

On Responsibilities and Ethics with the
Development and Fielding of AI: Reflections and Experiences

Dr. Eric Horvitz

Chief Scientific Officer, Microsoft

责任伦理与AI的发展与实践：反思与经验

Eric Horvitz博士 微软首席科学官

You have been championing AI ethics for a while. Tell us why corporations, governments and societies should care about AI ethics.

What worries you so much?

你一直在倡导AI伦理

告诉我们为什么公司、政府和社会应该关心道德规范？什么事让你这么担心？

1

00:00:00,009 --> 00:00:03,176

(Question)

4

00:00:19,440 --> 00:00:20,860

- So corporations, governments,

- 公司 政府

5

00:00:20,860 --> 00:00:23,080

and societies should care about AI ethics

社会应该关心AI伦理

6

00:00:23,080 --> 00:00:27,100

and more generally the

responsible development

更广泛地说 负责任地开发

7

00:00:27,100 --> 00:00:29,630

and fielding of AI technologies.

和应用AI技术

8

00:00:29,630 --> 00:00:33,750

Technology and its uses have
really been defining features

技术及其使用实际上已经定义了

9

00:00:33,750 --> 00:00:35,820

of our civilization as humans

我们人类的文明特性

10

00:00:35,820 --> 00:00:39,790

from stones, to fire,
to steam, to electricity

从石头 火 蒸汽到电力

11

00:00:39,790 --> 00:00:43,860

we've created and refined
technologies over the centuries.

几个世纪以来 我们创造并改进了技术

12

00:00:43,860 --> 00:00:46,410

And these technologies end up shaping us.

这些技术最终塑造了我们

13

00:00:46,410 --> 00:00:48,860

They influence who it is that we are

它们影响着我们是谁

14

00:00:48,860 --> 00:00:51,020

including our daily lives,

包括我们的日常生活

15

00:00:51,020 --> 00:00:56,020

our values, our work, and the
overall quality of our lives.

我们的价值观 我们的工作 以及我们生活的整体质量

16

00:00:56,190 --> 00:00:58,700

The technology itself is neutral

技术本身是中立的

17

00:00:58,700 --> 00:01:01,460

but we can harness
technologies in so many ways.

但是我们可以很多方面利用技术

18

00:01:01,460 --> 00:01:03,120

Some are valuable to society

有些技术对社会有价值

19

00:01:03,990 --> 00:01:08,280

and some are unfortunately
costly and destructive.

不幸的是 也有些技术让我们付出高昂代价并具有破坏性

20

00:01:08,280 --> 00:01:12,460

So decisions about how and
if we harness technology
所以关于我们如何以及是否以不同方式

21

00:01:12,460 --> 00:01:16,230

in various ways, touch on
questions of values and ethics,

利用技术的决定 则涉及到价值观念和道德层面问题

22

00:01:16,230 --> 00:01:18,153
especially when there are trade-offs.
尤其是在需要权衡取舍的情况下

23

00:01:19,060 --> 00:01:20,880
We're really at an inflection point
我们确实正处于一个转折点

24

00:01:20,880 --> 00:01:24,170
in the development and application
在我们称之为 AI 的

25

00:01:24,170 --> 00:01:28,430
of a constellation of
technologies we refer to as AI.
一系列技术的开发和应用方面

26

00:01:28,430 --> 00:01:31,500
And the upswing in AI is really fueled
而 AI 的崛起实际上

27

00:01:31,500 --> 00:01:34,770
by data, computation,
advances in algorithms
是基于数据 计算 机器学习算法的

28

00:01:34,770 --> 00:01:36,270
for machine learning,
进步

29

00:01:36,270 --> 00:01:39,160

and perception, and planning,
and natural language.

感知 规划和自然语言推动的

30

00:01:39,160 --> 00:01:43,080

And these advances promise
really incredible value to people

这些进步给人们带来了难以置信的价值

31

00:01:43,080 --> 00:01:46,740

in society, but with the
successes are coming concerns

但伴随着成功而来的是担忧

32

00:01:46,740 --> 00:01:48,780

and challenges, some quite new

和挑战 其中一些

33

00:01:49,680 --> 00:01:51,110

based on the potential effects

是基于这些技术

34

00:01:51,110 --> 00:01:55,170

that these technologies

can have on people's lives.

可能对人们生活产生的潜在影响而提出的新问题

35

00:01:55,170 --> 00:01:56,720

These concerns include safety

这些担忧包括

36

00:01:56,720 --> 00:01:59,593
and trustworthiness of AI systems.
AI 系统的安全性和可信度

37

00:02:00,500 --> 00:02:03,690
They also touch on
fairness and transparency
它们还涉及 AI 系统的公平性

38

00:02:03,690 --> 00:02:04,883
of the AI systems.
和透明度

39

00:02:06,650 --> 00:02:09,560
Many intentional as well
as inadvertent influences
许多有意或无意的 AI 影响

40

00:02:09,560 --> 00:02:12,563
once it is fielded in the open world.
特别是当它们充斥在开放的世界中时

41

00:02:13,450 --> 00:02:15,180
One way I think about this is that

我对此的看法是

42

00:02:15,180 --> 00:02:16,990

with the constellation of technologies

随着我们称之为 AI 的

43

00:02:16,990 --> 00:02:21,790

we call AI, we're coming to a

situation where automation is

一系列技术的出现 我们将面临这样一种情况：自动化

44

00:02:21,790 --> 00:02:26,220

for the first time edging

deeply into the realm

第一次深入到这个领域

45

00:02:26,220 --> 00:02:31,220

of human intellect, edging

into the world of capabilities

逐渐进入到过去我们期望

46

00:02:31,460 --> 00:02:36,460

that in the past, we would

expect humans to be doing

只有人类才具有的能力的世界

47

00:02:37,070 --> 00:02:40,170

and solely the providence

of humans, perception,

而只有人类才具有的智慧 感知

48

00:02:40,170 --> 00:02:42,360
reasoning, and understanding.
推理和理解力

49
00:02:42,360 --> 00:02:44,730
So as AI is coming into these advisory
所以当 AI 开始扮演那些具备

50
00:02:44,730 --> 00:02:47,143
and automated decision-making roles,
建议和自动决策能力角色的时候

51
00:02:49,180 --> 00:02:51,900
with interest and enthusiasm
of harnessing them in new ways
我们对能以新的方式利用它们产生了兴趣和热情

52
00:02:51,900 --> 00:02:54,780
for new accuracies and efficiencies,
它可以实现新的准确性和效率

53
00:02:54,780 --> 00:02:56,370
we have to consider the rough edges.
我们必须考虑到它的弱点和缺陷

54
00:02:56,370 --> 00:02:58,783
These bring up ethical issues.
这就引出了道德方面的问题

55
00:02:59,620 --> 00:03:02,780
We need to have discussions

and to carefully reflect
我们需要进行讨论 认真反思

56
00:03:02,780 --> 00:03:07,780
about best practices and
when necessary regulations.
最佳实践以及必要的法规

57
00:03:07,940 --> 00:03:11,030
Some of my key concerns,
my key worries include
我的主要担忧是 我的主要担心包括

58
00:03:12,060 --> 00:03:13,960
whether these systems will be applied
这些系统是否会

59
00:03:13,960 --> 00:03:17,680
in a fair way and an unbiased way versus
以公平和无偏见的方式应用

60
00:03:17,680 --> 00:03:21,640
in a way that will amplify
existing biases in our society,
而不是以一种将扩大我们社会中已有偏见的方式来应用

61
00:03:21,640 --> 00:03:24,430
the same society that provides the data.
正是同样的这个社会提供了原始的数据

62
00:03:24,430 --> 00:03:27,790

Data fueled classifiers are being used
数据驱动的分类正在被用于

63

00:03:27,790 --> 00:03:30,210
to guide high stakes decisions
指导高风险决策

64

00:03:30,210 --> 00:03:33,990
in healthcare, criminal
justice, and other areas.
如在医疗保健 刑事司法和其他领域

65

00:03:33,990 --> 00:03:35,970
And these biases can be buried quite deep
而这些偏差可能会被深埋在

66

00:03:35,970 --> 00:03:40,050
in the datasets leading to
unfair and inaccurate inferences.
数据集中 导致不公平和不准确的推论

67

00:03:40,050 --> 00:03:44,040
Other concerns include legal
issues regarding decisions made
其他担忧包括法律层面的问题

68

00:03:44,040 --> 00:03:46,710
by autonomous systems versus people
由自动系统而非人类所做的

69

00:03:46,710 --> 00:03:50,080

in terms of the goals and
the trade-offs they consider.
在考量目标以及相关取舍之间的决定

69
00:03:50,080 --> 00:03:52,180
There are definitely
challenges with the reliability
AI 系统的可靠性和安全性

70
00:03:52,180 --> 00:03:54,480
and safety of AI systems,
especially when they're used
肯定存在挑战 特别是当它们被用于

71
00:03:54,480 --> 00:03:57,623
in high stakes areas like medicine.
在医药等高风险领域时

72
00:03:59,360 --> 00:04:03,250
Other interesting challenges
include how people work
其他有趣的挑战包括人们如何工作

73
00:04:03,250 --> 00:04:05,750
or interact with AI systems,
或者与 AI 系统交互

74
00:04:05,750 --> 00:04:10,200
including whether or not
people understand what it is
包括人们是否理解是什么

75

00:04:11,040 --> 00:04:12,790

that led AI systems

导致 AI 系统

76

00:04:12,790 --> 00:04:17,550

to make particular conclusions
or recommendations.

做出特定的结论或建议

77

00:04:17,550 --> 00:04:21,470

On the economic side, there's a potential

在经济方面 有可能

78

00:04:21,470 --> 00:04:24,900

to displace workers from jobs

使工人失业

79

00:04:24,900 --> 00:04:29,140

and to amplify inequities and
wealth throughout the world.

并扩大全世界的不平等和贫富差距

80

00:04:29,140 --> 00:04:30,990

And there are rising concerns

有很多的担忧

81

00:04:30,990 --> 00:04:35,300

and discussions in other areas
as well beyond economics.

以及相关探讨发生在经济以外的其他领域

82

00:04:35,300 --> 00:04:38,810

For example, a critical discussion
topic is the potential threats

例如 一个重要的讨论主题是未来新形式的

83

00:04:38,810 --> 00:04:42,000

to civil liberties that
new forms of surveillance

监控可能对公民自由造成的

84

00:04:42,000 --> 00:04:43,853

could power.

潜在威胁

85

00:04:45,430 --> 00:04:47,870

Other questions that are coming up include
未来出现的其他问题还包括

86

00:04:47,870 --> 00:04:51,700

the role of AI and military applications,
AI的角色及其在军事方面的应用

87

00:04:51,700 --> 00:04:53,740

both in peace time and war.

无论是和平时期还是战争时期

88

00:04:53,740 --> 00:04:57,350

For example, with all the interest

例如 尽管有很多的热度

89

00:04:57,350 --> 00:05:02,110

on what AI might bring into
the world of defense,

在AI给国防领域带来的应用方面

90

00:05:02,110 --> 00:05:05,060

AI technologies could
actually be de-stabilizing

但AI技术实际上可能会降低稳定性

91

00:05:05,060 --> 00:05:08,700

by introducing more
complexity and uncertainty,

因其带来更多的复杂性和不确定性

92

00:05:08,700 --> 00:05:10,680

especially when AI systems

尤其是当多角度的AI系统

93

00:05:10,680 --> 00:05:14,530

on multiple sides interact
in unexpected ways

以意想不到的方式交互

94

00:05:14,530 --> 00:05:18,380

leading to poor outcomes,
such as unwanted escalations.
从而导致不良结果 如不必要的事态升级

95
00:05:18,380 --> 00:05:21,940
And there are also deep concerns
with machines being used
此外 人们深深担忧机器被用于

96
00:05:21,940 --> 00:05:24,653
in military settings to
make automated decisions.
在军事综合体中进行自动化决策

97
00:05:25,860 --> 00:05:30,560
So given the rise and the applicability
所以考虑到这些强大的

98
00:05:30,560 --> 00:05:32,070
of these powerful AI technologies,
AI技术的兴起和应用

99
00:05:32,070 --> 00:05:34,570
it's going to be important
to have discussions
重要的是要

100
00:05:34,570 --> 00:05:37,930
with academia and also across corporations
与学术界 企业和民间社会组织

101

00:05:37,930 --> 00:05:40,635
and civil society organizations
进行广泛讨论

102
00:05:40,635 --> 00:05:44,500
to bring people together to
understand these problems
让人们聚在一起了解这些问题

103
00:05:44,500 --> 00:05:47,450
and to share in each other's expertise.
并分享彼此的专业知识

104
00:05:47,450 --> 00:05:48,283
It's also going to be important
同样重要的是

105
00:05:48,283 --> 00:05:50,370
for government agencies to explore
政府机构进行探索

106
00:05:50,370 --> 00:05:53,150
and for multiple governments
across the world to coordinate
世界各地的政府进行协调合作

107
00:05:53,150 --> 00:05:57,640
on principles, and best
practices, and norms.
制定原则 最佳实践和规范

108

00:05:57,640 --> 00:05:59,550

We can do a lot more
than we're doing today.

我们能做的比现在要多得多

109

00:05:59,550 --> 00:06:03,030

And so, there are great
opportunities ahead

我们面临着巨大的机会共同探索

110

00:06:03,030 --> 00:06:05,280

for getting together on these technologies

随着这些技术

110

00:06:05,280 --> 00:06:07,883

as they edge into the
realm of human intellect.

逐渐浸入人类智慧领域

Can you tell us more about the Partnership on AI?

How can Chinese companies get more involved in the organization?

您能否详细介绍一下 AI 伙伴机构?

中国公司如何才能更多地参与这一组织?

111

00:06:17,440 --> 00:06:20,900

The partnership on AI is
a nonprofit organization

AI伙伴机构是一个非营利性组织

112

00:06:20,900 --> 00:06:24,010

that was formed four years ago.

成立于四年前

113

00:06:24,010 --> 00:06:27,070

The idea was to build an open platform

我们的初衷是想建立一个开放的平台

114

00:06:27,070 --> 00:06:31,360

for doing studies, developing

and sharing best practices,

来进行研究 发展和分享最佳实践

115

00:06:31,360 --> 00:06:33,270

bringing together the best ideas

汇集关于AI和人们

116

00:06:33,270 --> 00:06:36,030

on the influences of AI

and people in society.

在社会中的影响的最好的想法

117

00:06:36,030 --> 00:06:39,620

Now, the main impetus for

the initiative has been

现在 该倡议的主要动力是

118

00:06:39,620 --> 00:06:44,070

to bring together diverse

stakeholders, researchers

将不同的利益相关者 学术界的

119

00:06:44,070 --> 00:06:49,070

in academia, business

leaders, policy makers,

研究人员 商业领袖 政策制定者

120

00:06:49,660 --> 00:06:52,910

and others to get together

和其他人聚在一起

121

00:06:52,910 --> 00:06:56,550

and come to the same page

on these challenging issues.

在那些富有挑战性的问题上达成共识

122

00:06:56,550 --> 00:06:58,790

The partnership was initially founded

这个伙伴组织最初是由

123

00:06:58,790 --> 00:07:01,280

by concerned AI scientists

来自世界上一些

124

00:07:01,280 --> 00:07:03,580

from some of the world's

largest AI companies.

最大的AI公司的AI科学家创立的

125

00:07:04,640 --> 00:07:07,880

Back in 2015 or so, AI scientists,

早在2015年左右 AI科学家

126

00:07:07,880 --> 00:07:10,240

including myself, got together,

包括我自己在内 我们聚在一起

127

00:07:10,240 --> 00:07:13,190

and we shepherded our companies,
关照引导我们的公司

128

00:07:13,190 --> 00:07:14,980

and these include the founding companies,
其中包括创始公司

129

00:07:14,980 --> 00:07:19,210

Microsoft, Apple, Amazon, Facebook, Google
微软 苹果 亚马逊 Facebook 谷歌

130

00:07:19,210 --> 00:07:22,140

with DeepMind and IBM to join up
以及DeepMind和IBM

131

00:07:22,140 --> 00:07:23,760

with non-profit organizations,
非营利组织

132

00:07:23,760 --> 00:07:25,440

and these included academic teams,
其中包括学术团队

133

00:07:25,440 --> 00:07:29,330

civil society groups like
the ACLU, and nonprofit AI
公民社会团体 如美国公民自由联盟

134

00:07:29,330 --> 00:07:33,283

research teams like OpenAI
and the Allen AI Institute.

和非营利的AI研究团队 如OpenAI和艾伦AI研究所等

135

00:07:34,800 --> 00:07:37,010

The partnership is now led
by a board of directors

该伙伴组织现在由一个在营利性

136

00:07:37,010 --> 00:07:40,810

that's balanced between
for-profit and non-profit.

和非营利性之间保持平衡的董事会领导

137

00:07:40,810 --> 00:07:44,700

There are over a hundred
member organizations involved.

有一百多个成员组织参与其中

138

00:07:44,700 --> 00:07:48,520

You can read more about the
membership and the group

你可以在 AI 伙伴关系阅读更多关于成员资格

139

00:07:48,520 --> 00:07:51,823

at partnership on AI-- one word -- dot org.

和小组信息资讯 网址是 oneword.org

140

00:07:53,220 --> 00:07:56,680

It's great to have this group together.

很高兴能与这些团体在一起

141

00:07:56,680 --> 00:08:00,360

Baidu was the first Chinese
company to join the partnership.

百度是第一个加入合作伙伴组织的中国公司

142

00:08:00,360 --> 00:08:02,250

Chinese companies can get
involved by reaching out

中国企业可以通过该网站的

143

00:08:02,250 --> 00:08:05,730

through the membership link on that site.

会员链接参与其中

144

00:08:05,730 --> 00:08:08,490

You can go to that site too,
to read more about the tenets

你也可以去那个网站 阅读更多关于成员的

145

00:08:08,490 --> 00:08:10,930

or basic principles, starting with

基本原则 首先是

146

00:08:10,930 --> 00:08:15,300

the first basic founding
tenant of the organization.

从该组织的第一个基本创始成员开始

147

00:08:15,300 --> 00:08:18,810

We seek to ensure that AI

technologies benefit and empower
我们致力于确保AI技术普惠赋能

148

00:08:18,810 --> 00:08:21,210

as many people as possible.

尽可能多的人

149

00:08:21,210 --> 00:08:24,070

On the site, you can also

see high quality reports

在网站上 你还可以看到高质量的报告

150

00:08:24,070 --> 00:08:27,500

on criminal justice, facial recognition,

包括刑事司法 面部识别

151

00:08:27,500 --> 00:08:30,390

--the use of risk scores

in criminal justice--

在刑事司法中使用风险评分

152

00:08:30,390 --> 00:08:32,310

as well as some interesting technical work

还包括整个社区的一些有趣的

153

00:08:32,310 --> 00:08:34,040

across the community, like coming up

技术工作 比如

154

00:08:34,040 --> 00:08:37,803

with shared standards for documentation,

制定文档共享标准

155

00:08:39,080 --> 00:08:42,620

documenting the data, the
model, and its accuracy,

记录数据 模型及其准确性

156

00:08:42,620 --> 00:08:44,540

and performance, the maintenance needs

和性能 以及整个

157

00:08:44,540 --> 00:08:48,053

across the whole machine
learning lifecycle.

机器学习生命周期中的维护需求

158

00:08:49,190 --> 00:08:50,690

And so, looking ahead,
we'll be continuing to work

因此 展望未来 我们将继续

159

00:08:50,690 --> 00:08:52,810

on these key pillars and it'd be great

在这些关键领域中开展工作

160

00:08:52,810 --> 00:08:54,360

to have more people involved

让更多的人参与进来

161

00:08:54,360 --> 00:08:57,561

and more organizations

involved in the partnership.
让更多的组织成为伙伴是一件好事

You have been leading the Aether committee in Microsoft, overseeing the AI ethics work across the company. What lessons have you learned in the past few years that you may share with us?

你一直在领导微软的 Aether 委员会，监督整个公司的 AI 伦理工作。在过去的几年中，您有哪些经验教训可以与我们分享？

162
00:08:57,561 --> 00:09:00,811
(Question)

163
00:09:07,860 --> 00:09:12,830
It's been a journey working to set up
在微软建立

164
00:09:12,830 --> 00:09:16,900
and to lead the Aether
Committee at Microsoft.
和领导Aether人工智能伦理道德委员会是一段漫长的历程

165
00:09:16,900 --> 00:09:19,350
Aether stands for
Aether代表

166
00:09:19,350 --> 00:09:24,110
AI Ethics and Effects in
Engineering and Research.
AI Ethics and Effects in Engineering and Research (人工智能伦理道德及其在工程与研究
中的影响)

167
00:09:24,110 --> 00:09:26,150

I collaborated with Brad Smith and Harry Shum
四年前 我与布拉德·史密斯和沈向洋

168

00:09:26,150 --> 00:09:28,030
to set up the Aether committee
共同创立了道德委员会

169

00:09:28,030 --> 00:09:30,870
and its working groups four years ago.
及其工作小组

170

00:09:30,870 --> 00:09:33,610
The purpose of Aether is to
work across all of Microsoft
Aether的宗旨是 与微软全体员工

171

00:09:33,610 --> 00:09:38,610
and with its partners on
rising questions, challenges,
并与其合作伙伴一道 共同应对在AI和相关技术的

172

00:09:38,620 --> 00:09:41,360
and opportunities ahead
with the development
开发和应用过程中

173

00:09:41,360 --> 00:09:44,760
and fielding of AI and
related technologies.
不断出现的问题 挑战和机遇

174

00:09:44,760 --> 00:09:49,760

The group considers potential rough edges, concerns,
该小组考量潜在的缺陷和担忧

175

00:09:49,860 --> 00:09:51,930

and policy coming to the fore
以及随着新型的工具

176

00:09:51,930 --> 00:09:54,970

with new kinds of tools, applications,
新的应用程序的出现而产生的政策

177

00:09:54,970 --> 00:09:56,530

and thinks through best practices
并思考公司的

178

00:09:56,530 --> 00:09:58,263

and policies for the company.
最佳实践和政策

179

00:09:59,160 --> 00:10:03,960

It's interesting because
it's a company-wide process
这很有趣 因为这是全公司范围的工作

180

00:10:03,960 --> 00:10:05,780

and committee that works
委员会

181

00:10:05,780 --> 00:10:10,460

with corporate divisions as well as
与公司各部门以及

182

00:10:10,460 --> 00:10:14,060

with the leadership of the
company on company-wide policies.
公司领导层一起制定公司政策

183

00:10:14,060 --> 00:10:17,980

There's a main committee and
then a set of working groups
先设立一个主要委员会 然后是一些工作小组

184

00:10:17,980 --> 00:10:19,690

that are cauldrons of creativity
它们是思考具体议题的

185

00:10:19,690 --> 00:10:24,170

as to thinking through specific topics.
创造性源地

186

00:10:24,170 --> 00:10:25,810

There are six working groups now
现在有六个工作小组

187

00:10:25,810 --> 00:10:29,350

and they cover what we call
Microsoft's AI principles.
它们涵盖了微软AI原则的所有方面

188

00:10:29,350 --> 00:10:31,500

And these include privacy,
其中包括隐私

189
00:10:31,500 --> 00:10:34,750
and security, fairness and inclusiveness,
安全 公平 包容

190
00:10:34,750 --> 00:10:38,040
transparency, reliability and safety,
透明度 可靠性和安全性

191
00:10:38,040 --> 00:10:40,760
and others such as Human AI Collaboration
以及人类与AI的协作 (Human AI Collaboration)

192
00:10:40,760 --> 00:10:43,020
and a sensitive uses panel.
和敏感问题小组

193
00:10:43,020 --> 00:10:44,410
The sensitive uses panel actually
敏感小组实际上

194
00:10:44,410 --> 00:10:47,440
is a cross-company panel that deliberates
是一个跨公司各部门的小组

195
00:10:47,440 --> 00:10:49,800
on specific technologies
coming to the fore
专门讨论正在出现的特定技术

196

00:10:49,800 --> 00:10:54,800

and questions on ethics,
especially for high stakes areas.

及其引发的道德问题 特别是在高风险领域

197

00:10:55,300 --> 00:10:57,500

And that committee has
made recommendations

该委员会还向微软的

198

00:10:57,500 --> 00:10:59,770

to Microsoft's leadership team

领导团队提出了

199

00:10:59,770 --> 00:11:03,400

on policies moving forward
for how technologies are used

关于如何使用和共享技术的政策建议

200

00:11:03,400 --> 00:11:06,223

and shared whether they
should be gated or not,

包括这些技术是否应该被管控

201

00:11:07,080 --> 00:11:09,310

protected from general use, for example,

防止它们不被滥用 例如

202

00:11:09,310 --> 00:11:13,223

given the prospect that they

could be used in costly ways.
从它们是否会造成高昂代价的可能性的角度考量

203
00:11:14,390 --> 00:11:19,364
I've learned several
lessons over the years
这些年来我在建造Aether

204
00:11:19,364 --> 00:11:23,190
with building Aether
and leading the effort.
和领导的努力中学到了经验教训

205
00:11:23,190 --> 00:11:26,750
The first is that it takes
many teams coming together
首先 它需要很多团队团结在一起

206
00:11:26,750 --> 00:11:30,750
to make an effective,
responsible AI program
才能制作有效 负责任的AI程序

207
00:11:30,750 --> 00:11:32,940
at an organization like Microsoft,
在微软这样的组织中

208
00:11:32,940 --> 00:11:34,540
given the breadth of the company
要考虑到公司的

209

00:11:35,590 --> 00:11:39,130
and the span of its products and services
产品 服务

210
00:11:39,130 --> 00:11:40,363
and its businesses.
和业务的广度

211
00:11:41,240 --> 00:11:43,480
Second, I've learned it's important
其次 我了解到了

212
00:11:43,480 --> 00:11:47,330
really to have the top
leadership, right up to the CEO,
最高层领导的重要性 一直到首席执行官

213
00:11:47,330 --> 00:11:51,040
in our case, Satya
Nadella, being enthusiastic
比如我们的萨蒂亚·纳德拉

214
00:11:51,040 --> 00:11:53,133
and supportive, to have our backs.
他热情地全力支持我们

215
00:11:54,280 --> 00:11:57,670
Third, I've learned that
it can be quite difficult
第三 我学到, 有一定难度的是

216

00:11:57,670 --> 00:12:02,670
to go from the ongoing flurry of ideas
要梳理不断涌现的各种想法

217
00:12:05,760 --> 00:12:10,760
about ethics in AI into
concrete fielded policies
从 AI 中的伦理道德转化为具体领域的政策

218
00:12:10,760 --> 00:12:15,320
and technologies, making it a reality.
和技术 并使之成为现实

219
00:12:15,320 --> 00:12:19,470
And this includes work
on taking nascent ideas
这包括采纳新兴的想法

220
00:12:19,470 --> 00:12:23,650
and translating them into
real-world products and services,
并将其转化为现实世界的产品和服务

221
00:12:23,650 --> 00:12:26,860
Including: work on
governance, taking prototypes
包括治理方面的工作 建立原型样本

222
00:12:26,860 --> 00:12:30,893
making them hardened for
actual corporate use.
使它们更加坚固 以便于实际的企业使用

223

00:12:32,090 --> 00:12:33,490

You can learn more about
some of the efforts

您可以在我们的网站上

224

00:12:33,490 --> 00:12:37,780

that we've done and some of
the successes on our websites,

了解更多关于我们所做的努力以及取得的初步成绩

225

00:12:37,780 --> 00:12:42,420

including the sites that
GitHub and microsoft.com.

包括GitHub和microsoft.com

226

00:12:42,420 --> 00:12:46,400

On GitHub, I recommend people
take a look at InterpretML,

在GitHub上 我建议大家看看InterpretML

227

00:12:46,400 --> 00:12:49,340

a site with code and examples

那是关于构建更透明

228

00:12:49,340 --> 00:12:52,130

on building more transparent AI.

AI的代码和示例的网站

229

00:12:52,130 --> 00:12:55,530

Also, Fairlearn, a tool

还有 Fairlearn 作为

230

00:12:55,530 --> 00:12:59,170

and examples for how to
understand biases and fairness

如何理解AI应用程序中的数据集中的

231

00:12:59,170 --> 00:13:02,040

in our data sets in AI applications.

偏见和公平性的工具与案例

232

00:13:02,040 --> 00:13:04,560

You can learn more about
Aether more broadly

你可以更广泛地了解Aether

233

00:13:04,560 --> 00:13:07,870

by searching on our
approach to responsible AI

通过搜索微软在负责任的人工智能方面采取的方法

234

00:13:07,870 --> 00:13:10,010

at Microsoft and read about the principles

并阅读我们将新兴想法

235

00:13:10,010 --> 00:13:11,760

and some of the ways we reduced

转化为实践时

236

00:13:11,760 --> 00:13:15,250

nascent ideas into practices.

所遵循的原则

237

00:13:15,250 --> 00:13:17,900

Just to end on a few other
lessons that I've learned,
最后讲几个我学到的教训

238

00:13:19,200 --> 00:13:22,060

it's clear to me that
leadership must be ongoing.
我很清楚 领导力必须是持续的

239

00:13:22,060 --> 00:13:25,180

You can't just start
something and have it pop
开始一个事业, 不能听之任之

240

00:13:25,180 --> 00:13:28,960

or just spontaneously
combust across the company.
期望它能在整个公司中自然生长

241

00:13:28,960 --> 00:13:33,283

Leadership involvement must
be consistent and continual.
领导层的参与必须始终如一和持续不断

242

00:13:34,320 --> 00:13:35,930

It's also never too late to start.
即使现在开始也永远不会太晚

243

00:13:35,930 --> 00:13:38,243
I wish we founded Aether earlier.
我希望我们更早地建立Aether

244
00:13:40,380 --> 00:13:42,760
And finally, let me say
that I'm really proud
最后，我想说我真的很自豪

245
00:13:42,760 --> 00:13:45,900
to see how far we've gotten
看到我们在这么短的时间里

246
00:13:45,900 --> 00:13:48,630
in so short a time, how Aether has grown.
取得了许多成绩
看到 Aether 在是如何成长的

247
00:13:48,630 --> 00:13:51,084
We've created other
organizations to support
我们还创建了其他相关机构

248
00:13:51,084 --> 00:13:53,143
the effort as well.
来支持这些工作

249
00:13:54,830 --> 00:13:56,130
[Pause] I want to extend my thanks
我要感谢

250

00:13:56,130 --> 00:13:58,020
to Harry and to the organizing committee
沈向洋和组委会

251
00:13:58,020 --> 00:14:01,550
for bringing this great workshop together.
感谢他们举办了这次了不起的研讨会

252
00:14:01,550 --> 00:14:03,350
It looks like there are
great topics being discussed.
看起来有很好的话题在讨论

253
00:14:03,350 --> 00:14:07,690
There's so much to do, and
it's best to work together,
要做的事情太多
能一起工作真的了不起

254
00:14:07,690 --> 00:14:12,090
whether we're separated by
oceans, by organizations,
不管我们是远隔重洋
还是来自不同机构

255
00:14:12,090 --> 00:14:13,470
or by governments.
或是来自不同的政府

256
00:14:13,470 --> 00:14:14,528
Thanks very much.

非常谢谢

257

00:14:14,528 --> 00:14:17,695

(gentle upbeat music)