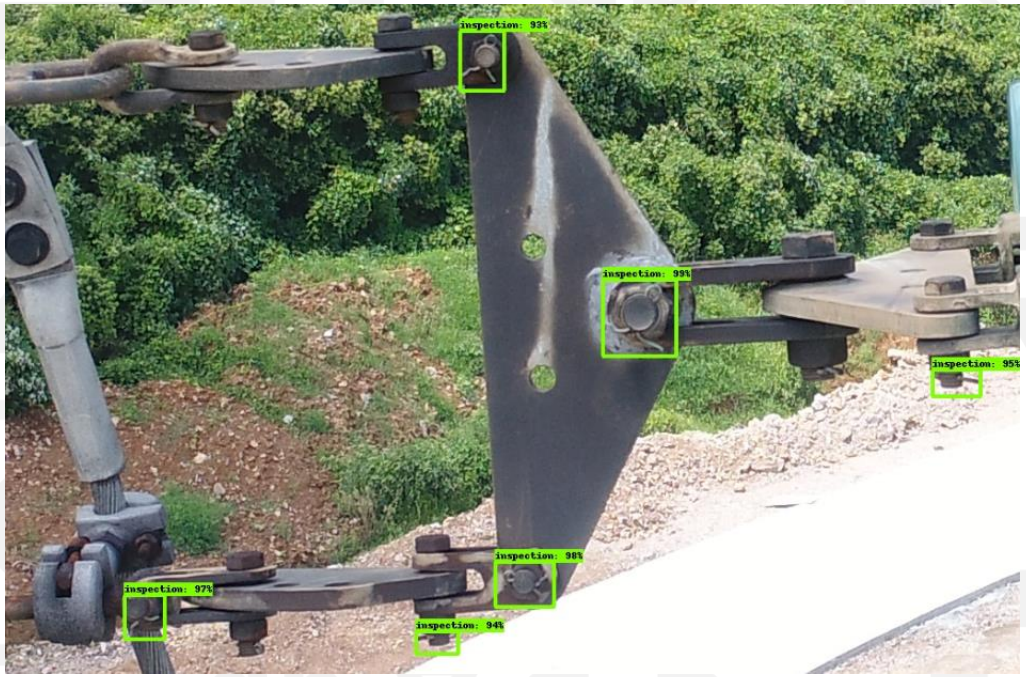
- Tower base related defects
  - **vegetation cover**, **water immersion**, debris accumulation, etc.
- Tower related defects
  - **bird nest**
- Insulator related defects
  - **insulator blow,** insulator damage, insulator aging
- Fitting related defects
  - **defect detection**: corona ring skew**,** connector corrosion, bolt missing
  - **fitting detection**: damper, connector, corona ring, spacer, bolt
- Subsidiary facilities related defects
  - **tower plate fading**, tower plate corrosion, tower plate missing, etc.

Challenges：complex shooting environment

Lenovo Data Intelligence

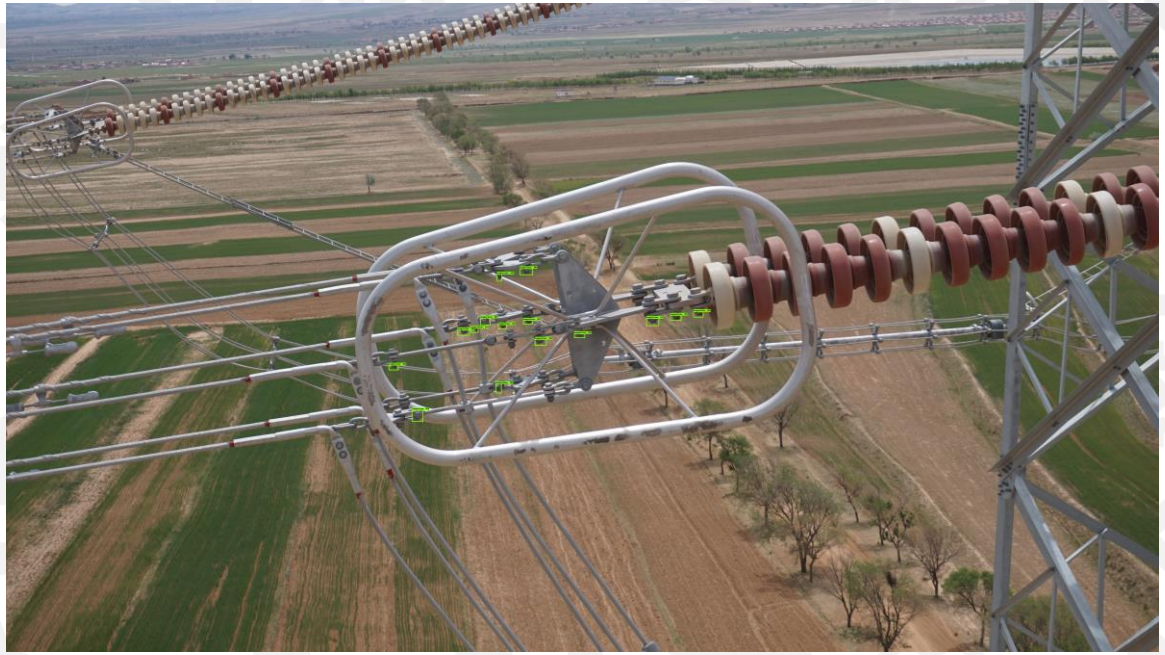
Challenges：small object

zoom in



Challenges： multiple objects

Lenovo Data Intelligence

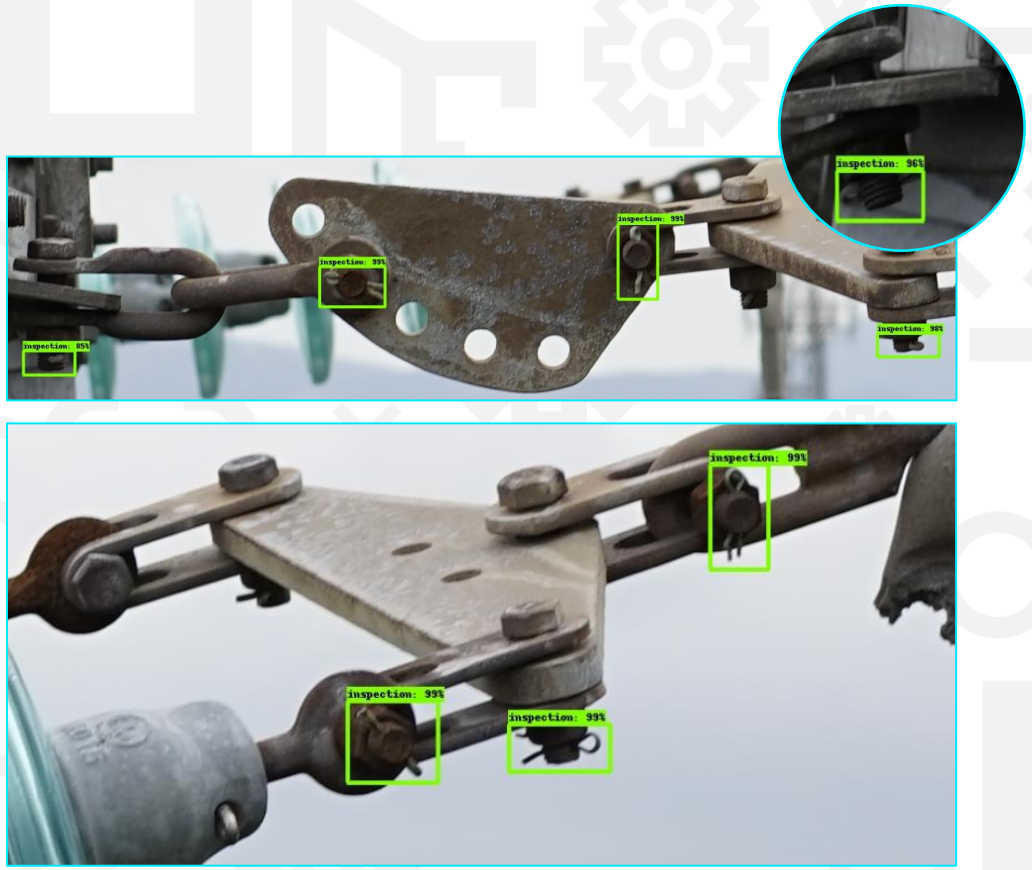


zoom in

Challenges：limited labelled data

Lenovo Data Intelligence

zoom in

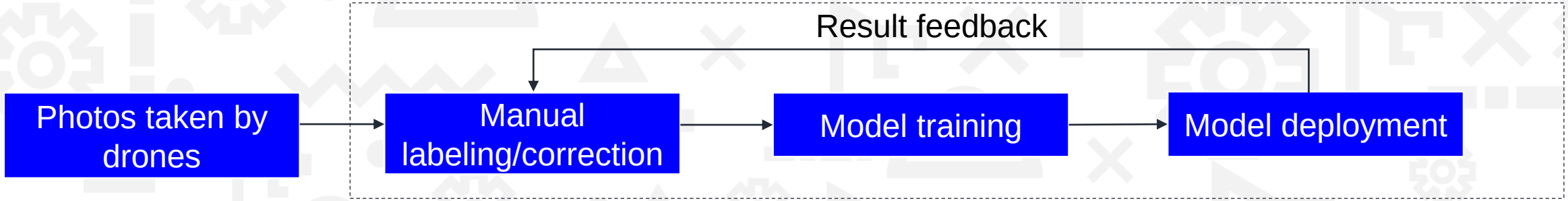


Project highlights

Lenovo Data Intelligence

Automated labeling platform

– 200 labeled images of bird nest provided by customers were expanded to over 7,000



Power line inspection overall performance (selected tasks)

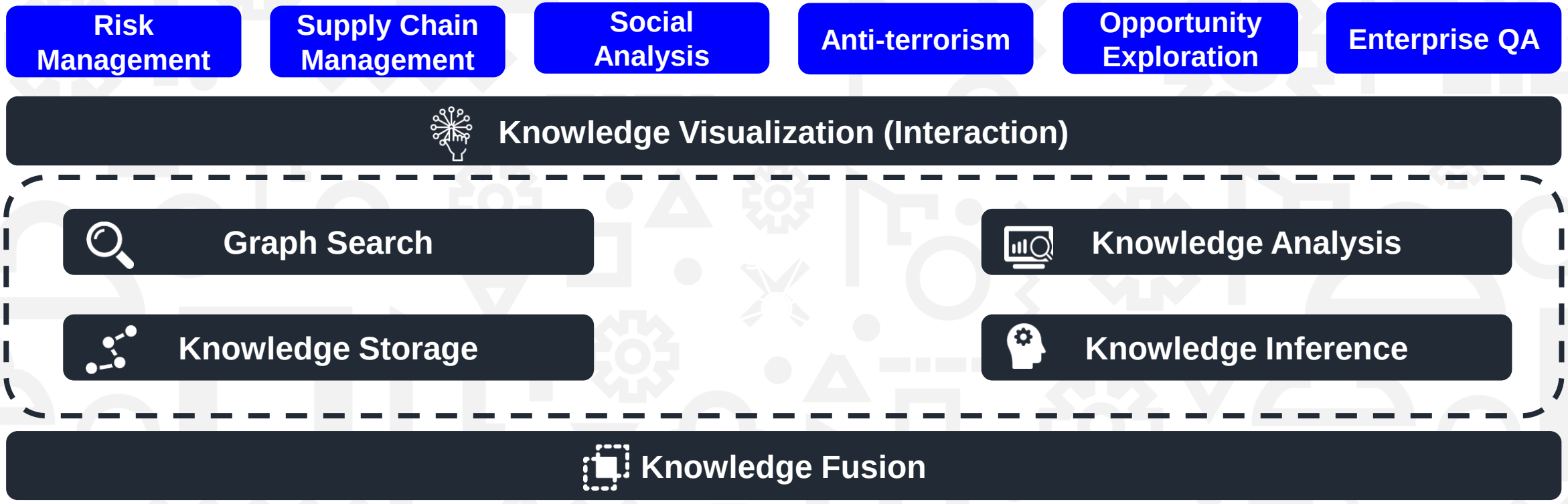
Task	mAP	Precision	Recall	Industry Level
Bird Nest Detection	86.8%	90%	80%	about 60%
Split Pin Detection	76.13%	80%	89%	N/A

LeapAI Knowledge Graph

Lenovo Data Intelligence

Intelligent & Interactive Enterprise Level Knowledge Base

AI-enabled Transformation, Risk Management, Decision Making, Enterprise QA



Petroleum Knowledge Graph

Lenovo Data Intelligence

Helping customer building a knowledge hub, digitalizing **millions** of engineering documents for the past 20 years.

DocReader Assistant	Smart Search	Smart Recommendation	Knowledge Q&A
Risk Alert	Full text search	User Profile	Knowledge graph
Key indicators	In-doc Search	Similarity search	Specialized Q&A
Nested clauses	Advance Search	customization	Process Q&A
.....	.....	.....	.....

Knowledge Intelligence

RDBMS **Enterprise Knowledge Base**

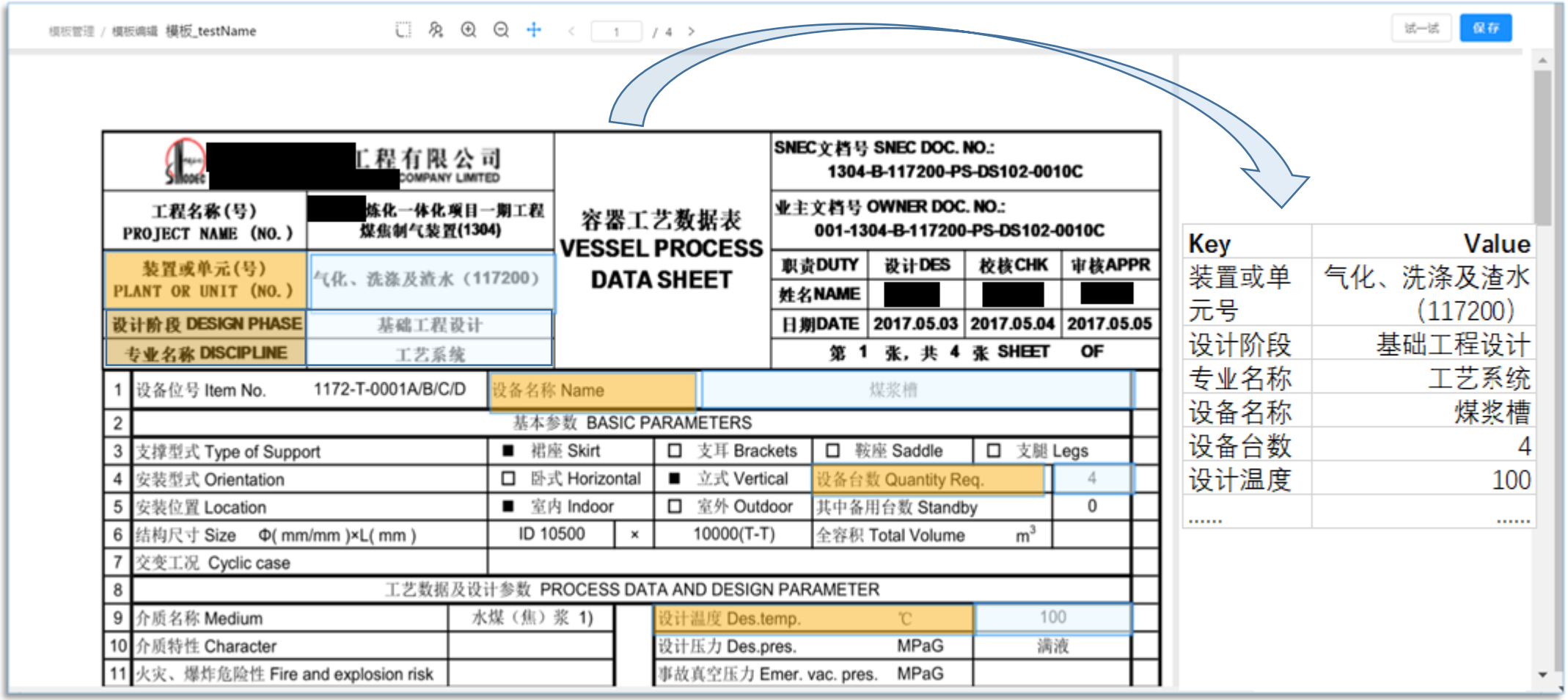
Enterprise multi-source unstructured data

ERP CRM Engineering Doc Public website Patent Wikipedia Third-party Publication

Petroleum Industry Knowledge Graph

Lenovo Data Intelligence

DocReader Assistant: In order to build knowledge graph, entities and relationships should be extracted from different types of documents, ex. word, excel, PPT, txt, pdf, image files, CAD files, etc.



Petroleum Industry Knowledge Graph

Lenovo Data Intelligence

Smart Search:

Supporting millions of internal & external documents, from RDBMS, ERP, CRM etc.

Supporting various types of docs, ex. word, excel, PPT, txt, pdf, image files, CAD files, etc.

Main Topic Extraction

Natural Language Processing,
5-10 topics of one document

Main Topic

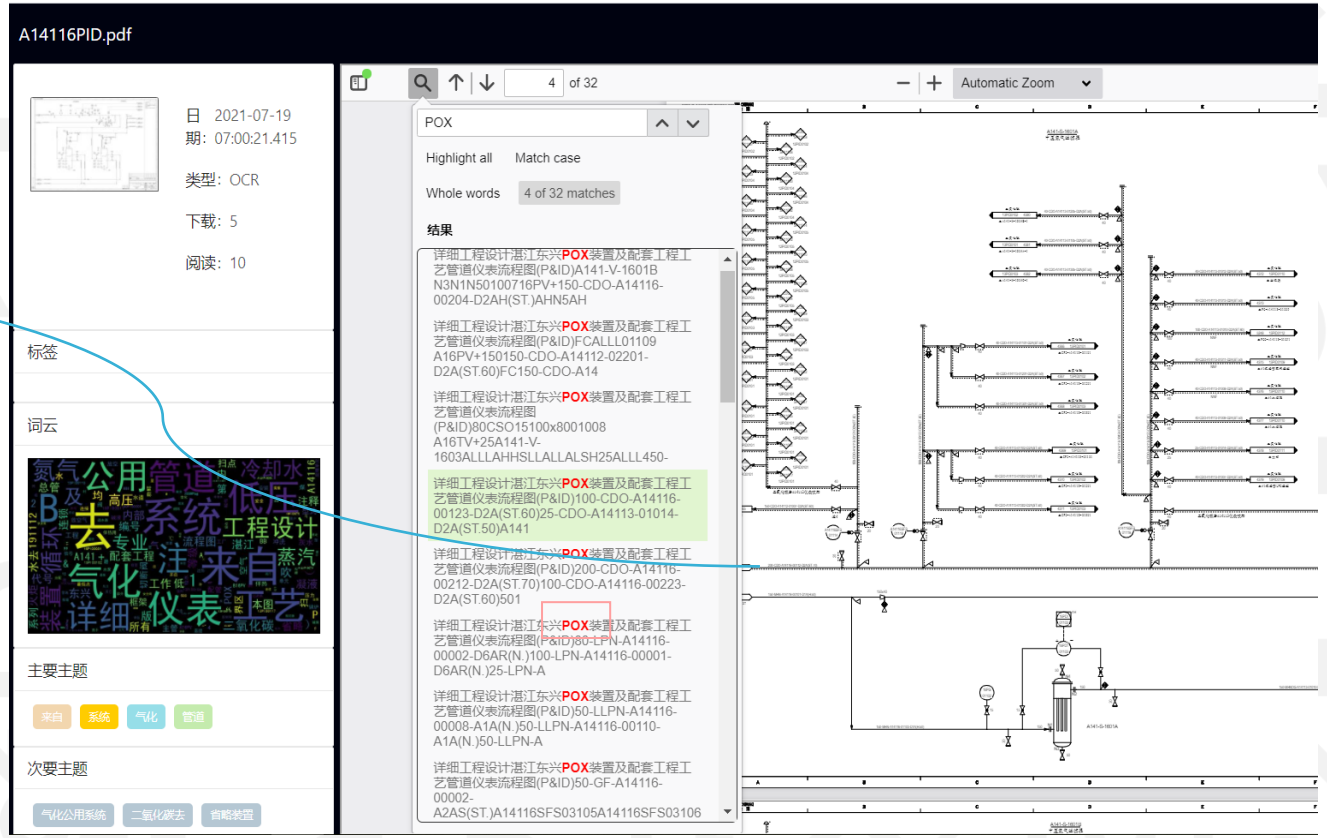
管道 POX 材料 设备

Secondary Topic

防潮层材料 保温材料 设备和管道 工程设计

Keyword Extraction

Extract keywords from article, word clouds etc.



Petroleum Industry Knowledge Graph

Lenovo Data Intelligence

Knowledge Recommendation: The platform supports knowledge recommendation based on knowledge graph, and provides customized services such as semantic associated recommendation and knowledge push.

Knowledge Graph Recommendation

Based on the knowledge graph technology, relevant documents with the same subject words can be pushed regarding to node similarity.

Semantic Association Recommendation

Based on the deep learning model, the document features are embedded in quantization and associated documents are pushed based on features.

User Profile Recommendation

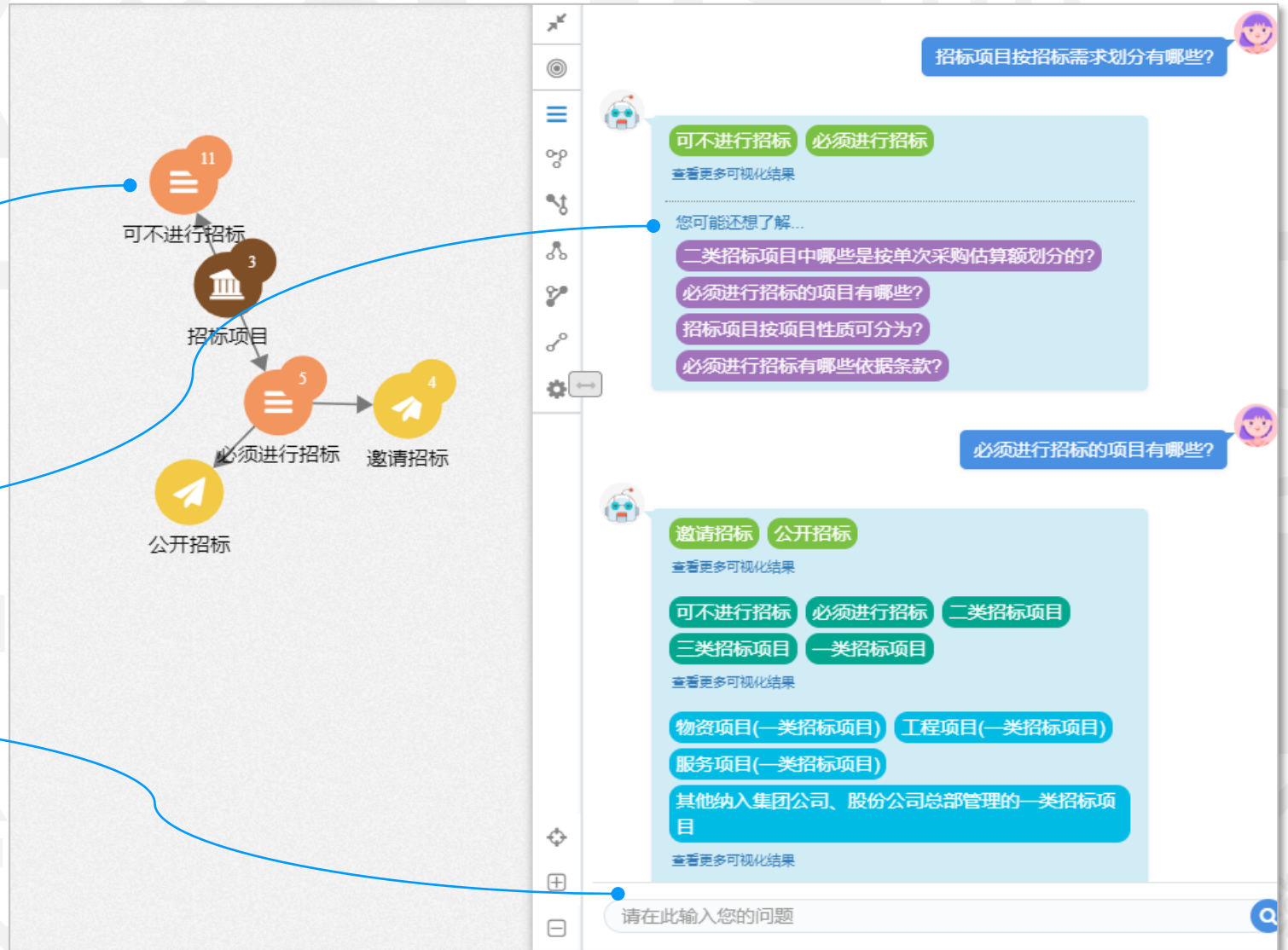
Recommend related documents to users according to their profiles and historical click behavior

Petroleum Industry Knowledge Graph

Lenovo Data Intelligence

Knowledge Q&A: Based on the industrial knowledge base, interactive knowledge Q&A applications can be built to enable users to quickly and accurately obtain knowledge, so as to improve the efficiency of knowledge acquisition.

- Interactive Q&A
- Visualization the relationships
- Similar question recommendation
- Natural language query



Fintech Cases

Lenovo Data Intelligence

Marketing

- Retail bank customer personalized marketing
- Micro loans recommendation
- SMS marketing of financial products
- Credit card installment recommendation
- Financial product recommendation system

Graph Data Mining

- Anti-money laundering
- Relationship mining

Risk Management

- Consumer loan default prediction
- Fraud transaction detection

Financial NLP

- Sentimental analysis of financial news

Lenovo Data Intelligence

Thanks

Different is better

Jun Luo
jluo1@lenovo.com

Machine Intelligence Center
Lenovo Data Intelligence Business Group