

语音用户界面设计启示录

The Revelation of Voice User Interface Design

王一行

Senior interaction designer at Alibaba





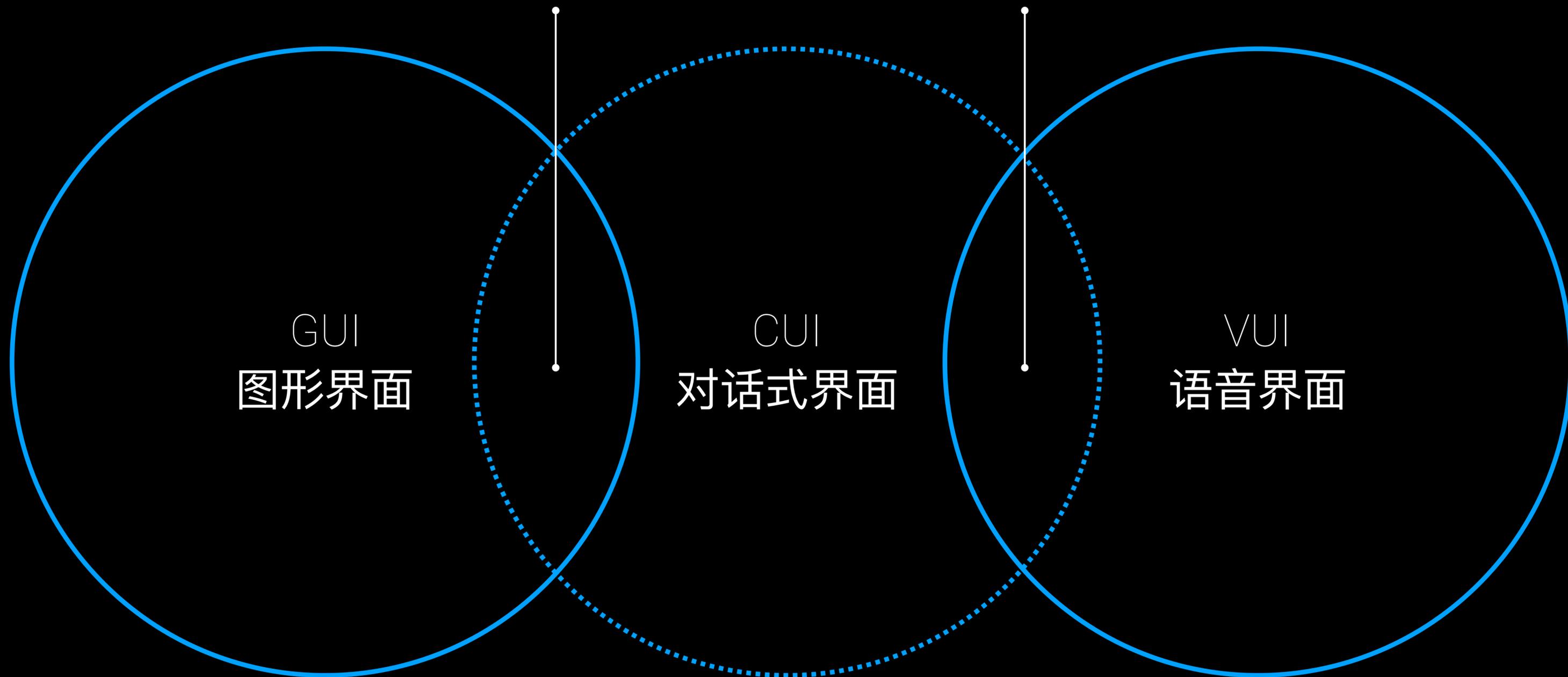
不能脱离场景谈设计
No context no design

语音用户界面是什么

What is voice user interface

Chatbot
界面式对话机器人

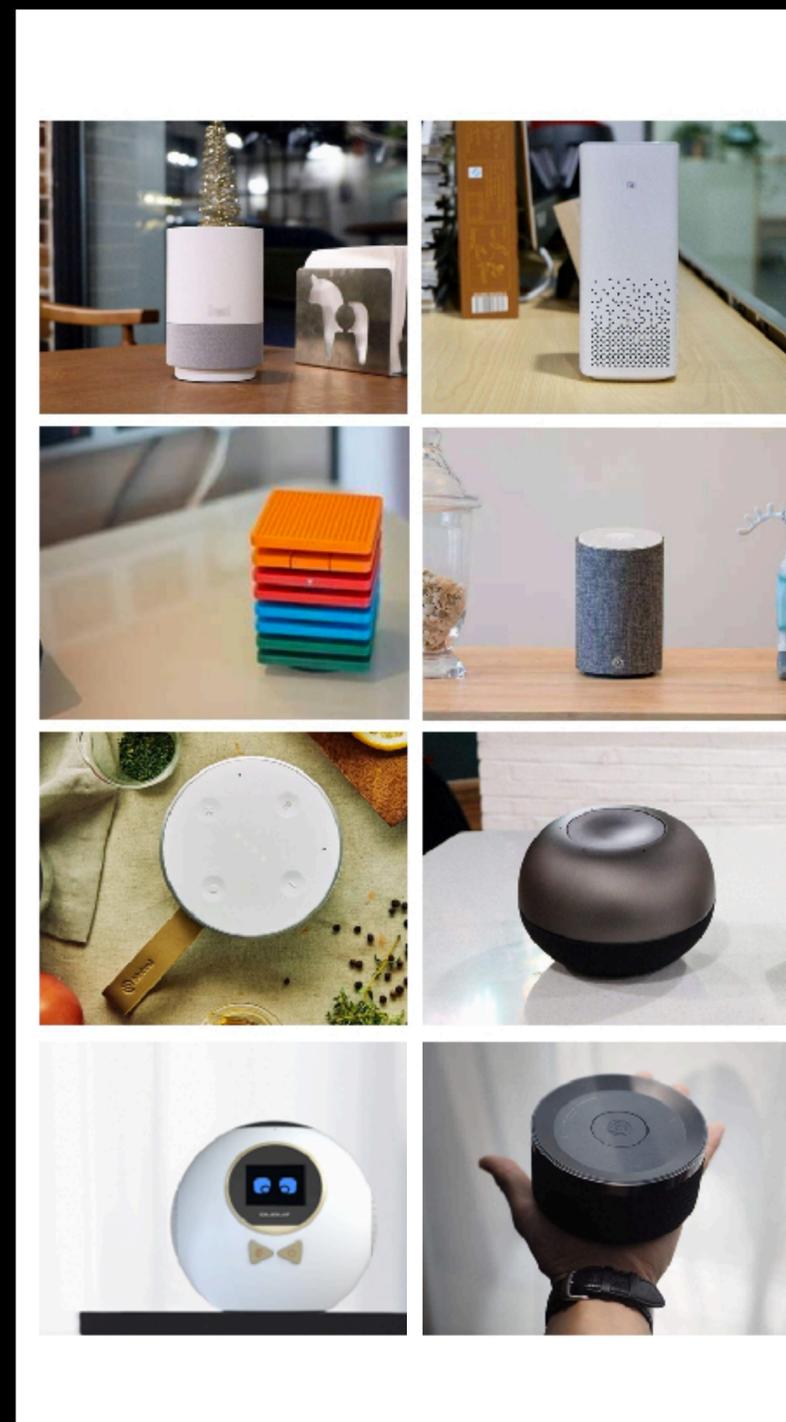
Virtual Voice Assistant
智能语音助理



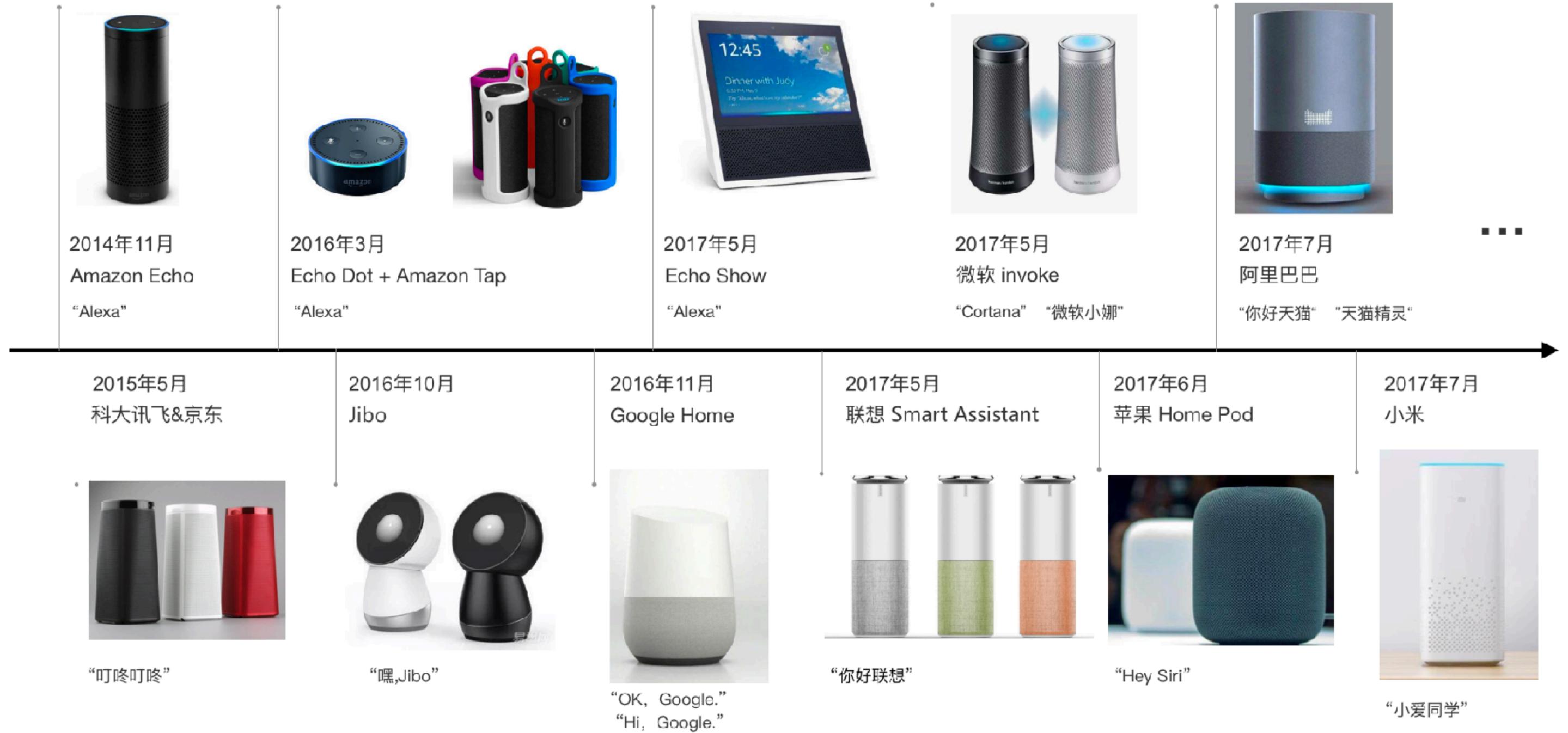
Chatbot 界面式对话机器人



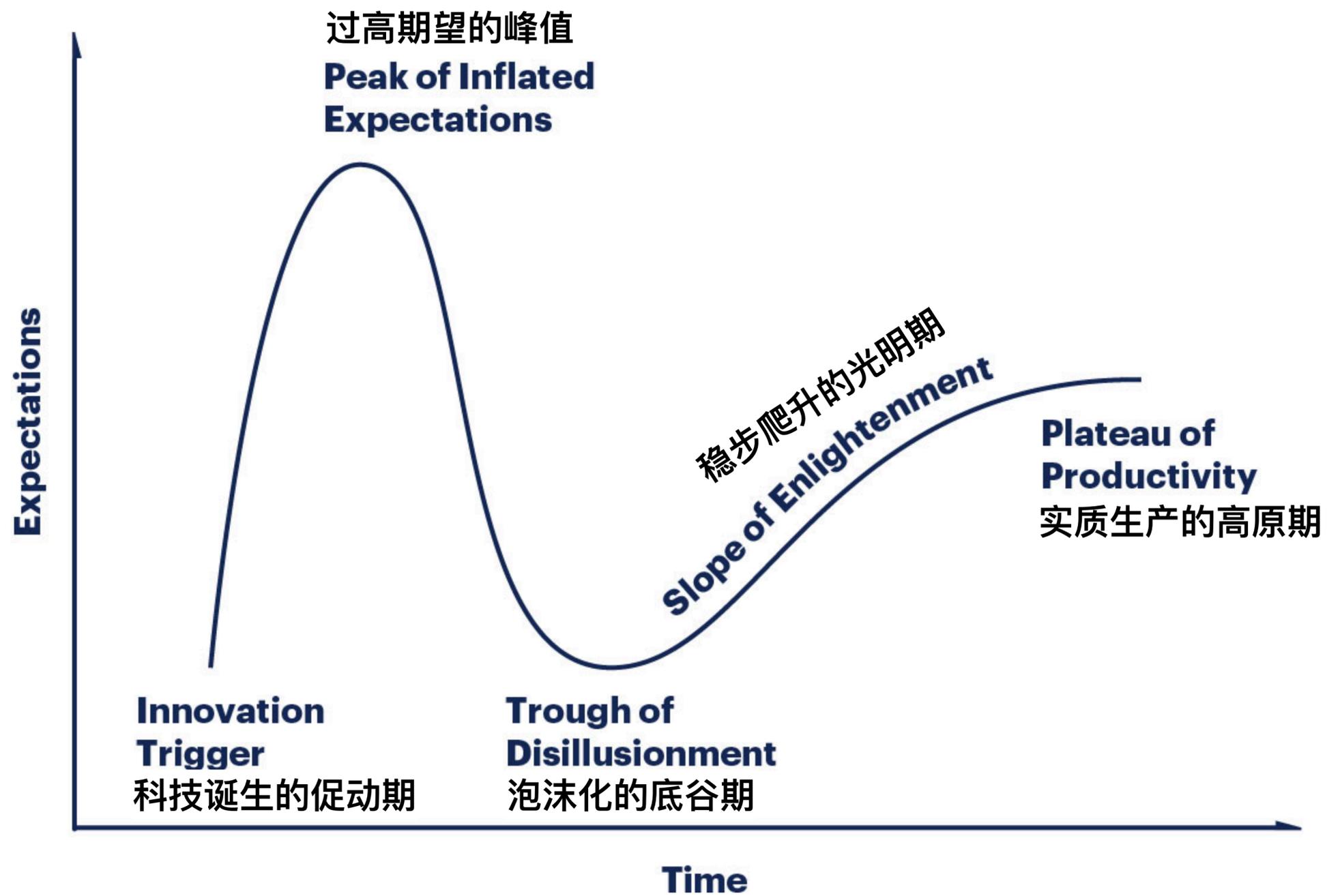
Virtual Voice Assistant 智能语音助理

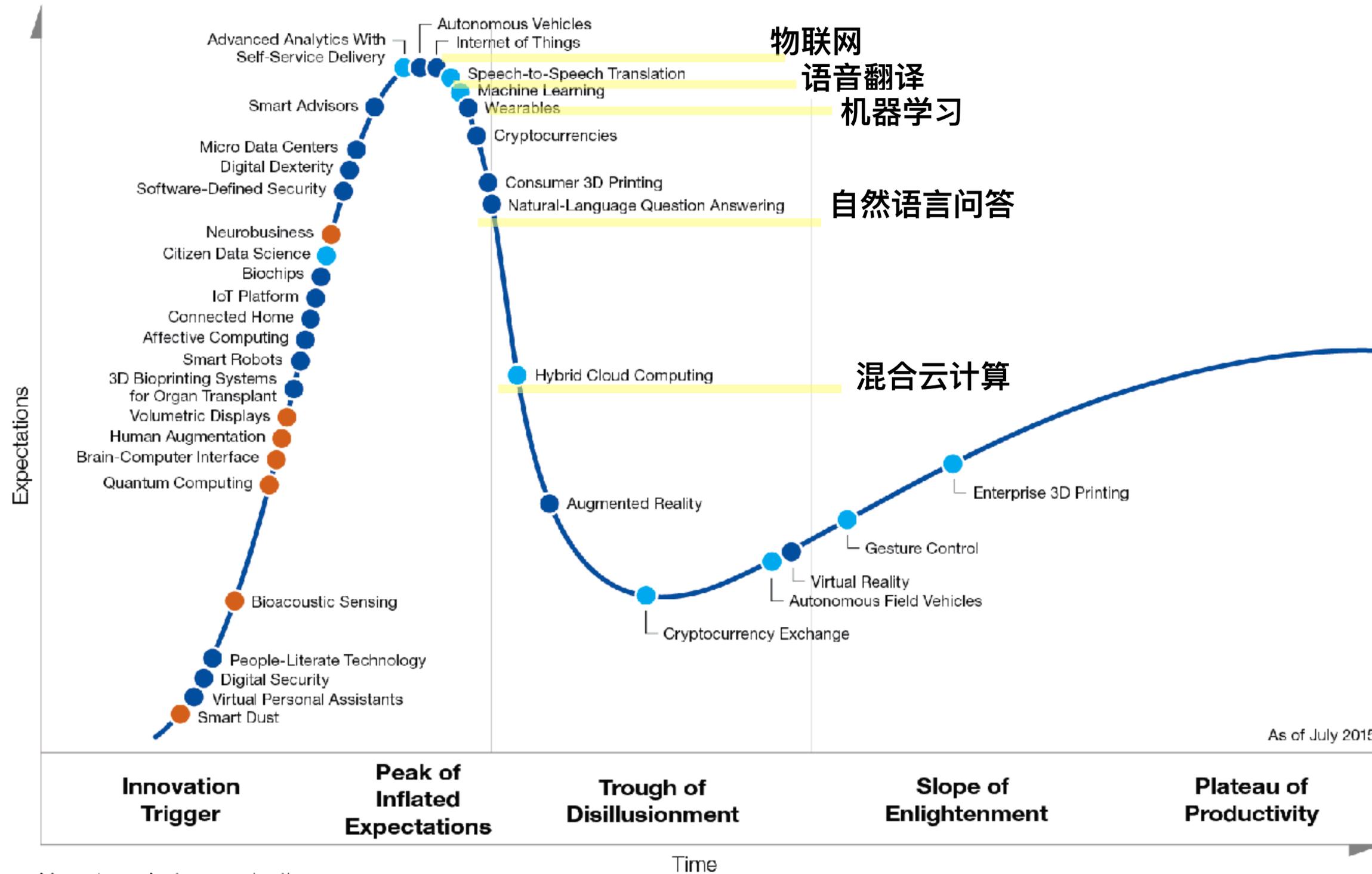


迎面而来的智能音箱



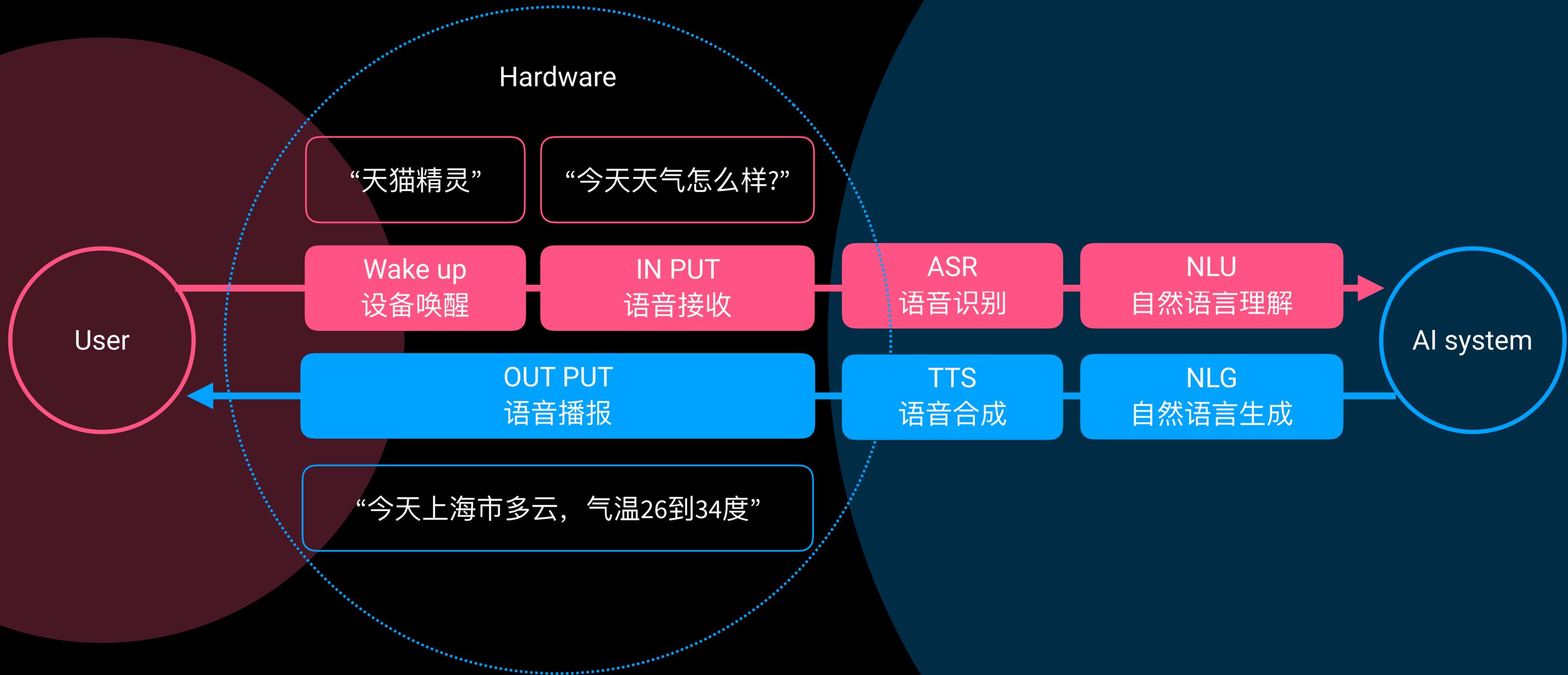
The Hype Cycle 技术成熟度曲线





Years to mainstream adoption:

- less than 2 years
- 2 to 5 years
- 5 to 10 years
- more than 10 years
- obsolete before plateau

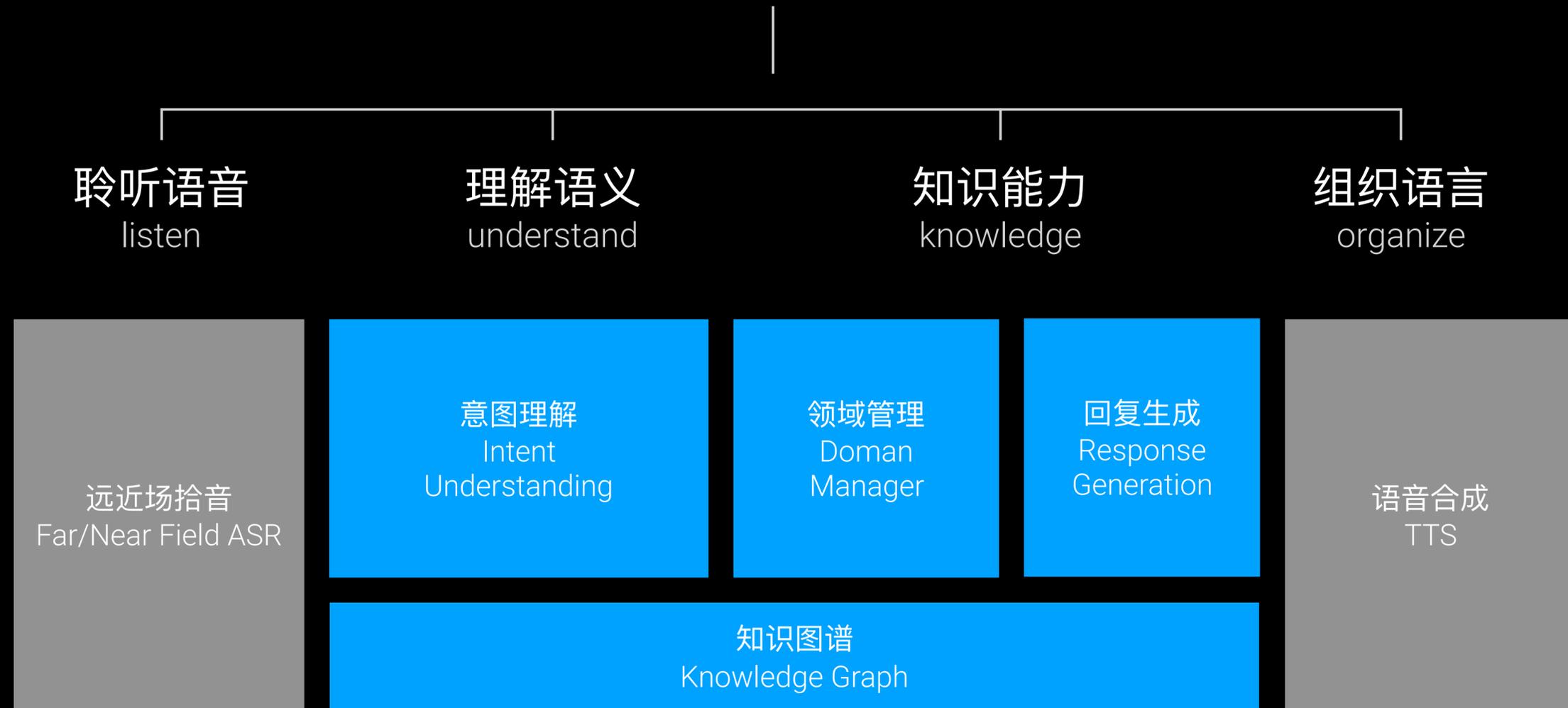


人工智能初级阶段如何做好设计

Design at primary stage of AI

产生交谈

Have a conversation



用户知道什么时候该说话，什么时候不该说话吗？

When

他们知道什么时候可以开始说话吗？

When to start

他们会说什么？系统如何回复？

What to say and how to reply

他们在哪里会感到困惑和犹豫？

Puzzled

产生交谈

Have a conversation

聆听语音
listen

理解语义
understand

知识能力
knowledge

组织语言
organize

交互流程设计

意图模式设计

行为引导设计

知识结构设计

人格情绪设计

设计VUI要做的事情

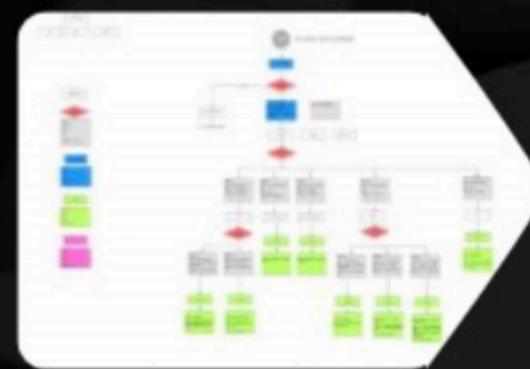
voice interface is...



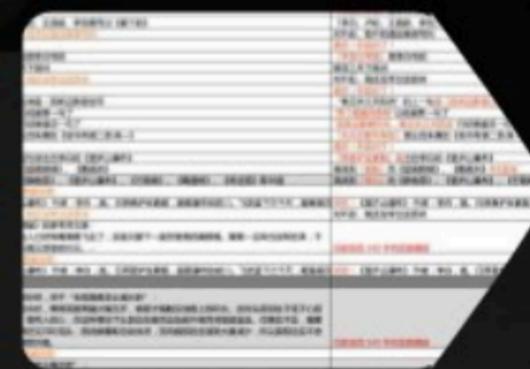
需求的挖掘
Intents



语料的收集
Corpus



对话流程的梳理
Flow



语音内容的设计
Contents



语音体验的测试
Testing

An aerial photograph of a dense forest of evergreen trees. A single tree with bright yellow leaves stands out prominently in the middle-left area of the frame. The overall lighting is dim, suggesting a low sun or dusk, creating a dark and moody atmosphere.

挖掘意图

Intent mining

“明天天气怎么样？”

“雨还要下多久？”

角色扮演，还原真实情境
Role playing, restoring real situations

Participant Observation







目标信息

辅助引导

用户



目标信息

辅助引导

用户



目标信息

辅助引导

用户



目标信息

辅助引导

用户

情境模态

时间

物理环境

所在地

用户模态

认知

情绪状态

感知

心理状态

人口统计学

感官

体能

设备模态

设备类型

软/硬件

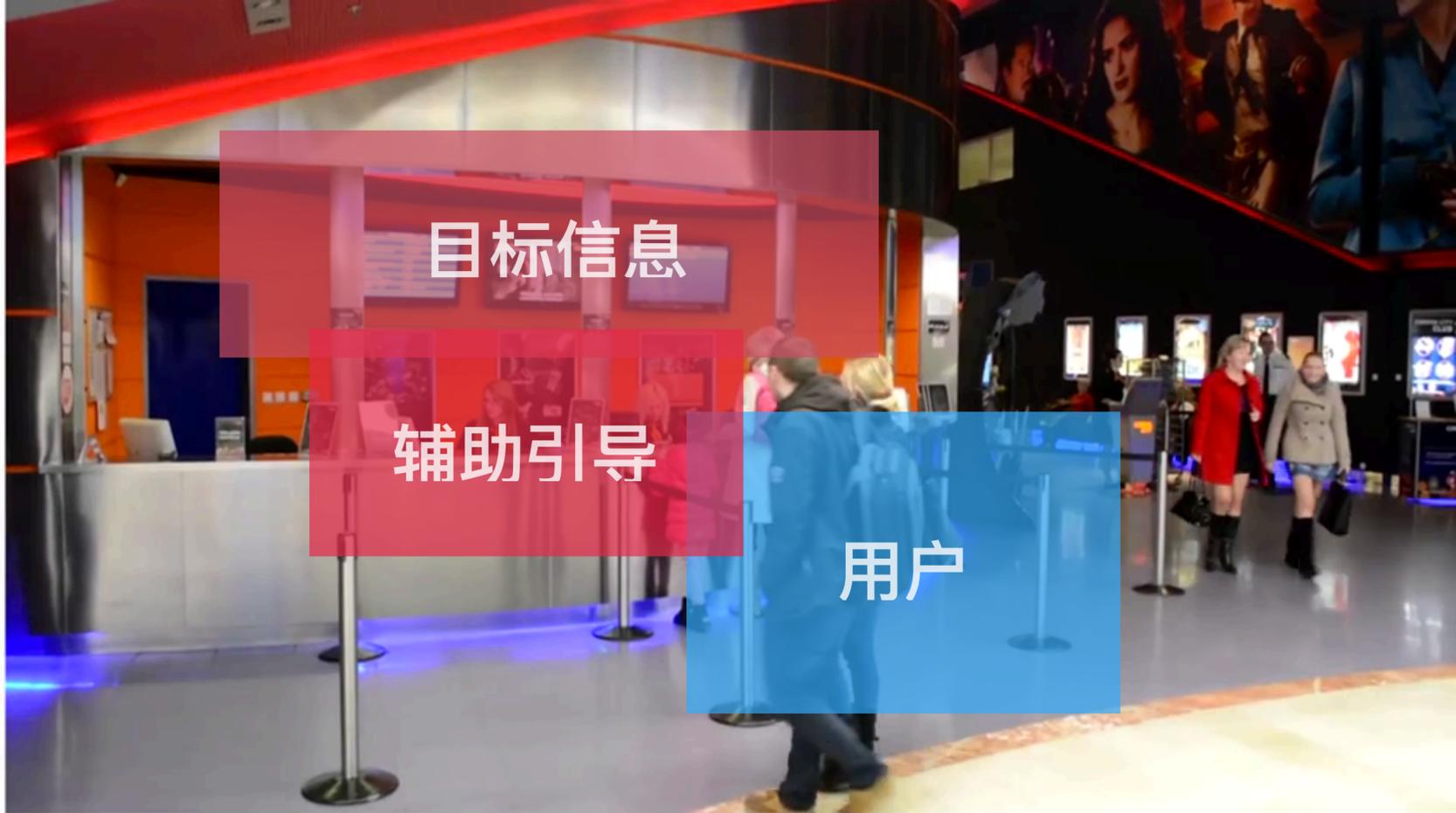
输入输出





结合行为的引导





根据场景逐步引导用户以语音说出需求

查询意图下的引导



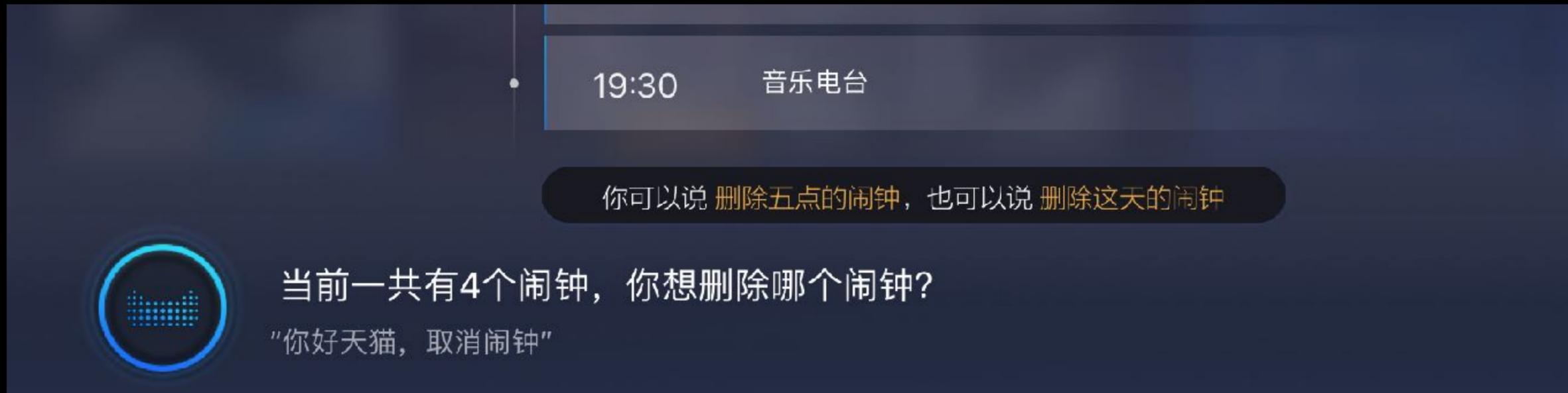
19:00 钢琴

你可以说 设一个闹钟，也可以说 删除闹钟

您今天一共有8个闹钟
“今天的闹钟”

Detailed description: This UI mockup illustrates a search intent scenario. At the top, a list item shows '19:00 钢琴'. Below it, a dark rounded rectangle contains the text '你可以说 设一个闹钟，也可以说 删除闹钟'. On the left, a circular icon with a grid of dots is next to the text '您今天一共有8个闹钟' and '“今天的闹钟”'.

取消意图下的引导



19:30 音乐电台

你可以说 删除五点的闹钟，也可以说 删除这天的闹钟

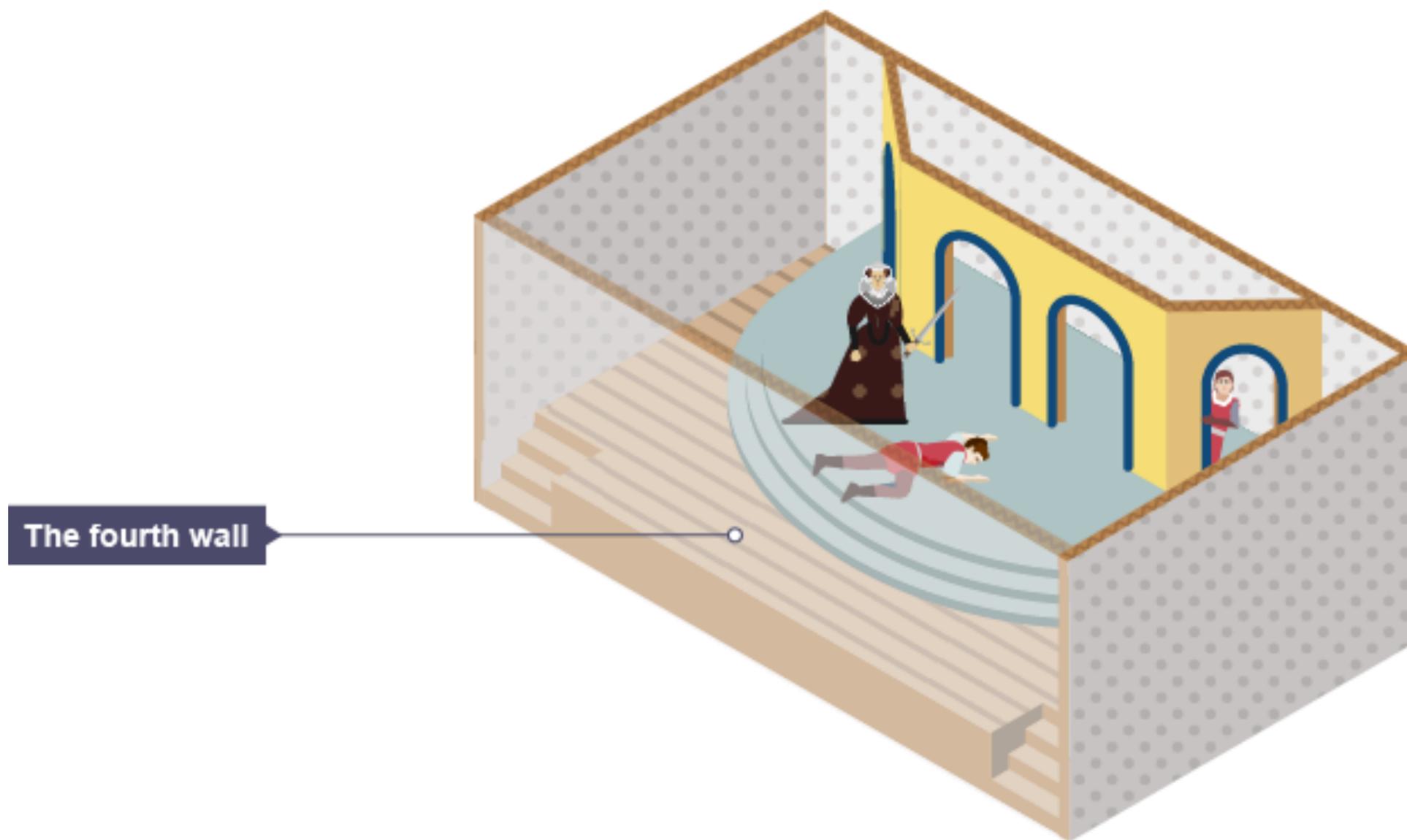
当前一共有4个闹钟，你想删除哪个闹钟？
“你好天猫，取消闹钟”

Detailed description: This UI mockup illustrates a cancellation intent scenario. At the top, a list item shows '19:30 音乐电台'. Below it, a dark rounded rectangle contains the text '你可以说 删除五点的闹钟，也可以说 删除这天的闹钟'. On the left, a circular icon with a grid of dots is next to the text '当前一共有4个闹钟，你想删除哪个闹钟？' and '“你好天猫，取消闹钟”'.

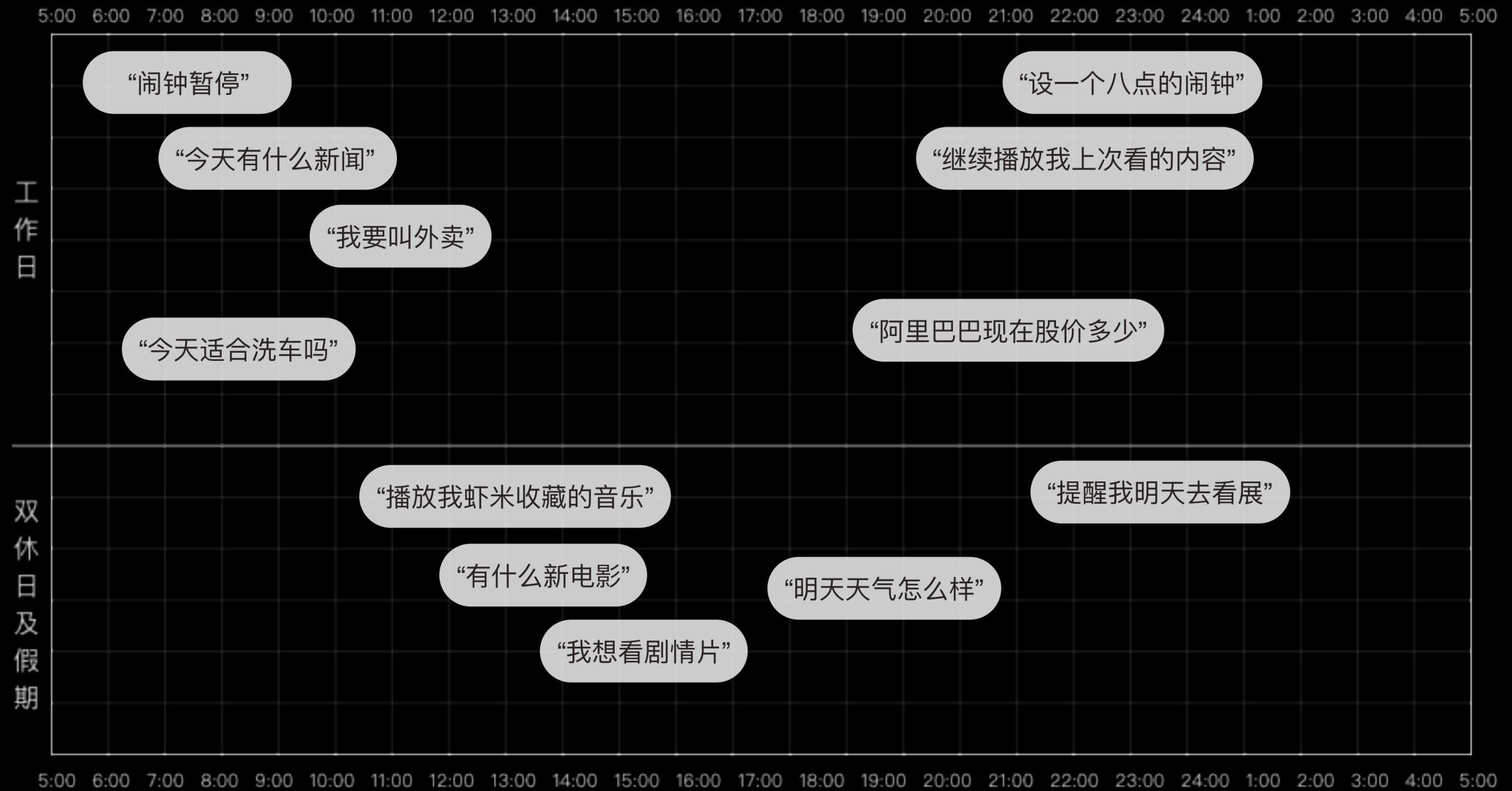


主动性
Initiative

“第四面墙”



基于真实生活的 Intents





创造意识错觉

用户的一切行为

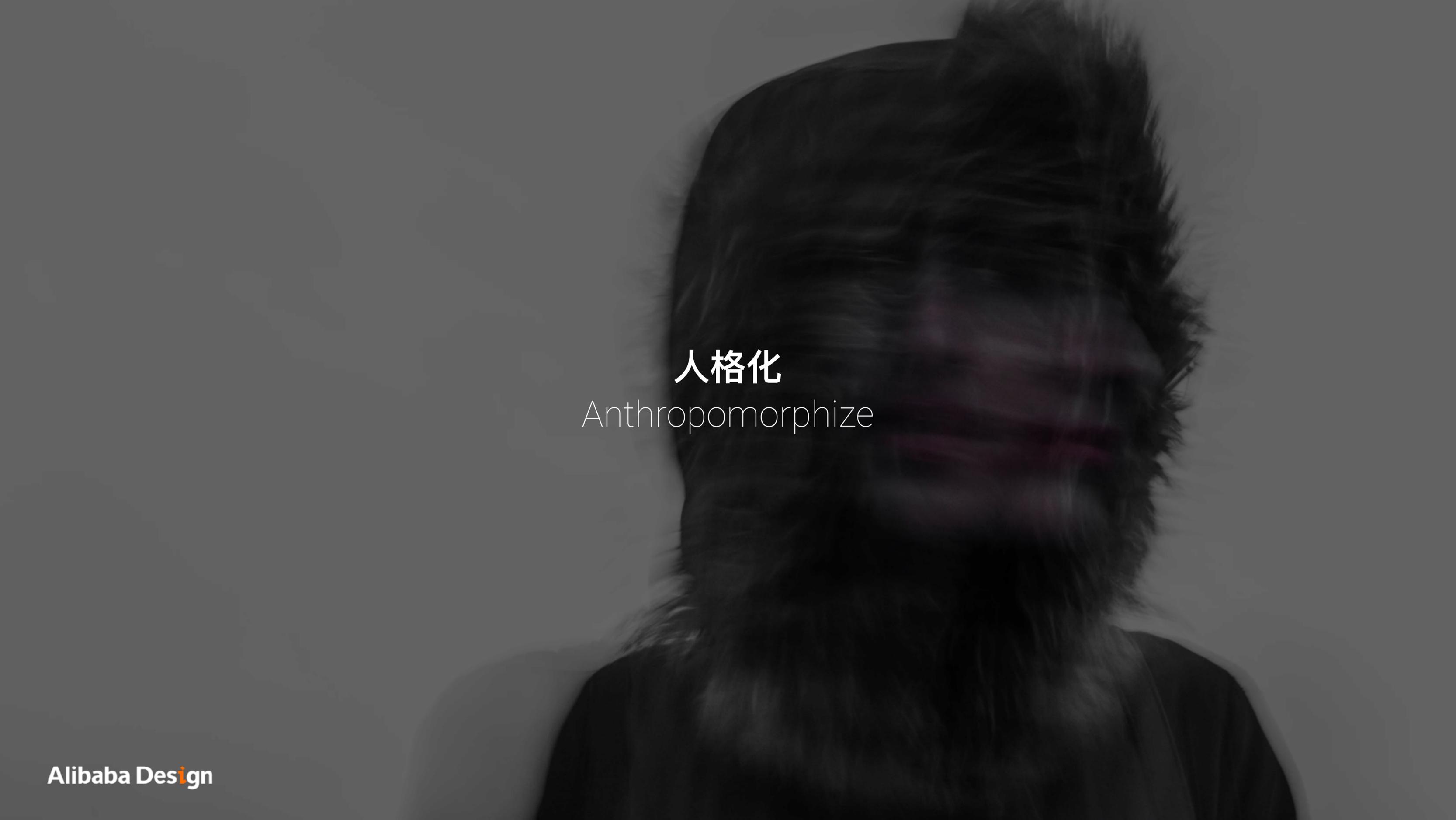
用户所处的实际时间和空间

不同使用者所处情境与行为的比较

哈里·戈特利布 Harry Gottlieb

Founder of Jellyvision Lab & Jackbox Games





人格化

Anthropomorphize

江小白
我有一瓶酒 有话对你说

我们那些
共同的记忆
是最好的
下酒菜



江小白
我有一瓶酒 有话对你说

我们
总是发现
以前的自己
有点傻

梦梦
发表于...



江小白
我有一瓶酒 有话对你说

不想做一个
多牛逼的人
只想做一个
不可替代
的人

的 顶
发表于...



江小白
我有一瓶酒 有话对你说

最想说的话
在眼睛里
草稿箱里，梦里
和酒里

刘亮
发表于...



江小白
我有一瓶酒 有话对你说

江小白
我有一瓶酒 有话对你说

江小白
我有一瓶酒 有话对你说

还有...的事情



江小白

Straight 单纯高粱酒
重庆江记酒庄有限公司

酒精度:

江小白

白鹿庄园

我有一瓶酒 有话对你说

话说四海之内
皆兄弟
然而
四公里之内
却不联系

【我不喜欢苍老师】

发表于重庆科技学院

江小白

白鹿庄园

我有一瓶酒 有话对你说

不说错话
不做错事
青春白走一回

[better-atongm]

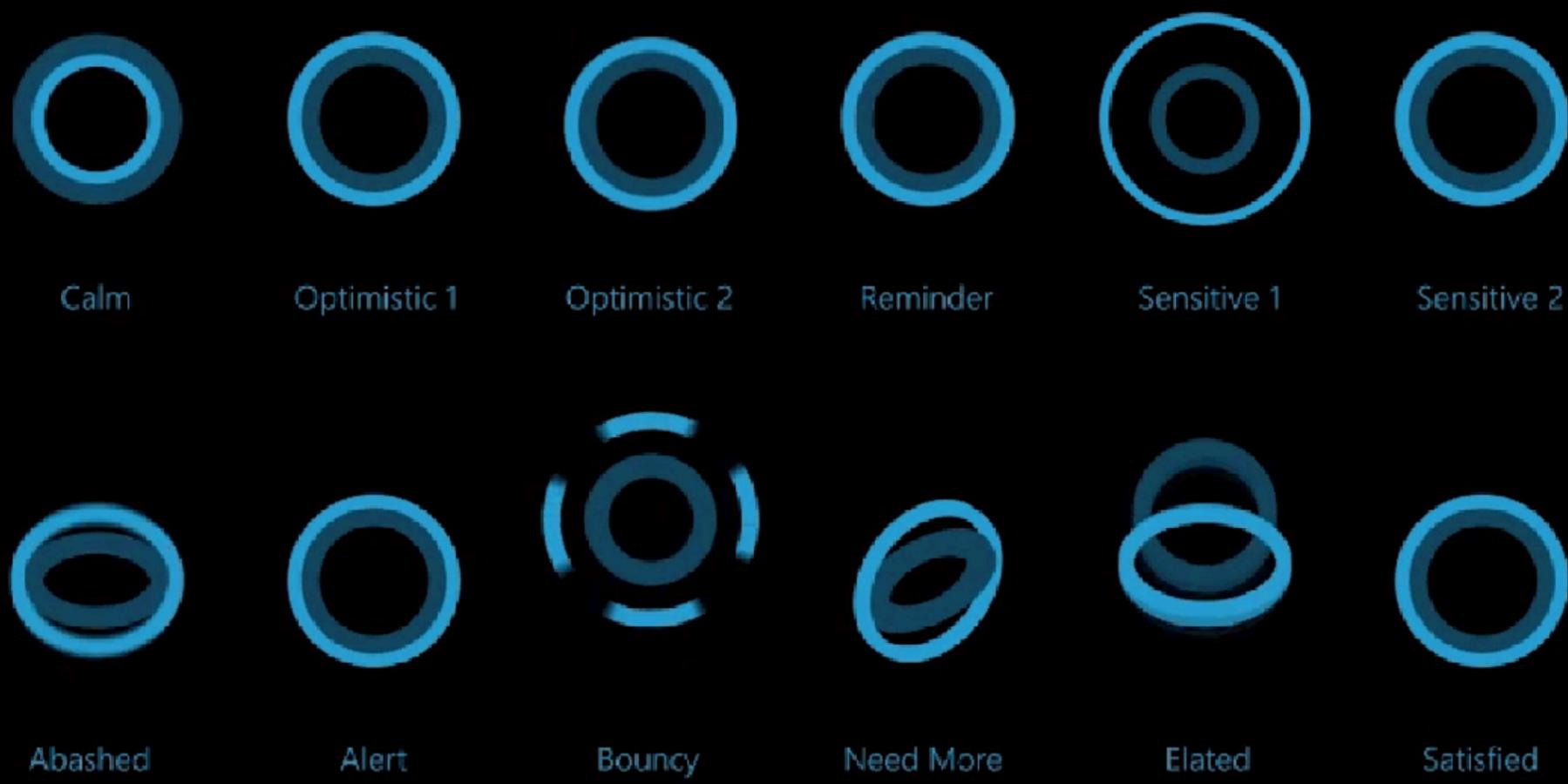
发表于新市场村

A woman with long brown hair and glasses, wearing a black top, is looking at a large screen. The screen displays a reflection of her face, which is slightly offset and dimmed, creating a visual metaphor for consistency in user models. The background is a plain, light-colored wall.

人物模型的一致性使人们能够在与它们沟通时得以预测接下来会发生什么。

凯瑟琳·斯比斯特 Katherine Isbister

HCI and game researcher



我们很清楚Cortana并不是人类。用纯数字化的表现形式可以更真实地展现她的本质。

德波拉·哈里森 Deborah Harrison

Cortana Editorial Writer







多模态
Multimodality

感官	感知	理解	情感	关联	信任	行为
视觉	+	+	+	+	+	+
听觉	+	+	+	+	-	+
触觉	+	-	+	+	+	+
嗅觉	+	-	+	+	-	+
味觉	+	-	+	+	-	+

新技术带来的多模态交互体验



SeeFront



Tmall Genie



Ultrahaptics



Aromyx



新的内容带来新的心智负荷

New content added new mental load

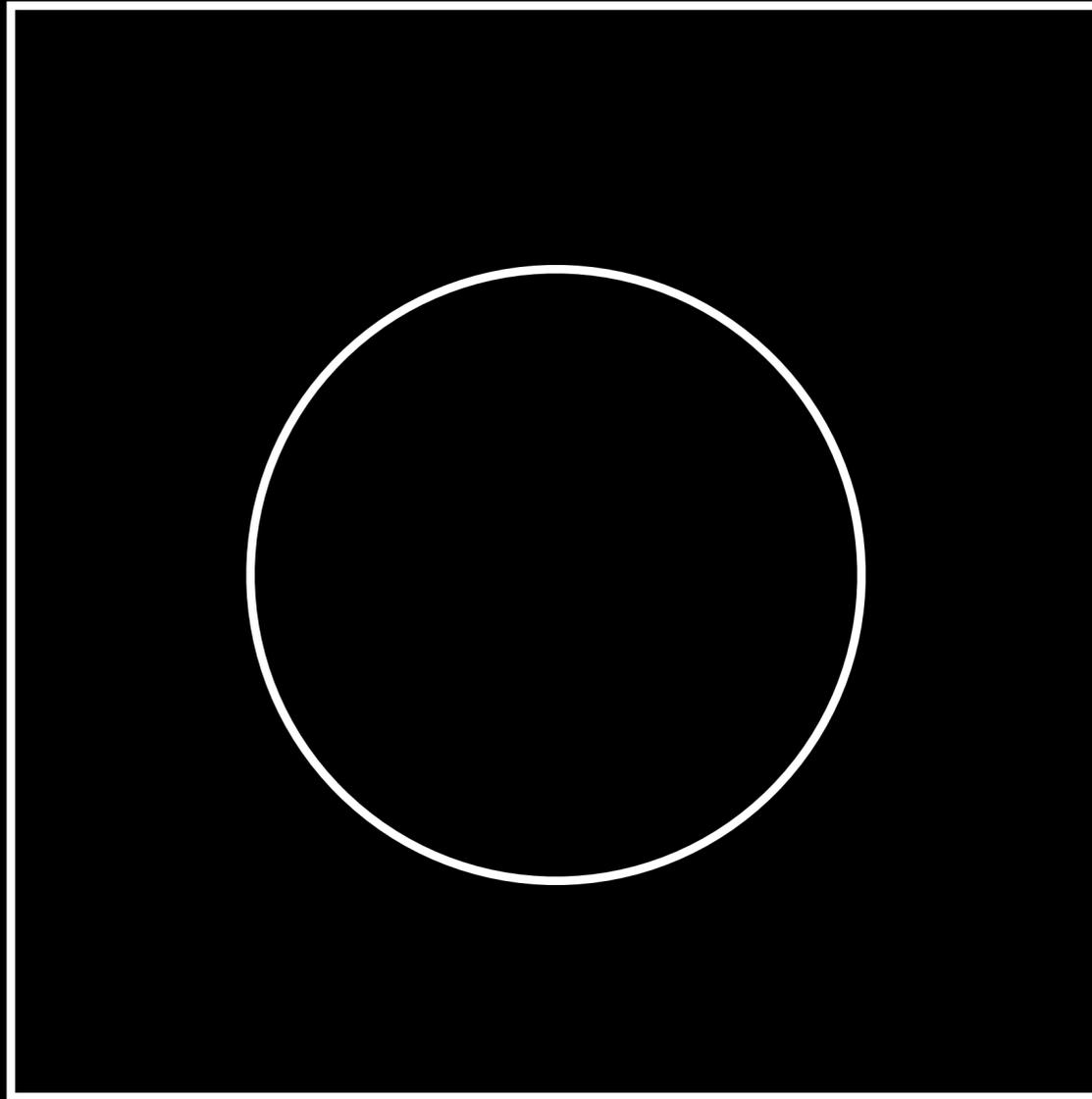


操作者可提供的
心智资源



任务要求的
心智资源

操作者可提供 \geq 任务要求



任务要求的
心智资源

操作者可提供的
心智资源

任务要求的
心智资源

任务要求的
心智资源

操作者可提供的
心智资源

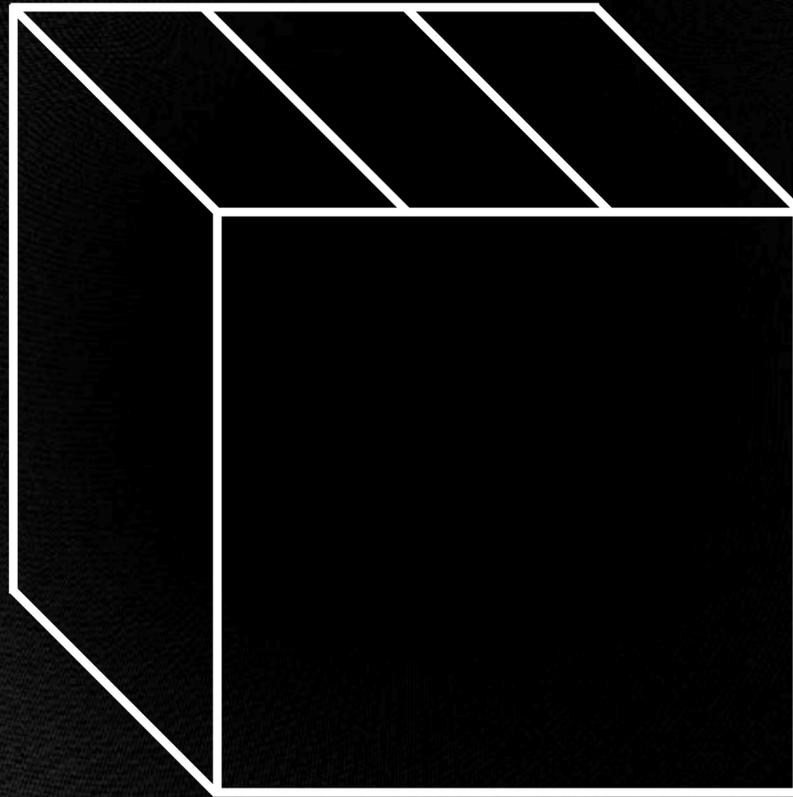
信息处理

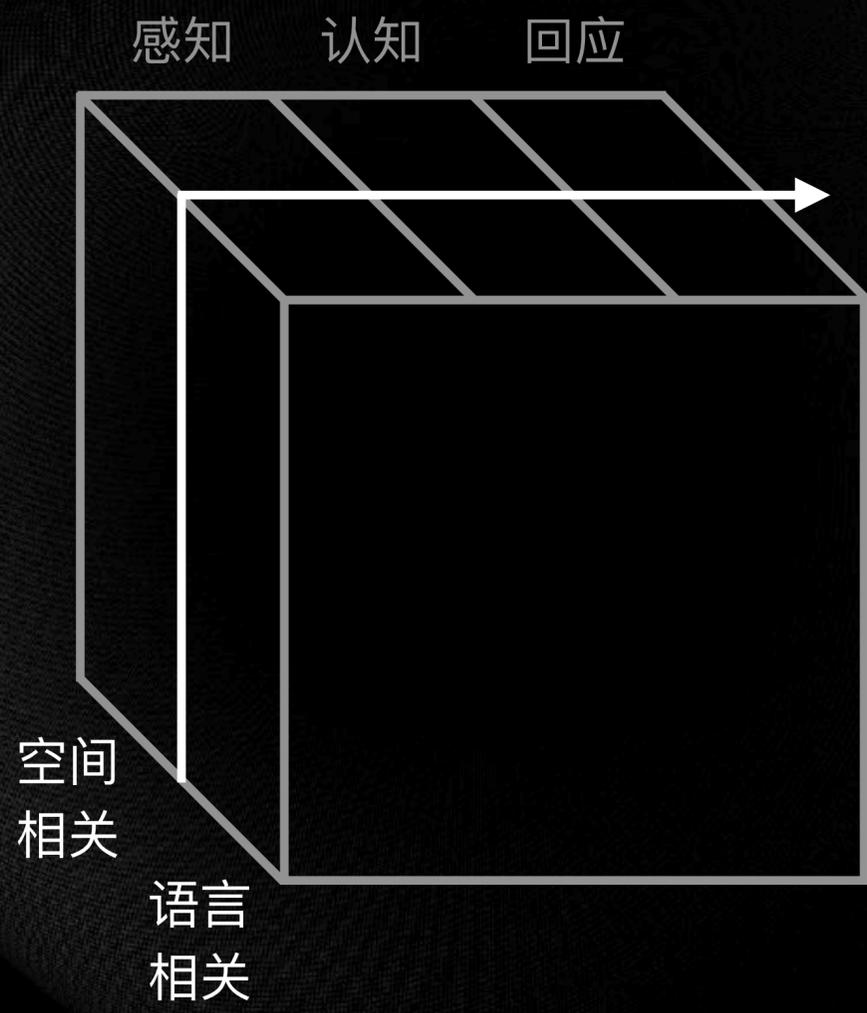
感知

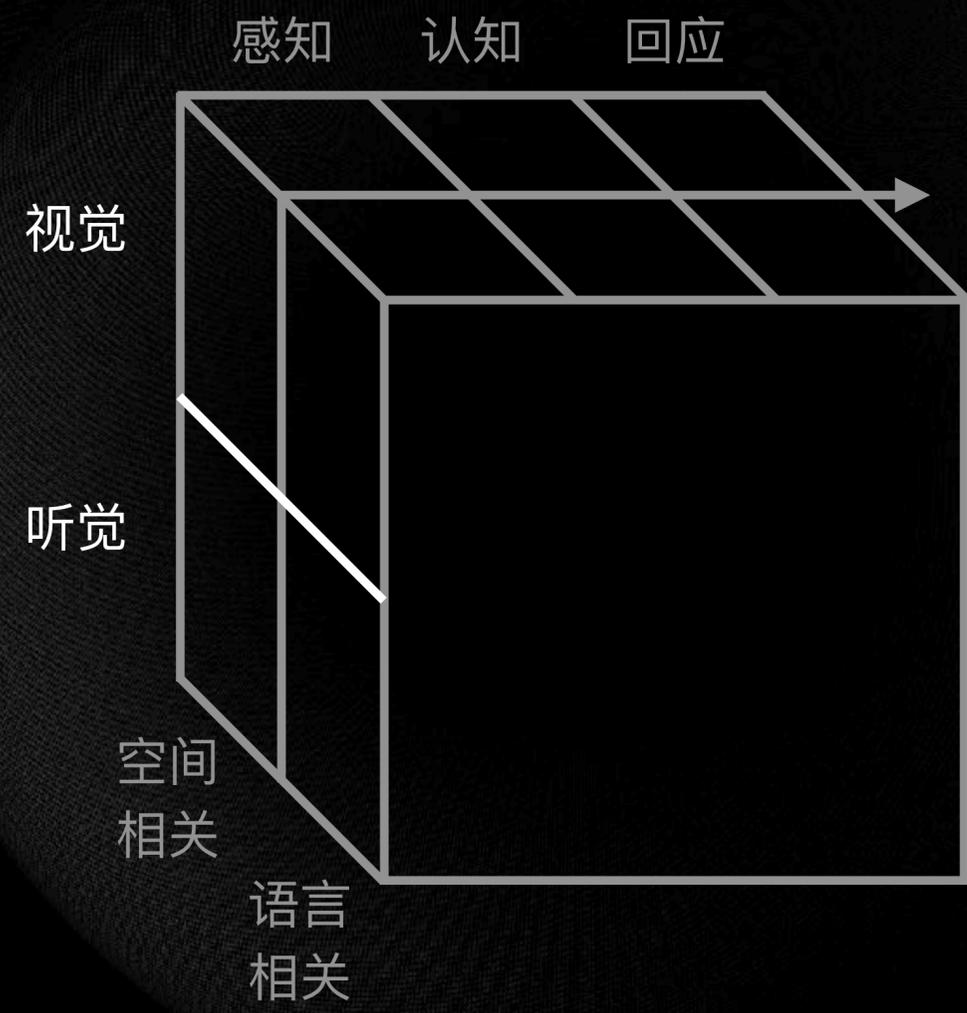
认知

回应

感知 认知 回应







感知

空间视觉
语言视觉

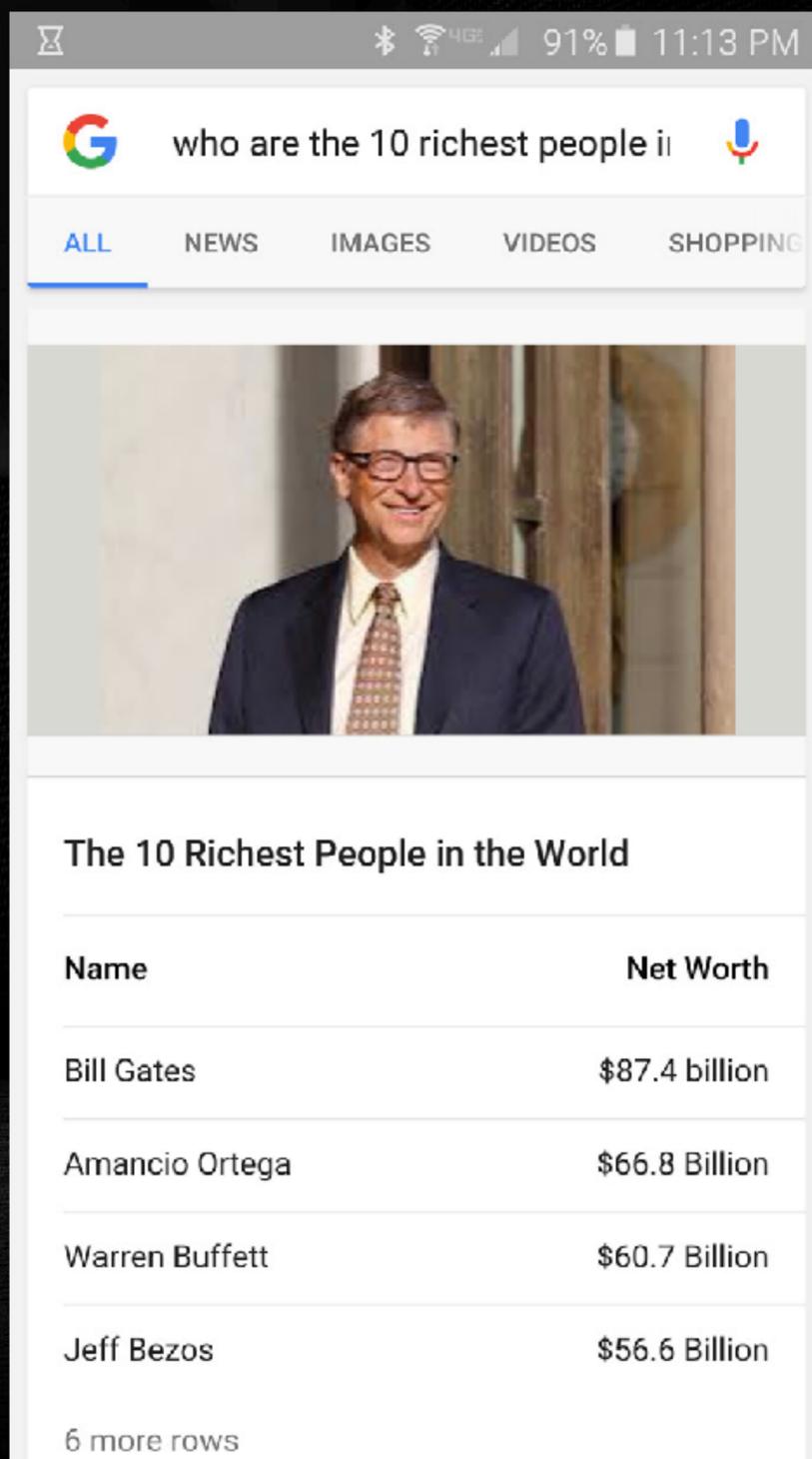
空间听觉
语言听觉

认知

空间认知
语言认知

回应

空间回应
语言回应



语言视觉

语言认知

屏显
文本

阅读
理解

任务A

任务B

语言视觉

语言认知

屏显
文本

干扰

阅读
理解

语言视觉

语言认知

语言视觉

1

0

语言认知

0

1

红 黄 蓝 绿

多模态干扰矩阵

任务A		感知				认知		回应	
		空间视觉	语言视觉	空间听觉	语言听觉	空间认知	语言认知	空间回应	语言回应
任务B	空间视觉	0.8	0.6	0.6	0.4	0.7	0.5	0.4	0.2
	语言视觉		0.8	0.4	0.6	0.5	0.7	0.2	0.4
	空间听觉			0.8	0.4	0.7	0.5	0.4	0.2
	语言听觉				0.8	0.5	0.7	0.2	0.4
	空间认知					0.8	0.6	0.6	0.4
	语言认知						0.8	0.4	0.6
	空间回应							0.8	0.6
	语言回应								0.8
	语言回应								

灯效
LIGHT



音效
SOUND



屏显
GUI

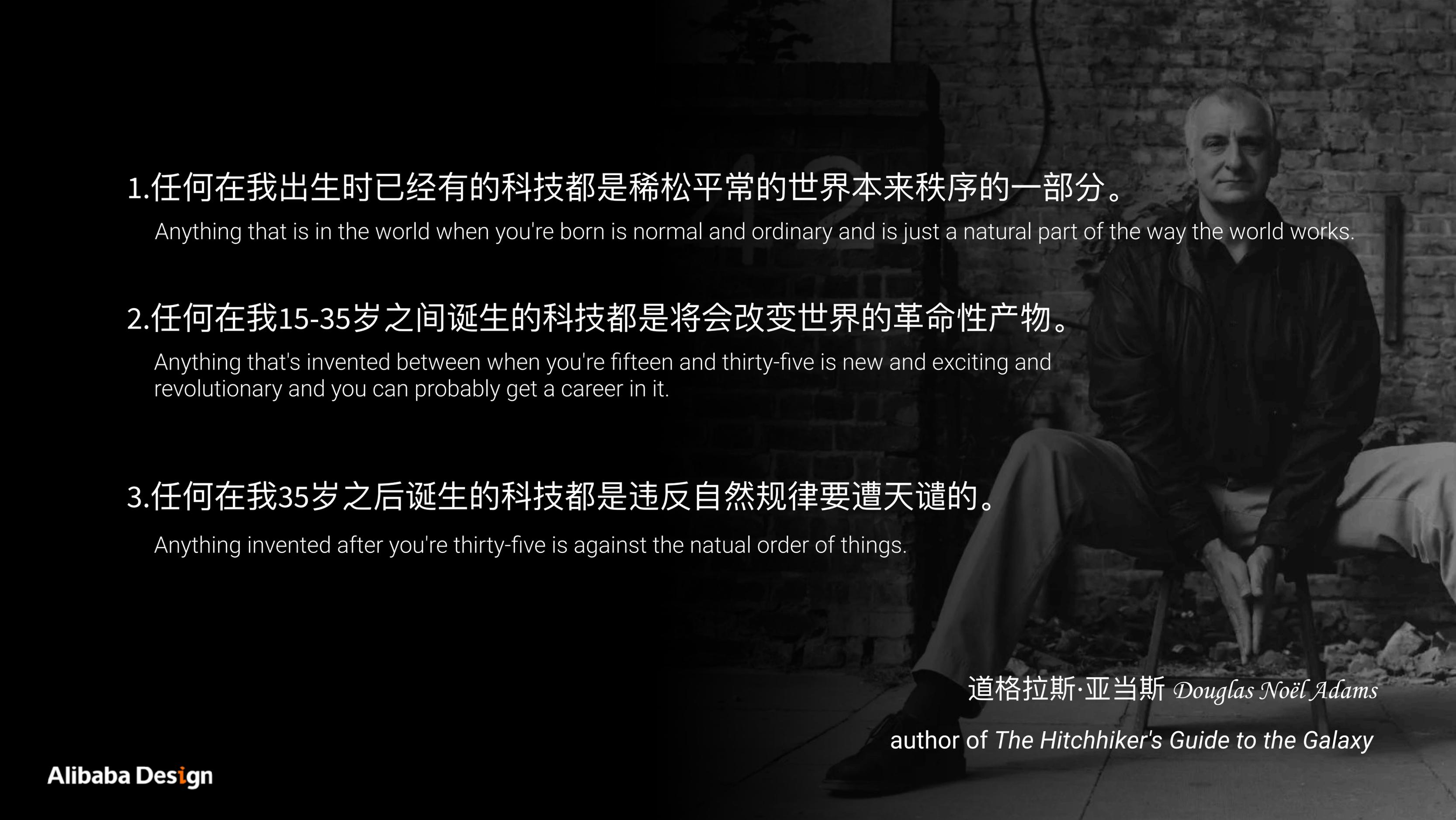


人+模态+场景



One More Thing

普惠科技下，设计的教育意义

A black and white photograph of Douglas Noël Adams sitting on a wooden chair in a room with stone walls. He is wearing a dark jacket and light-colored trousers. The text is overlaid on the left side of the image.

1.任何在我出生时已经有的科技都是稀松平常的世界本来秩序的一部分。

Anything that is in the world when you're born is normal and ordinary and is just a natural part of the way the world works.

2.任何在我15-35岁之间诞生的科技都是将会改变世界的革命性产物。

Anything that's invented between when you're fifteen and thirty-five is new and exciting and revolutionary and you can probably get a career in it.

3.任何在我35岁之后诞生的科技都是违反自然规律要遭天谴的。

Anything invented after you're thirty-five is against the natural order of things.

道格拉斯·亚当斯 *Douglas Noël Adams*

author of *The Hitchhiker's Guide to the Galaxy*

Thank you.

