

# Lenovo Enterprise Analytics Platform LeapAI and Case Studies

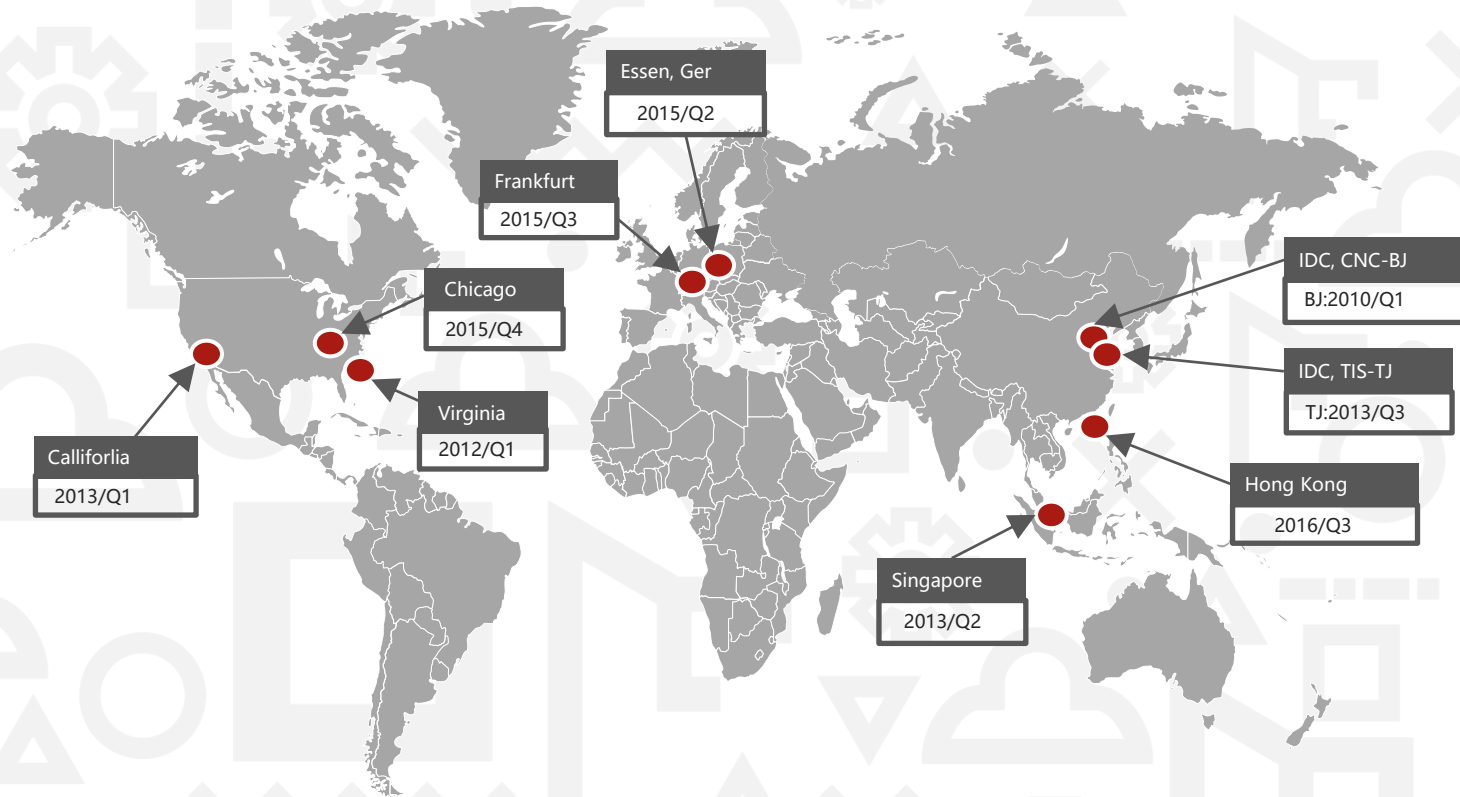
● Jun Luo

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**

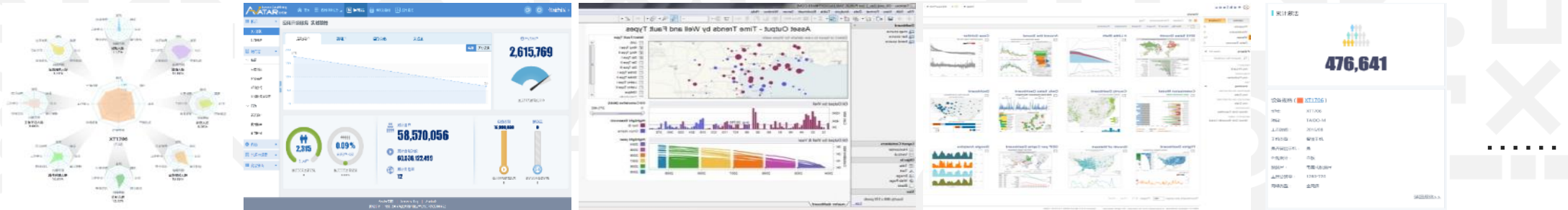
New data **>30TB/day**

Task processed **>15K/day**

Availability **>99.99%**

# Lenovo Digital Transformation

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



Precise insight

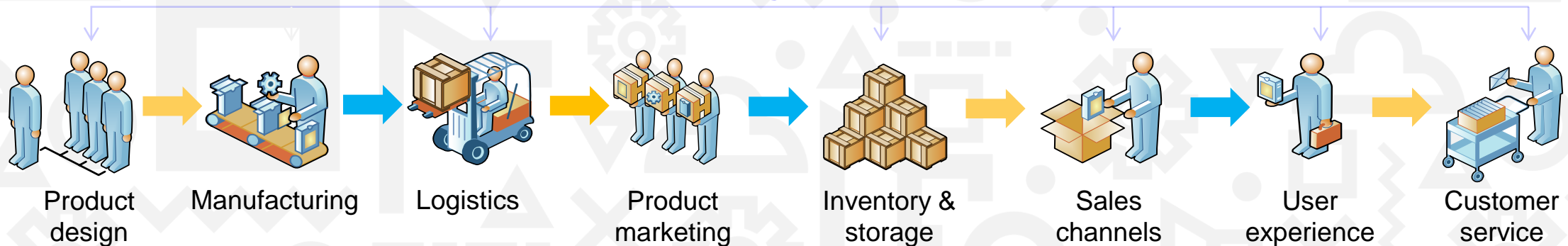
Improve supply chain management

Improve product marketing ROI

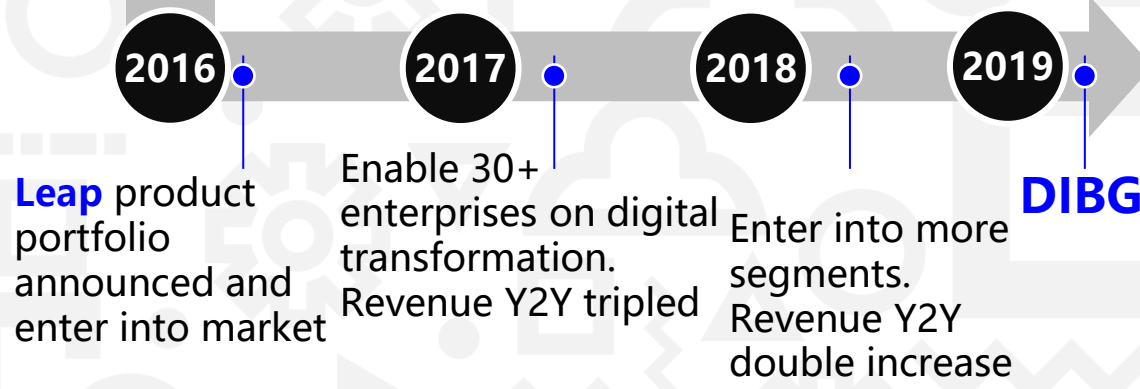
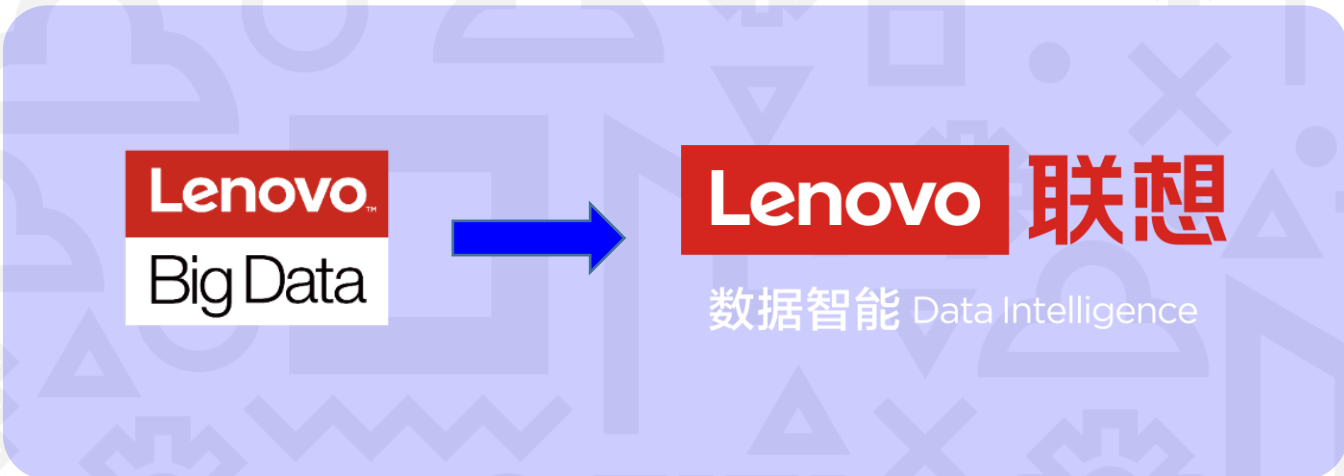
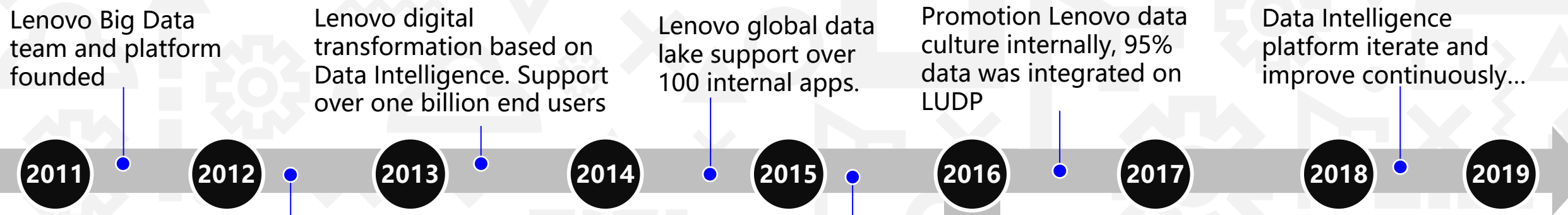
Establish global insight & analysis capabilities

Increase customer satisfaction

## Data-driven through whole value chain



# Lenovo Way to Data Intelligence



# We Offer End-to-End Data Intelligence Solution

DI Application



**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	Security
Architecture	MasterData

AI Platform



**BD Platform LeapHD**

DW	MPP	Parallel Computing	No SQL	Cluster Management	
Parallel Filesystem	HDFS	Resource Management	Memory DB	Graph DB	Temporal DB

**Data Trusted**

Authentication	Application
BIOS	Platform
Network	Server

Big Data Platform



**IoT Data & Computing - LeapIOT**

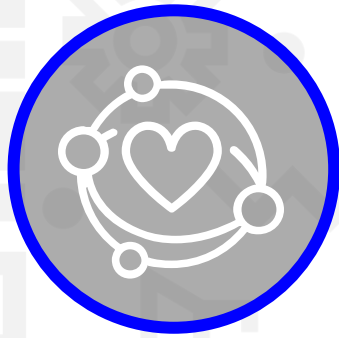
**IT System Data Integration - DataHub**

CRM	SCM	ERP	MES	PLM	...
DMS	CAE	MRP	MDC	OA	...

IoT & Data Collection

# 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.

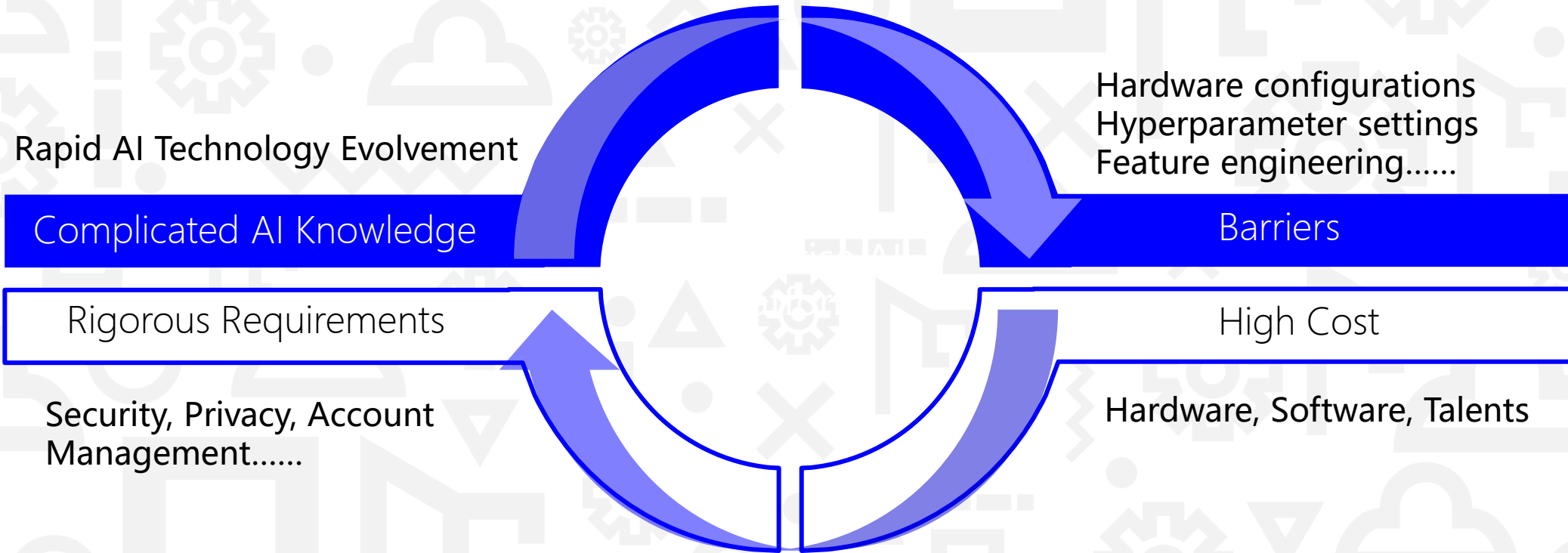


## LeapAI

**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





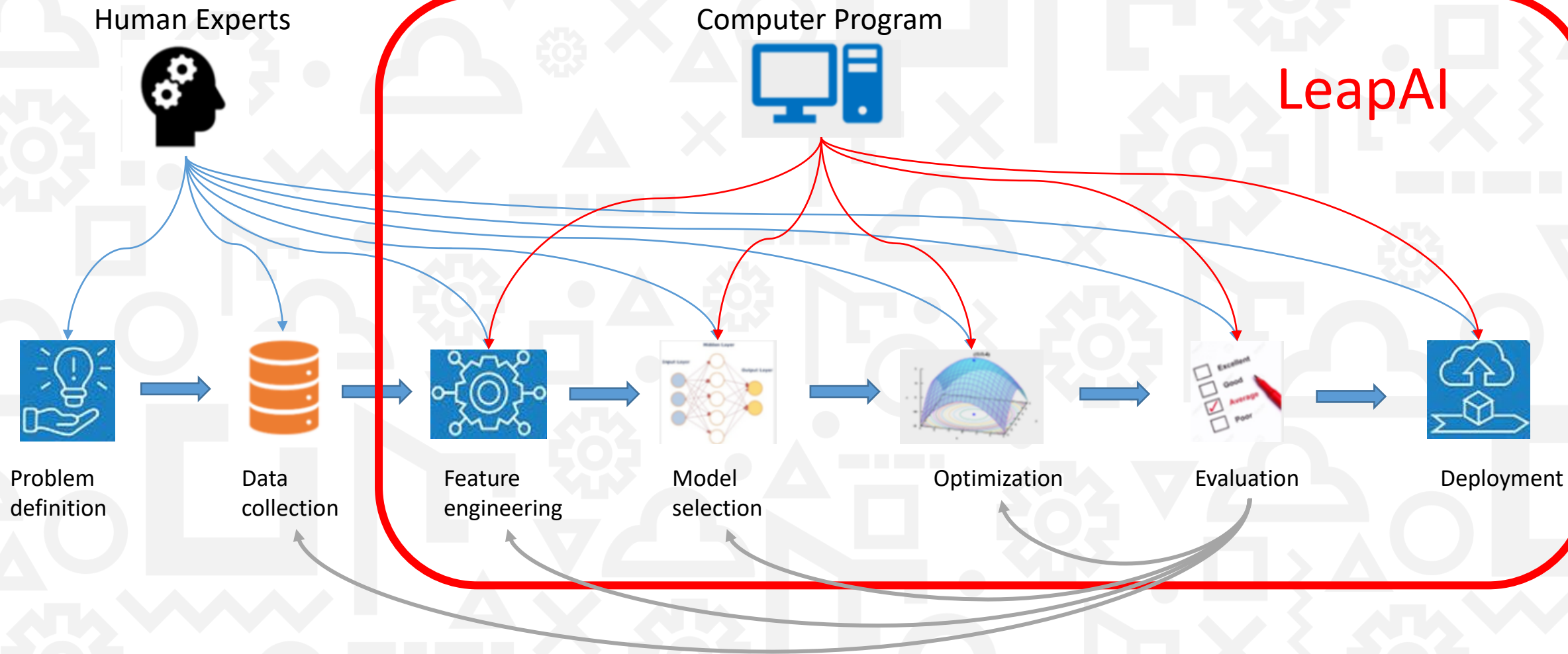
Fish **LeapAI** Net



**Lower barriers for landing AI cases**  
**Enable customers to build their own AI abilities**

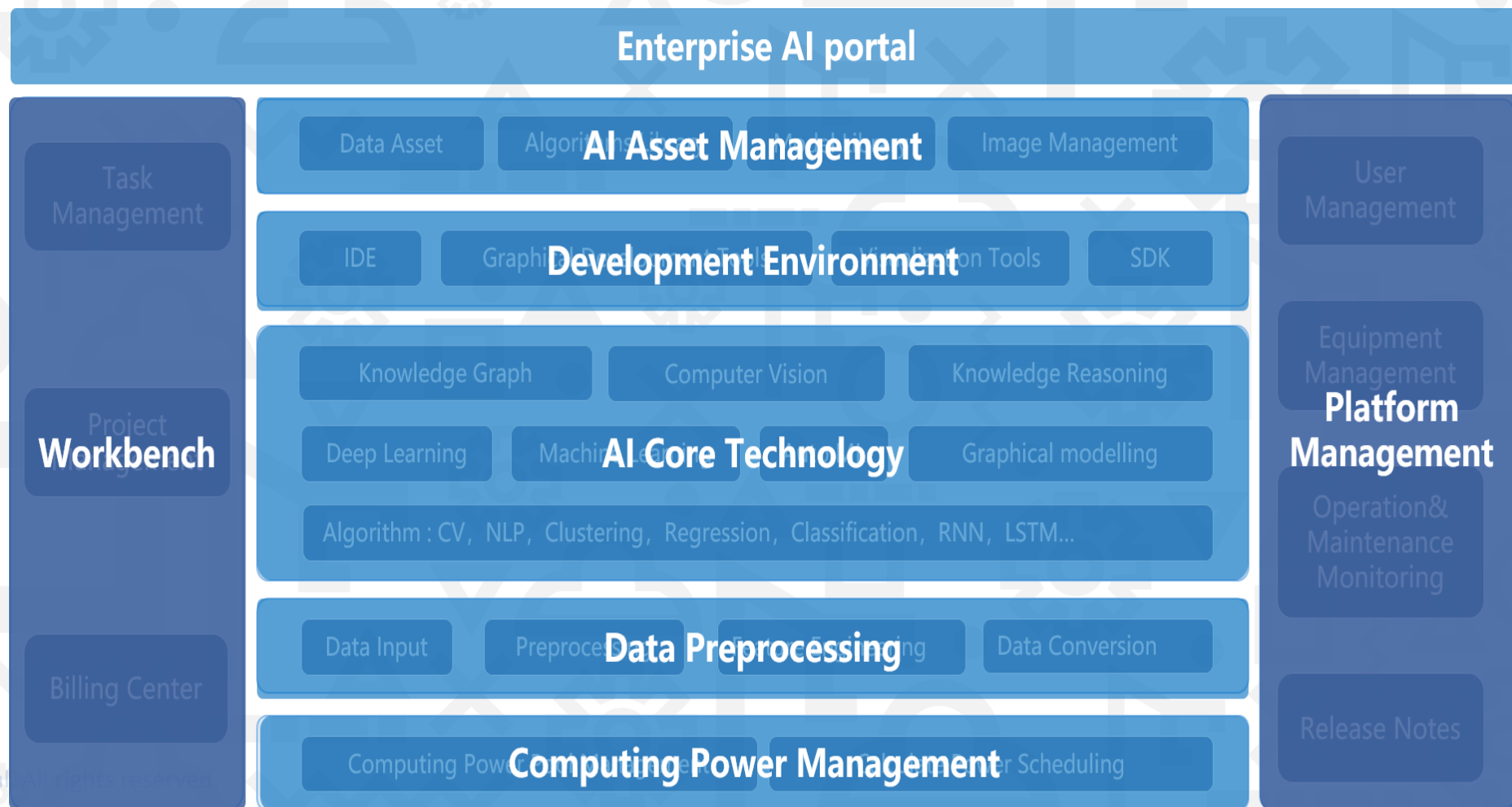


# Machine Learning Workflow



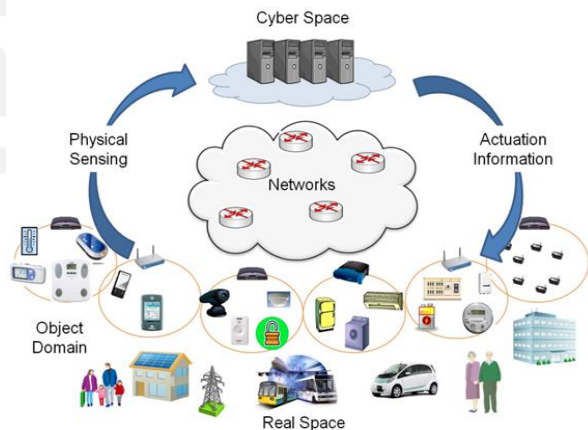
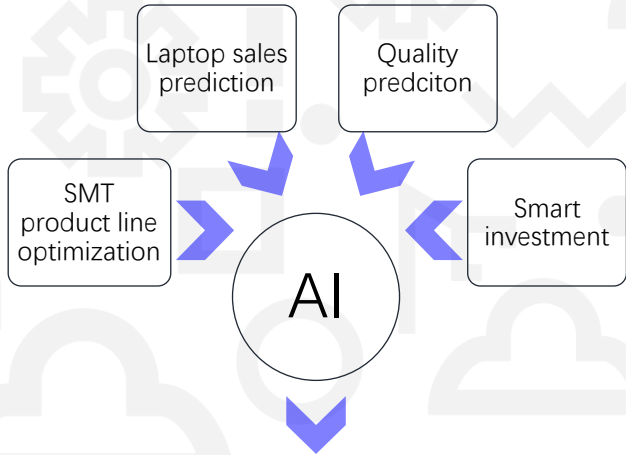
# LeapAI Platform

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

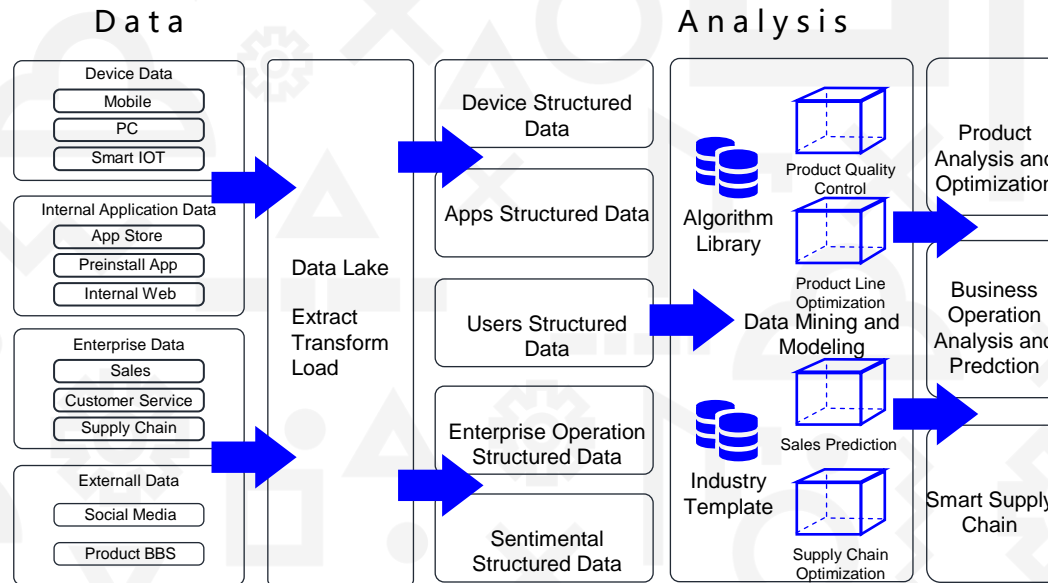


# Lenovo Internal AI Cases: LUDP AI

## Challenges



## Data Analysis



## Applications

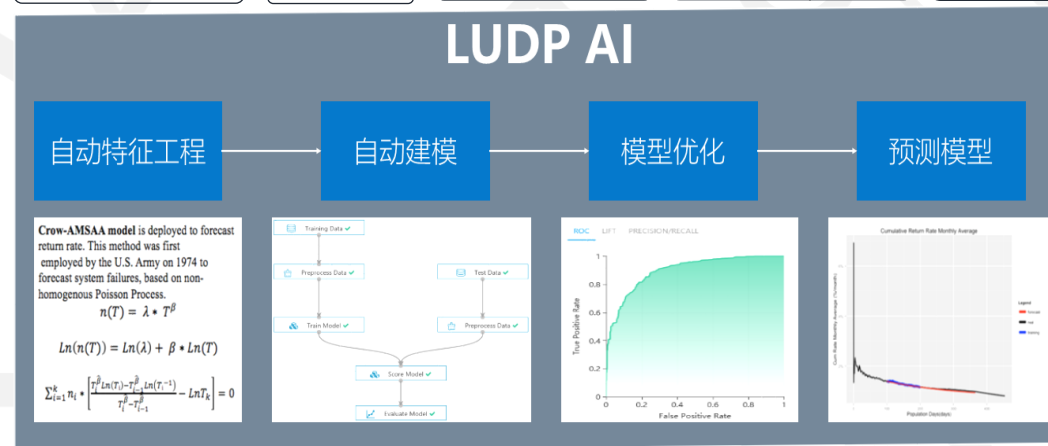
- Dozens of business domain
- Application scenarios **100+**
- Users **1000+**

Supply chain prediction

Quality prediction

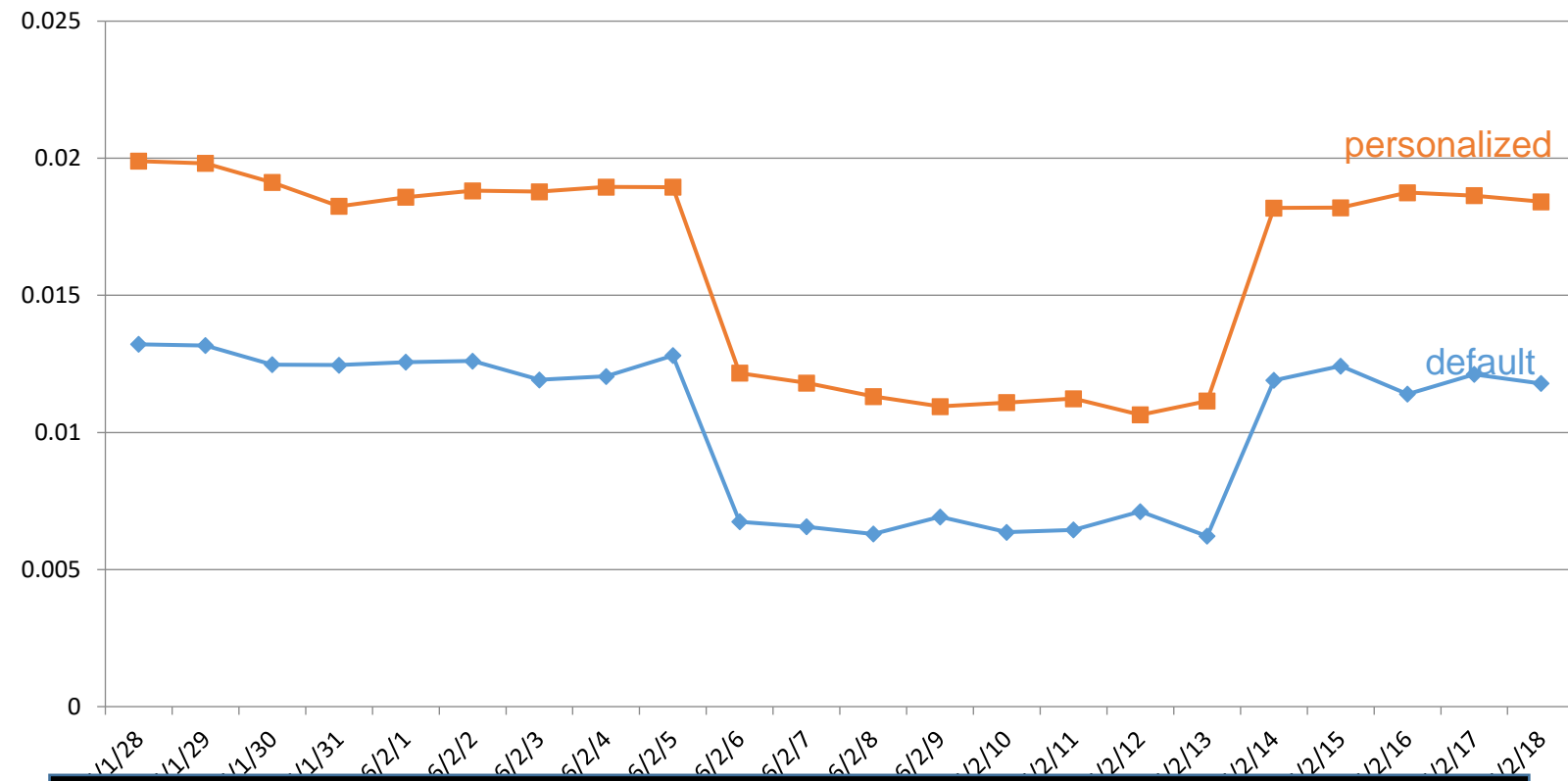
Product line optimization

## LUDP AI



# Lenovo Mobile Big Data Project-Personalized Game App Recommendation

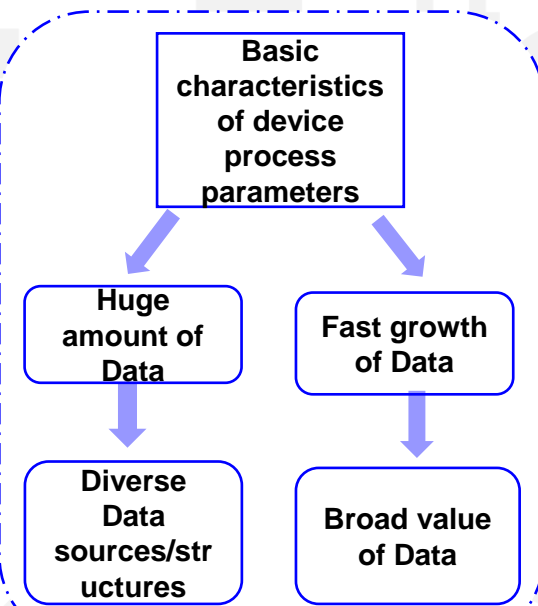
- **Background:** Lenovo app store has thousands of game apps. Originally, all users have the same ranking order of game apps.
- **Goal:** Provide different game apps ranking order for different users in order to increase the click through rate.
- **Difficulties:** millions of features and users, sparse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



# Oil-Refinery Process Optimization

## 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
- ...

Process parameter correlation analysis

Indicator anomaly monitoring

Index parameter optimization

Warning of key points

## Big Data analysis

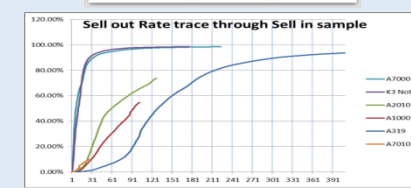
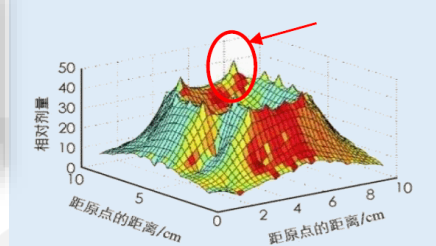
The Big Data analysis section is organized into three rows and three columns:

- Row 1:**
  - Data Acquisition:** Shows a dashboard with various data points and a flow diagram.
  - Cleaning & standardization:** Illustrates the process of handling '原始数据' (original data) into '相关数据' (related data) using a neural network diagram.
  - Correlation algorithm analysis:** Displays '正相关系数' (positive correlation coefficient) and '负相关系数' (negative correlation coefficient) bar charts, along with a '线性&非线性相关系数算法' (linear & non-linear correlation coefficient algorithm) diagram.
- Row 2:**
  - Correlation analysis and feature selection:** Shows '正相关系数' and '负相关系数' bar charts.
  - Prediction model construction:** Features an 'SVM预测模型' (SVM prediction model) diagram and a scatter plot.
  - Indicator anomaly judgment:** Includes an '异常指标报警' (abnormal indicator alarm) table and a flowchart.
- Row 3:**
  - Raw material clustering analysis:** Shows '原料数据' (raw material data) and '主成分降维' (principal component dimensionality reduction) charts.
  - Material classification model:** Illustrates '分类模型的数据源' (classification model data source), 'SVM分类器' (SVM classifier), and '分类结果评估' (classification result evaluation).
  - Sample database of operations:** Shows '待优化目标强相关变量整理' (整理待优化目标强相关变量) and '操作样本库生成' (operation sample library generation).
  - Parameter optimization:** Displays '不同类别原料目标参数最优值计算' (optimal value calculation for different raw material categories).

Data Acquisition      Model training      Model prediction

## Analysis results

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

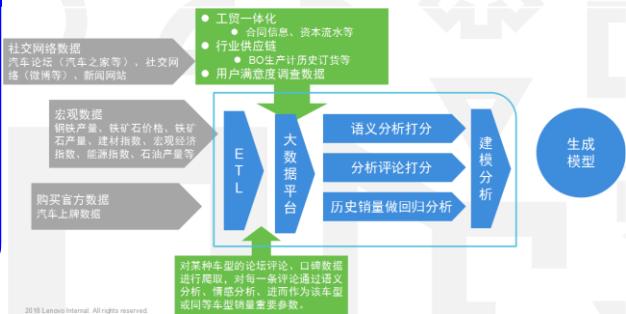
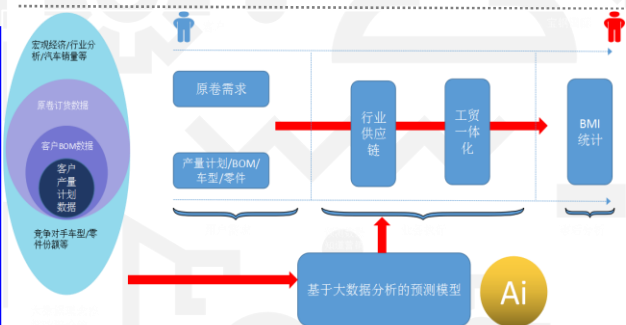


# Steel Demand Prediction

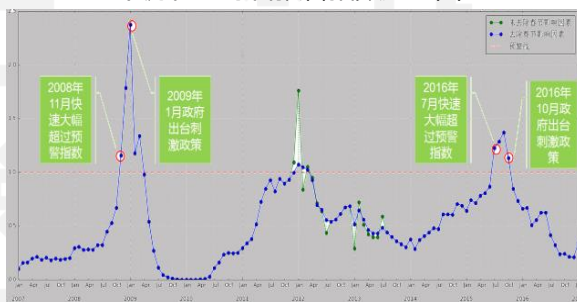
## Challenges

- Internal data are not enough
- Demands change abruptly

## Solutions

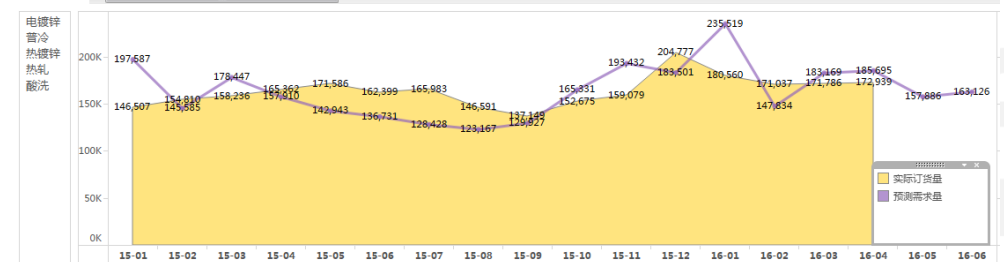


下游行业综合预警指数追踪图



Steel demand prediction of east china region (ton)

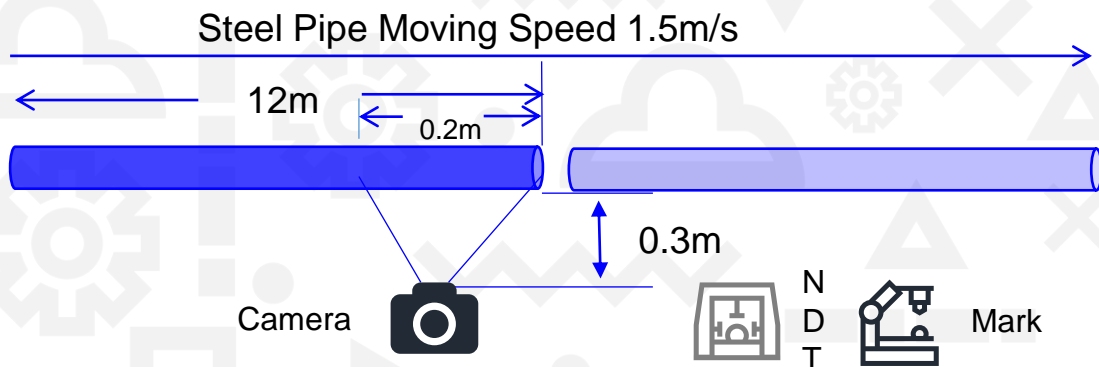
PSR	2015-01			2015-02			2015-03			2015-04			2015-05		
	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率	预测需求量	实际订货量	准确率
总计	197,587	146,507	92.91%	145,585	154,810	89.33%	178,447	158,236	92.86%	157,910	165,362	89.11%	142,943	171,586	88.06%
普冷	84,088	62,350	95.08%	66,481	70,694	91.69%	78,154	70,189	97.80%	72,498	75,919	91.37%	60,804	72,888	90.28%
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	81.18%
电镀锌	66,976	49,662	95.45%	48,457	51,506	99.62%	56,903	52,231	94.94%	50,545	52,930	84.97%	45,989	55,204	82.21%
酸洗	27,217	20,181	98.10%	16,819	17,885	73.89%	21,898	19,418	90.00%	20,301	21,259	83.65%	16,453	19,750	81.28%
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	82.14%



## Results

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

# Steel Pipe Defects Detection

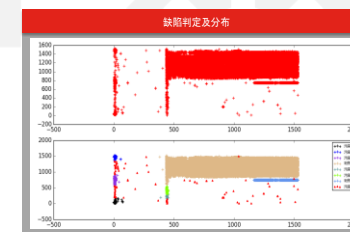


## Process Data Collection

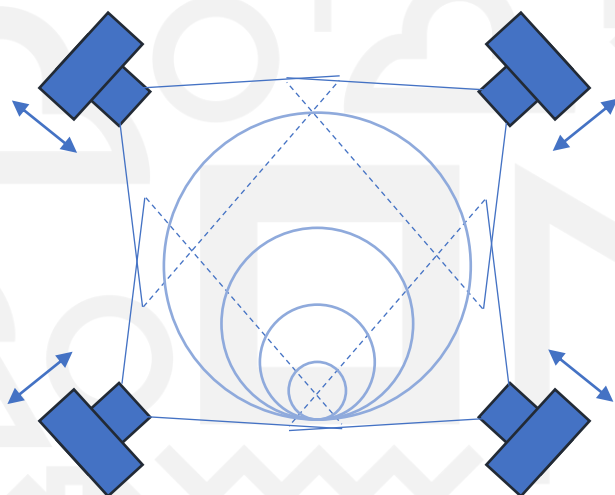
时间	位置	缺陷类型	缺陷长度	缺陷深度	缺陷面积	缺陷体积
21:0	1.217	42.283	1178.676	9.834	2.84	
19:0	1.677	40.962	1179.777	4.215	2.86	
4:0	3.089	43.164	1177.793	2.81	2.86	
5:0	3.714	101.479	1058.513	5.907	2.86	
6:0	5.094	1102.501	105.962	10.396	13.82	
10:0	5.157	566.310	14.479	491.578	641.47	
11:0	3.243	81.397	577.618	41.586	30.71	
12:0	5.404	1257.886	40.113	4.215	4.83	
13:0	5.404	100.000	1057.367	19.545	11.61	
14:0	5.994	1182.054	13.382	1360.513	45.10	

## System Functions

### 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%**

# Inspecting Power Lines Defects with Drones

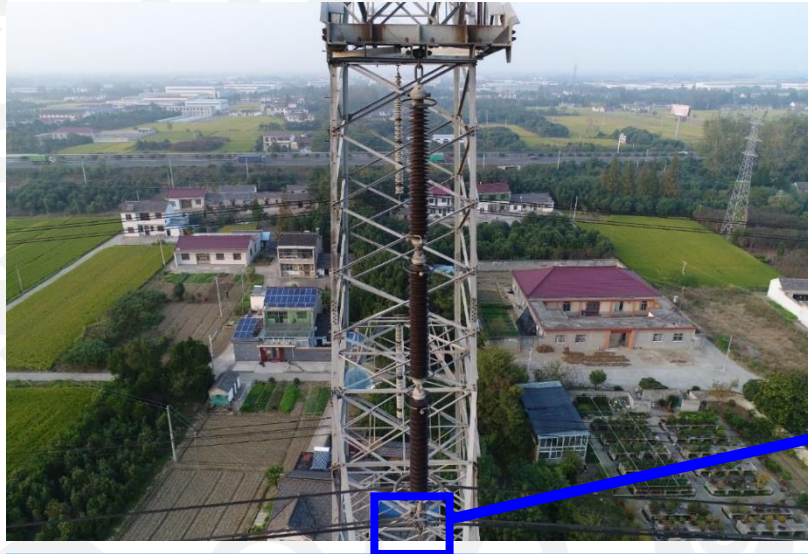


# Power line defects category

- 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.



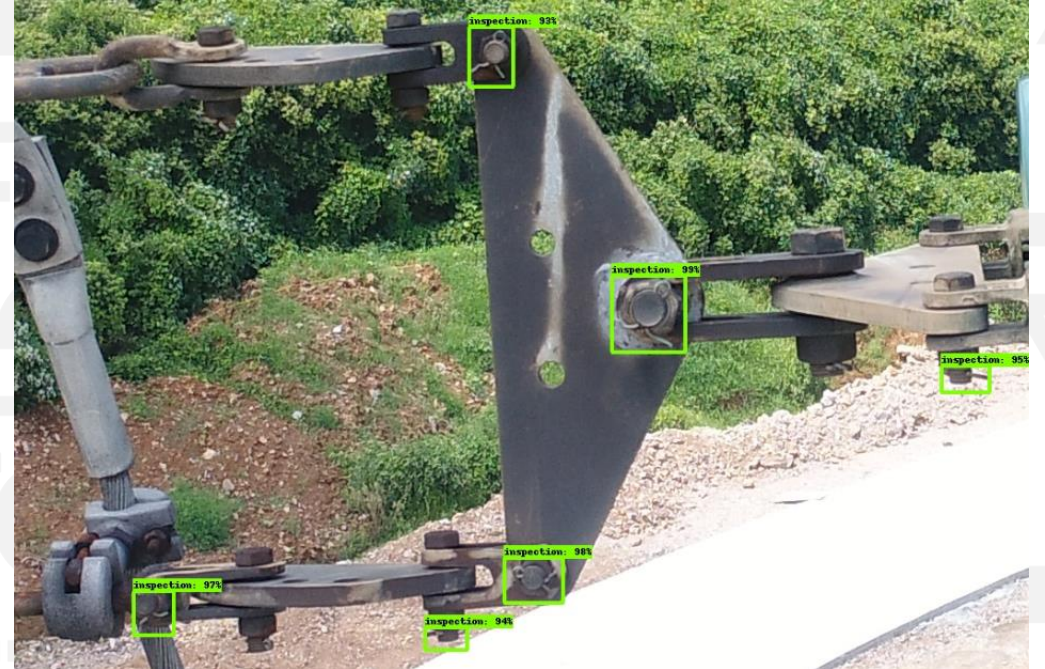
# Challenges: complex shooting environment



# Challenges: small object



zoom in



# Challenges: multiple objects



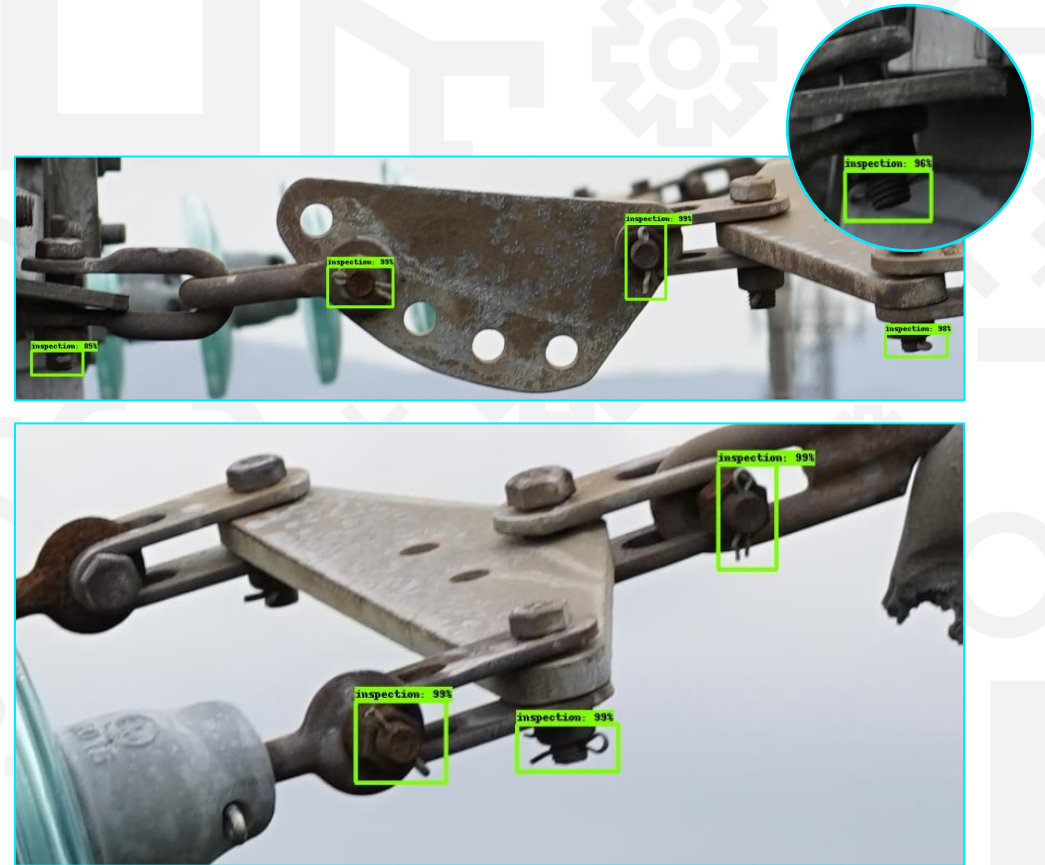
zoom in



# Challenges: limited labelled data



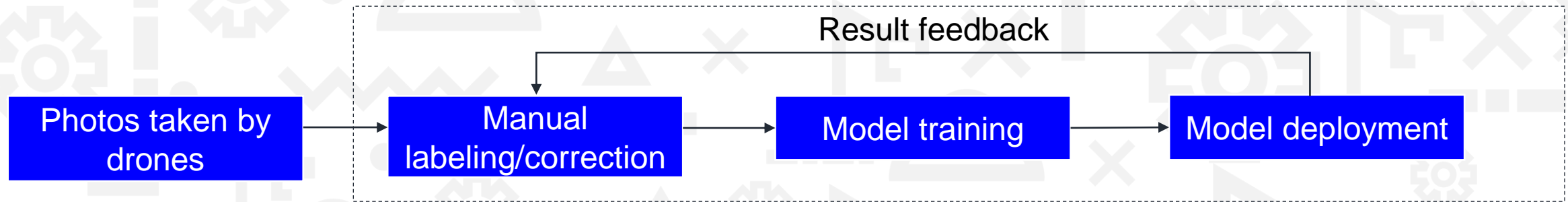
zoom in



# Project highlights

## Automated labeling platform

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



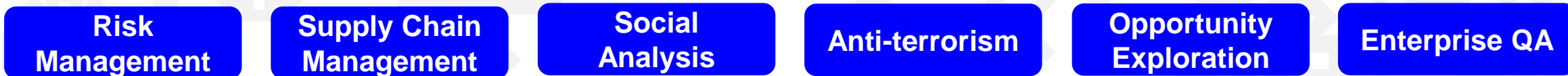
## Power line inspection overall performance (selected tasks)

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

## Intelligent & Interactive Enterprise Level Knowledge Base

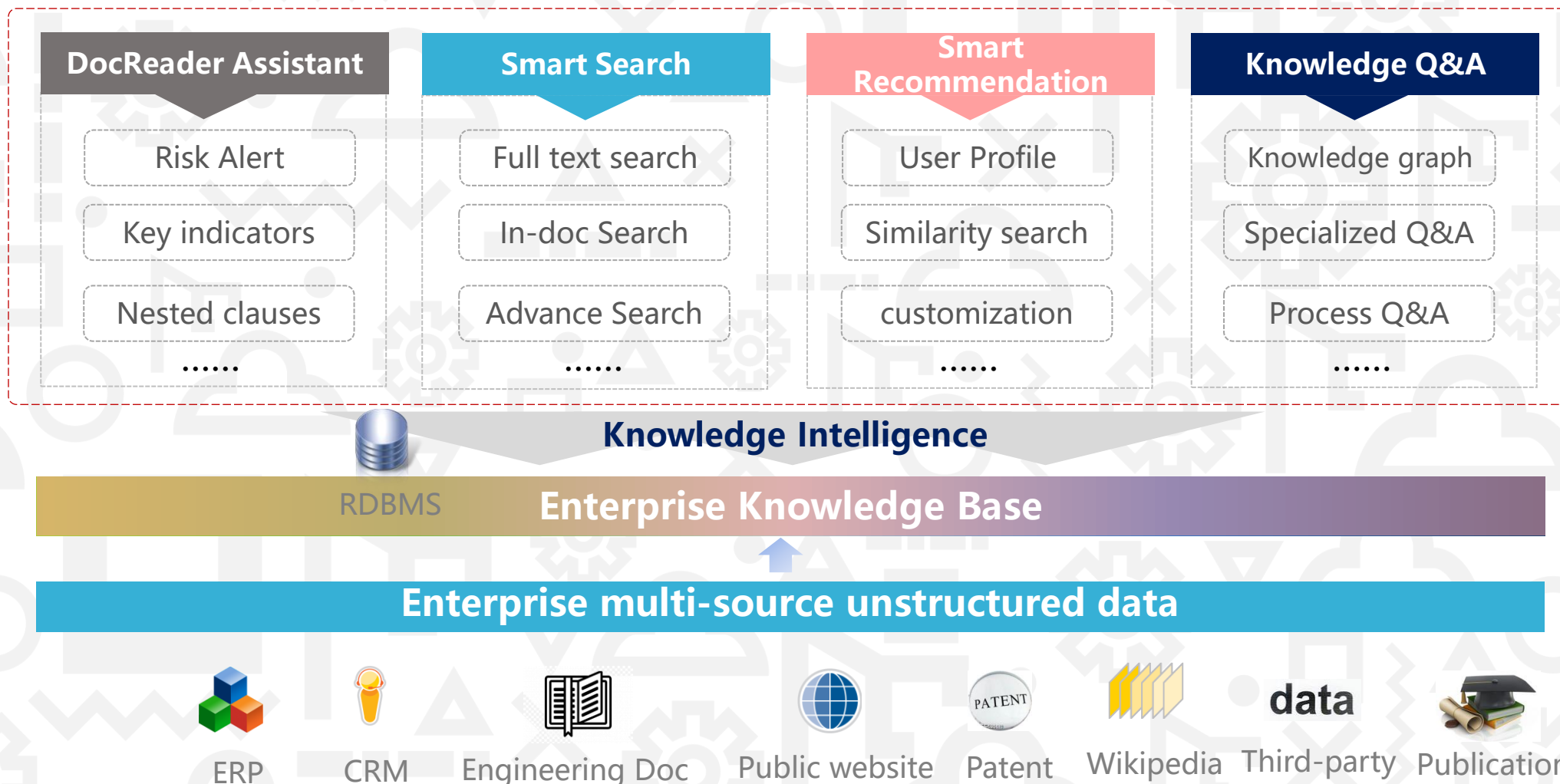
AI-enabled Transformation, Risk Management, Decision Making, Enterprise QA





# Petroleum Knowledge Graph


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



# Petroleum Industry Knowledge Graph

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.

模板管理 / 模板编辑 模板\_testName

 有限公司 COMPANY LIMITED		<b>容器工艺数据表</b> <b>VESSEL PROCESS DATA SHEET</b>	SNEC文档号 SNEC DOC. NO.: 1304-B-117200-PS-DS102-0010C			
工程名称(号) PROJECT NAME (NO.)	炼化一体化项目一期工程 煤焦制气装置(1304)		业主文档号 OWNER DOC. NO.: 001-1304-B-117200-PS-DS102-0010C			
装置或单元(号) PLANT OR UNIT (NO.)	气化、洗涤及渣水 (117200)		职责DUTY	设计DES	校核CHK	审核APPR
设计阶段 DESIGN PHASE	基础工程设计		姓名NAME			
专业名称 DISCIPLINE	工艺系统		日期DATE	2017.05.03	2017.05.04	2017.05.05
第 1 张, 共 4 张 SHEET OF						
1	设备位号 Item No.	1172-T-0001A/B/C/D	设备名称 Name	煤浆槽		
2	基本参数 BASIC PARAMETERS					
3	支撑型式 Type of Support	<input checked="" type="checkbox"/> 裙座 Skirt <input type="checkbox"/> 支耳 Brackets <input type="checkbox"/> 鞍座 Saddle <input type="checkbox"/> 支腿 Legs				
4	安装型式 Orientation	<input type="checkbox"/> 卧式 Horizontal <input checked="" type="checkbox"/> 立式 Vertical	设备台数 Quantity Req.	4		
5	安装位置 Location	<input checked="" type="checkbox"/> 室内 Indoor <input type="checkbox"/> 室外 Outdoor	其中备用台数 Standby	0		
6	结构尺寸 Size $\Phi$ (mm/mm) $\times$ L(mm)	ID 10500	$\times$	10000(T-T)	全容积 Total Volume m <sup>3</sup>	
7	交变工况 Cyclic case					
8	工艺数据及设计参数 PROCESS DATA AND DESIGN PARAMETER					
9	介质名称 Medium	水煤(焦)浆 1)	设计温度 Des.temp. °C	100		
10	介质特性 Character		设计压力 Des.pres. MPaG	满液		
11	火灾、爆炸危险性 Fire and explosion risk		事故真空压力 Emer. vac. pres. MPaG			

Key	Value
装置或单元号	气化、洗涤及渣水 (117200)
设计阶段	基础工程设计
专业名称	工艺系统
设备名称	煤浆槽
设备台数	4
设计温度	100
.....	.....

# Petroleum Industry Knowledge Graph

## 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.

**Document Types:** Word (W), Excel (X), PowerPoint (P), PDF, DWG, TXT, JPG, PNG.

**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.

**Search Results (A14116PID.pdf):**  
Date: 2021-07-19, 07:00:21.415  
Type: OCR  
Downloads: 5  
Views: 10

**Search Query:** POX  
4 of 32 matches

**Search Results:**  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)A141-V-1601B-N3N1N50100716PV+150-CDO-A14116-00204-D2A(H)ST.JAHN5AH  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)FCALLL01109-A16PV+150150-CDO-A14112-02201-D2A(ST.60)FC150-CDO-A14  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)80CSO15100x8001008-A16TV+25A141-V-1603ALLAHHSLALLLALSH25ALLL450-  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)100-CDO-A14116-00123-D2A(ST.60)J25-CDO-A14113-01014-D2A(ST.50)A141  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)200-CDO-A14116-00212-D2A(ST.70)100-CDO-A14116-00223-D2A(ST.60)501  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)80-CPN-A14116-00002-D6AR(N)100-LPN-A14116-00001-D6AR(N)J25-LPN-A  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)50-LLPN-A14116-00008-A1A(N)J50-LLPN-A14116-00110-A1A(N)J50-LLPN-A  
详细工程设计湛江东兴POX装置及配套工程工艺管道仪表流程图(P&ID)50-GF-A14116-00002-A2AS(ST)A14116FS03105A14116SF03106

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

### 聚乙烯催化剂技术

搜索文档

聚乙烯催化剂技术是中石化最大的品种，也是石化行业最重要的产品（LLDPE），也是石化行业最重要的产品（LLDPE），也是石化行业最重要的产品（LLDPE），也是石化行业最重要的产品（LLDPE）。

#### 相关知识推荐

- 1 msds 聚乙烯  256阅读
- 2 聚乙烯醇
- 3 三类催化剂及其催化作用—酸碱催...
- 4 聚氯乙烯和聚乙烯薄膜区别
- 5 酶生物催化剂PPT课件
- 6 聚乙烯树脂(PE)
- 7 催化剂表征PPT课件
- 8 聚乙烯国标目录
- 9 催化剂的制备方法PPT课件
- 10 沉淀法制备催化剂PPT课件

## 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

**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

The screenshot displays a knowledge graph on the left and a chat interface on the right. The knowledge graph shows a central node '招标项目' (Bidding Project) with three outgoing arrows to nodes '可不进行招标' (11), '必须进行招标' (5), and '邀请招标' (4). The '可不进行招标' node has an arrow pointing to '公开招标' (Public Bidding). The '必须进行招标' node has an arrow pointing to '公开招标'. The '邀请招标' node has an arrow pointing to '公开招标'. The chat interface shows a user asking '招标项目按招标需求划分有哪些?' (What are the bidding projects divided by bidding requirements?). The system responds with '可不进行招标' and '必须进行招标', followed by a link to '查看更多可视化结果' (View more visualization results). Below this, it suggests '您可能还想了解...' (You may also want to know...) and lists several related questions: '二类招标项目中哪些是按单次采购估算额划分的?' (Which of the second-class bidding projects are divided by single purchase estimate amount?), '必须进行招标的项目有哪些?' (Which projects must be bid?), '招标项目按项目性质可分为?' (How are bidding projects divided by project nature?), and '必须进行招标有哪些依据条款?' (Which provisions are the basis for mandatory bidding?). A second user question asks '必须进行招标的项目有哪些?' (Which projects must be bid?). The system responds with '邀请招标' and '公开招标', followed by a link to '查看更多可视化结果'. It then lists categories: '可不进行招标', '必须进行招标', '二类招标项目', '三类招标项目', and '一类招标项目'. Below this, it lists specific project types: '物资项目(一类招标项目)', '工程项目(一类招标项目)', '服务项目(一类招标项目)', and '其他纳入集团公司、股份公司总部管理的一类招标项目'. A final link '查看更多可视化结果' is provided. At the bottom, there is an input field with the placeholder text '请在此输入您的问题' (Please enter your question here).

## 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

**Thanks**  
**Different is better**

Jun Luo  
[jluo1@lenovo.com](mailto:jluo1@lenovo.com)

Machine Intelligence Center  
Lenovo Data Intelligence Business Group