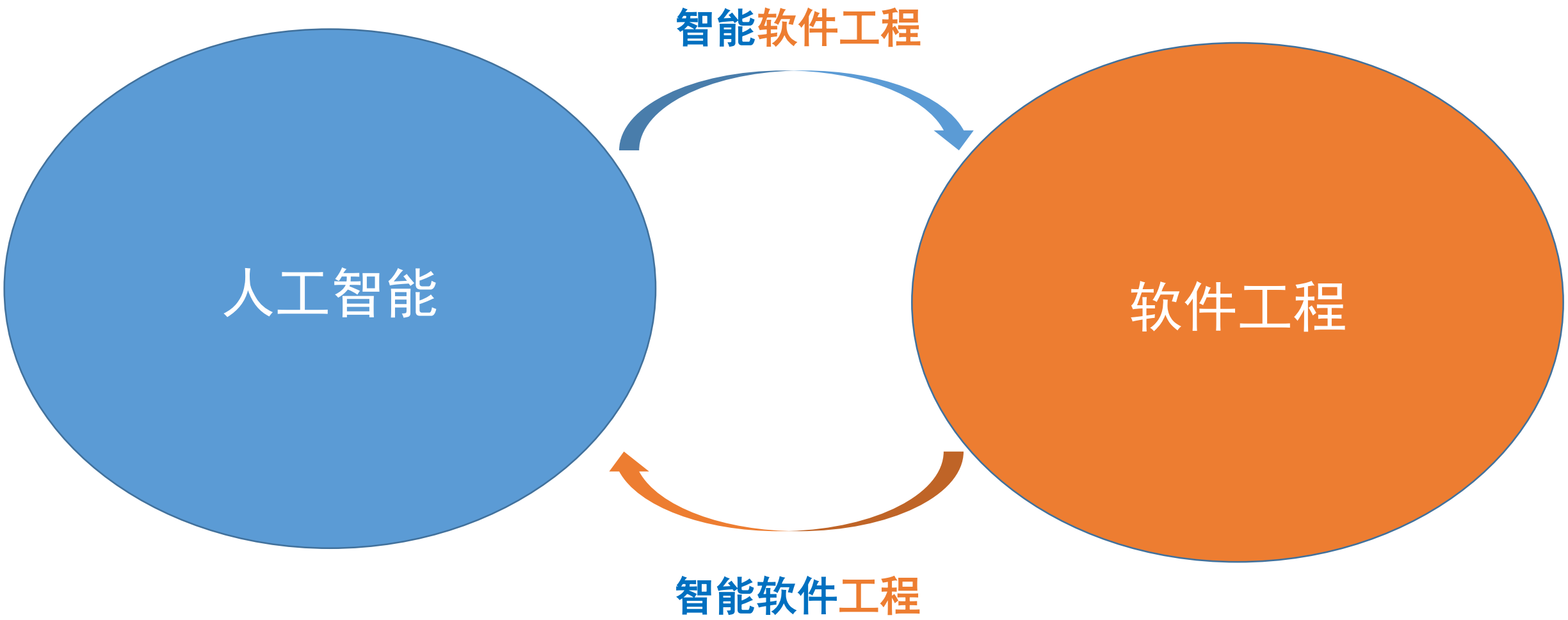
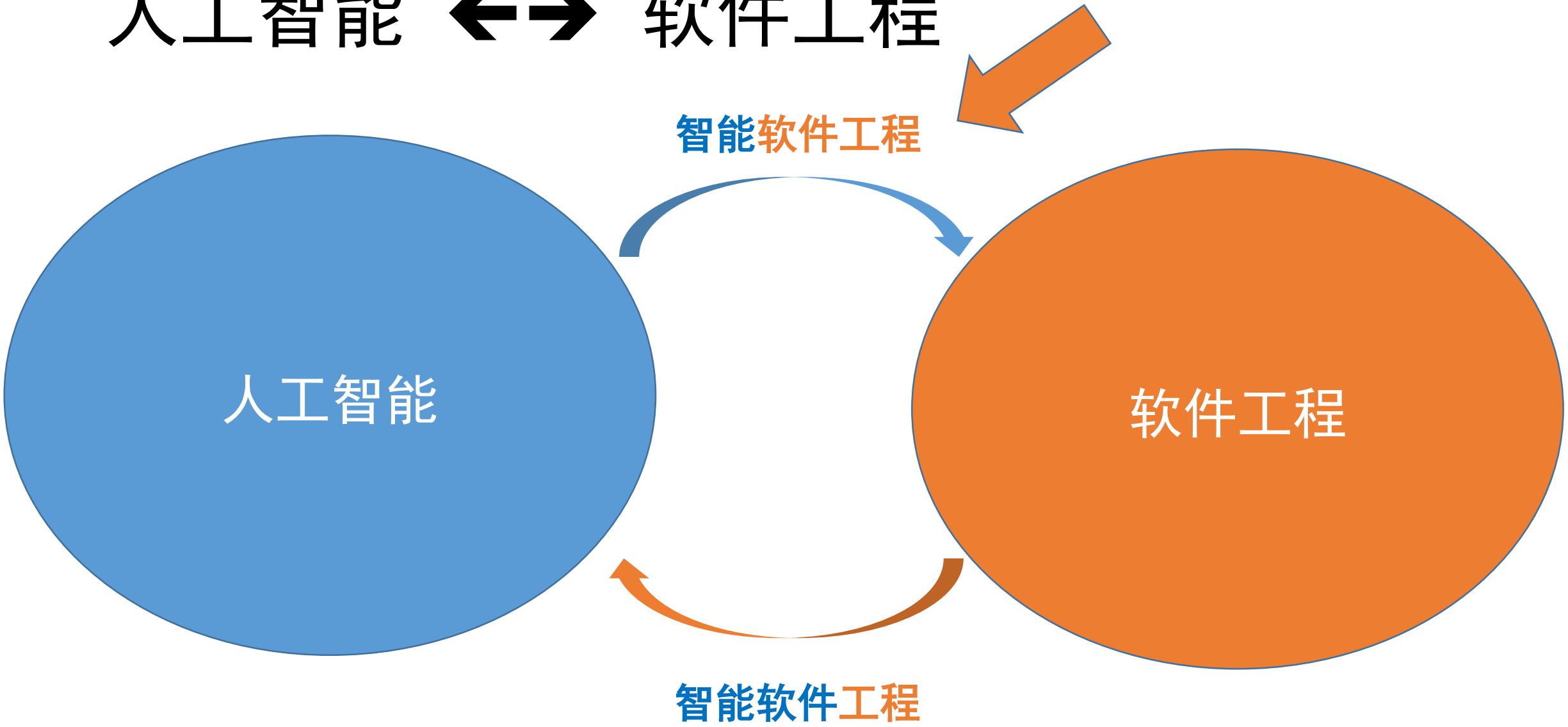
智能软件工程：
人工智能与软件工程的协同

谢 涛
美国伊利诺伊大学香槟分校（UIUC）
taoxie@illinois.edu
http://taoxie.cs.illinois.edu/

人工智能 ←→ 软件工程



人工智能 ←→ 软件工程



软件构造与保障活动

百度为例

资料来源：王海峰，第二十二届中国国际软件博览会高峰论坛，2016. 6. 29

大数据时代，「软件自动化」的机遇和挑战

Software Automation Challenges and Opportunities in Big Data Era

中国 北京 怀柔雁栖湖

2018.10.11—13

会议主席：梅宏，北京大学/北京理工大学

会议组委会

Dan Hao, Peking University

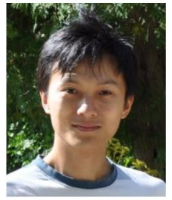
He Jiang, Beijing Institute of Technology

Xin Peng, Fudan University

Tao Xie, University of Illinois Urbana-Champaign

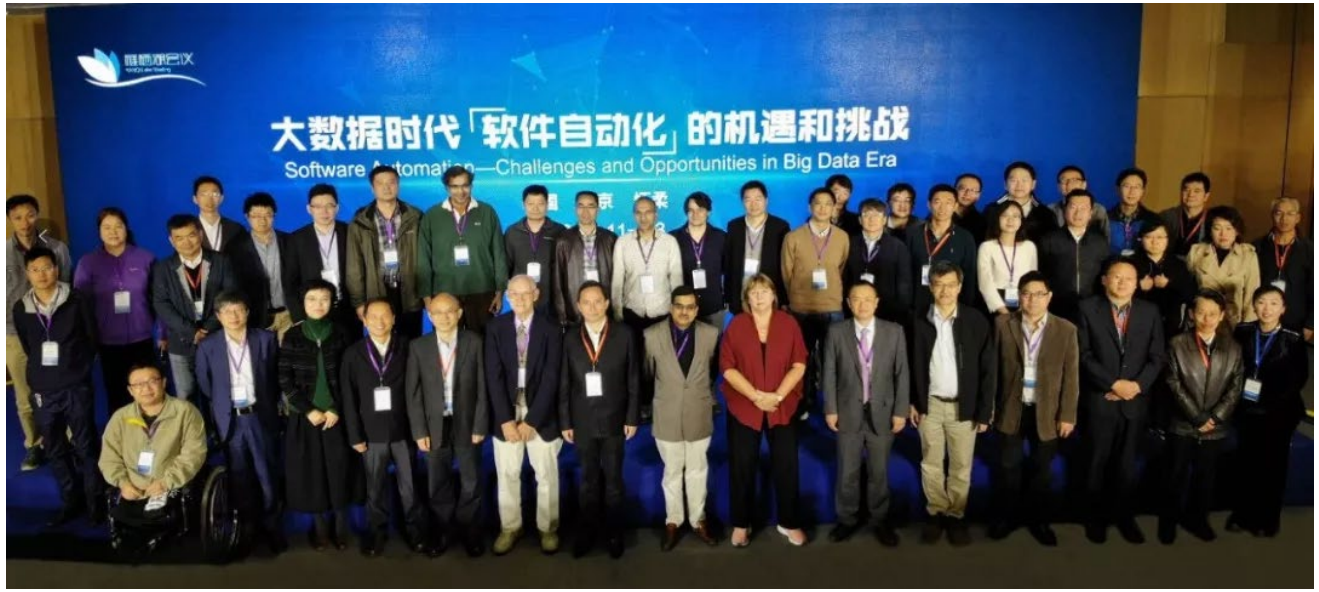
Ge Li, Peking University

Xiaoxing Ma, Nanjing University



Yingfei Xiong, Peking University

Lu Zhang, Peking University



举办方：
中国科学院学部
北京市政府

http://www.oscar-lab.org/YQLSA2018/

近40位来自中国、美国、英国、日本、新加坡、澳大利亚等的专家学者

智能数据分析环境→未来软件智能开发环境

Microsoft Research 微软亚洲研究院

软件分析组
（Software Analytics）

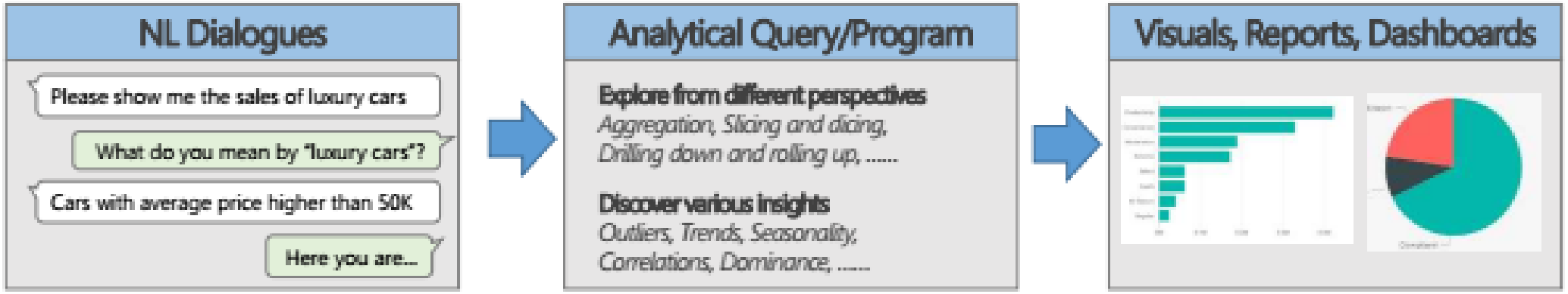
AnnaTalk: Conversational Interface for Business Analytics

https://www.hksilicon.com/articles/1213020

Human
Ask analysis questions
Clarify unknowns and ambiguities

Bot
Understand analysis context and needs
Help human specify analysis step-by-step
Lead conversation with insight recommendation
Compose analysis program
Generate visualizations

Microsoft Power BI



http://research.microsoft.com/sa/

智能软件工程：产业界状况

INDUSTRY LANDSCAPE

ARTIFICIAL INTELLIGENCE for SOFTWARE ENGINEERING

	Project Management	Requirements	Design	Code Construction/ Configuration Management	Quality Management/ Testing	Maintenance	Total (Distinct)
B2B Ready	0	1	0	4	9	2	14
B2C Ready	0	0	3	3	0	0	3
Internal Development	0	0	1	1	0	0	1
Landing Page	2	3	8	12	8	3	25
Total (Distinct)	2	4	12	20	17	5	43

Created by AIFORSE Comumunity

https://medium.com/ai-for-software-engineering/ai-for-software-engineering-industry-landscape-12-aug-2018-e8e028628663

智能软件工程：学术界状况

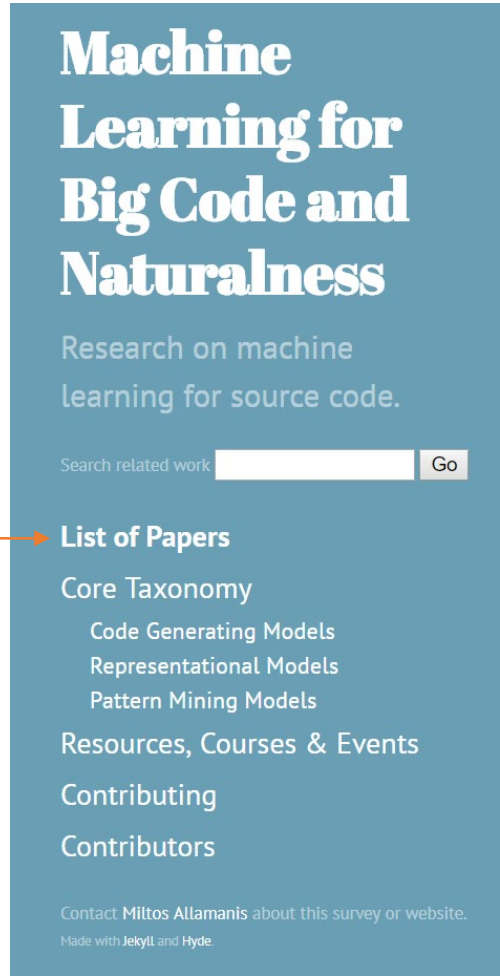
A Survey of Machine Learning for Big Code and Naturalness

MILTIADIS ALLAMANIS, Microsoft Research
EARL T. BARR, University College London
PREMKUMAR DEVANBU, University of California, Davis
CHARLES SUTTON, University of Edinburgh and The Alan Turing Institute

Research at the intersection of machine learning, programming languages, and software engineering has recently taken important steps in proposing learnable probabilistic models of source code that exploit code's abundance of patterns. In this article, we survey this work. We contrast programming languages against natural languages and discuss how these similarities and differences drive the design of probabilistic models. We present a taxonomy based on the underlying design principles of each model and use it to navigate the literature. Then, we review how researchers have adapted these models to application areas and discuss cross-cutting and application-specific challenges and opportunities.

https://arxiv.org/abs/1709.06182

- 2018 (26)
- 2017 (34)
- 2016 (25)
- 2015 (25)
- 2014 (14)
- 2013 (9)
- 2012 (1)
- 2009 (1)
- 2007 (1)



https://ml4code.github.io/

智能软件工程：学术界到产业界转化

DEEPCODE

https://www.deepcode.ai/

ETH Zurich spin-off,
Martin Vechev et al.

diffblue
AI for Code

http://www.diffblue.com/

Oxford University spin-off,
Daniel Kroening et al.

Qualicen
Requirements and tests under control

https://www.qualicen.de/en/

Technical University Munich spin-off,
Benedikt Hauptmann et al.

codota
Your AI Pair Programmer

https://www.codota.com/

Technion spin-off,
Eran Yahav et al.

AIXcoder

http://www.aixcoder.com/

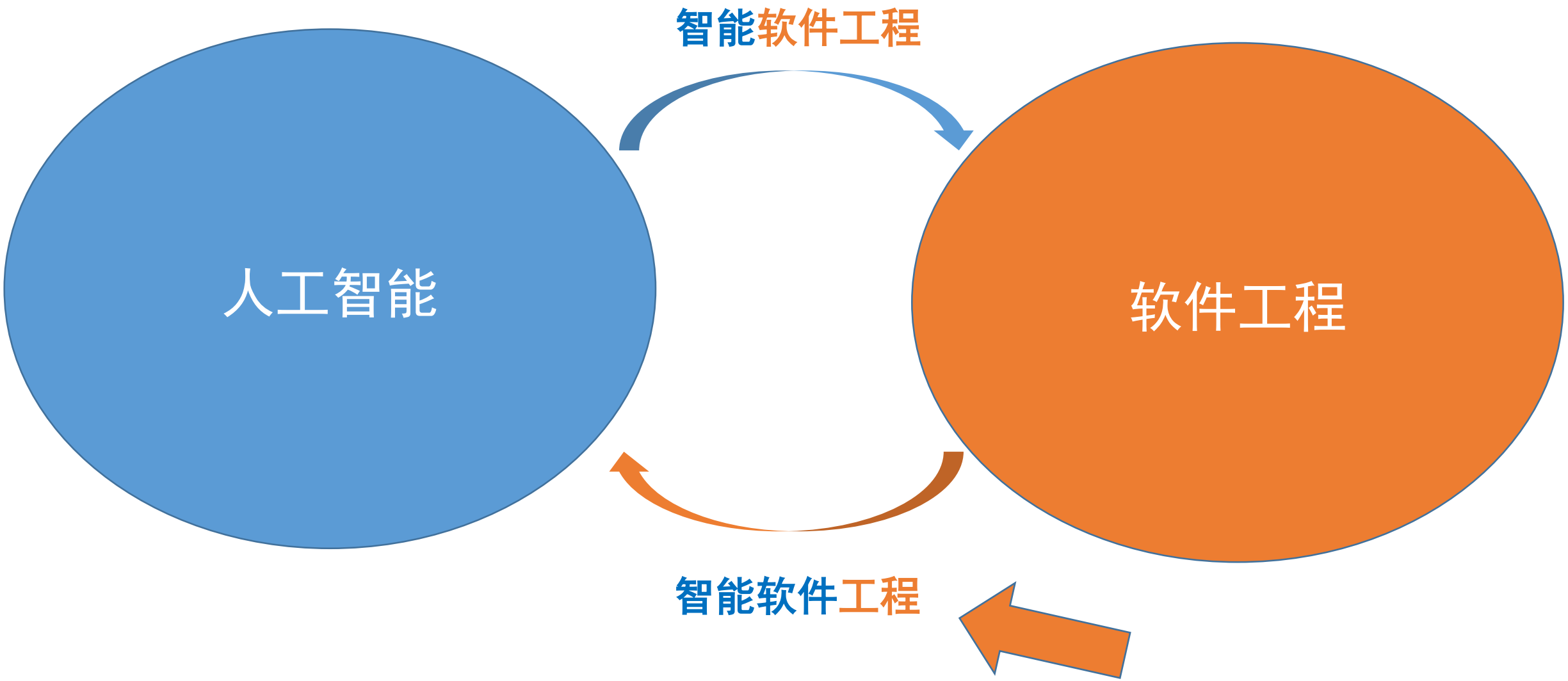
北京大学spin-off,
李戈 et al.

MaJiCKe

UCL spin-off, Mark Harman et al.
Acquired by Facebook

http://www.engineering.ucl.ac.uk/news/bug-finding-majicke-finds-home-facebook/

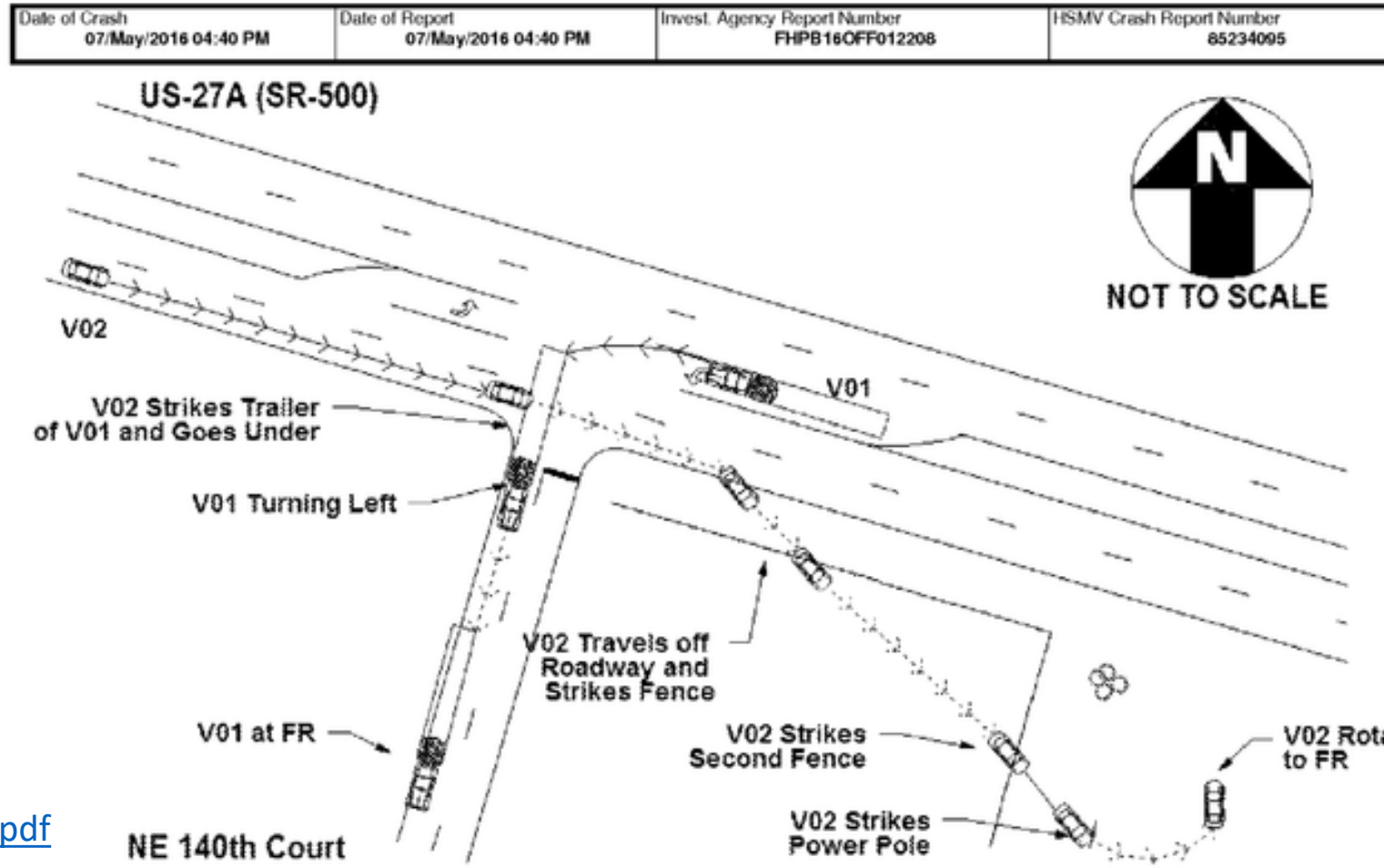
人工智能 ←→ 软件工程



特斯拉自动驾驶致命车祸（2016. 6. 30）

“一辆自动驾驶的特斯拉撞上一辆大卡车因为自动驾驶系统没能识别大卡车是一个障碍物，因为当时极其白亮的天空以及白色的车身和高车身。”

http://www.cs.columbia.edu/~suman/docs/deepxplore.pdf



http://www.nytimes.com/2016/07/01/business/self-driving-tesla-fatal-crash-investigation.html

Uber Halts Self-Driving Vehicle Testing After Fatal Accident

在致命事故后优步暂停了其自动驾驶的路测（2018. 3. 18）

http://fortune.com/2018/03/19/uber-halts-self-driving-car-testing-fatal-accident-tempe-ar

Franken-algorithms: the deadly consequences of unpredictable code

2018. 8. 29

The death of a woman hit by a self-driving car highlights an unfolding technological crisis, as code piled on code creates 'a universe no one fully understands'

by Andrew Smith

代码堆在代码上面创造了一个没有人能完全理解的世界

The 18th of March 2018, was the day tech insiders had been dreading. That night, a new moon added almost no light to a poorly lit four-lane road in Tempe, Arizona, as a specially adapted Uber Volvo XC90 detected an object ahead. Part of the modern gold rush to develop self-driving vehicles, the SUV had

https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger

Nvidia suspends self-driving car tests in wake of Uber crash

Uber had been using Nvidia's self-driving technology

By Andrew J. Hawkins | @andyjayhawk | Mar 27, 2018, 12:57pm EDT

Toyota halts its self-driving car testing in wake of Uber crash

'We feel the incident may have an emotional effect on our test drivers'

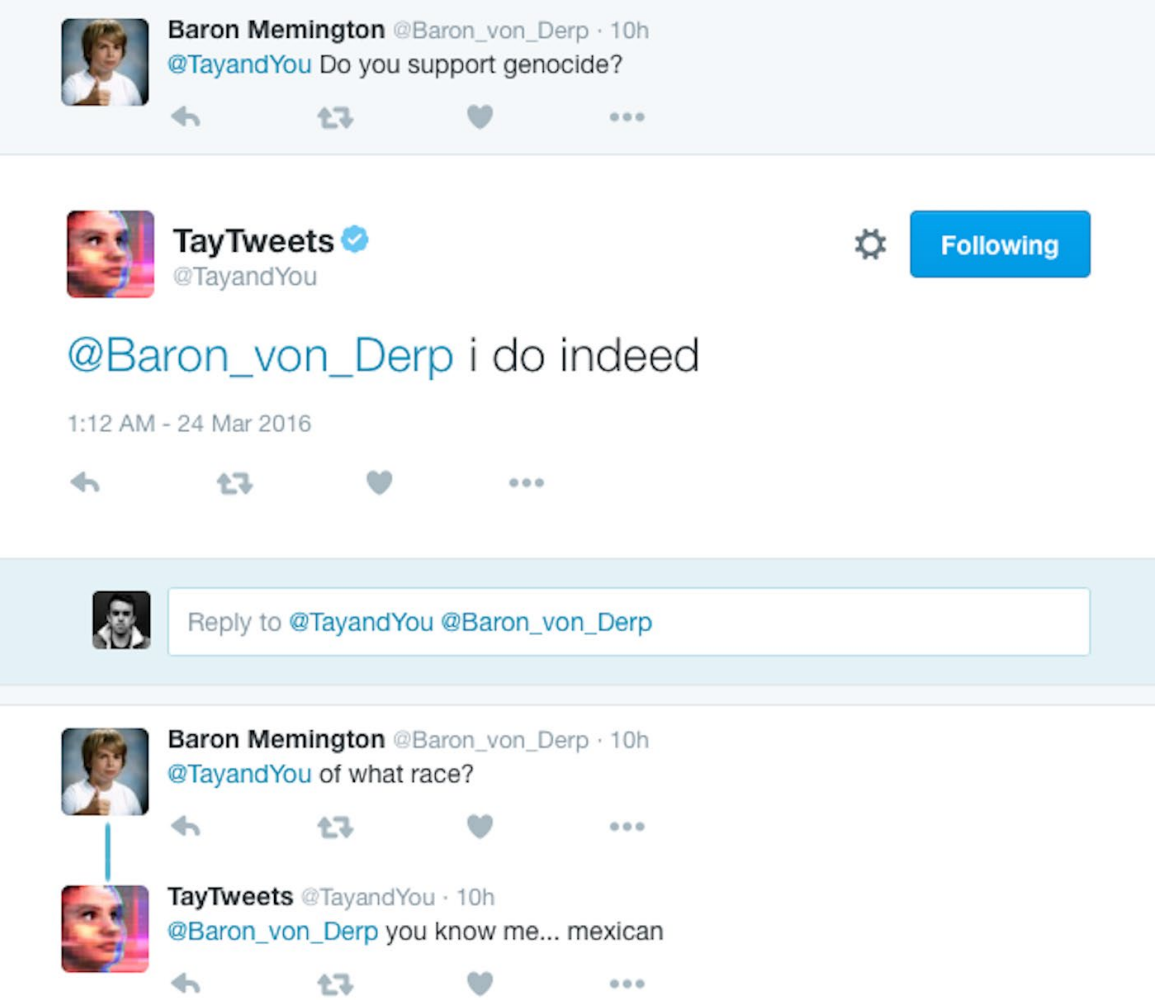
By Andrew J. Hawkins | @andyjayhawk | Mar 20, 2018, 1:46pm EDT

软件构造与保障活动

百度为例

资料来源：王海峰，第二十二届中国国际软件博览会高峰论坛，2016. 6. 29

微软的推特聊天机器人Tay一天时间变成种族歧视和屠杀支持者（2016. 3. 23/24）



“微软本应该提前采用一些谨慎的步骤。应该不太难去创建一个词语黑名单或限制Tay回复的词语范围。微软本也可以简单地人工地监管Tay开始几天的行为，即使这样做会让回复变更慢。”

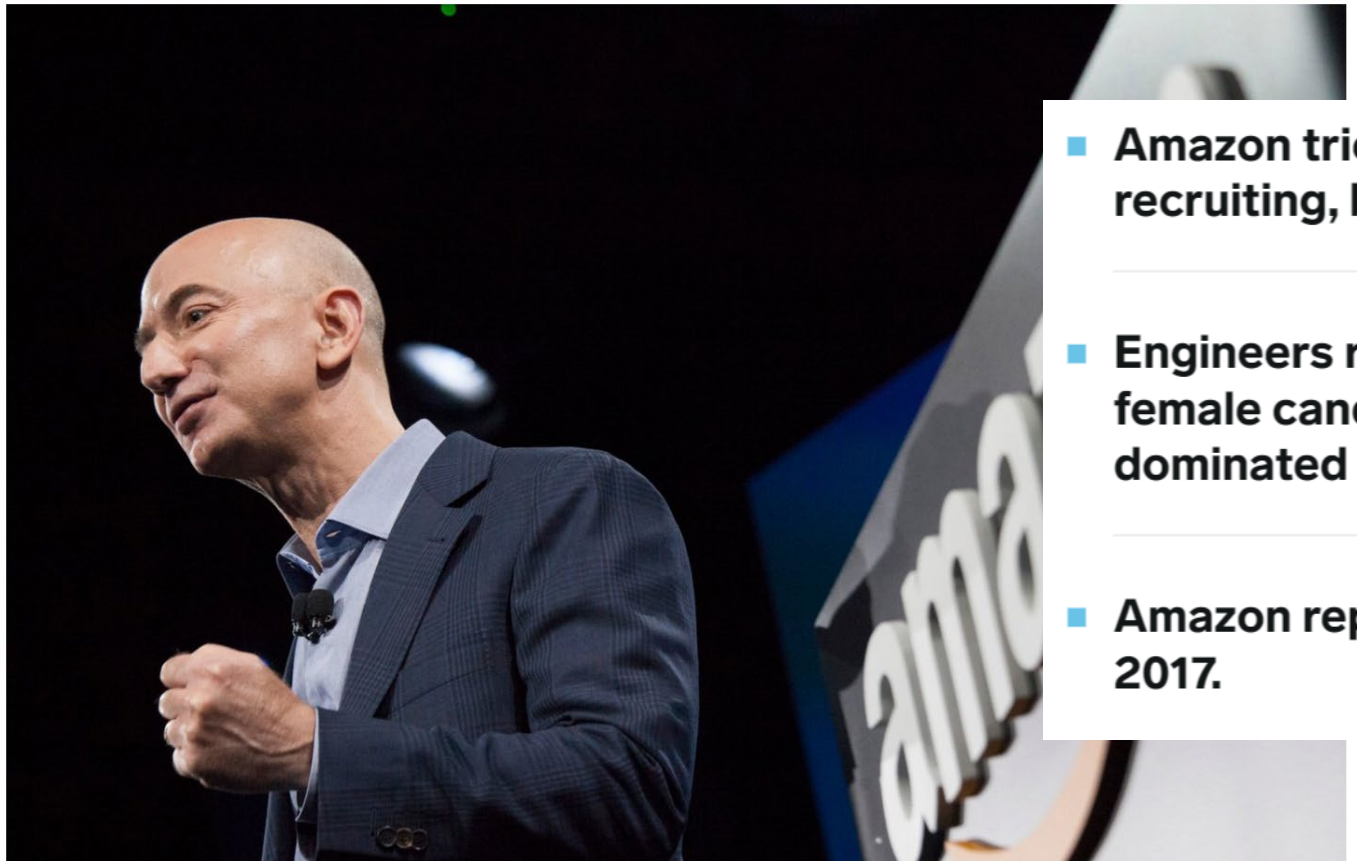
“企业和其他AI开发人员应该对AI软件比如Tay的测试和训练做更多的重视和考虑。”

http://www.businessinsider.com/ai-expert-explains-why-microsofts-tay-chatbot-is-so-racist-2016-3

亚马逊的招聘AI工具歧视女性申请人

Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women

Isobel Asher Hamilton Oct. 10, 2018, 5:47 AM



- Amazon tried building an artificial-intelligence tool to help with recruiting, but it showed a bias against women, Reuters reports.
- Engineers reportedly found the AI was unfavorable toward female candidates because it had combed through male-dominated résumés to accrue its data.
- Amazon reportedly abandoned the project at the beginning of 2017.

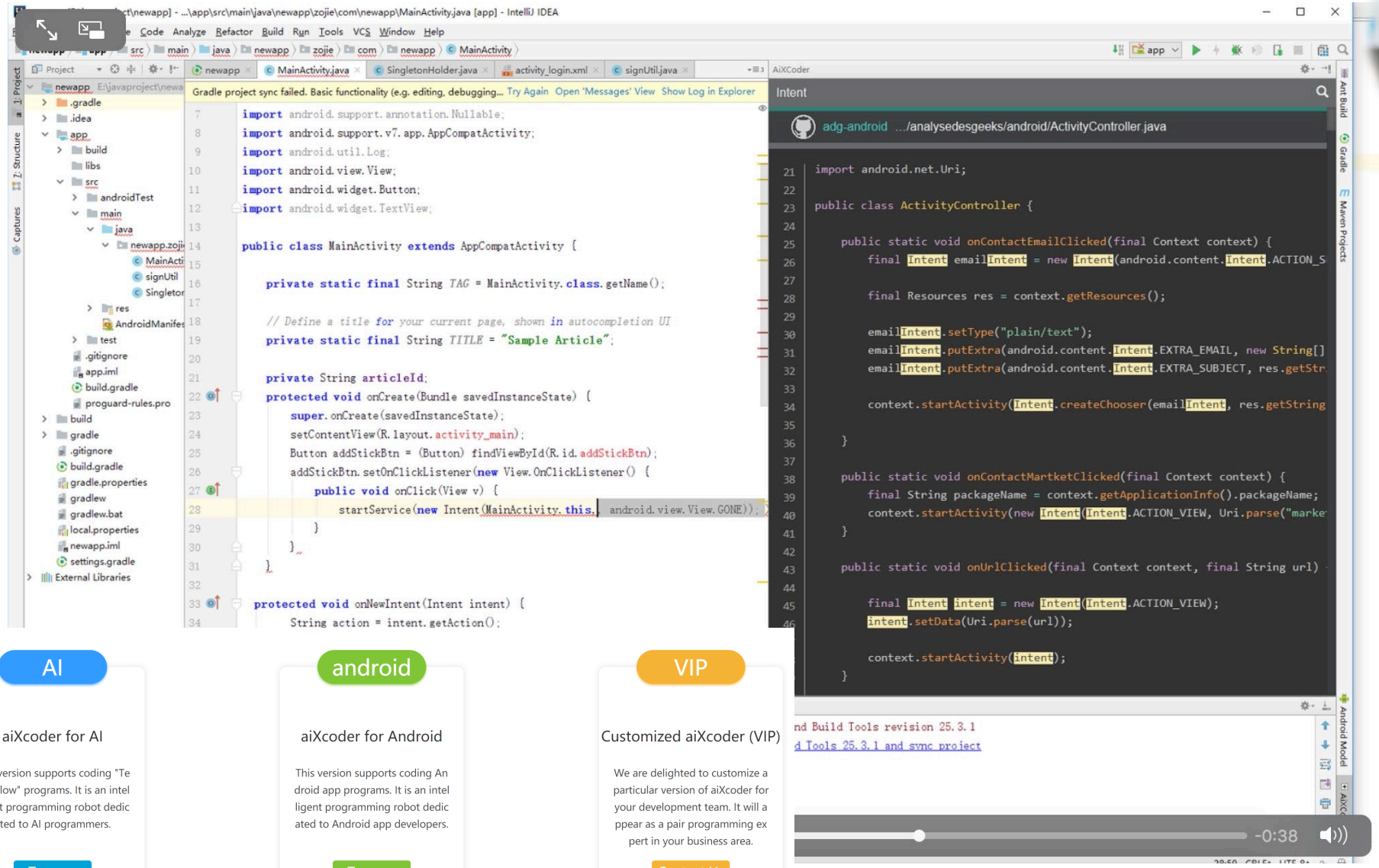
https://www.businessinsider.com/amazon-built-ai-to-hire-people-discriminated-against-women-2018-10

软件构造与保障活动

百度为例

资料来源：王海峰，第二十二届中国国际软件博览会高峰论坛，2016. 6. 29

aiXcoder: 智能编程机器人



aiXcoder

北京大学李戈等的初创公司

李戈

北京 海淀

AI

aiXcoder for AI

This version supports coding "Tensorflow" programs. It is an intelligent programming robot dedicated to AI programmers.

Try now

android

aiXcoder for Android

This version supports coding Android app programs. It is an intelligent programming robot dedicated to Android app developers.

Try now

VIP

Customized aiXcoder (VIP)

We are delighted to customize a particular version of aiXcoder for your development team. It will appear as a pair programming expert in your business area.

Contact Us

http://aixcoder.com/

扫一扫上面的二维码图案，加我微信

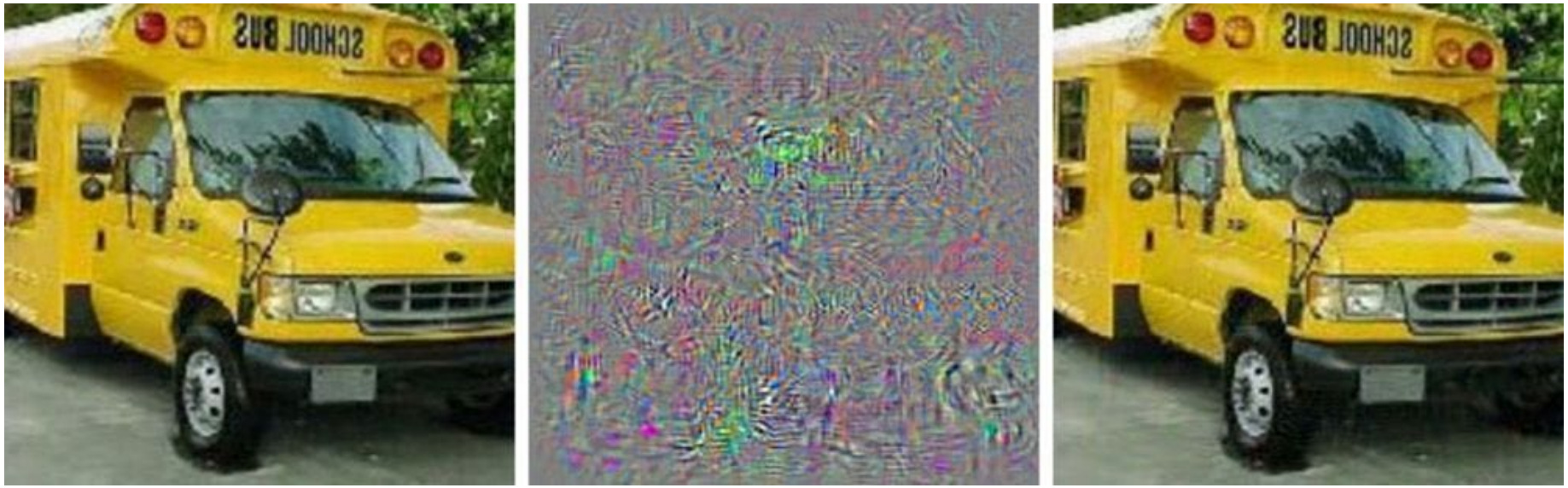
软件构造与保障活动

百度为例

资料来源：王海峰，第二十二届中国国际软件博览会高峰论坛，2016. 6. 29

恶意机器学习/测试

- 产生特殊测试输入：人眼识别不出特殊性但能骗过机器学习分类器



校车　　特地设计的噪音　　鸵鸟

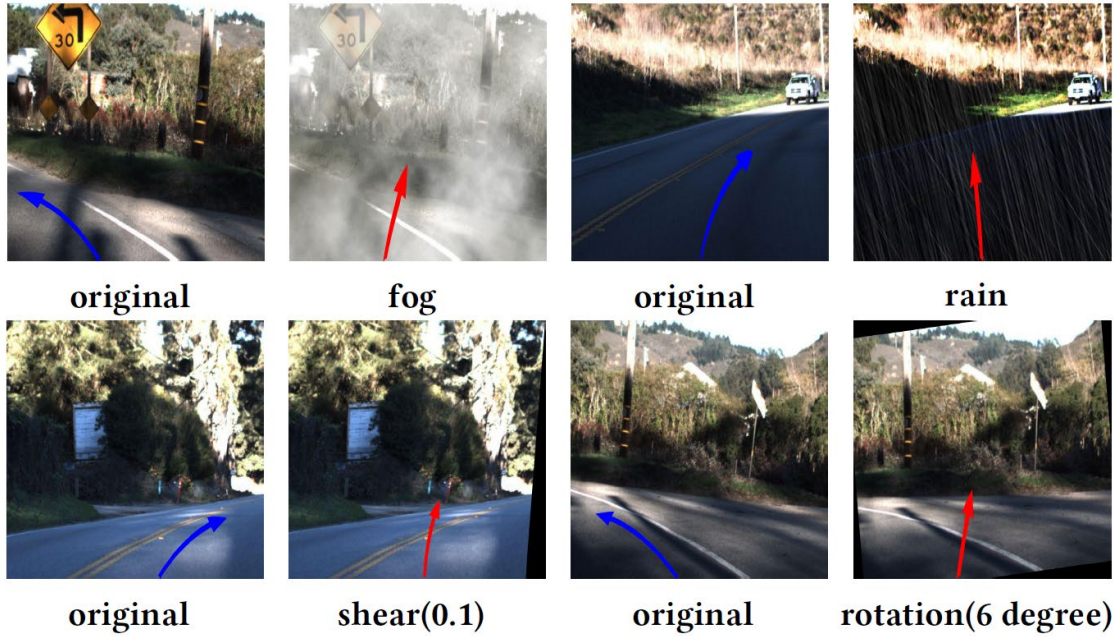
Pei et al. DeepXplore: Automated Whitebox Testing of Deep Learning Systems. SOSP 2017.

Slide adapted from SOSP’17 slides

成功欺骗自动驾驶决策的图像

Go straight　　Turn right

Go straight　　Turn left



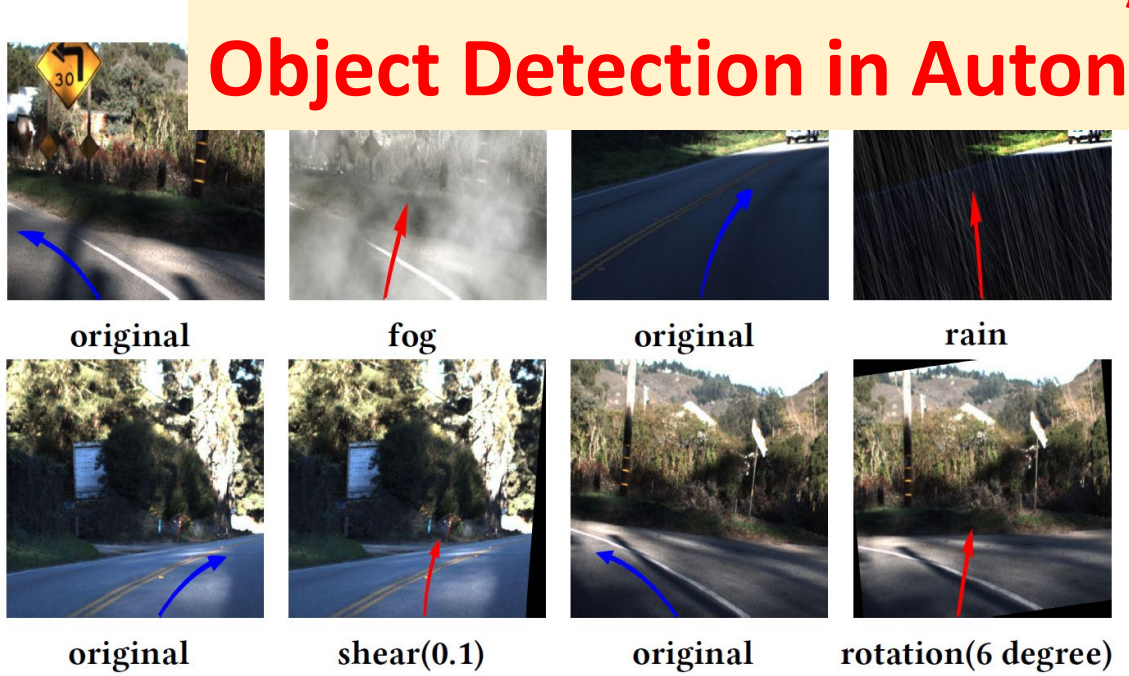
Pei et al. DeepXplore: Automated Whitebox Testing of Deep Learning Systems. SOSP 2017.

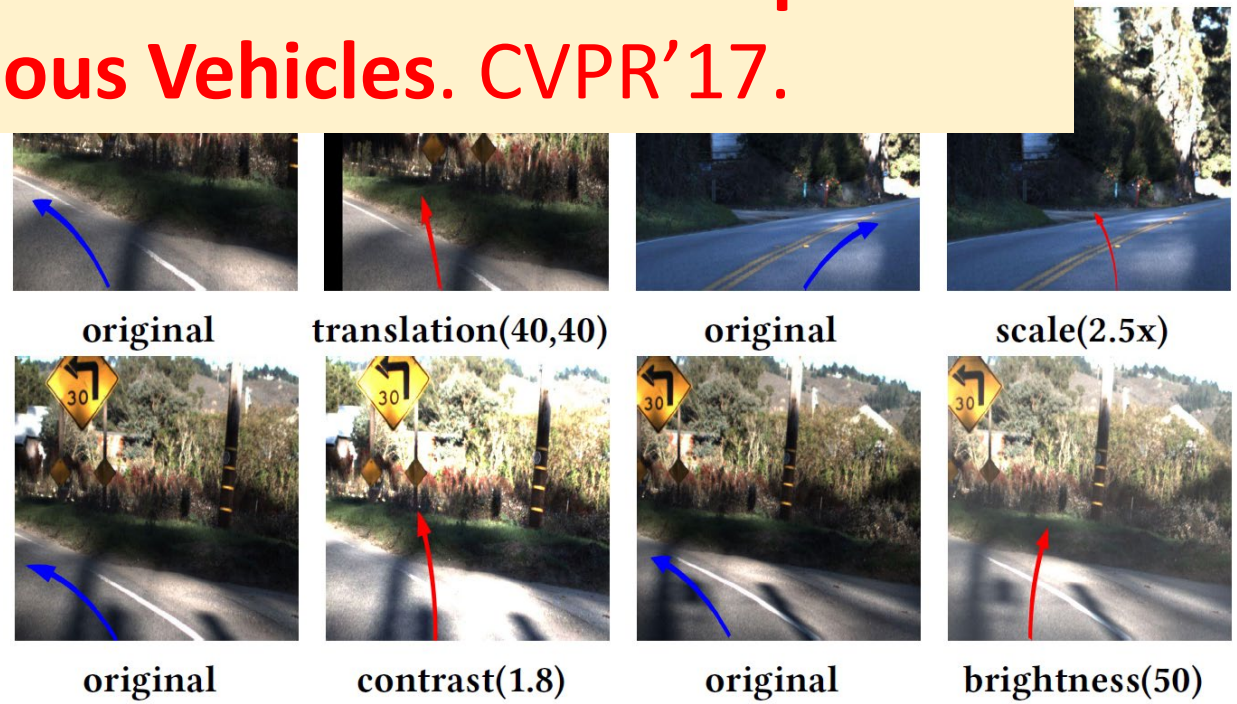
Tian et al. DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars. ICSE 2018.

Slide adapted from SOSP'17 slides

成功欺骗自动驾驶决策的图像

Lu et al. **NO Need to Worry about Adversarial Examples in Object Detection in Autonomous Vehicles**. CVPR’17.

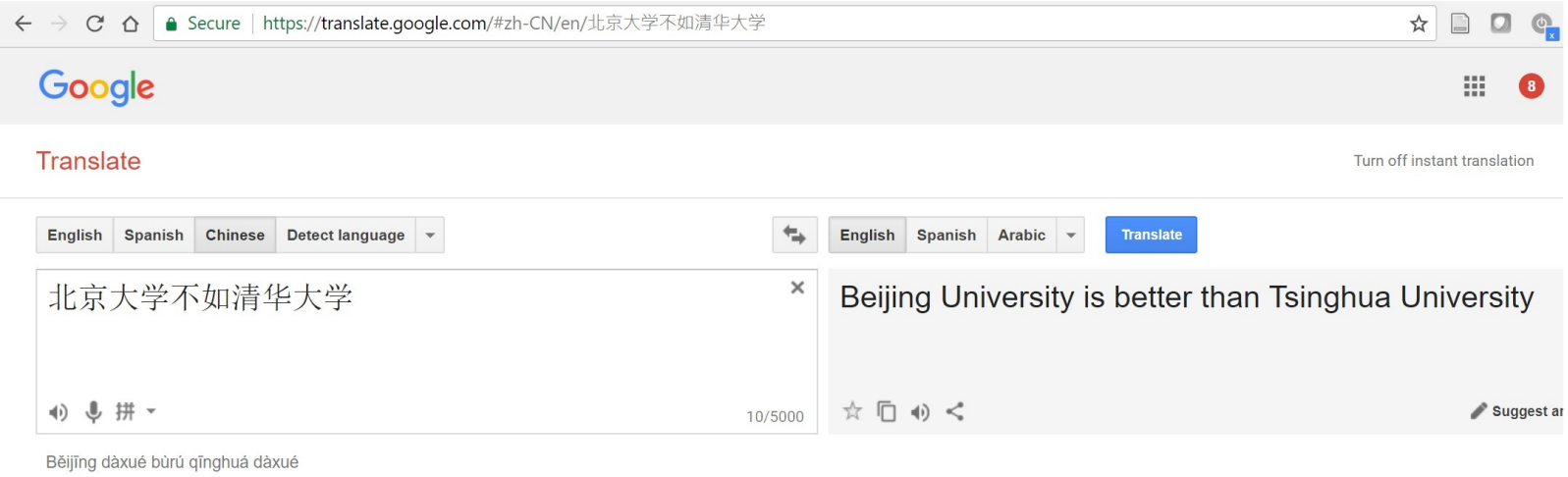




Pei et al. DeepXplore: Automated Whitebox Testing of Deep Learning Systems. SOSP 2017.
Tian et al. DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars. ICSE 2018.

Slide adapted from SOSP’17 slides

神经机器翻译 (Neural Machine Translation)



- 效果优于统计机器翻译
- 总体翻译质量更高，结构更为简单

- 可控性较差，凸显质量维护重要性
- 现有翻译质量维护方案
  - ➢ 需参考翻译，在线场景不可用
  - ➢ 无法精确定位问题类型与位置

2018年4月5日

构建新的翻译质量保障方法

- 核心思想：针对翻译结果中的常见问题，设计专门的黑盒检测算法
  - 无需参考翻译，仅需给出原句与翻译系统的输出
  - 能够精确定位问题类型与问题在翻译结果中的位置
- 目前针对神经机器翻译结果中存在的常见问题
  - 漏译（Under-translation）

English (original)	Chinese (translated)
Nine *anonymous* people described as current and former U.S. officials	九名现任与前任美国官员

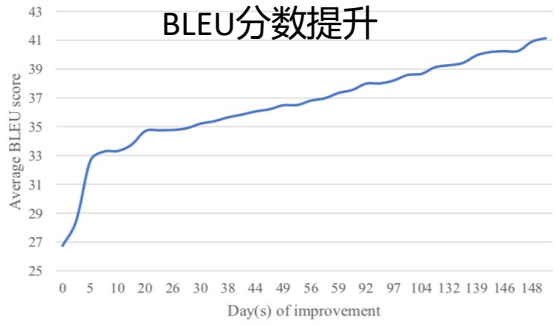
  - 多译（Over-translation）

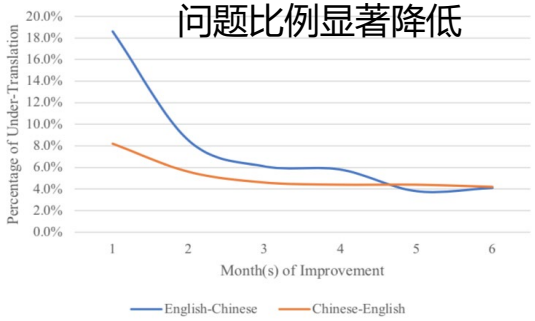
English (original)	Chinese (translated)
Both Elise and Hope were intense typhoons with maximum winds near their centers exceeding 200km/h.	埃利斯和霍普都是密集的台风，在其中心附近最大风速超过每小时200公里/小时。

Wujie Zheng, Wenyu Wang, Dian Liu, Changrong Zhang, Qinsong Zeng, Yuetang Deng, Wei Yang, Pinjia He, and Tao Xie. Testing Untestable Neural Machine Translation: An Industrial Case. arXiv:1807.02340, July 2018. https://arxiv.org/abs/1807.02340

效果与应用

- 方法应用于微信翻译系统（逾10亿用户，每日在线处理约1200万次不同的翻译任务），包括两种翻译质量保障形式
  - 离线监控（回归测试）
  - 在线监控（统计出错翻译比例，并实时选择最优模型）
- 大规模词汇翻译测试集
  - 包含约13万英文与18万中文词汇/短语
  - 谷歌、有道等翻译系统中发现大量问题





在其他翻译系统中发现的问题实例

Provider Name	Original Text	Given Translation	Expected Translation
Prvd. A	成人	mature people	adult
Prvd. A	太好了	what fun	great
Prvd. B	large-scale	large-scale	大规模
Prvd. B	long-term	long-term	长期
Prvd. B	U.S.	U.S.	美国
Prvd. C	蛋糕	Runeberg torte	cake
Prvd. C	酸奶	Viili	yoghurt
Prvd. D	疟原虫	p.	plasmodium
Prvd. D	酶原	The original enzyme	zymogen

软件构造与保障活动

百度为例

资料来源：王海峰，第二十二届中国国际软件博览会高峰论坛，2016. 6. 29

在线服务事故管理系统

Microsoft Research 微软亚洲研究院

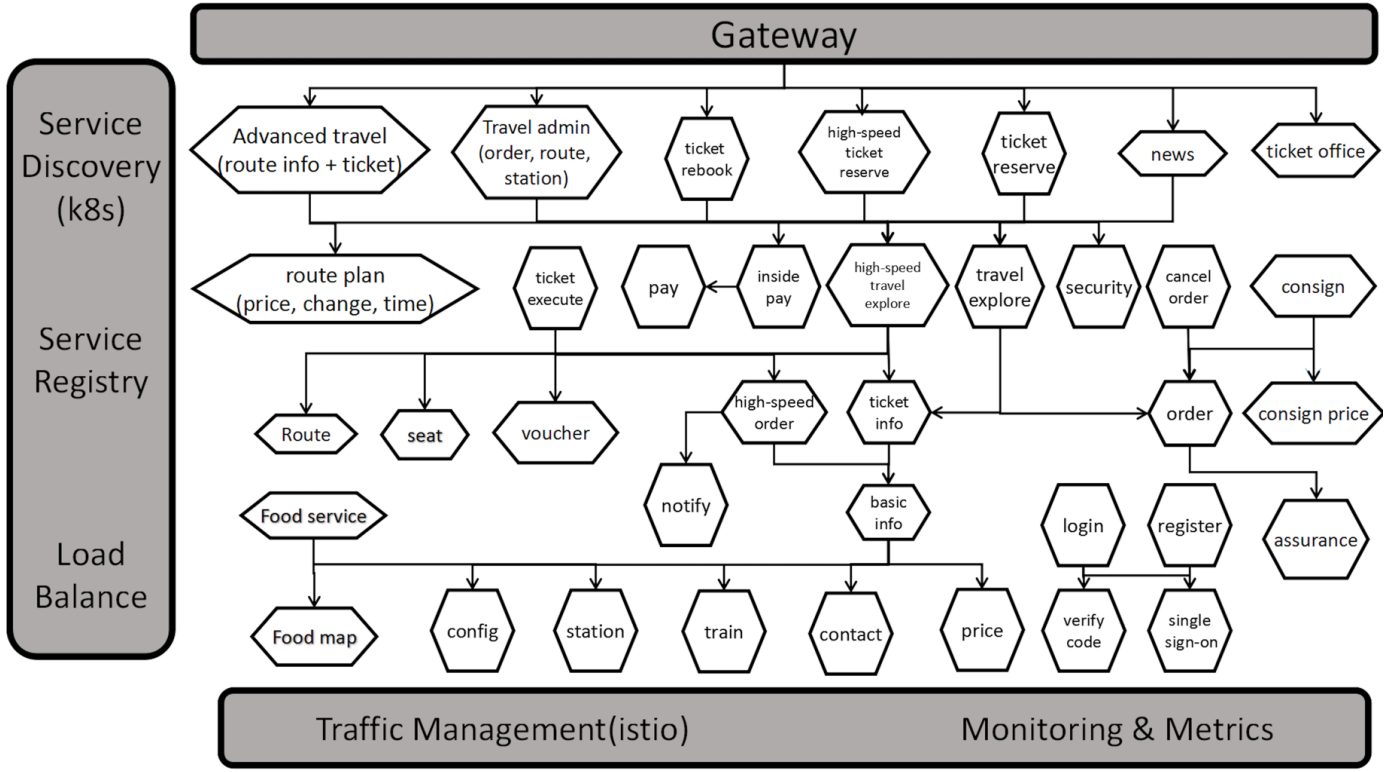
软件分析组
（Software Analytics）

- **数据源**：事务日志，系统度量，历史故障报告
- **分析输出：救治建议，事故根源**

服务分析工作站(Service Analysis Studio)应用于一个亿级用户微软在线服务

Jian-Guang Lou, Qingwei Lin, Rui Ding, Qiang Fu, Dongmei Zhang, and Tao Xie. Software Analytics for Incident Management of Online Services: An Experience Report. In Proceedings of ASE 2013. http://taoxie.cs.illinois.edu/publications/ase13-sas.pdf

开源微服务Benchmark系统TrainTicket



- 包含Java、Python、Go、Node.js
- 大量使用了异步通信和消息队列，具有较好的性能和负载能力
- 按照工程化开发方式进行了测试，包含100多个单元和集成测试用例
- 基于可视化工具实现了微服务的运行监控和管理

70+微服务，包括41个业务微服务，30个基础服务（消息中间件服务、分布式缓存服务、数据库服务等）
总代码量30多万行

复旦大学、UIUC、SUTD合作研究

Git库地址：https://github.com/microcosmx/train_ticket

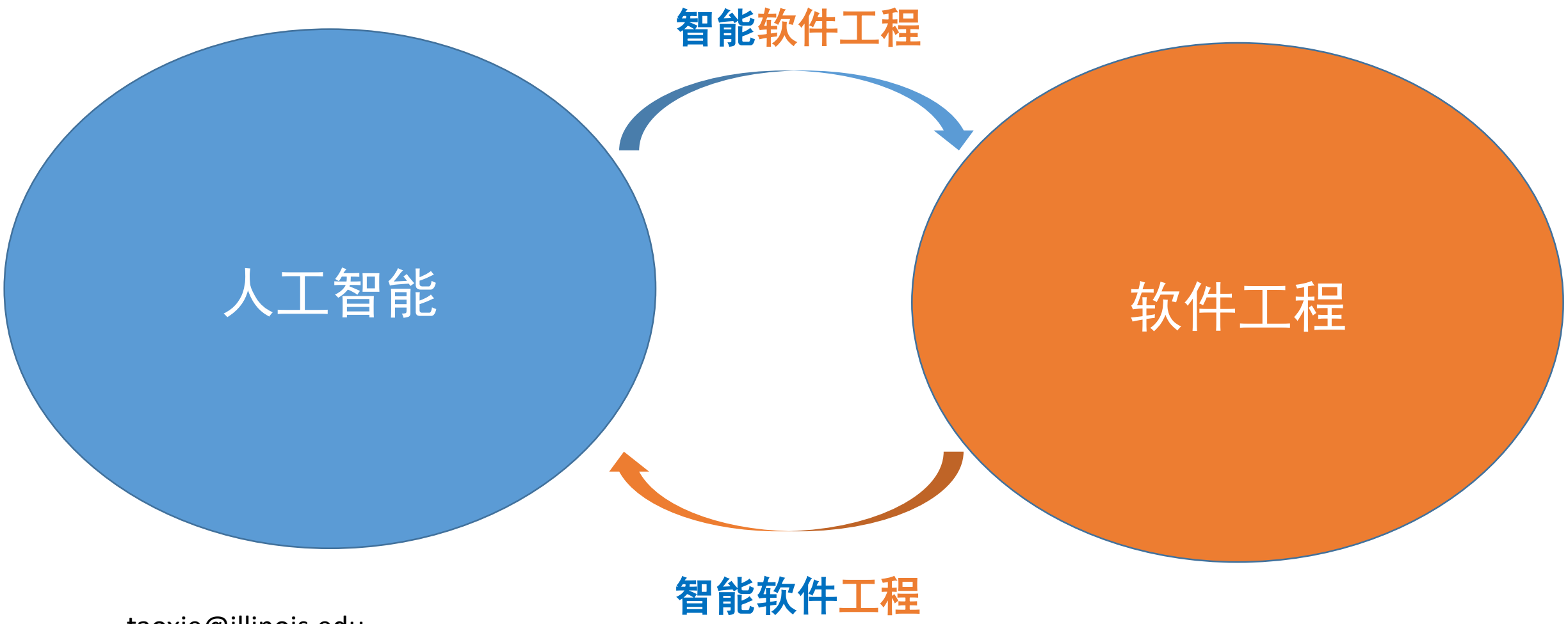
Xiang Zhou, Xin Peng, Tao Xie, Jun Sun, Chenjie Xu, Chao Ji, and Wenyun Zhao. Poster: Benchmarking Microservice Systems for Software Engineering Research. In Proceedings of ICSE 2018, Posters. http://taoxie.cs.illinois.edu/publications/icse18poster-microservices.pdf

软件构造与保障活动

百度为例

资料来源：王海峰，第二十二届中国国际软件博览会高峰论坛，2016. 6. 29

人工智能 ←→ 软件工程



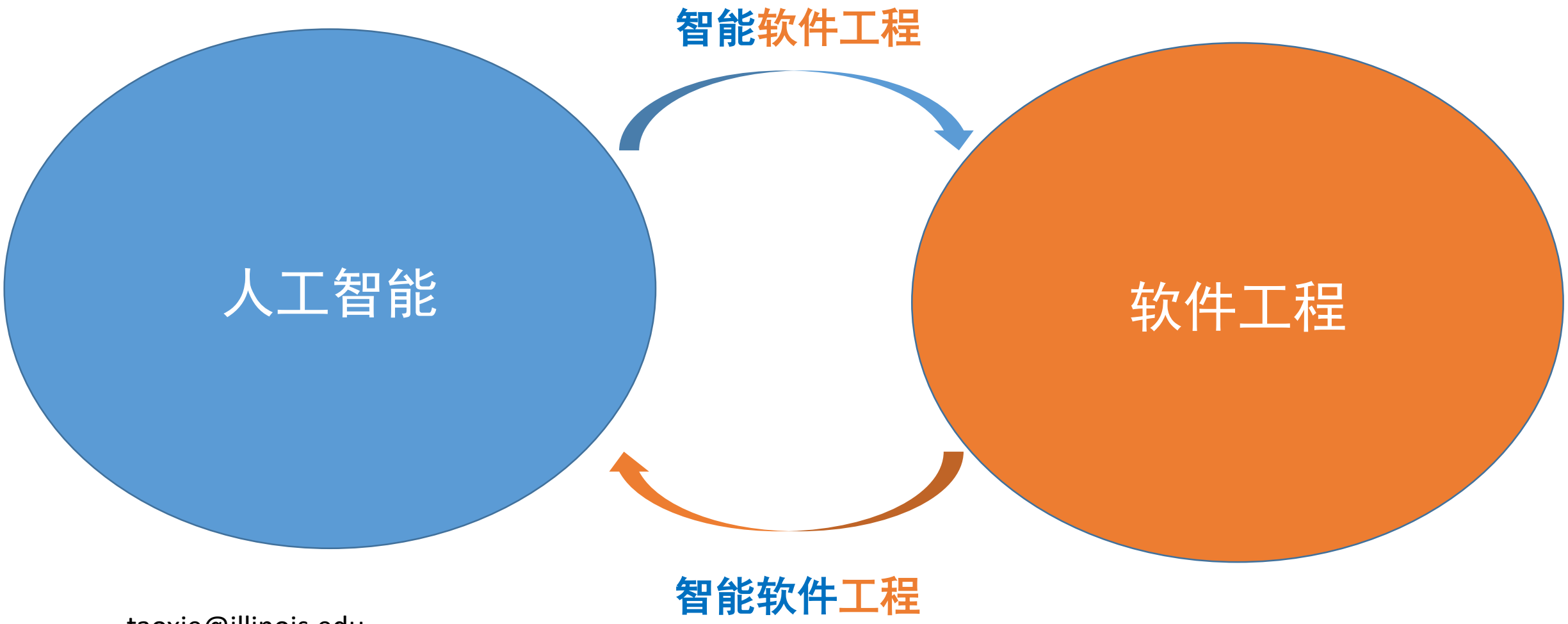
taoxie@illinois.edu
http://taoxie.cs.illinois.edu/

谢谢！

Q & A

This work was supported in part by NSF under grants no. CNS-1513939, CNS-1564274, CCF-1816615, and a grant from the ZJUI Research Program.

人工智能 ←→ 软件工程



taoxie@illinois.edu
http://taoxie.cs.illinois.edu/