

21天实战人工智能系列

人工智能产品经理最佳实践

主讲：张子良

Part

1

人机交互模型

核心问题：过程模型和关键技术



1 章节目标



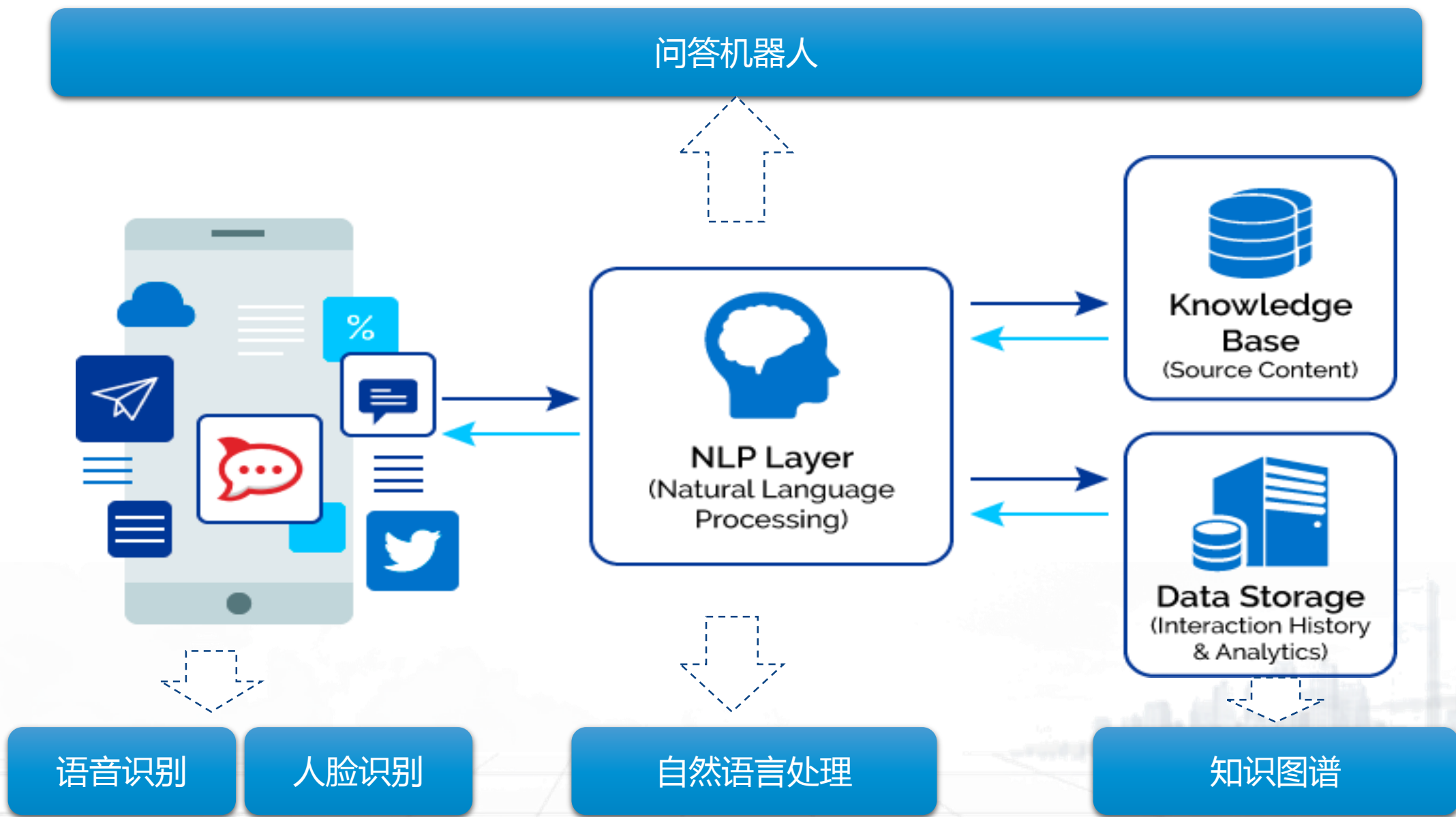
了解人工智能都有哪些关键技术？

了解人工智能关键技术的分类？

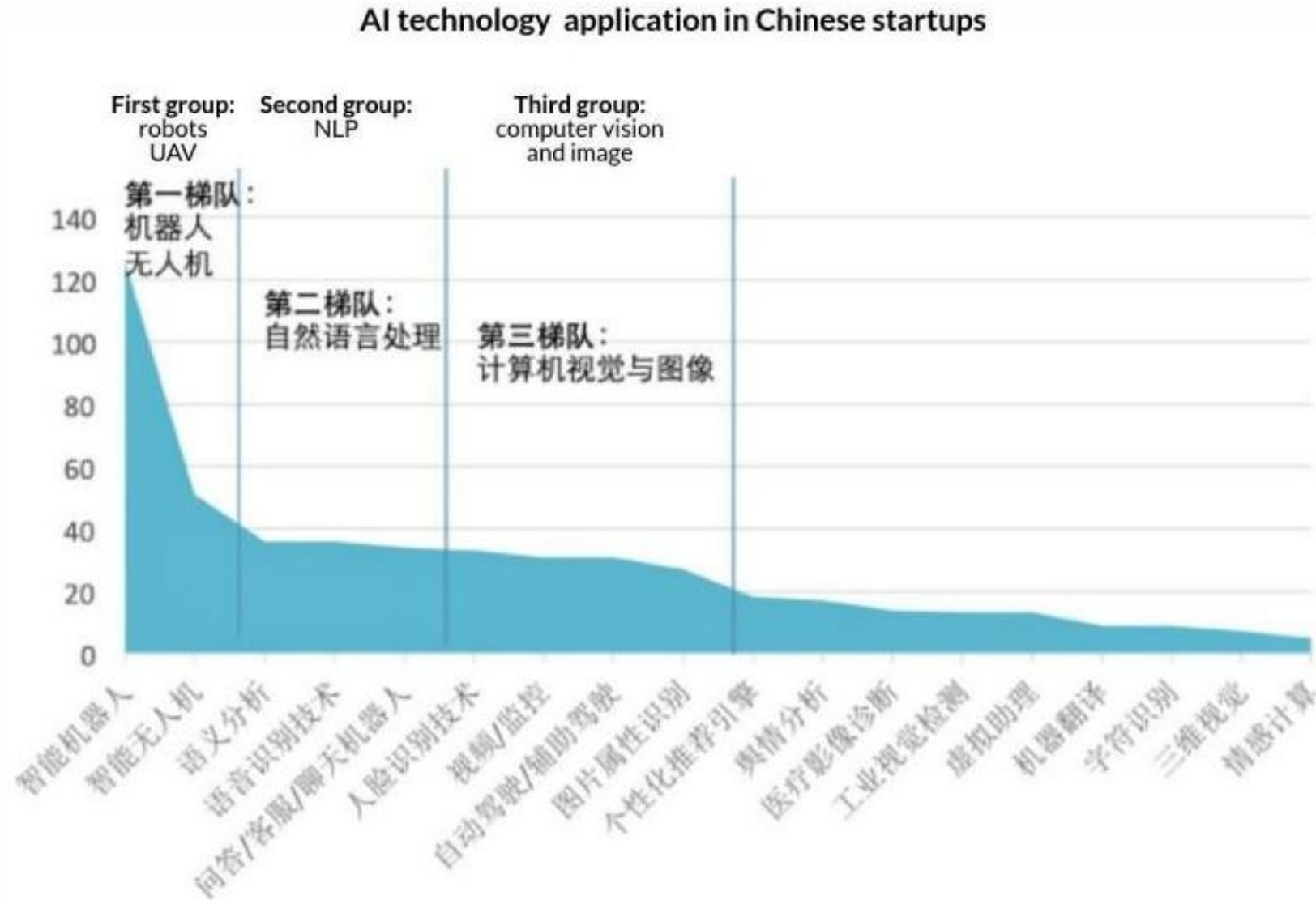
了解人工智能关键技术的协同？

...

1.1 人机交互模型

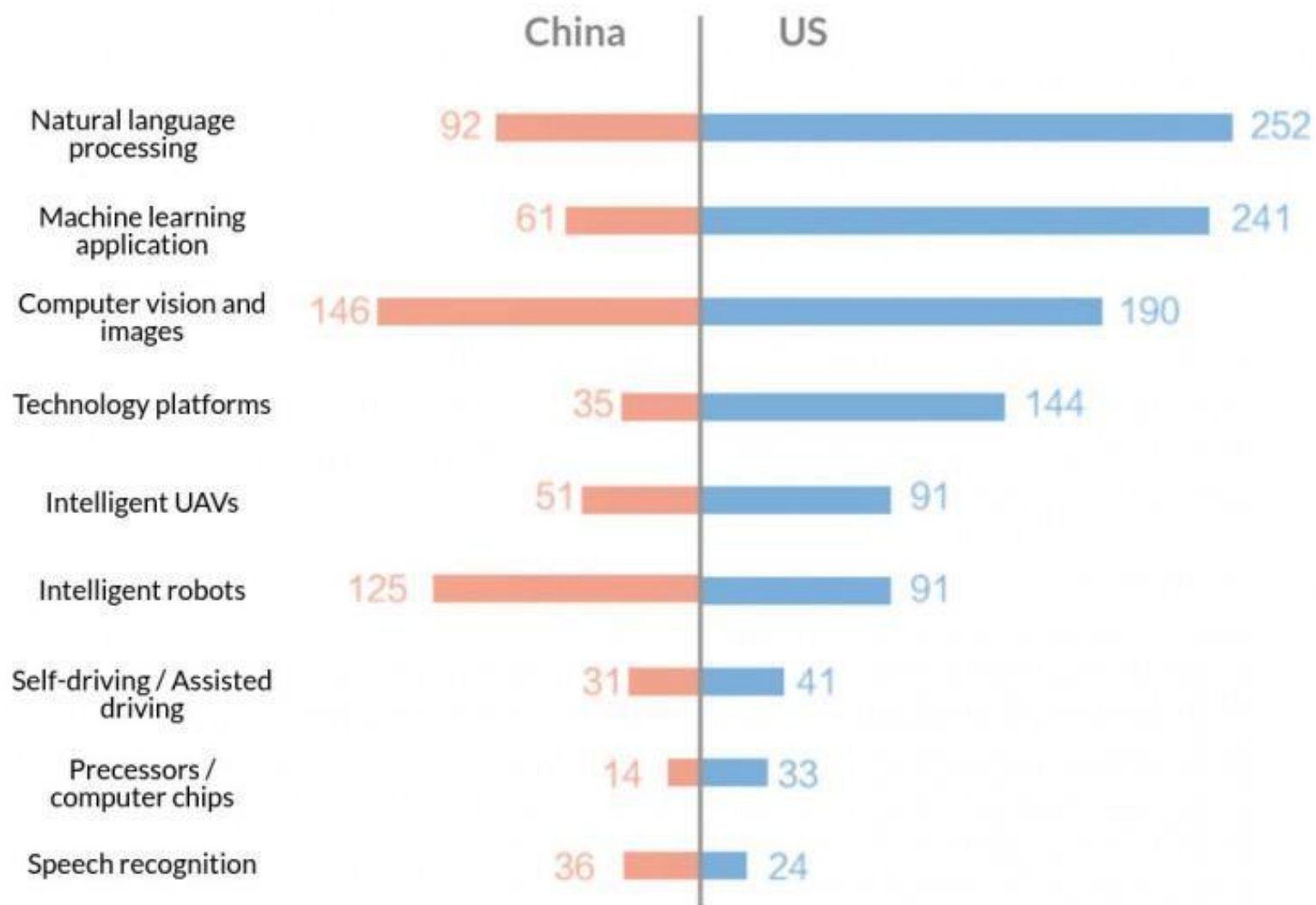


1.2 人工智能技术的梯队划分



List of fields from top to bottom: AI robots, UAV, semantic analysis, Q&A/customer/chatbots, face recognition technology, video/supervision, automatic/assisted driving, image attribute recognition, public sentiment analysis, medical image diagnosis, industrial visual examination, virtual assistants, machine translation, character recognition, 3D vision, affective computing.

1.3 中美人工智能企业布局对比



(Numbers indicate the number of enterprises)

1.4 想一想

**人工智能技术你都了解哪些，
他们发展的怎么样？**

Part

2

智能语音

核心问题：技术框架、用户痛点、算法模型、开源项目



2 章节目标



了解智能语音的技术框架？

了解智能语音的技术痛点？

了解语音识别的算法模型？

熟悉语音识别的开源框架？

2.1 智能语音的技术框架

语音交互

语音识别

孤立词语音识别

连续词语音识别

大词表连续语音识别

语音合成

语言处理

韵律处理

声学处理

语义理解

词法分析

句法分析

语义分析

语用分析

语言生成

知识问答

对话管理

2.2 智能语音的技术痛点

语音交互-痛点

语音识别

远场识别

方言识别

噪音干扰

持续对话

语音合成

情感适配

语调适配

语速适配

语义理解

精准分词

词性标注

语法理论

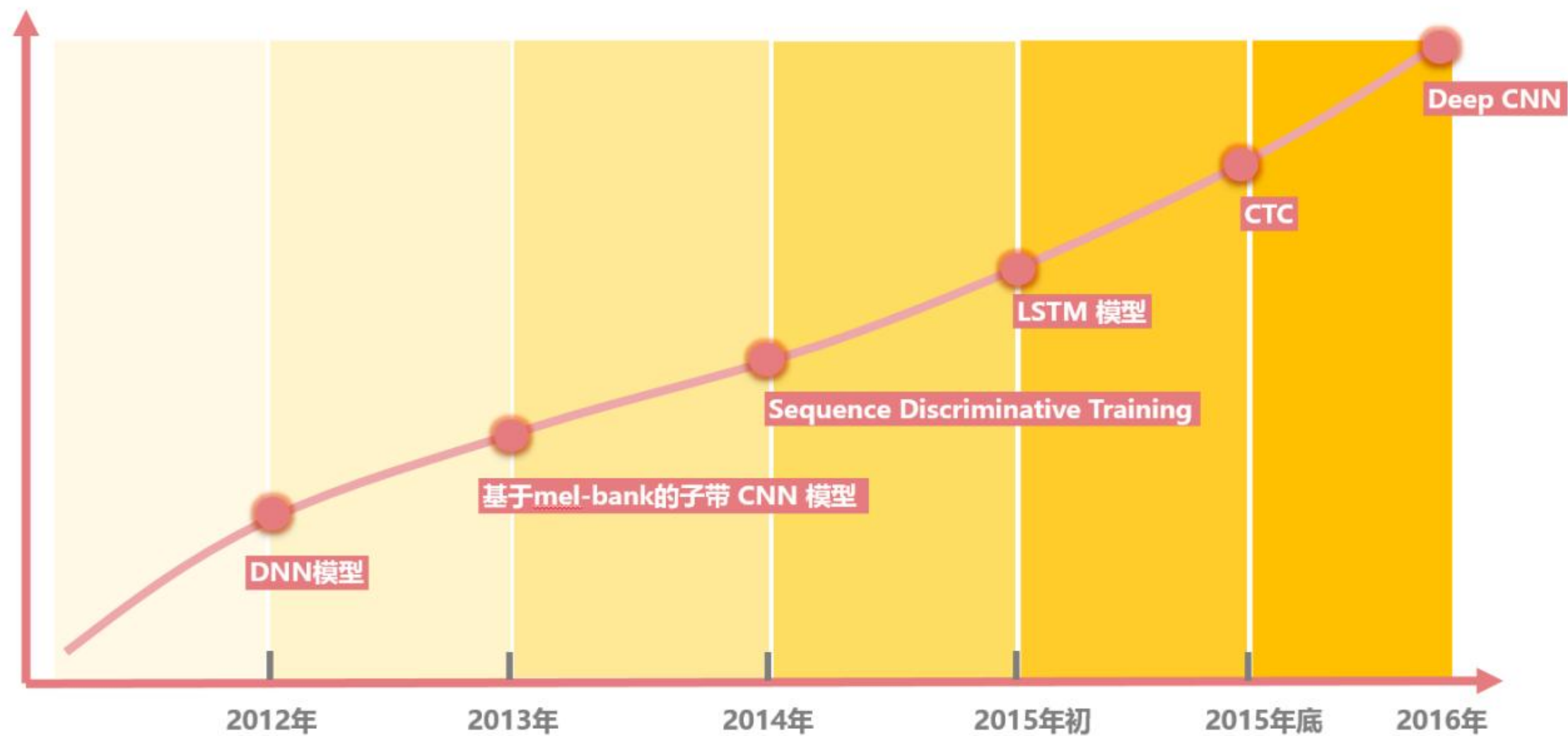
词义消歧

语言生成

知识图谱

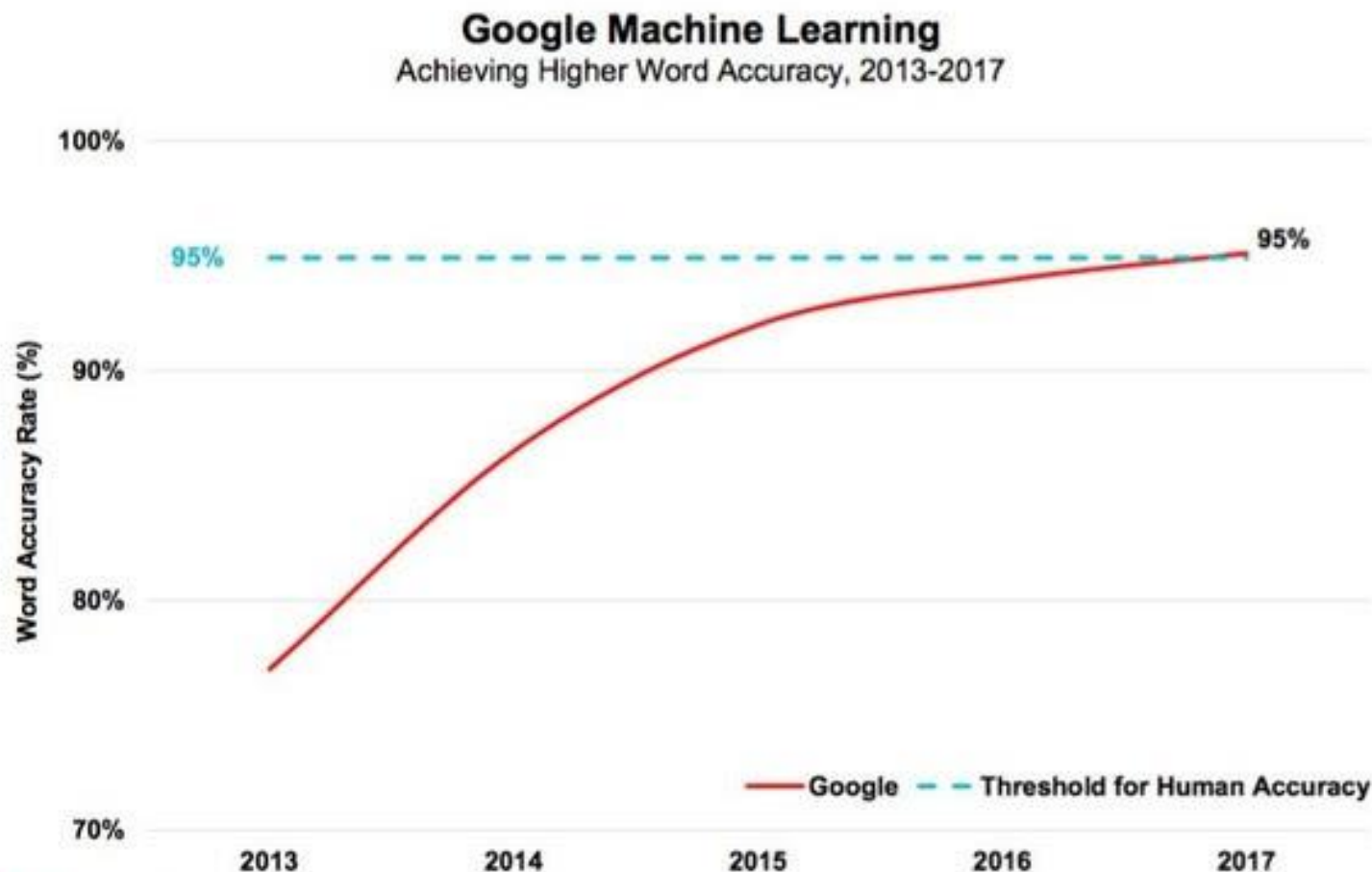
主动学习

2.3 智能语音的技术演进：语音识别模型



2.4 智能语音的技术演进：语音识别错误率

...Voice-Based Platform *Back-Ends* =
Voice Recognition Accuracy Continues to Improve



2.5 语音识别的开源框架

Toolkit	Programming languages	Development activity	Tutorials and examples	Community	Trained models
CMU Sphinx	Java, C, Python, others	+++	+++	+++	English plus 10 other languages
Kaldi	C++, Python	+++	++	+++	Subset of English
HTK	C, Python	++	+++	++	None
Julius	C, Python	++	++	+	Japanese
ISIP	C++	++	++	+	Digits only

2.6 想一想

**语音识别技术真的达到
突破了吗？**

Part

3

自然语言处理

核心问题：技术框架、用户痛点、算法模型、开源项目



3 章节目标



了解自然语言处理的技术框架？

了解自然语言处理的技术痛点？

了解自然语言处理的算法模型？

熟悉自然语言处理的开源框架？

3.1 自然语言处理的技术框架

自然语言处理

词法分析

中文分词

词性标注

实体识别

词义消歧

句法分析

短语结构句法体系

依存结构句法体系

语义分析

浅层语义分析

深层语义分析

语用分析

意义分析

语境分析

3.2 自然语言处理的技术痛点

NLP-痛点难点

分词

北京大学/好
北京/大学好

词性标注

我

代词

喜欢

动词

唱歌

名词

语法理论

结构理论

变换理论

特征理论

配价理论

指向理论

范畴理论

认知理论

语用理论

移位理论

约束理论

词义消歧

组合歧义

人名歧义

语用歧义

语境歧义

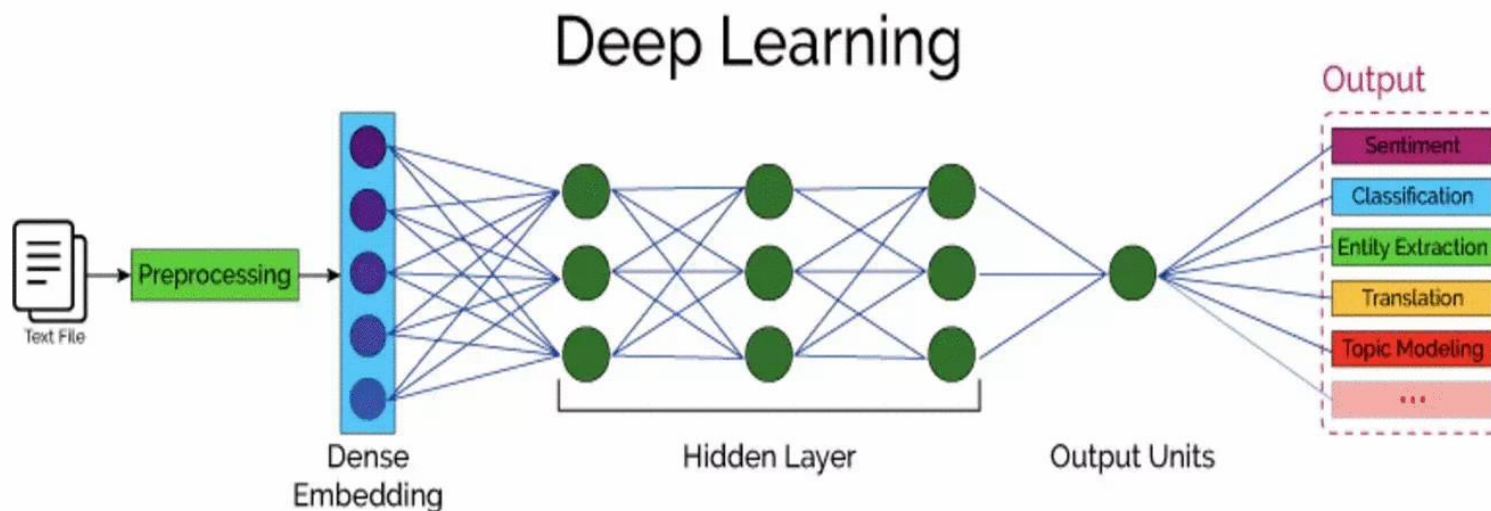
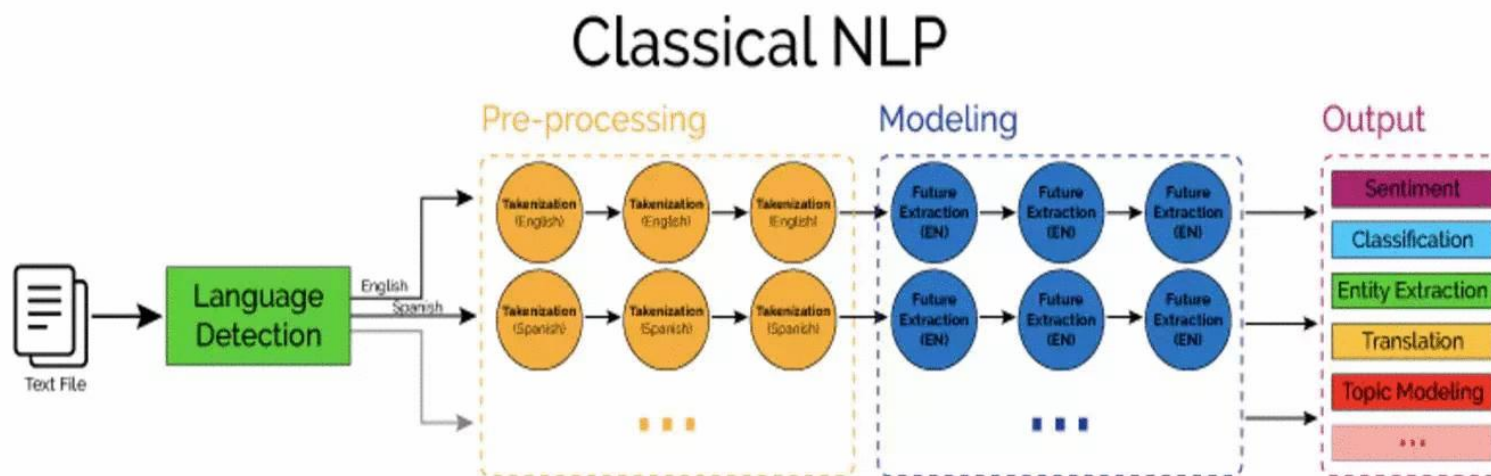
真歧义

分词歧义

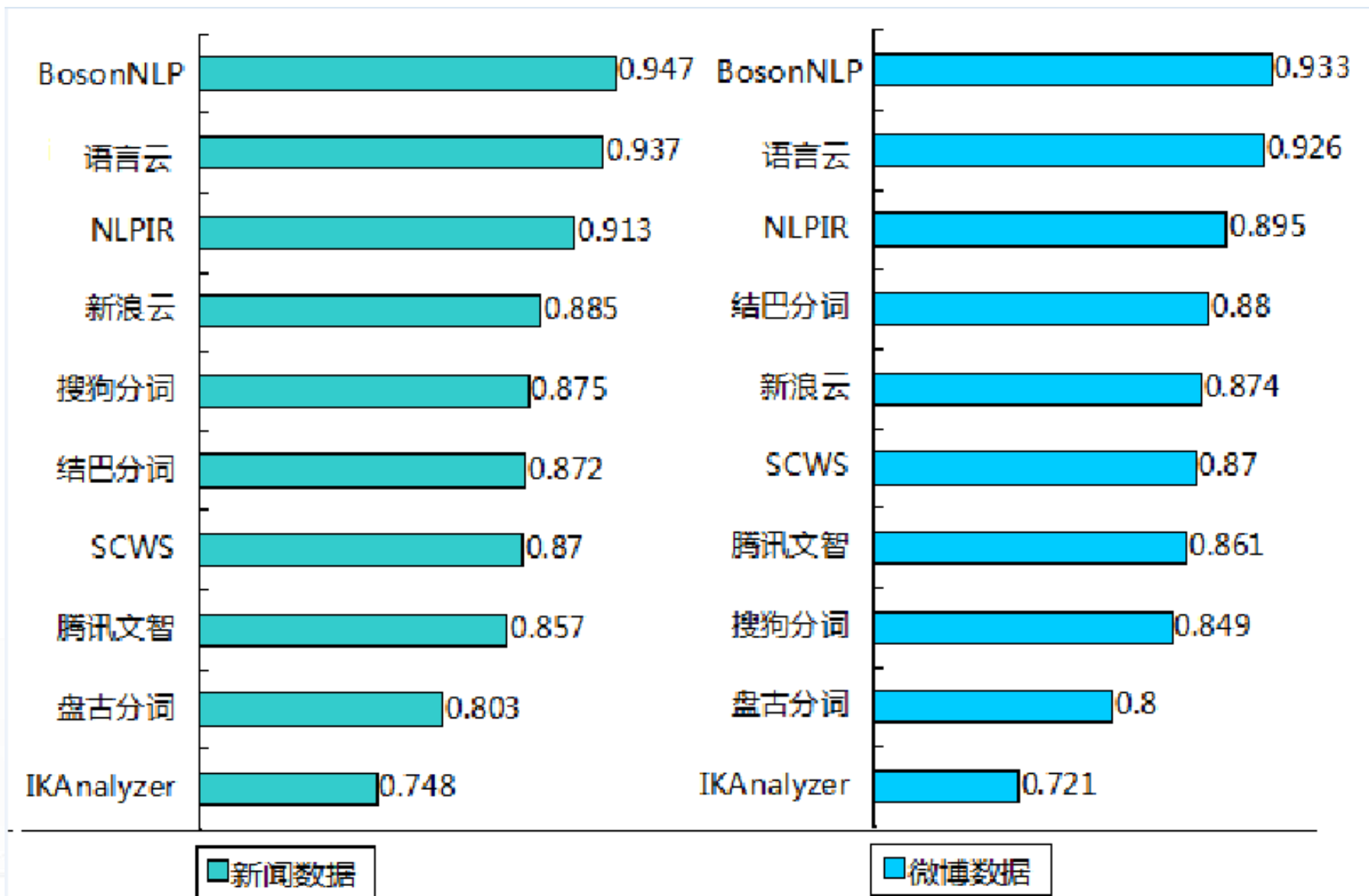
3.2 自然语言处理的技术痛点-语义理解



3.3 自然语言处理算法模型的技术演进



3.4 自然语言处理的开源项目



3.5 想一想

自然语言处理技术真实

的发展水平到底如何？

Part

4

人脸识别

核心问题：技术框架、用户痛点、算法模型、开源项目



4 章节目标



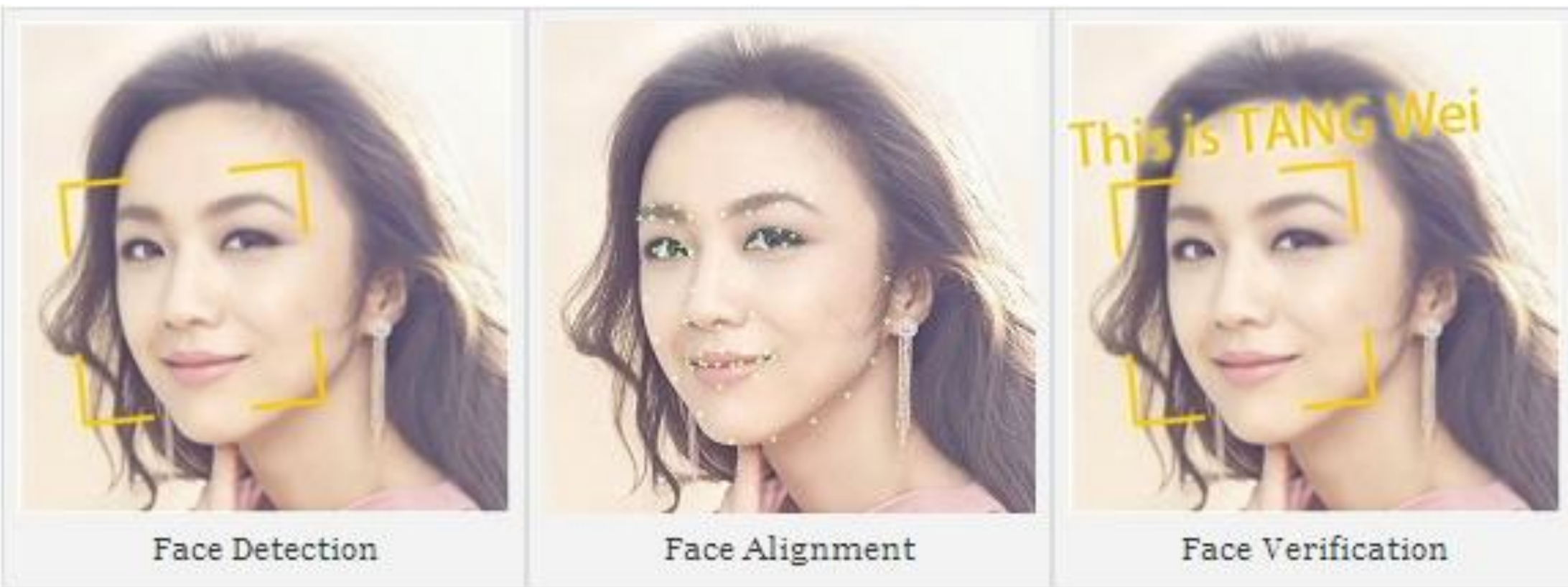
了解人脸识别的技术框架？

了解人脸识别的技术痛点？

了解人脸识别的算法模型？

熟悉人脸识别的开源项目？

4.1 人脸识别的技术框架

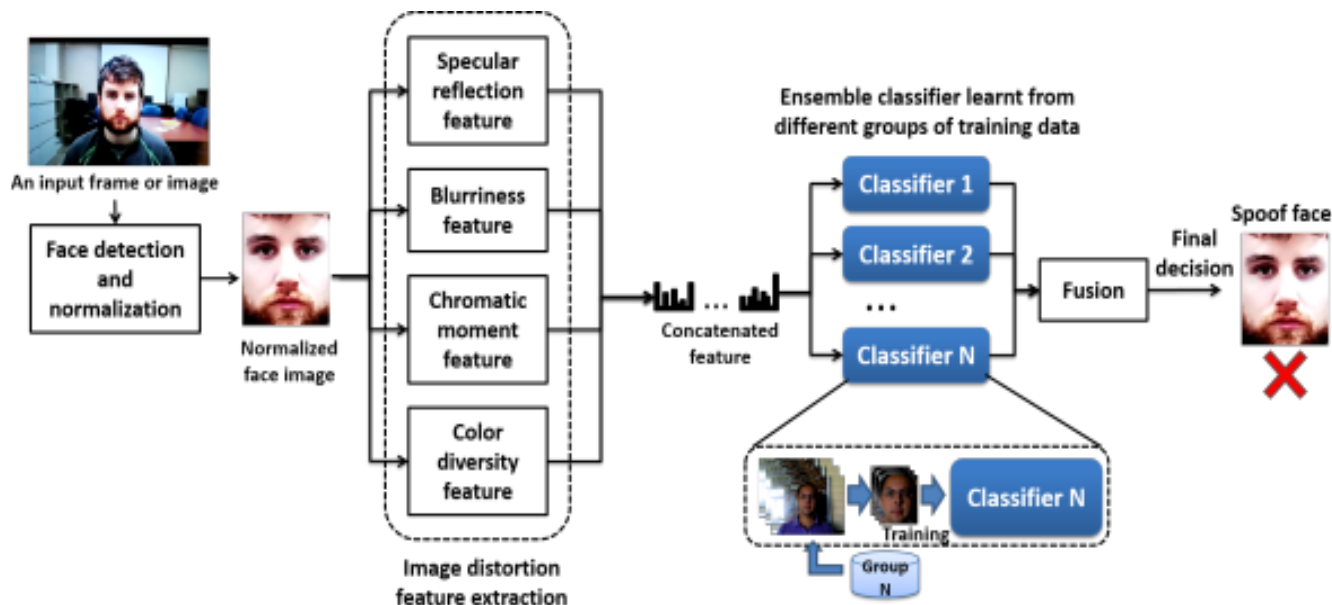


人脸检测

人脸对齐

人脸验证

4.2 人脸识别的技术痛点：活体检测



特征构成

- 镜面反射
- 模糊程度
- 图像色度和对比度退化
- 颜色多样性

判别方法

- 121维向量
- 2值SVM

数据库	公开时间	人数	视频数	采集硬件	攻击类型	人种	性别	年龄
NUAA	2010	15	24个真人33个伪装	Web-cam(640*480)	打印照片	亚洲人	80%男20%女	20-30
IdiapREPLAY-ATTACK	2012	50	200个真人1000个伪装	MacBook自带摄像头(320*240)	打印LCD显示视频回放	76%白22%黄2%黑	86%男14%女	20-40
CASIAFASD	2012	50	150个真人450个伪装	高中低三种不同质量的摄像头	打印剪切照片视频回放	亚洲人	86%男14%女	20-35
MSUMFSD	2014	55	110个真人330个伪装	MacBook Air自带(640*480)谷歌Nexus5手机摄像头(720*480)	打印照片视频回放	70%白28%黄2%黑	63%男37%女	20-60

4.3 人脸识别的算法模型

Simile classifiers ¹¹	0.8472 ± 0.0041
Attribute and Simile classifiers ¹¹	0.8554 ± 0.0035
Multiple LE + comp ¹⁴	0.8445 ± 0.0046
Associate - Predict ¹⁸	0.9057 ± 0.0056
Tom-vs-Pete ²³	0.9310 ± 0.0135
Tom-vs-Pete + Attribute ²³	0.9330 ± 0.0128
combined Joint Bayesian ²⁶	0.9242 ± 0.0108
high-dim LBP ²⁷	0.9517 ± 0.0113
DFD ³³	0.8402 ± 0.0044
TL Joint Bayesian ³⁴	0.9633 ± 0.0108
face.com r2011b ¹⁹	0.9130 ± 0.0030
Face++ ⁴⁰	0.9950 ± 0.0036
DeepFace-ensemble ⁴¹	0.9735 ± 0.0025
ConvNet-RBM ⁴²	0.9252 ± 0.0038
POOF-gradhist ⁴⁴	0.9313 ± 0.0040
POOF-HOG ⁴⁴	0.9280 ± 0.0047
FR+FCN ⁴⁵	0.9645 ± 0.0025
DeepID ⁴⁶	0.9745 ± 0.0026
GaussianFace ⁴⁷	0.9852 ± 0.0066
DeepID2 ⁴⁸	0.9915 ± 0.0013
TCIT ⁵³	0.9333 ± 0.0124
DeepID2+ ⁵⁵	0.9947 ± 0.0012
betaface.com ⁵⁶	0.9953 ± 0.0009
DeepID3 ⁵⁷	0.9953 ± 0.0010
insky.so ⁵⁹	0.9551 ± 0.0013
Uni-Ubi ⁶⁰	0.9900 ± 0.0032
FaceNet ⁶²	0.9963 ± 0.0009
Baidu ⁶⁴	0.9977 ± 0.0006
AuthenMetric ⁶⁵	0.9977 ± 0.0009
MMDFR ⁶⁷	0.9902 ± 0.0019
CW-DNA-1 ⁷⁰	0.9950 ± 0.0022
Faceall ⁷¹	0.9967 ± 0.0007
JustMeTalk ⁷²	0.9887 ± 0.0016
Facevisa ⁷⁴	0.9955 ± 0.0014

pose+shape+expression augmentation ⁷⁵	0.9807 ± 0.0060
ColorReco ⁷⁶	0.9940 ± 0.0022
Asaphus ⁷⁷	0.9815 ± 0.0039
Daream ⁷⁸	0.9968 ± 0.0009
Dahua-FacelImage ⁸⁰	0.9978 ± 0.0007
Skytop Gaia ⁸²	0.9630 ± 0.0023
CNN-3DMM estimation ⁸³	0.9235 ± 0.0129
Samtech Facequest ⁸⁴	0.9971 ± 0.0018
XYZ Robot ⁸⁷	0.9895 ± 0.0020
THU CV-AI Lab ⁸⁸	0.9973 ± 0.0008
dlib ⁹⁰	0.9938 ± 0.0027
Aureus ⁹¹	0.9920 ± 0.0030
YouTu Lab, Tencent ⁶³	0.9980 ± 0.0023
Orion Star ⁹²	0.9965 ± 0.0032
Yuntu WiseSight ⁹³	0.9943 ± 0.0045
PingAn AI Lab ⁸⁹	0.9980 ± 0.0016
Turing123 ⁹⁴	0.9940 ± 0.0040
Hisign ⁹⁵	0.9968 ± 0.0030
VisionLabs V2.0 ³⁸	0.9978 ± 0.0007
Deepmark ⁹⁶	0.9923 ± 0.0016
Force Infosystems ⁹⁷	0.9973 ± 0.0028
ReadSense ⁹⁸	0.9982 ± 0.0007
CM-CV&AR ⁹⁹	0.9963 ± 0.0039
sensingtech ¹⁰⁰	0.9970 ± 0.0008
Glasssix ¹⁰¹	0.9983 ± 0.0018
icarevision ¹⁰²	0.9977 ± 0.0030
Easen Electron ⁸¹	0.9983 ± 0.0006
yunshitu ¹⁰³	0.9975 ± 0.0006
RemarkFace ¹⁰⁴	0.9963 ± 0.0030

4.3 人脸识别的算法模型

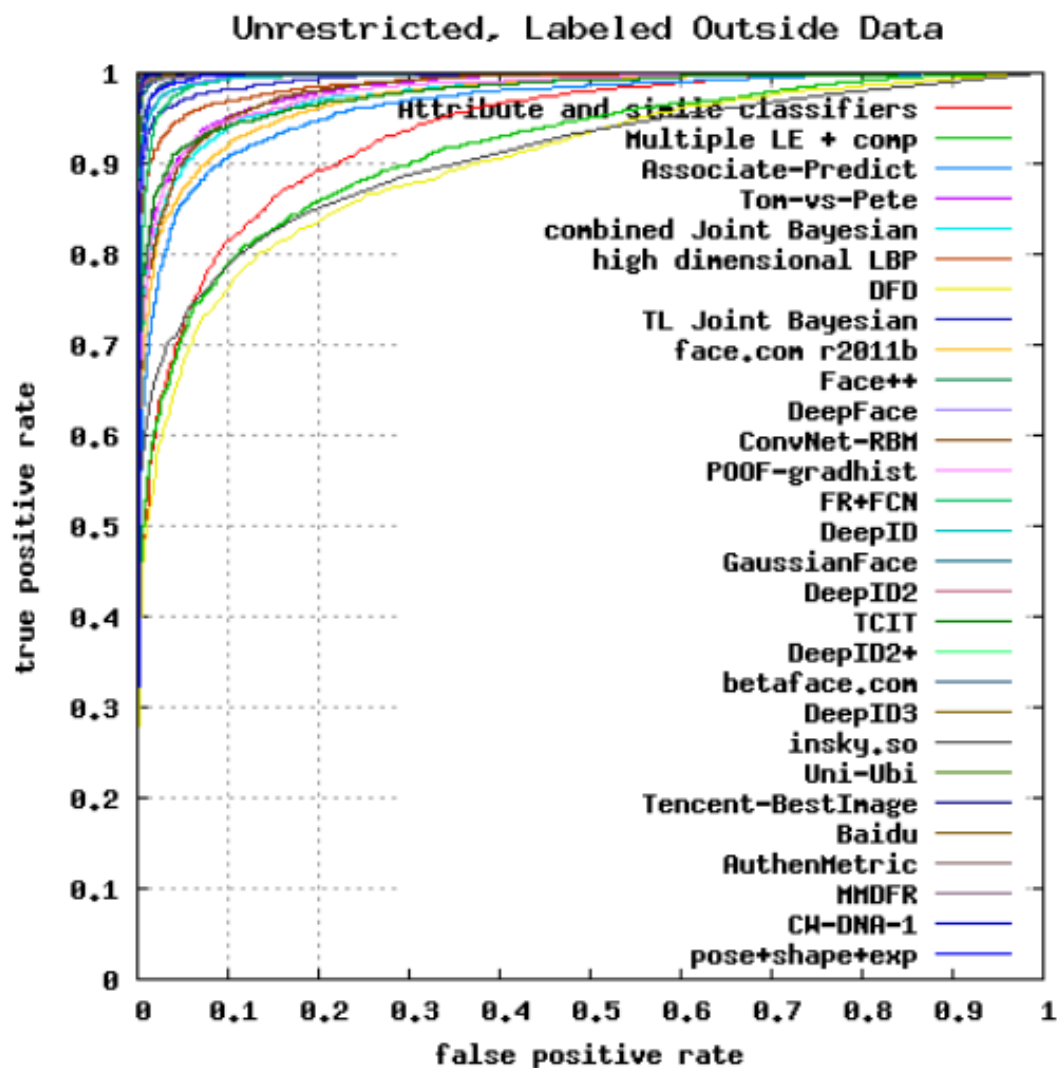


Figure 6: ROC curves averaged over 10 folds of View 2.

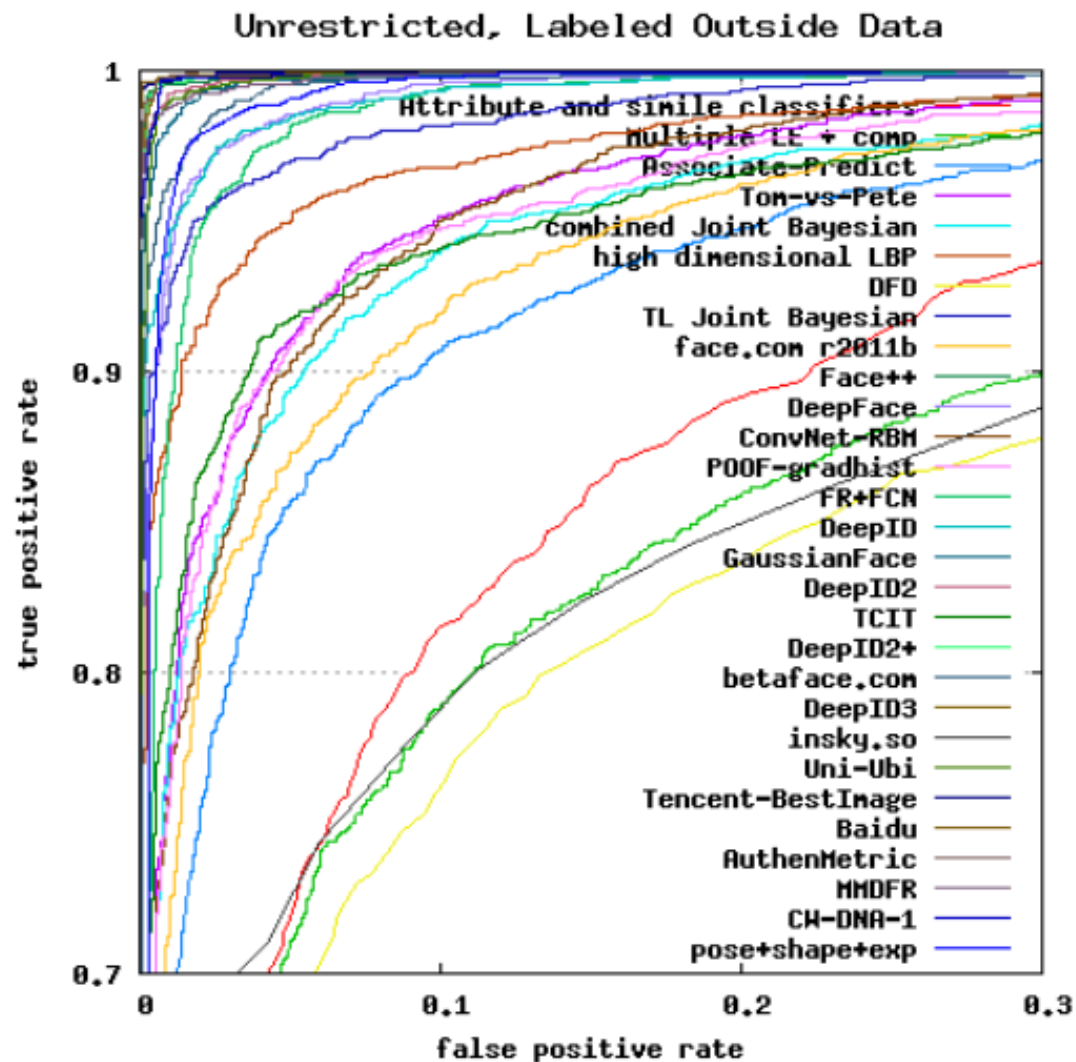
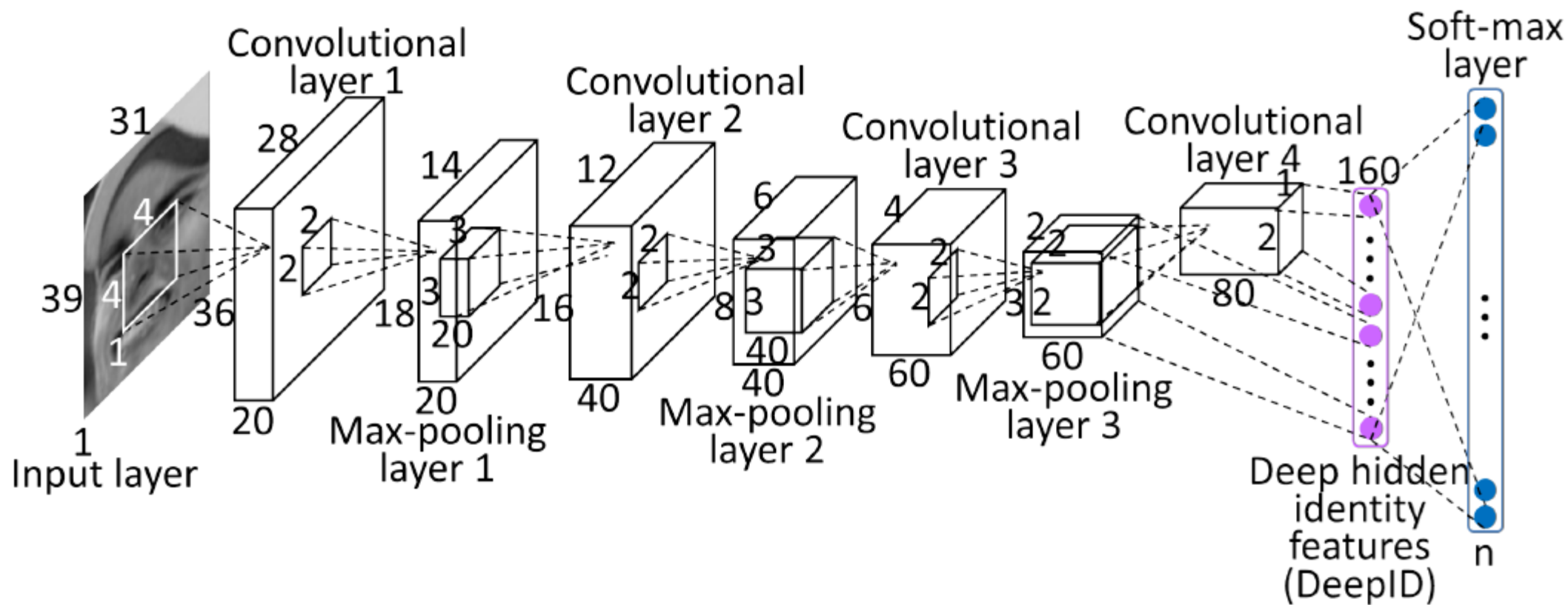


Figure 6: ROC curves averaged over 10 folds of View 2.

4.3 人脸识别的算法模型：DeepID



4.4 人脸识别的开源项目

人脸识别

数据库	描述	用途
WebFace	10k+人, 约500K张图片	非限制场景
FaceScrub	530人, 约100k张图片	非限制场景
YouTube Face	1,595个人 3,425段视频	非限制场景、视频
LFW	5k+人脸, 超过10K张图片	标准的人脸识别数据集
MultiPIE	337个人的不同姿态、表情、光照的人脸图像, 共750k+人脸图像	限制场景人脸识别
MegaFace	690k不同的人的1000k人脸图像	新的人脸识别评测集合
CAS-PEAL	1040个人的30k+张人脸图像 主要包含姿态、表情、光照变化	限制场景下人脸识别
Pubfig	200个人的58k+人脸图像	非限制场景下的人脸识别
MS-CELEB-1M	100K不同的人, 10M人脸图像	非限制场景下的人脸识别

人脸检测

数据库	描述	用途
FDDB	2845张图片中的5171张脸	标准人脸检测测评集
Caltech10k Web Faces	10k+人脸, 提供双眼和嘴巴的坐标位置	人脸点检测

人脸其他属性

数据库	描述	用途
IMDB-WIKI	包含: IMDB中20k+个名人的460k+张图片和维基百科62k+张图片, 共523k+张图片	名人年龄、性别
Adience	包含2k+个人的26k+张人脸图像	人脸性别, 人脸年龄段(8组)
CACD2000	2k名人160k张人脸图片	人脸年龄
CeleBrayA	200k张人脸图像40多种人脸属性	人脸属性识别

4.5 想一想

计算机视觉技术

还有哪些潜在应用？

Part

5

问答机器人

核心问题：技术框架、人工智能标记语言、开源项目



5 章节目标



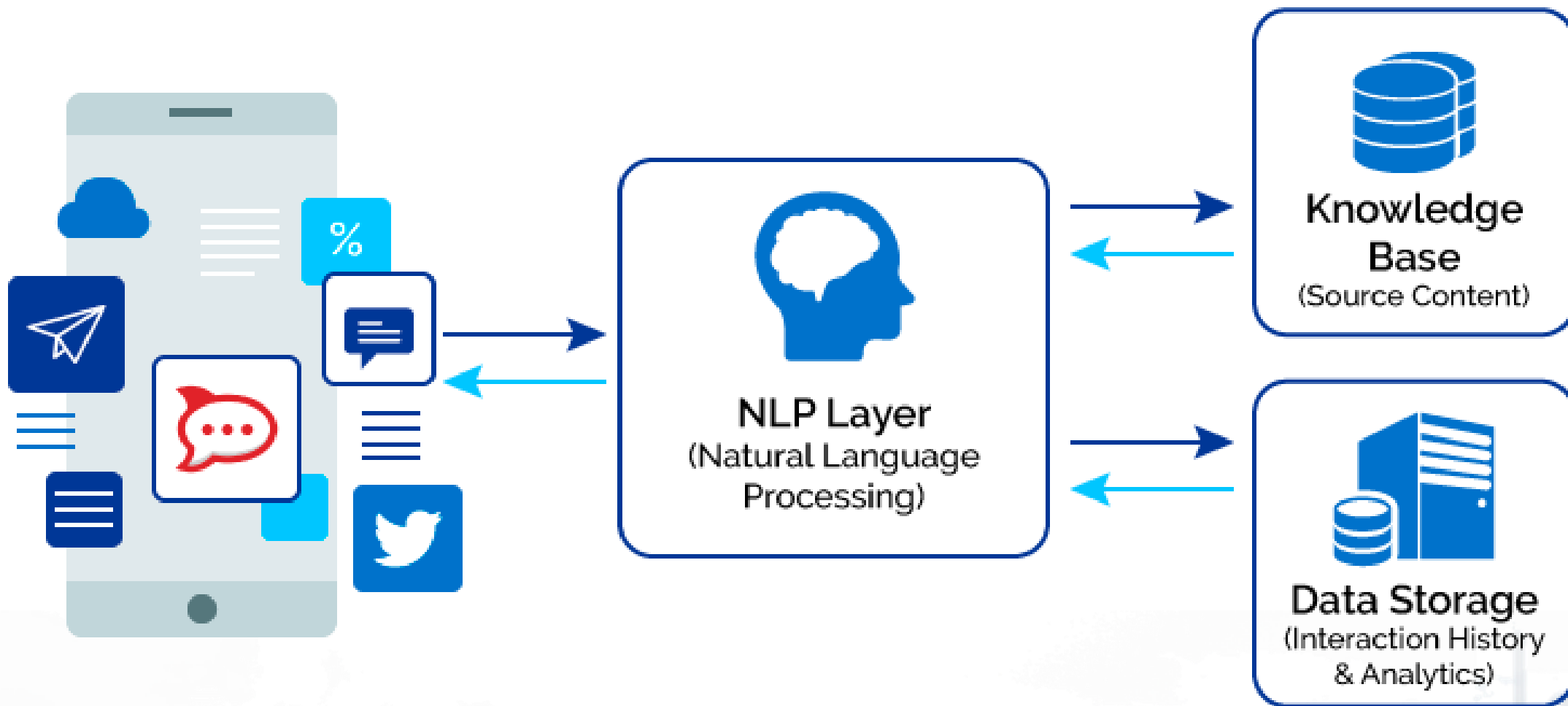
了解问答机器人的工作机制？

了解问答机器人的业务流程？

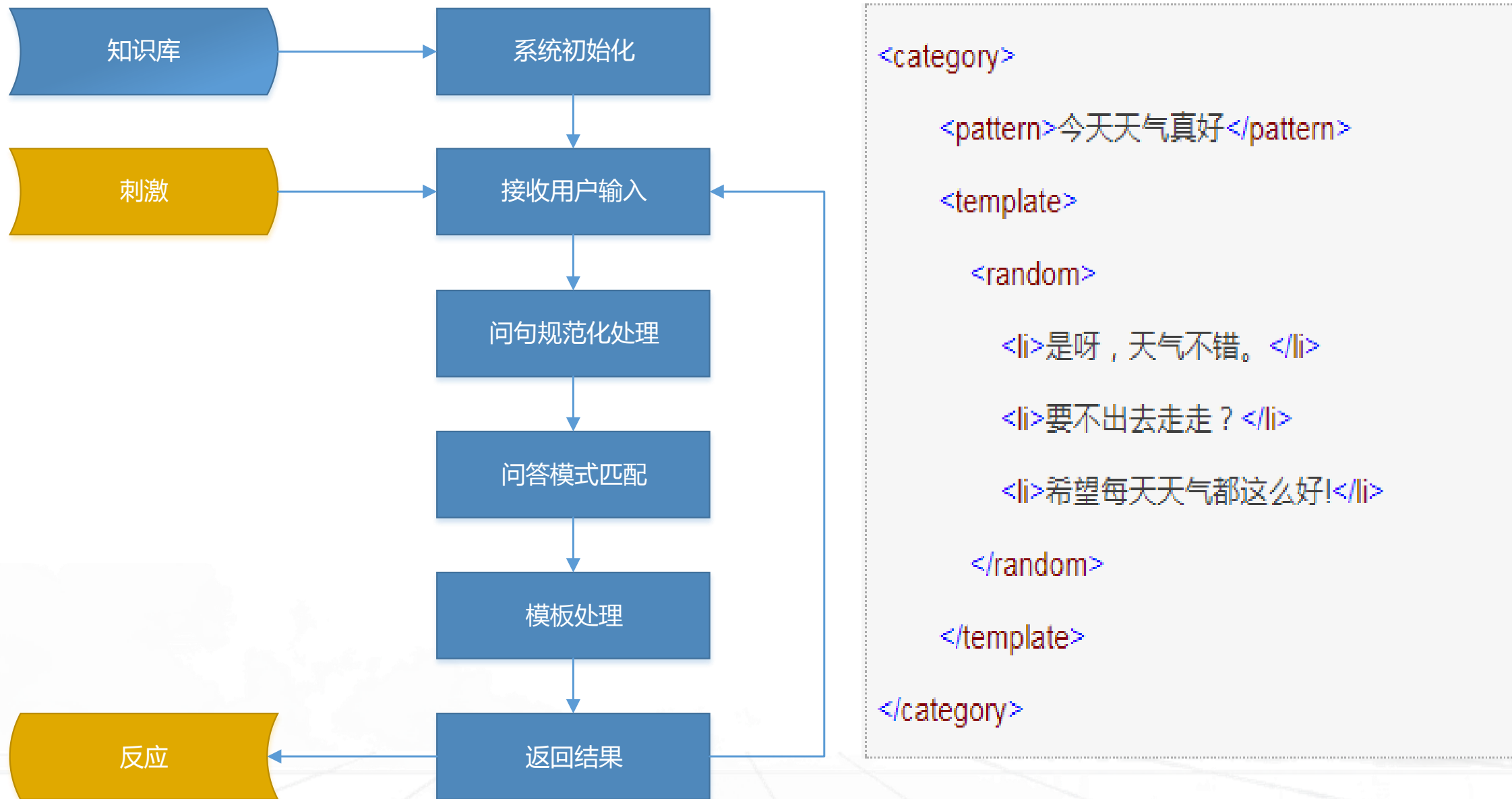
了解人工智能标记语言AIML？

熟悉问答机器人开源项目？

5.1 问答机器人的工作机制



5.2 人工智能标记语言 (AIML)



5.3 问答机器人的开源框架

← → ↻ 🏠 ⓘ www.alicebot.org/aiml.html



A.L.I.C.E.

ARTIFICIAL INTELLIGENCE FOUNDATION
PROMOTING THE DEVELOPMENT AND ADOPTION OF
ALICE AND AIML FREE SOFTWARE

A.L.I.C.E. Silver Edition

DAVE E.S.L. bot

C.L.A.U.D.I.O Personality Test

GET SITEPAL AVATARS

Free Live Chat with the award winning A. I. chat robot **A. L. I. C. E.**

Get Started

[Chat with A.L.I.C.E.](#)

[Chat with Fake Kirk](#)

[What is AIML?](#)
[Foundation Bot Directory](#)
[Bot Industry Survey](#)
[AIML Overview](#)

Software

[Downloads](#)
[Bot Hosting](#)
[AIML Sets](#)
[AIML 1.1 Specification](#)
[AIML 2.0 Working Draft](#)
[Documentation](#)
[Superbot - New!](#)

Links

[ESL](#)
[Books](#)
[Film and TV](#)
[Recent Press](#)
[Popular Culture](#)
[Web Ontologies](#)
[ALICE and A.I. History](#)
[Scholarly Research and Teaching](#)

Site Info

AIML: Artificial Intelligence Markup Language

AIML (Artificial Intelligence Markup Language) is an XML-compliant language that's easy to learn, and makes it possible for you to begin customizing an Alicebot or creating one from scratch within minutes.

The most important units of AIML are:

- `<aiml>`: the tag that begins and ends an AIML document
- `<category>`: the tag that marks a "unit of knowledge" in an Alicebot's knowledge base
- `<pattern>`: used to contain a simple pattern that matches what a user may say or type to an Alicebot
- `<template>`: contains the response to a user input

There are also 20 or so additional more tags often found in AIML files, and it's possible to create your own so-called "custom predicates". Right now, a beginner's guide to AIML can be found in the [AIML Primer](#).

The free **A.L.I.C.E. AIML** includes a knowledge base of approximately 41,000 categories. Here's an example of one of them:

```
<category>
  <pattern>WHAT ARE YOU</pattern>
  <template>
    <think><set name="topic">Me</set></think>
    I am the latest result in artificial intelligence,
    which can reproduce the capabilities of the human brain
    with greater speed and accuracy.
  </template>
</category>
```

(The opening and closing `<aiml>` tags are not shown here, because this is an excerpt from the middle of a document.)

Everything between `<category>` and `</category>` is -- you guessed it -- a category. A category can have one pattern and one template. (It can also contain a `<that>` tag, but we won't get into that here.)

The pattern shown will match *only* the exact phrase "what are you" (capitalization is ignored).

But it's possible that this category may be invoked by another category, using the `<srail>` tag (not shown) and the principle of **reductionism**.

5.4 想一想

问答机器人如何实现从闲聊到

精准问答的转变？

Part

6

知识图谱

核心问题：技术框架、典型应用、开源项目



6 章节目标



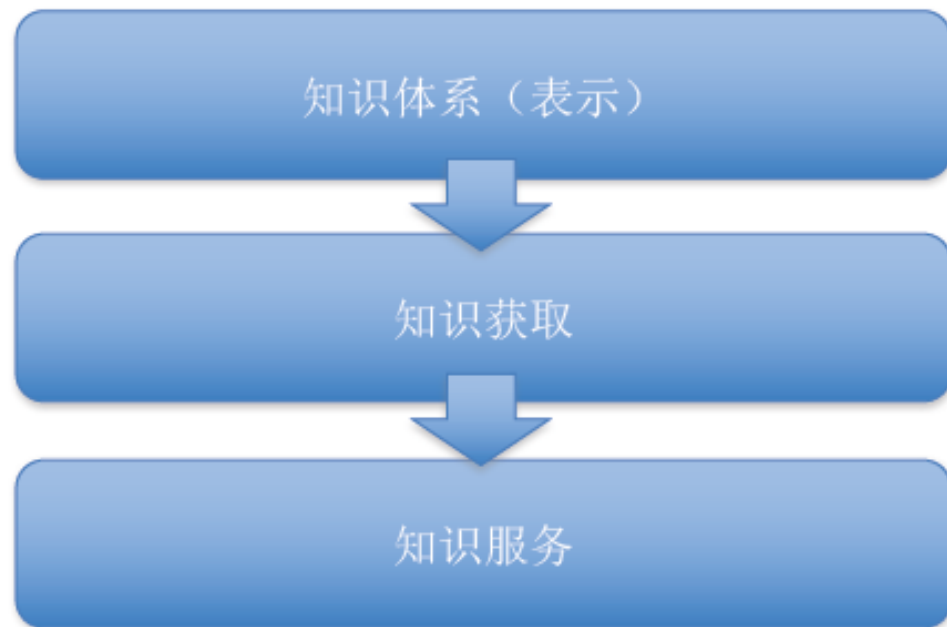
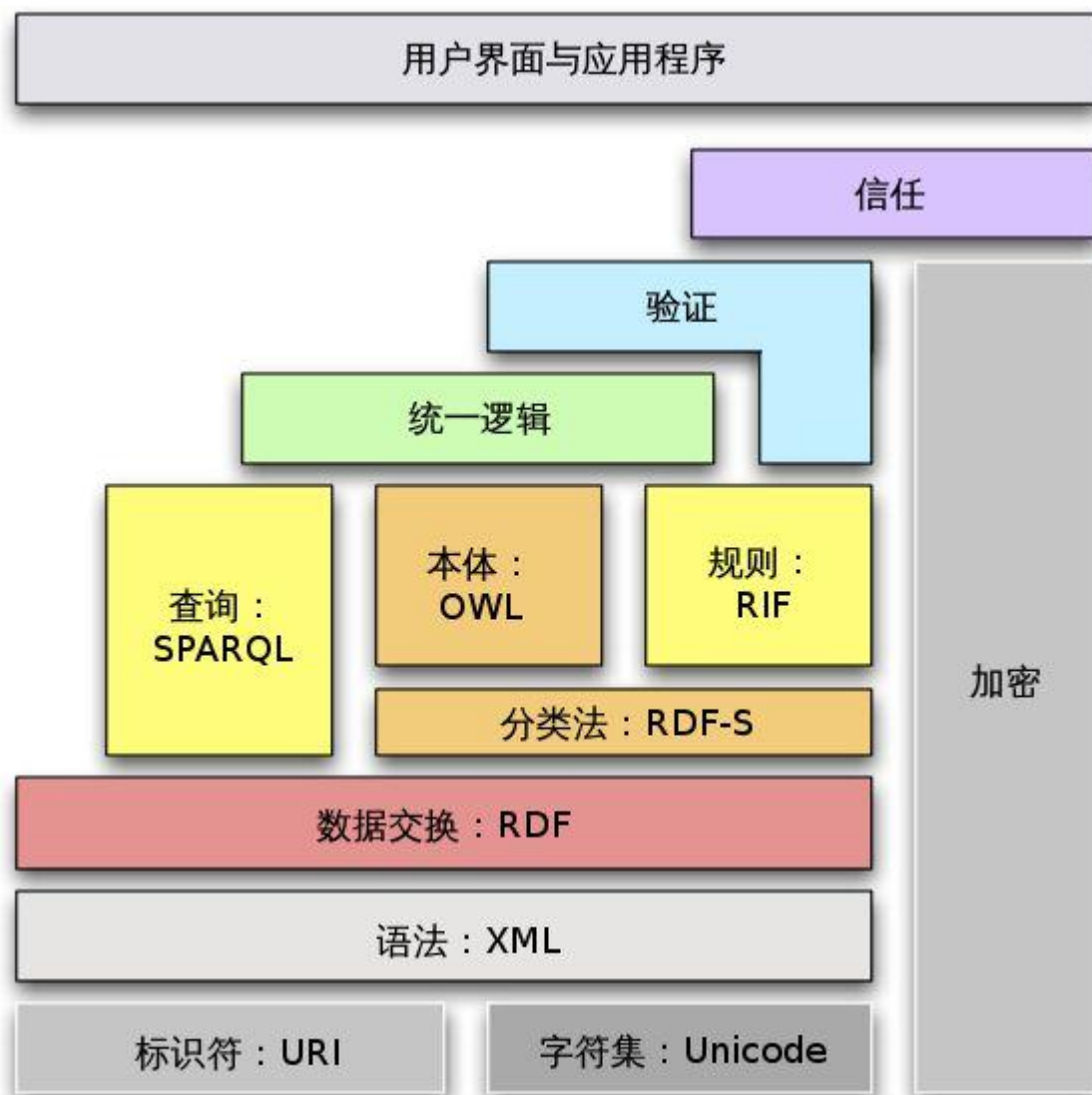
了解知识图谱的技术框架？

了解知识图谱的典型应用？

了解知识图谱发展的痛点？

熟悉知识图谱的开放资源？

6.1 知识图谱的技术框架

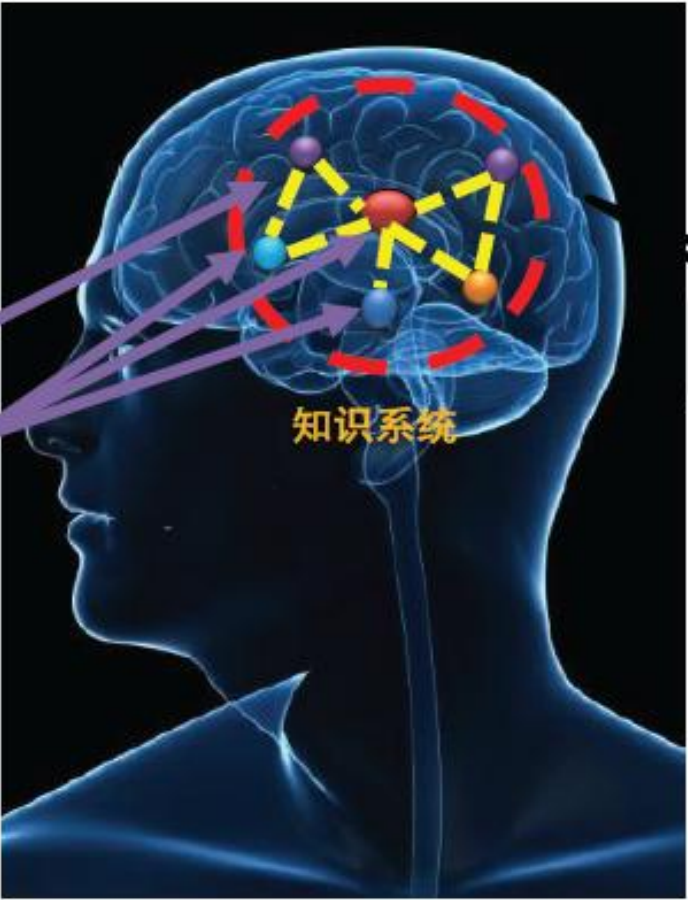


6.2 知识图谱的典型应用：自然语言理解



语义单元识别

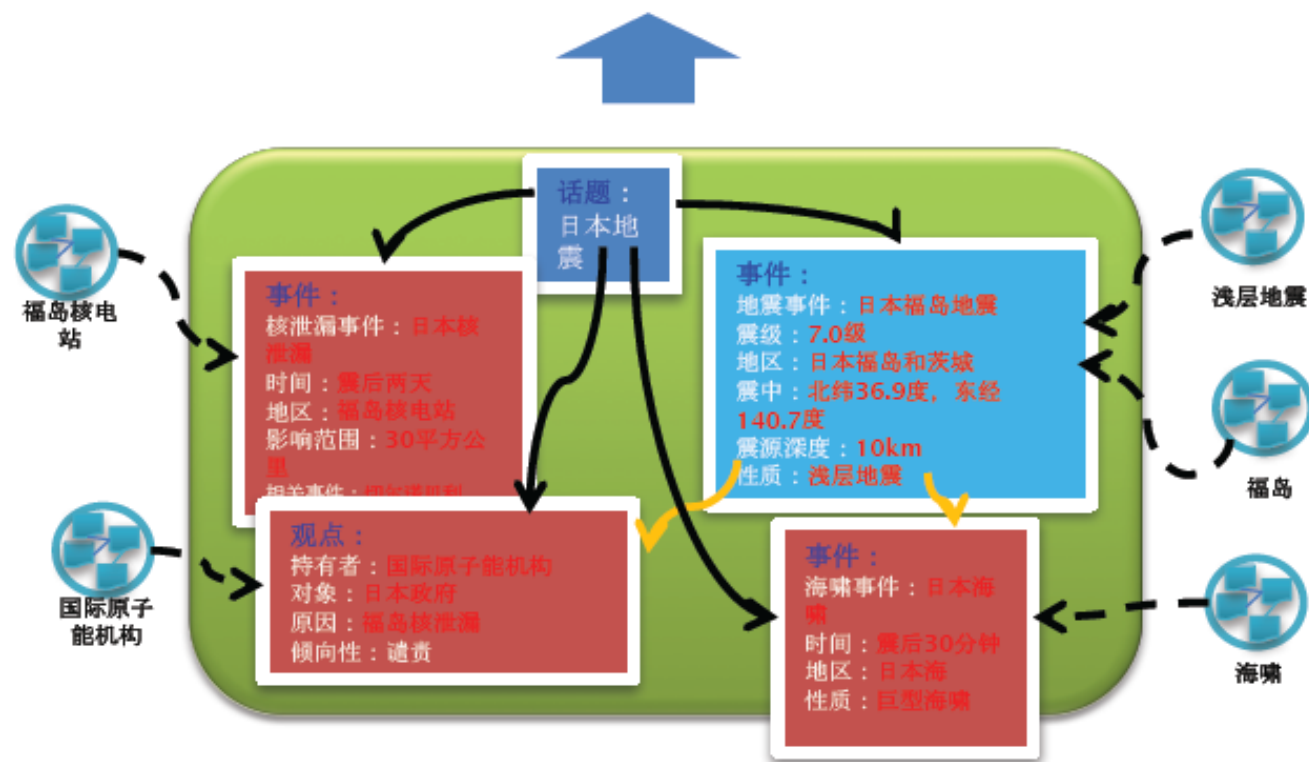
知识激活与关联



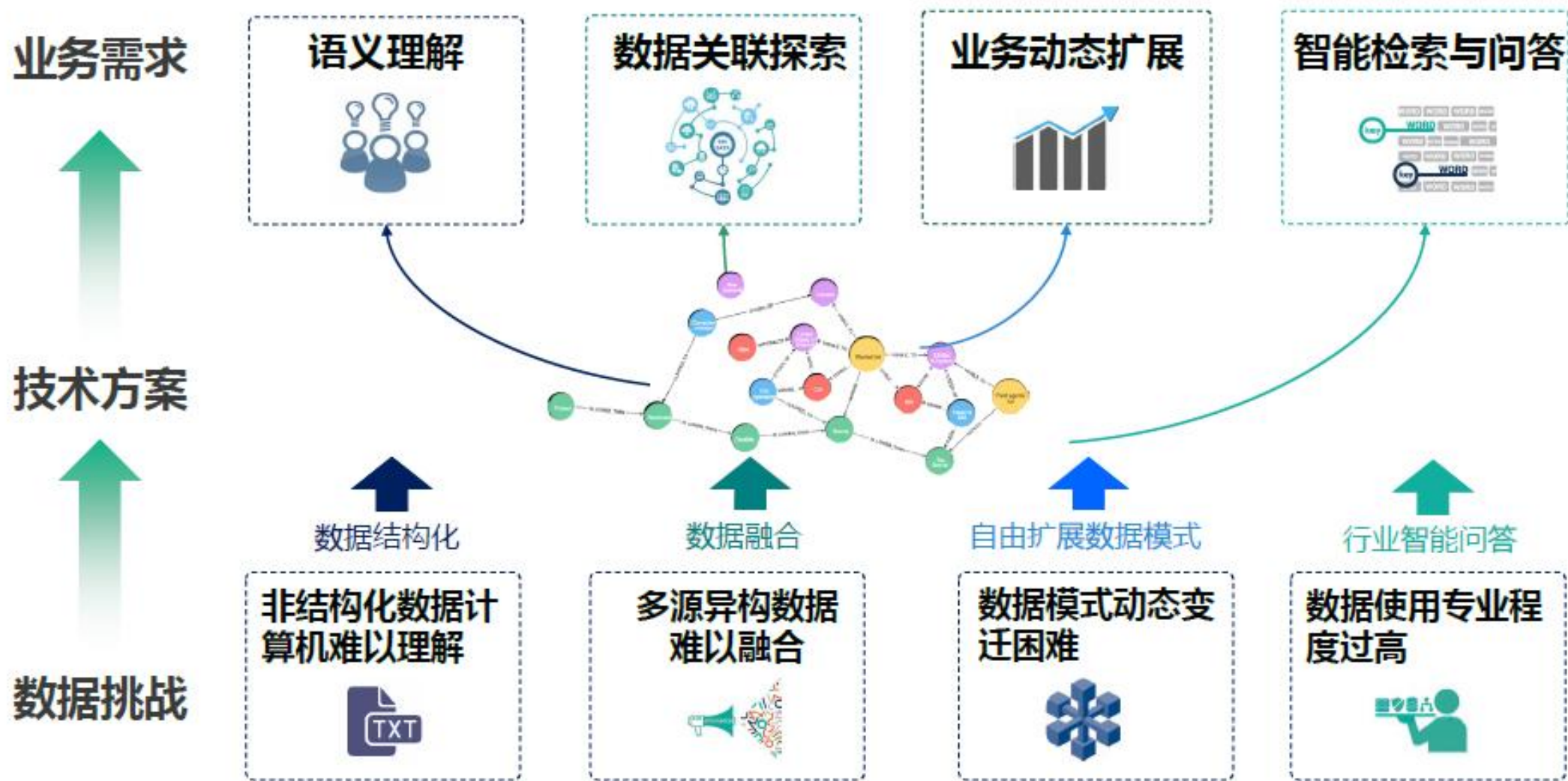
知识推理

6.2 知识图谱的典型应用：自然语言理解

2011年4月11日17点16分，日本东北部的福岛和茨城地区发生里氏7.0级强烈地震（震中北纬36.9度、东经140.7度，即福岛西南30公里左右的地方，震源深度10公里，属于浅层地震）。当局已经发布海啸预警。震后约30分钟后在日本海地区发生巨型海啸，同时造成福岛核电站出现核泄漏。震后第十天，国际原子能机构对于日本政府反应迟钝进行了谴责。



6.3 知识图谱的技术痛点



6.4 知识图谱的开源项目

■ 语言知识图谱

- [WordNet](#) : 155, 327个单词，同义词集117,597个，同义词集之间由22种关系连接

■ 事实性知识图谱

- [OpenCyc](#) : 23.9万个实体，1.5万个关系属性，209.3万个事实三元组
- [Freebase](#) : 4000多万实体，上万个属性关系，24多亿个事实三元组
- [DBpedia](#) : 400多万实体，48,293种属性关系，10亿个事实三元组
- [YAGO2](#) : 980万实体，超过100个属性关系，1亿多个事实三元组
- [百度百科](#) : 词条数1000万个
- [互动百科](#) : 800万词条，5万个分类，68亿文字

6.4 知识图谱的开源项目

■ 领域知识图谱

- **Kinships** : 描述人物之间的亲属关系, 104个实体, 26种关系, 10,800个三元组
- **UMLS** : 医学领域, 描述医学概念之间的联系, 135个实体, 49种关系, 6,800个三元组。
- **Cora** : 2,497个实体, 7种关系, 39,255个三元组

■ 机器自动构建的知识图谱

- **NELL** : 519万实体, 306种关系, 5亿候选三元组
- **Knowledge Vault**: 4500万实体, 4469种关系, 2.7亿三元组

建设通用知识图谱

还是领域知识图谱？

Part

7

自动驾驶

核心问题：能力模型、技术框架和算法模型



7 章节目标



了解自动驾驶的能力模型？

了解自动驾驶的生态格局？

了解自动驾驶的关键技术？

熟悉自动驾驶的开源项目？

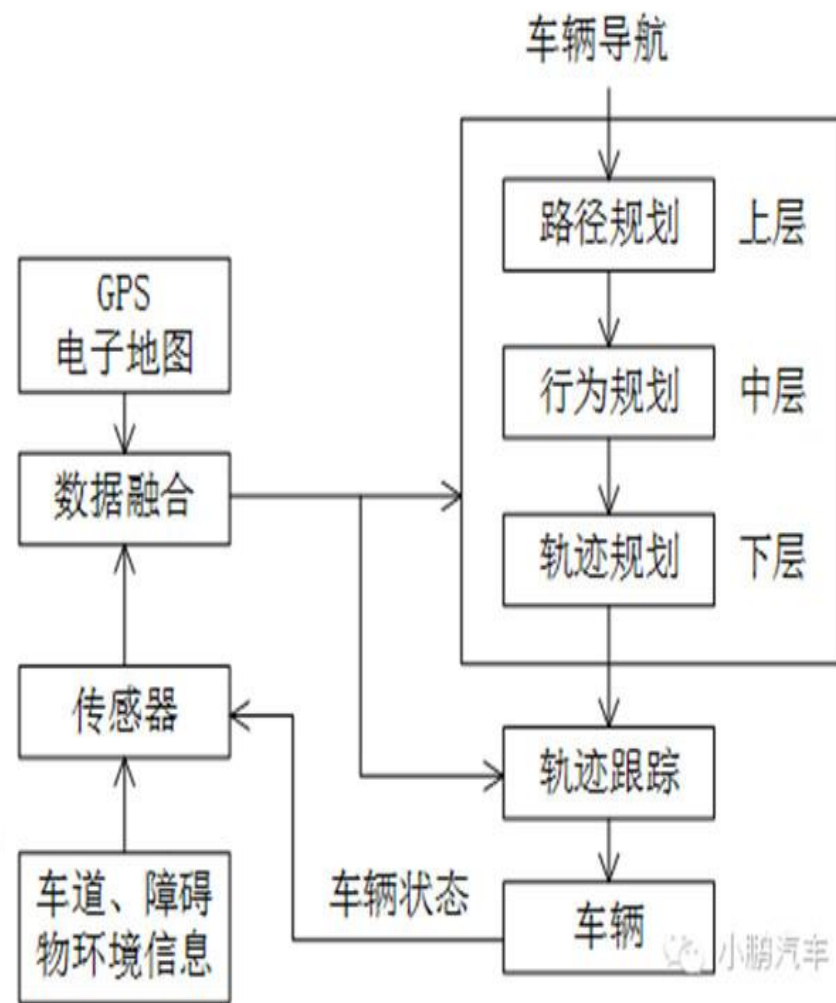
7.1 自动驾驶能力划分

自动驾驶分级		称呼 (SAE)	SAE定义	主体			
NHTSA	SAE			驾驶操作	周边监控	支援	系统作用域
0	0	无自动化	由人类驾驶者全权操作汽车，在行驶过程中可以得到警告和保护系统的辅助。	人类驾驶者	人类驾驶者	人类驾驶者	无
1	1	驾驶支援	通过驾驶环境对方向盘和加减速中的一项操作提供驾驶支援，其他的驾驶动作都由人类驾驶员进行操作。	人类驾驶者 系统			部分
2	2	部分自动化	通过驾驶环境对方向盘和加减速中的多项操作提供驾驶支援，其他的驾驶动作都由人类驾驶员进行操作。	系统			
3	3	有条件自动化	由无人驾驶系统完成所有的驾驶操作。根据系统请求，人类驾驶者提供适当的应答。		系统		
4	4	高度自动化	由无人驾驶系统完成所有的驾驶操作。根据系统请求，人类驾驶者不一定需要对所有的系统请求作出应答，限定道路和环境条件等。	系统	系统	系统	
	5	完全自动化	由无人驾驶系统完成所有的驾驶操作。人类驾驶者在可能的情况下接管。在所有的道路和环境条件下驾驶。				全域

7.2 中国自动驾驶生态格局



7.3 自动驾驶关键技术：环境感知、路径规划、车辆控制、车载操作系统



7.4 自动驾驶开源项目：车辆平台、硬件平台、软件平台、云端数据

Apollo技术框架



7.5 想一想

自动驾驶真的能够

实现吗？



Thanks