

# 用户体验 面面观

——方法、工具与实践

Observing the User Experience:  
A Practitioner's Guide  
to User Research

(美) Mike Kuniavsky 著

汤海译

李鸿审校



清华大学出版社

哇！我们这些年提炼和使用的如此之多的用户体验研究方法，如今都在这本书中得以清楚地呈现和描述！

——Christian Rohrer，雅虎用户体验研究经理

我为这本书着迷！这本内容广泛的指南通过一种独特和创新的方式来阐述用户体验研究，行文通俗易懂，能让读者拥有一种很好的阅读体验。

——Kelly Braun，易趣可用性经理

您了解您的用户吗？设计师心目中的用户和实际情况中的用户之间存在着很大的鸿沟，这是阻碍产品或服务走向成功的最大问题。《用户体验面面观》有助于了解用户需求和产品需求，从而消除这一鸿沟，让用户能够真正使用设计师所创造的产品或服务。

本书充满实际经验和大量实用信息，展现了一个完整的方法工具箱，旨在帮助设计师和开发人员从用户的角度来看待问题。本书深入介绍了13种用户体验研究方法，这些方法是开发卓越产品的基础，适用于Web应用、软件和移动产品。此外，本书还考量了真实世界中的软件开发过程，涉及了预算紧张、日程紧密以及现有流程这些因素。

## 本书特色

- 解释如何在保持原创、新颖和独特的基础上增强产品的可用性
- 适合任何有志于创新产品（服务）的人员阅读
- 从研究角度结合真实事例，从纲的角度引发读者思考
- 针对如何低成本、高速度进行用户研究和如何通过强有力的方式展示结果提供真知灼见
- 向读者提供工具并增强其信心，鼓励他们根据自己的设计来执行用户研究，并优化用户体验来满足产品和用户的独特需求

## 关于作者

**迈克·库涅夫斯基**是旧金山用户体验咨询公司Adaptive Path的创始合伙人。他从1994年开始从事开发商业网站，是获奖搜索引擎HotBot的交互设计师。他的设计作品和著作发表于很多出版物，包括WebMonkey，ID Magazine，《连线》，《华尔街日报》和《洛杉矶时报》。



**MK**

MORGAN KAUFMANN

MORGAN KAUFMANN PUBLISHERS

AN IMPRINT OF ELSEVIER

www.mkp.com

ISBN 978-7-302-22353-5



9 787302 223535 >

定价：69.00元



好设计，更有爱，关注微信



idesign123



# 用户体验面面观 ——方法、工具与实践

(美) Mike Kuniavsky 著

清华大学出版社  
北 京



好设计,更有爱,关注微信



idesign123

## 内 容 简 介

这是一本专注于用户研究和用户体验的经典读物,同时也是一本容易上手的实战手册,从实践者的角度,着重讨论和阐述了用户研究的重要性、主要的用户研究方法和工具,同时借助于实例介绍了相关的应用。全书共3部分18章,深度剖析了何为优秀的用户设计,用户体验包括哪些研究方法和工具,如何得出和分析用户体验调查结果等。

本书适合设计人员和高等院校设计类学生阅读和参考,也适合任何一个希望有所建树的各界策划和设计人员阅读。

Observing the User Experience: A Practitioner's Guide to User Research, 1st Edition

Mike Kuniavsky

ISBN: 9781558609235

Copyright © 2003 by Elsevier. All rights reserved.

Authorized Simplified Chinese translation edition published by Elsevier (Singapore) Pte Ltd Press and Tsinghua University.

ISBN: 9789812724816

Copyright © 2010 by Elsevier (Singapore) Pte Ltd and Tsinghua University Press. All rights reserved.

Published in China by Tsinghua University Press under special arrangement with Elsevier (Singapore) Pte Ltd.. This edition is authorized for sale in China only, excluding Hong Kong SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由 Elsevier (Singapore) Pte Ltd.授予清华大学出版社在中国大陆地区(不包括香港、澳门特别行政区以及台湾地区)出版与发行。未经许可之出口,视为违反著作权法,将受法律之制裁。

本书封底贴有 Elsevier 防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

北京市版权局著作权合同登记号 图字:01-2009-7992

图书在版编目(CIP)数据

用户体验面面观——方法、工具与实践/(美)库涅夫斯基(Kuniavsky, M.)著;汤海译;李鸿审校.  
—北京:清华大学出版社,2010.5

书名原文: Observing the User Experience: A Practitioner's Guide to User Research

ISBN 978-7-302-22353-5

I. ①用… II. ①库… ②汤… ③李… III. ①用户界面—程序设计 IV. TP311.1

中国版本图书馆 CIP 数据核字(2010)第 054173 号

责任编辑:文开琪

责任印制:孟凡玉

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京市清华园校印厂

装 订 者:三河市新茂装订有限公司

经 销:全国新华书店

开 本:185×230 印 张:28.75 插 页:2 字 数:705 千字

版 次:2010 年 5 月第 1 版 印 次:2010 年 5 月第 1 次印刷

印 数:1~4000

定 价:69.00 元

产品编号:034685-01



好设计,更有爱,关注微信

idesign123

## Mike Kuniavsky 致中国读者

I am deeply honored to have this book translated into Chinese. Chinese speakers are one of the world's fastest growing populations of Internet users. Up to December 31 in 2009, 384 million Chinese people go online, 86 million for the first time.

With every new group of users, especially one this large, the Internet changes. A products or service that worked for one group won't work for another. What worked last year may not work next year.

In this dynamic environment, designers, developers and managers are challenged to keep up with the shifting needs of their audience, usually with limited resources and under time pressure.

I believe that close observation of many different people from a variety of perspectives is the key to designing products that are popular and successful, while reducing unnecessary engineering and design. Because of the speed of technology adoption, the needs of Chinese language technology markets may be the worlds greatest. Fortunately, the opportunities for observing people in this environment are enormous. Users are excited to explore the possibilities of the technology and talk about it while companies are driven to find out how they can make their products more profitable, more popular and more effective.

It's a great time to be working in this environment. I hope you find it as exciting as I do, and I hope you find the techniques described in this book useful.

我深感荣幸本书被译成中文。说中文的人是世界上增长最快的互联网用户。截至 2009 年 12 月 31 日，中国网民达 3.84 亿，其中新增用户 8600 万。

每产生一批新用户，尤其是如此庞大的用户，互联网都会发生变化。以前对一个群体能发挥作用的产品或服务不再适用于另一个群体。去年能发挥作用的明年可能就无法发挥作用。

这种环境充满活力，设计师、开发人员和管理者面临着要跟上受众不断变换需求的挑战，而且通常资源有限，还要承受时间压力。

我相信，从各种角度密切观察不同的人，是设计既能受到欢迎、产品又能获得成功的关键，同时能减少不必要的工程和设计。由于技术采用的速度，中文语言的科技市场需求可能是世界上最大的。幸运的是，在这种环境下观察人的机会很多。



用户很兴奋能够探索技术的可能性，并进行讨论，而公司会受到驱动，来发现如何能够让产品更赚钱，更受欢迎、更有效。

在这种环境中工作，现在是好时机。我希望你和我一样觉得兴奋，我也希望你觉得本书中描述的方法有用。

(Cassidy Curtis, otherthings.com 版权所有)



# 互联网大佬推荐的必读书

周鸿祎、蔡文胜、雷军、李开复、马化腾、张小龙等互联网大佬及腾讯、阿里等高级产品经理推荐的书籍，互联网从业者及产品经理必读。










书籍资源	
	<div>1. IT 不再重要</div> <div>书籍简介：云计算与信息整合，蔡文胜强力推荐（购买了数千本送人）。</div> <div>豆瓣评分：7.6</div>
	<div>2. 精益创业</div> <div>书籍简介：偏重管理，介绍如何快速开发最精简的产品得到对用户最有价值的认知，李开复推荐。</div> <div>豆瓣评分：8.5</div>
	<div>3. 女人的起源</div> <div>书籍简介：做产品对人性的把握最重要，女性占据了一般的用户，张小龙推荐。</div> <div>豆瓣评分：7.6</div>
	<div>4. 认知盈余</div> <div>书籍简介：这是一个分享的时代，马化腾称作者为“互联网革命最伟大的思考者”。</div> <div>豆瓣评分：7.4</div>
	<div>5. 至关重要的设计</div> <div>书籍简介：前苹果首席设计师罗伯特·布伦纳写的设计书籍，雷军推荐。</div> <div>豆瓣评分：7.2</div>
	<div>6. 简约至上</div> <div>书籍简介：提出了交互设计的四条策略：删除、组织、隐藏、转移，很实用。</div> <div>豆瓣评分：8.4</div>
	<div>7. 点石成金：访客至上的网页设计秘籍</div> <div>书籍简介：设计应当不让用户思考，对网站设计的可用性提出了不少建议。</div> <div>豆瓣评分：8.5</div>
	<div>8. 学会提问：批判性思维指南</div> <div>书籍简介：锻炼批判性思维。</div> <div>豆瓣评分：8.7</div>



好设计，更有爱，关注微信



idesign123

	<p>9. 精益求精：卓越的互联网产品设计与管理</p> <p>书籍简介：偏重理论，案例相对较少，不适合初级产品经理、产品助理。</p> <p>豆瓣评分：7.3</p>
	<p>10. 掌握需求过程</p> <p>书籍简介：偏重需求挖掘。</p> <p>豆瓣评分：7.8</p>
	<p>11. 网站设计解构 - 有效的交互设计框架和模式</p> <p>书籍简介：介绍网站框架体系，适合 Web 产品经理。</p> <p>豆瓣评分：7.0</p>
	<p>12. 瞬间之美：WEB 界面设计如何让用户心动</p> <p>书籍简介：介绍 Web 界面设计，适合 Web 产品经理。</p> <p>豆瓣评分：8.0</p>
	<p>13. GUI 设计禁忌 2.10</p> <p>书籍简介：偏重于客户端产品经理。</p> <p>豆瓣评分：7.4</p>
	<p>14. About Face 3 交互设计精髓</p> <p>书籍简介：建议客户端产品经理仔细看看</p> <p>豆瓣评分：8.8</p>
	<p>15. 用户体验度量</p> <p>书籍简介：有一定用户群的产品经理可以看看</p> <p>豆瓣评分：8.1</p>
	<p>16. Web 导航设计</p> <p>书籍简介：偏重 Web 设计，但客户端产品经理也可以看看。</p> <p>豆瓣评分：8.3</p>
	<p>17. 锦绣蓝图：怎样规划令人流连忘返的网站</p> <p>书籍简介：Web 信息架构必读</p> <p>豆瓣评分：8.0</p>



好设计,更有爱,关注微信



idesign123



18. 赢在用户—WEB 人物角色创建和应用实践指南

书籍简介：创建人物角色的方式做用研。  
豆瓣评分：7.9



19. 重来

书籍简介：可以一看，周鸿祎等不少大佬都推荐过。  
豆瓣评分：8.4



20. 故事的领导力

书籍简介：锻炼沟通能力。  
豆瓣评分：7.3



21. 人月神话

书籍简介：软件工程经典著作。  
豆瓣评分：8.5



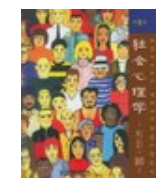
22. 情感化设计

书籍简介：做产品，理解情感很重要。  
豆瓣评分：8.0



23. 经济学的思维方式

书籍简介：形成经济学的思维模式，很有用。  
豆瓣评分：8.8



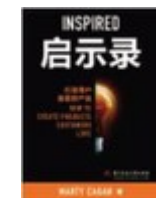
24. 社会心理学

书籍简介：做产品就是找到用户的心理诉求，然后满足他们。  
豆瓣评分：8.9



25. 用户体验面面观

书籍简介：有方法、工具和实践。  
豆瓣评分：8.1



26. 启示录-打造用户喜爱的产品

书籍简介：适合有一定产品经验的，非常棒的一本书。  
豆瓣评分：8.5





	<div>27. 世界是平的</div> <div>书籍简介：了解世界的扁平化趋势</div> <div>豆瓣评分：7.7</div>
	<div>28. 设计中的设计</div> <div>书籍简介：无印良品 设计师原研哉的书。</div> <div>豆瓣评分：8.7</div>
	<div>29. 三双鞋</div> <div>书籍简介：美捷步创始人谢家华自述，感受如何把体验做到极致。</div> <div>豆瓣评分：8.3</div>
	<div>30. 产品经理手册</div> <div>书籍简介：偏重传统行业，但有利于拓展思维和视野。</div> <div>豆瓣评分：7.7</div>
	<div>31. 结网</div> <div>书籍简介：偏重商业战略和思维，理论性较强。</div> <div>豆瓣评分：8.0</div>
	<div>32. 人人都是产品经理</div> <div>书籍简介：侧重产品流程、技巧，适合初学者了解产品经理，但阿里的风格太强。</div> <div>豆瓣评分：7.8</div>
	<div>33. 失控</div> <div>书籍简介：思考人类进化，预言了维基百科等，将生物进化引入机器，大部头一般人很难啃。</div> <div>豆瓣评分：8.9</div>
	<div>34. 用户体验的要素</div> <div>书籍简介：将开发流程五个层次：战略层、范围层、结构层、框架层、表现层，纲领性书籍，值得一读。</div> <div>豆瓣评分：8.1</div>

更多互联网书籍下载：

<http://pan.baidu.com/share/link?shareid=4290411135&uk=151033183#dir/path=/互联网大佬推荐的必读书>



好设计，更有爱，关注微信



idesign123

## 推荐序

我和迈克·库涅夫斯基相识于1994，那时我正在为我的第一本书《设计 Web 图形》(Designing Web Graphics)做研究。我找到一个叫 hothothot.com 的网站，想把它描述出来，因为它概念好，执行好，做得也漂亮。那个时候，实际上非常难找到有任何美感的网站！在互联网上发现这个漂亮的网站之后，我才知道 hothothot.com 网站的设计公司实际上就在离我不远的地方。我和迈克进行了个人交流，我们俩开始了职业上的亲密关系和友谊。迈克搬到了湾区，而我待在南加州，那些年网络世界起起伏伏，但我们始终保持着联系。我之前从未读过迈克写的任何东西，一直到邮箱中收到本书的手稿。我有点担心我会不喜欢他写的书，因为以前就有人写过这个主题，而且有几本写得还不错。我不敢肯定这本书是不是属于“没有主见”的书，也不敢肯定书中有没有原创的东西。

读完之后，我能证明市场上没有这样的书。而且我想用所有可能的最佳方式来表达我的意思！本书采用对话风格，就好像迈克坐在你旁边一样。他坐在旁边的好处就是他有成堆的真实网站测试和研究经验。大多数其他作者都把自己当成咨询师或者专家，但并没有像迈克一样真正进行过自己的项目。迈克并不是告诉你他是全能教条专家，只是像老兵一样愿意分享自己的战争经历和胜利。

这本书非常好。我知道你会认为我有偏见，但有一个原因让我从一开始就喜欢迈克——他了解他从事的工作，他还用幽默、慷慨和智慧来分享他知道的東西。

——琳达·温曼，作家及 lynda.com 公司创始人



好设计，更有爱，关注微信



idesign123

# 译者序：处处皆体验

翻译书就像设计开发产品一样，要从了解出版社的商业需求和读者需求开始。

和出版社商定期限和要求等之后，接下来就是了解读者对象。

本书主要是针对用户体验入门者，还是中高级人员？如果是前者，该如何措辞？语序如何安排？他们是否能明白术语？有哪些术语？如果是后者，情况又当如何？要考虑的因素不一而足。

翻译不同于其他产品，译者可创造的空间有限，因为原书作者思想和写作风格已经对翻译做出了限制，前述大部分问题都需要仔细阅读原著来找到答案。但这并不代表译者不用考虑读者的想法。

从我的阅读和分析来看，本书最适合中级读者，同时又能很好满足新手和高级人员的需要。新手可以快速明白什么是用户体验，有哪些方法可用，用户体验能发现什么问题。高级用户可以把本书当作手册，检查自己是否正确理解了用户体验，是否正确灵活使用了方法，是否大胆创新了方法。而中级人员可以利用本书迅速完成从新手到高级人员的转变过程，练习并掌握这些方法的要领。作者更特别强调书中提到的方法用于发现和定义问题，而不是解决问题，因为不可能去解决没发现的问题。

原著通俗易懂，实例丰富，读起来每每让人娓娓不倦。对于很多读者而言，作者 Mike Kuniavsky 虽然没有如雷贯耳的名头，但他绝对是用户体验领域的先行者、实践者和行家，为很多公司设计和开发过产品，他同时还是 Adaptive Path 的创始人之一。

在翻译过程中，Mike Kuniavsky 给予我很多帮助，在此向他表示感谢。在此过程中，我也帮他纠正了书中的一些错误:)

感谢清华大学出版社文开琪的大力帮助。

感谢李鸿的火眼金睛，帮助我发现并弥补了译稿中的疏漏。

感谢 Joe 博士带我接触用户体验。

感谢 Little Lion、大灰狼和卡西带来的快乐和帮助。





## 用户体验面面观——方法、工具与实践

感谢分我烟抽的同事。

感谢大好人帮我处理书中的图片。

感谢父母和其他家人一直对我默默无私的付出。

书中如有错误和疏漏之处，还请各位读者赐教讨论。我的邮件地址是  
haitang@gmail.com。

希望本书能为从事用户体验工作的同行带来一些帮助。

汤 海

2010.5



# 前言

## 为什么要阅读本书？

因为某种原因，你选择了本书。你渴望知道谁在使用你的产品。你也觉得你是对的。确定客户，了解他们的期望，调查他们的需求，这些都是设计任何一个产品的起点。你的客户不是你，他们的思考方式和你不一样，做事情的方式不像你，也没有你有的期望和假设。如果他们和你一样，他们就不是你的客户，而是你的竞争对手。

本书用来帮助你弥补对用户所了解的情况和他们的真实情况之间的差异。本书不是学术论文，而是思想的工具箱，可以用来了解人们会如何体验你为他们所创建的环境。这些方法——来自人机交互、营销和许多社会科学——有助于知道用户是谁，并站在他们的角度来想问题。

另外，本书还提到了创建可用产品这个行业。它认同产品开发存在于商业风险的复杂环境中，此时真实世界限制因素的推拉不允许采用理想解决方案。可用性是一个棘手的综合性概念，充满了复杂、不确定性和办公室政治。如果能起到目的，本书能帮助你解决一些混乱，理清思路，并深入了解如何让产品比以前更好用，从而让世界更好。

## 读者对象

在某种方式上，如果你负责产品提供的用户体验，本书就是为你而准备的。在今天的软件和 Web 开发世界中，可以是战壕中的任何人，而且实际上，这种责任会随着项目进展从这个人的转移到那个人。基本上，如果你发现自己负责回答最终用户会如何看待正在制作的东西，或者负责回答最终用户与产品如何交互——或者甚至是他们应该怎么样与产品交互——本书就是为你而准备的。

也就是说，本书是为下列人员设计的：

- 项目经理，想知道如何确定下一版本软件的规格





## 用户体验面面观——方法、工具与实践

- 界面设计师，需要知道如何让人们最佳理解你为他们设计的任务
- 营销经理，想知道人们在产品中找到了哪些最有价值的东西
- 信息架构师，需要知道哪种组织机制能起作用
- 程序员，创建用户界面并努力解释模糊规格
- 咨询师，努力让客户产品做得更好

无论是哪种职位，你都想知道使用产品的人如何认知产品，他们希望得到什么，他们需要得到什么，以及他们是否能使用产品。

### 本书包含哪些内容？

本书分为三大部分。第 I 部分(第 1 章~第 4 章)叙述了最终用户研究为什么有好处、商业拉力如何在用户体验方面互相竞争，该部分展现了创建平衡、可用和能带来收入的产品哲学思想。

第 I 部分包含简短的一章，说明 15 分钟内教你掌握第二天即可开始进行用户研究的方法。真的是这样。

第 II 部分(第 5 章~第 16 章)是一本烹饪手册，深入说明 12 种了解人们需求、希望和能力的方法。每章叙述方法的内容都自成一体，全面引导你何时进行研究、如何进行研究以及如何理解结果。

第 III 部分(第 17 章和第 18 章)叙述如何处理结果以及如何利用结果改变公司的工作方式。该部分告诉你如何向公司推广以用户为中心的设计理念，如何让公司更好地运营并带来更多收入。

### 本书不包含哪些内容？

本书最主要的内容是如何定义问题。所有方法都是为了能更好地了解人们及其问题。本书不关注如何解决这些问题。当然，有时，好的问题定义能使解决方案很明显，但这不是本书的主要目的。

我深信，设计体验时没有绝对的规则来说明什么是对，什么是错。每个产品都受到不同约束来定义“正确”。和股票投资组合管理应用工具相比，为学龄前儿童设计的视





频游戏受到不同限制。想对两者采用相同的规则无疑很荒谬。这就是为什么没有指南能解决这些方法所定义的问题。没有“前 10 大”清单，也没有“法则”，也没有启发方法。很多优秀书籍都对如何解决交互问题提供了很好的想法，这些书提供了敏锐的合成解决方案，而且大部分时间都是正确的，但本书例外。

## 致谢

首先感谢人机交互(HCI)社区的英雄们，他们做了大量工作来创建人机交互领域，他们的工作激励我进入这一领域：斯图尔特·卡德、托马斯·莫兰、艾伦·纽厄尔、赫伯特·西蒙、雅克布·尼尔森和唐·诺曼。

戴夫·基拉斯和杰伊·埃尔克顿教我在大学学会了 GOMS<sup>①</sup>，我当时有些抵触，但现在有点明白了。有点明白了。

大部分用户体验知识，我都是在两家令人惊奇的公司学到的，这两家公司就是 Presence 和 HotWired。在 Presence，我明白了设计是一门生意，而在 HotWired 我明白了生意是一种设计。特别感谢 Presence 公司的汤姆·索兰尼尔、迈克·拉文、斯科特·布里洛和托德·库尔特，HotWired 公司的芭芭拉·库尔特，托！、埃里克·阿迪加德、埃里克·伊顿、乔纳森·路易、道格·鲍曼、凯文·莱昂斯、泰勒、纳达夫·萨维和纳达夫·萨维帮我用不同方式思考交互，但 HotWired 96' 级的每个人都很神奇。是的，每个人！

额外感谢埃里克·阿迪加德、凯文·亨塞克、马歇尔·普拉特和 Webmonkey 帮。

我的客户给我热情支持，我才有机会实践、学习、探索并调整这些方法。没有他们的支持和输入，本书不会这么丰富。

感谢在写作本书期间给我建议的所有人：大卫·皮特斯、维多利亚·贝洛蒂、卢·罗森菲尔德、约翰·希普尔、乔恩·卡茨、乔纳森·冈特利特、肯特·达尔戈仁、苏珊娜·吉布斯、克里斯·米勒、贝丝·莱伯、菲尔·贝雷特和 Flow Interactive 公司的梅里埃尔·耶茨、克里斯蒂娜·沃德柯、金·拉丁、肯·斯莫尔、乔希·格罗斯-尼克尔、珍·罗宾斯和埃弗里·科恩。

技术编辑给了我很好的建议，但我只听取了一部分。总是不听取卡罗琳·斯奈德、

① GOMS 模型指 Goals(目标), Operations(操作), Methods(方法)和 Selection Rules(选择规则)。





贾里德·布雷特曼、克里斯·诺德和昌西·威尔逊的建议，这纯粹是我的错。

非常感谢琳达·温曼为本书写序。

本书一切都源于理查德·科曼问我想不想写本完全不同的书，我说“不，我不想写那种书，但有本书我真的想写。”他倾听了我的想法，然后我就开始写。洛里·勒琼打电话给我，让我保持智慧和一些幽默，他还有惊人的洞察力，让我保持前进。露西·摩西阅读了还不适合送去印刷发行的草稿。乔斯林·利夫是我的第二双眼睛，他阅读了所有文字，并保持文字简单。戴安娜·塞拉和摩根·考夫曼(Morgan Kaufmann)公司的所有人都帮我成功完成本书，他们总是保持温暖的微笑。

在我消失的时间里，我在 Adaptive Path 的商业伙伴，莱恩·贝克尔、贾尼丝·弗雷泽、杰西·詹姆斯·加勒特、彼得·梅尔霍尔兹、杰夫·维恩和英迪·杨都非常有耐心，而且非常支持我。杰夫和莱恩甚至尽力在业余时间编辑本书。我都想不出来我还愿意跟其他什么团队共事。

我写书的地方是整个过程的关键部分。琼妮·多比尔、汉恩罗·科波和茱莉亚·多比尔给我在 Dagersheim 安排了一个居所。苏珊·阿格斯在 Castleton 提供了 Black Anguses。在整个过程中，旧金山使命区和 Lower Haight 的咖啡馆就是我的办公室、厨房和娱乐场所(特别感谢 Mission Grounds, Café Macondo 和 Crepevine)。

和几位朋友的友谊对写作本书发挥了重要作用：安德烈亚和斯科特拉我去加州，我在那里接触到万维网。摩西，露西和费利克斯给我建议、食物、电视，并陪我天马行空地胡侃。吉姆用小狗分散我的注意力。比安卡给予我爱、沙发和烤奶酪。罗宾写来邮件。吉纳维芙打来电话。

莫莉·赖特·斯廷森和我分享她的生活，还分享她 80 箱的爱心猴，比我希望的更多。

本书要献给我的父母，他们总是坚信我能更聪明。

从第一个字母——几乎——到最后一个字，本书完全用 OpenOffice 软件写成。它只崩溃过几次。真的！

陪伴我写作的音乐有 The Magnetic Fields, The Klezmatics, Photek, Tindersticks, Badmarsh and Shri, Boards of Canada, DJ Food, Nick Cave 和 Black Box Recorder。



# 目 录

## 第 I 部分 为什么说研究有好处? 研究如何融入产品开发

第 1 章 “台风”传说 .....	3	迭代开发的优点 .....	27
“台风”简介 .....	4	迭代开发问题 .....	28
第 2 章 马上进行可用性测试 .....	8	迭代开发与用户研究 .....	29
小型可用性测试 .....	8	案例: 日程安排服务 .....	30
你学到了什么? .....	12	第一轮 .....	30
下一步 .....	13	第二轮 .....	32
第 3 章 迭代开发与平衡用户需求 .....	14	第三轮 .....	32
最终用户眼中的成功 .....	15	第四轮 .....	33
功能性 .....	15	第五、六、七轮 .....	34
有效性 .....	16	第八轮 .....	35
愿望 .....	16	第 4 章 用户体验 .....	36
广告商眼中的成功 .....	17	信息架构 .....	36
流量 .....	18	信息架构师 .....	38
知名度 .....	19	信息架构师的信息需求 .....	38
公司定义的成功 .....	19	有用工具及方法 .....	39
利润 .....	20	交互设计 .....	40
推广 .....	20	交互设计师 .....	40
平衡系统: 迭代开发 .....	23	交互设计师的信息需求 .....	41
迭代开发怎么不起作用 .....	24	有用工具 .....	41
迭代螺旋 .....	25	形象设计 .....	42
		形象设计师 .....	43





形象设计师的信息需求 .....	43	用户体验研究 .....	44
有用的工具 .....	43		

## 第 II 部分 用户体验研究方法

<b>第 5 章 研究计划</b> .....	47	访谈结构 .....	96
目标 .....	47	非引导性访谈 .....	97
用具体问题扩充一般问题 .....	52	常见问题 .....	102
提示 .....	53	何时打破规则 .....	104
日程表 .....	53	<b>第 7 章 用户背景资料</b> .....	105
从开始阶段开始研究 .....	54	何时建立用户背景资料 .....	106
从中间阶段开始研究 .....	55	如何建立用户背景资料 .....	107
将研究问题组织成项目 .....	56	初步研究 .....	108
跨多个项目进行提问 .....	59	列出属性 .....	110
计划格式 .....	60	分簇属性 .....	117
预算 .....	61	围绕族群创建人 .....	119
X 公司研究计划 .....	62	确定优先级别 .....	122
摘要 .....	63	讲故事 .....	122
维护 .....	67	利用背景资料 .....	123
<b>第 6 章 通用工具：招募与访谈</b> .....	68	文档 .....	124
招募 .....	68	共享 .....	125
挑选受众 .....	69	利用背景资料进行开发 .....	126
找到受众 .....	71	定期更新 .....	126
从家人和朋友入手 .....	72	实例 .....	127
日程安排 .....	84	<b>第 8 章 情境调查、任务分析</b> .....	130
容易犯的错 .....	89	和卡片分类 .....	130
与专业招募师合作 .....	92	情境调查 .....	131
访谈 .....	96	什么时候适用情境调查 .....	131



情境调查过程.....	133	实例 .....	202
如何分析情境调查数据 .....	143	附录 A：讨论指南 .....	204
任务分析 .....	149	<b>第 10 章 可用性测试</b> .....	211
何时适用任务分析 .....	149	何时测试.....	211
如何进行任务分析 .....	150	迭代测试过程实例：Webmonkey	
如何分析结果.....	152	2.0 全局导航.....	212
卡片分类 .....	157	如何测试.....	215
卡片分类何时进行 .....	158	准备 .....	215
卡片分类过程.....	158	实施访谈 .....	232
卡片分类分析 .....	159	如何分析.....	238
<b>第 9 章 焦点小组</b> .....	164	收集观察结果 .....	239
何时适用焦点小组 .....	165	组织观察结果 .....	241
焦点小组的长处 .....	165	提取趋势 .....	241
焦点小组的不足 .....	166	实例 .....	242
四种焦点小组 .....	167	<b>第 11 章 调查</b> .....	248
如何实施焦点小组 .....	169	何时开展调查 .....	249
确定日程表.....	170	如何实施调查 .....	249
挑选目标受众 .....	170	设定日程表 .....	250
招募 .....	172	编写调查 .....	251
定义范围.....	173	实施调查 .....	268
选择主题.....	174	如何分析调查答复 .....	277
编写指南.....	175	统计 .....	278
讨论指南模板 .....	176	比较 .....	282
实施小组.....	181	判断误差 .....	285
焦点小组分析 .....	195	度量误差 .....	287
收集数据.....	196	做出结论 .....	288
提取趋势.....	198	跟进和持续研究 .....	291
做出假设.....	201		





跟进定性研究.....	291	日志文件中有的和没有的.....	327
跟踪调查.....	292	日志和 cookie.....	331
精细调查.....	292	一些有用指标.....	333
前/后调查.....	292	聚合度量.....	333
实例：新闻网站调查.....	294	基于会话的统计.....	335
<b>第 12 章 持续关系.....</b>	<b>299</b>	基于用户的统计.....	335
背景.....	300	点击流分析.....	336
日记.....	301	从数据中提取知识.....	337
非结构化日记.....	302	<b>第 14 章 竞争性研究.....</b>	<b>342</b>
结构化日记.....	305	竞争性分析何时有效.....	342
管理答复.....	309	竞争性研究方法.....	343
如何分析日记.....	310	确定竞争对手.....	344
咨询委员会.....	313	勾画竞争对手.....	345
委员会成员.....	314	竞争性分析方法.....	348
和委员会共事.....	315	分析竞争性研究.....	352
Beta 版测试.....	318	根据竞争性研究采取行动.....	354
嵌套.....	319	实例：ZDNet 网站快速评估.....	354
看得更远.....	320	<b>第 15 章 其他艰苦工作：公开信息</b>	
<b>第 13 章 日志文件和客户支持.....</b>	<b>321</b>	和专业顾问.....	358
客户支持过程.....	322	公开信息.....	358
收集意见.....	323	独立分析.....	359
阅读意见.....	323	流量/人口统计.....	360
组织和分析意见.....	324	市场调研.....	361
编码意见.....	325	出版物.....	363
汇总意见.....	326	聘请专家.....	365
分析意见.....	326	选择适当的时机.....	365
日志文件.....	327	找到专家.....	366





设定期望.....	370	平行研究.....	380
<b>第 16 章 新兴方法</b> .....	374	参与性设计.....	381
方法变体.....	374	组合使用.....	382
虚拟焦点小组.....	374	焦点小组和日记.....	382
提名小组.....	376	观察性访谈和可用性测试.....	384
摩擦小组.....	377	调查和焦点小组.....	385
虚拟可用性测试.....	378	日志文件和可用性测试.....	386
眼动跟踪.....	379	任务分析和可用性测试.....	386

## 第 III 部分 交流研究结果

<b>第 17 章 报告和演示</b> .....	391	展示结果.....	416
准备工作.....	391	保持耐心，但要坚持.....	417
了解受众.....	391	鼓励具体用户体验角色.....	417
了解过程.....	393	理由.....	419
报告.....	394	以用户为中心的过程的原因.....	419
挑选格式并精心组织.....	394	度量有效性.....	421
报告样本：在线贺卡公司的 可用性测试.....	395	可用性定价.....	424
演示.....	401	如果太困难了怎么办？.....	426
展示给特定群体.....	403	唯一的方向.....	428
常见问题.....	407	<b>附录 A 用户研究实验室的预算</b> .....	429
<b>第 18 章 打造以用户为中心的 企业文化</b> .....	410	视频设备.....	429
集成.....	411	可选设备.....	431
了解现有过程.....	411	研究实验室作为目的地.....	431
从小处着手.....	413	<b>附录 B 常见调查问题</b> .....	432
尽早邀约利益相关者.....	415	人口统计问题.....	432
		互联网/互联网使用问题.....	433





## 用户体验面面观——方法、工具与实践

产品使用问题 .....	435	焦点小组 .....	438
技术能力问题 .....	436	可用性测试 .....	439
附录 C 观察员指南 .....	438	参考书目 .....	441





# 第1部分 为什么说研究有好处？ 研究如何融入产品开发



好设计，更有爱，关注微信



idesign123





# 第1章 “台风”传说

有时，有些东西要经过很长时间才会“显山露水”，露出“其庐山真面目”：熟悉的街区有条近道，两个朋友之间的关系，父母实际并没有那么糟糕。不可思议的事情需要经过时间的沉淀和思考，才能恍然大悟，认为是理所当然。

网站和用户研究的关系也如此。虽然互联网发展历史不长，但是人们在很长一段时期内都觉得产品未完成就交给客户是无法想象的奢望，也纯属多余。那时，互联网设计圈关注品牌(“确保品牌标志有蓝色弧线!”)、定位(“我们是卫浴产品的 amazon.com!”)、或者做率先进入市场的开拓者。至于调查和分析用户需求所涉及的相关费用，则不包含在预算内。如果网站或者产品大致能用，就表明它有用(因而会受到欢迎，能赚钱，赢得开发者预期的其他正面效果。)向用户征求意见与此无关，而且可能会破坏设计才华。

最近的发展历史清楚证明这种模式是错误的。率先进入市场，光有蓝弧和网上购物车远远不够。如今，必须要有人们真正想要的产品，产品必须满足人们需要，能够供人们实际使用。这便意味着用户研究。用户研究是了解设计对受众如何产生影响的过程。设计阶段开始前的调查、焦点小组和其他形式的用户研究会造成两种不同结果：要么是有用、可用和成功的网站(或者任何经过设计的产品)，要么只是一次使参与人员沮丧的不赢利的尝试。

如今，“产品应当迎合受众的需要”这一理念似乎显而易见。但并非一直如此。回首 20 世纪 90 年代中期的互联网世界，那时绝顶聪明和理性的人(包括我自己)都无法想象去设计用户不喜欢的产品。下面就来说说不考虑用户会发生什么情况。





## “台风”简介

1996年，日子振奋人心，PointCast 红极一时。PointCast 提供的独特服务能把俗气的屏幕保护程序转变成由广告驱动的新闻和股票，它是新闻股票服务界的黑马。这项服务吸引了上万用户，其发明者成为圈内所有杂志的封面人物。信息主动送到人们面前，而不是人们必须索取才会出现，这是一个全新的概念，人们还迅速创造了一个时髦词来概括这种服务，即推送(push)技术，代表着未来。不久，所有人都加入推送潮流，创建推送服务。

有一家公司加入了这场角逐，我来说一说。这里的描述基于真实情况，但为了增加故事的趣味性，细节做了更改(当然也是为了保护无辜者)。我们把这家公司叫本伽利(Bengali)。对于在线新闻和信息服务，本伽利有过几次相当成功的案例，它现在信心满满，准备充分，渴望接受新挑战。他们想创造具有革命性的东西，要挑战人们对媒体的假设，想创造未来的电视、电台或出版社。他们认为，借由推送服务通过互联网开创新媒体的梦想肯定会大获成功。

他们要将其命名为“台风”(Typhoon)，而“台风”最后却让 PointCast 蒙羞。本伽利进入臭鼬工厂(skunkworks)模式，完全封闭式开发台风，使用的开发人员个个聪明绝顶，而且只有在万事俱备的情况下，才会发布产品。

PointCast 杀手的开发需要大量工作。开发人员秘密工作了一年，从不与公司以外的人讨论此事(公司内部也几乎无人知道)。“未来社会如何与媒体互动？”开发小组根据这些想法建立了未来媒体的远景。他们质疑所有对媒体的假设。每个答案又带来更多问题，而每个问题都需要设想未来的其他方面。

软件和远景一起成长，一起变化。最终的产品精细、复杂，完美融合一起。但过了一年，本伽利才准备发布。

准备发布时，开发人员把“台风”展示给公司高层看。高管虽然不否认对获得的成果印象深刻，但仍然有些不安。有些高管想知道哪些人会





注意：

他们挑了十二个人，邀请他们进行一对一用户测试。这些用户来了，一次一位，做了几天测试。

是受众。另一些高管问开发团队人们如何使用“台风”。虽然开发团队对一切问题都有答案(一年时间，他们已经形成了一套非常完善的程序及其用法的模型)，但他们承认大部分答案都需要验证，因为许多终端用户都还没有用过“台风”。开发团队经验丰富，也做了一些内部评估，所以他们觉得设计非常接近用户期望。但为了消除高管的疑惑，检查粗糙的地方，开发团队决定在发布之前再进行一些用户研究。

他们挑了十二个人，邀请他们进行一对一用户测试。这些用户来了，一次一位，做了几天测试。测试计划是请用户坐下，说说他们对产品的初步看法，然后执行几项不同任务。

测试是一场灾难。以下内容摘自其中一次对话的完整笔录。

#### 可用性测试笔录

- 如果图片看起来就是这种形式，会有些让人困惑，因为图片也许应该有点儿用途……我搞不懂。
- 下方所有单词几乎都无法阅读。
- 如果这里是想说些我应该能明白的东西，那么我觉得很难搞清楚它想表达什么。
- 这些标题我一个也看不懂。
- 我不知道这些东西是什么构成的。
- 我知道，如果单击它们，它们会执行一些事情，但是……
- 目前为止，还没什么东西吸引我去单击它们。
- 还有，没有状态栏，所以无法确定网页什么时候加载完。
- 我不知道这些数字和那个东西是否有关系……我真的不知道。
- 我希望这个链接能链接到新闻。
- 我估计下载速度肯定超过了 28.8 [调制解调器]。
- 好像有点慢。
- 好像不是我要找的东西。
- 我真的很好奇下面这些数字代表什么。
- 它们可能什么也不是。





- 我只想知道是什么内容。
- 好吧，我真的不想跟着那个……
- 我等著有文字形式告诉我是什么内容。
- 什么也没有，我不知道是什么内容。
- 我甚至搞不清楚网页是不是载入成功了。
- 哦，原来在这里……
- 我按住它的时候，希望它能停在那里不动，但最终它还是消失了。

即使没听到用户谈论的内容，沮丧、混乱也显而易见。测试者使用“台风”时，要么没按照开发者期望的方式，要么虽然与开发者期望的方式相同，但“台风”的表现又未达到预期。从可用性角度来看，“台风”基本无法使用。

还有一点与前述情况一样糟，甚至更糟：没有一个测试者知道“台风”的用途。很明显，人们并不想用“台风”，因为他们对它的用途一无所知。如果产品难以理解，可用性从何谈起？！

为了让“台风”用处更清楚，为了帮助人们使用“台风”，开发团队乱成一团，抓狂解决问题，但是开发团队并不清楚从何处着手。产品发布迫在眉睫，他们只能够解决一些最明显的问题。许多问题还留在那里，开发团队的信心开始动摇，甚至紧张地怀疑还存在更多问题。

“台风”发布了，毫无悬念地一片混乱。无论是新闻媒体，还是其最初用户，都不知道它如何使用。“台风”带来的流量显著少于本伽利的预期，尽管广告宣传很强势，但用户数量还是不可遏制地下滑。

随着开发团队开始开发下一个版本的“台风”，开发方向变得更不清晰。根本问题不断出现。对范围、受众、目的和功能的争论不绝于耳。本来确定的事情突然变得不可确定。高管们很快对开发团队搞定“台风”的能力失去信心。辩论仍在继续。开发团队改进了“台风”，但仍然无法留住访问者，放弃使用“台风”的用户持续增加，几个月之后，其他更有传统软件背景的人取代了项目负责人。新项目负责人想把“台风”改造成更普通的新闻服务，但新产品的规格说明搞得“台风”看上去就像公司的其



他产品，而这和“台风”的假定前提恰恰相反。开发团队要求增加人手进行深入变更，此要求遭到拒绝。对如何吸引访问者的观点开始分化，最终产生分歧。开发团队觉得被背叛，觉得创造力被浪费了，管理层已经抛弃了好点子。

新项目开工4个月后，最初的开发团队成员放弃了项目，转投到公司的其他项目。两个月后，项目悄然关闭，并被计入公司损失，最终完全消失。现在，剩下的只有T恤。

可悲的是，这是一个真实的故事。虽然细节有更改，但核心情况都是真的。它就发生在我身边。我见证了“台风”的产生，执行过测试，亲眼看着它分崩离析。多年之后，参与此项目的一些人心中仍然隐隐作痛，因为他们那么多努力、那么多想法、那么多创新，最后都付之东流。

但理由充分。这是一个很糟糕的产品。糟糕的原因不是代码质量(代码非常紧凑)，也不是核心创新(创新真实且有理)。真正使其沦为“糟糕产品”的，是因为“台风”是一个不了解其受众的优秀产品。产品虽好，但无法符合需求，无法满足愿望，也不关注受众的能力，因此不会是好产品，代码或者视觉设计再好，也枉然。

本书讲的是了解受众，并利用知识创造优秀软件。它能帮助您避免出现像“台风”一样的情况，而同时又能保持创造力，这种创造力带来的产品富有创新精神，让人激动，很独特，还能赚钱。用户体验研究是一系列工具的合集，用于洞察人们的需求和能力。本书提到的其核心(本书所支持的理念)与创建解决方案无关，而是侧重于问题的定义。最终目标不是仅仅让人们觉得高兴，而是通过让人们觉得高兴而创造成功的产品。了解人们的问题之后，就不容易为错误问题甚至根本没有问题的情况提供解决方案。



## 第2章 马上进行可用性测试

基础的用户研究简单、快速而且高效。对任何产品都能实施某种形式的用户体验研究。问题是你是否想要亲自动手。想要知道答案，办法只有一个。试一试就知道了。本章将教你如何进行一种快速简单的用户研究方法，即邀请你的朋友及家人进行可用性测试。通过15分钟阅读加上几小时聆听，你就能更好地了解你的客户，更清楚产品哪些地方不好用。

### 小型可用性测试

#### 提示：

如果没有正在开发的产品，也没有半功能原型，那么想充分利用这种方法就为时过早。此时应该采用适用于仍停留在概念阶段的研究方法，如情景调查法或者焦点小组法。本书第8章和第9章将分别讨论这两种方法。

**可用性测试(usability)**能说明受众是否能使用产品。它有助于确定人们使用网站时存在的问题，暴露出不好用的界面和容易混淆的语言。可用性测试通常作为大型研究系列的一部分，涉及准备和分析工作。本书第5章~第16章将介绍这些内容。但从快速展示产品和经济的角度而言，邀请朋友和家人进行可用性测试最合适。这样做，可以用最小花费获得产品的直接反馈。如果你上午阅读本章，那么到当天工作结束的时候就可以找人讨论讨论，第二天便能重新思考一些产品功能了。但如果你是第一次进行用户研究，还是先给自己一两天时间准备一下。

进行可用性测试过程可以分为四个主要步骤。

- (1) 定义受众及其目标。
- (2) 创建达成目标的任务。
- (3) 物色合适的人选。
- (4) 观察人们执行任务的过程。

#### 1. 定义受众及其目标

“评估总是始于‘为什么要有这个东西？’”

——戴夫·亨德里，助理教授，华盛顿大学信息学院，私下交流



出于某种原因，你正在开发某个产品。你觉得你的想法能让世界上有些人过得更好。也许能帮他们买到更便宜的东西。也许能帮他们获得其他途径得不到的信息。也许能帮他们联系到其他人。也许能让他们感到开心。

不管什么原因，你都在开发自己觉得会给特定人群带来价值的东西。要能获得这些价值，他们必须做点事情。通常他们要做这么几件事。在购物网站上，人们要做的是“找到东西，购买，并订阅新闻邮件”。在婚姻介绍网站上，人们要做的是“找到她，给她留言，发送留言，然后阅读回信”。

因此，对于可用性测试，首先要搞清楚网站的使用对象。对于你预期的使用网站最多的用户，如何描述他们？他们和其他人有什么区别？区别在于他们的年龄、兴趣和问题吗？实际可能包含前面所有情况，可能还有更多情况。

例如，假设要检查一家网上刀叉餐具商店的浏览可用性和购买用户体验。可以快速创建网站受众的受众定义。目标受众如下：

要买餐具的人

但这不够具体。我祖母也会为家庭野餐定期购买塑料餐叉，但她不会从网上购买。所以受众定义范围要稍微广泛一点。目标用户受众如下：

- 要买高端刀叉餐具的人
- 注重价值的人
- 希望有广泛选择的人
- 精通计算机并接受电子商务的人
- 不属于专业刀叉餐具买家的人

接下来，要找出关键产品功能特性。把产品内容写下来。为什么人们要使用它？为什么它对用户有价值？如果在一次嘈杂的聚会上，给你 30 秒时间把你把自己的网站介绍给以前从未听说过这个网站的人，你会说些什么？把这些都写下来。

全北美的人都可以通过 [forkopolis.com](http://forkopolis.com) 网站购买刀叉餐具，它的选择范围最广，以包含所有主要奢侈品牌和设计师的产品为特色。由于很容易找到特定风格和数量，所以买家能快速且以低价更换单个损坏的茶匙，也可以购买足以供整个餐馆使用的银制餐具。





### 2. 创建达成目标的任务

现在把网站五个最重要功能写下来。除了前面提到的功能，人们还应该做什么？在购物网站上，人们显然要能买东西。但不管是否准确清楚自己要买什么，他们都应该能买到东西。此外，也许他们还要能发现促销商品和超值商品。列份清单，用一两句话描述一下每个功能。

- (1) 通过风格找到特定刀叉。
- (2) 购买单件刀叉。
- (3) 购买整套刀叉。
- (4) 发现特惠刀叉。
- (5) 找到退货信息。

从使用者角度出发，用两三句话描述一下人们使用某项功能的情况，即所谓任务。如果“通过风格找到特定刀叉”属于功能之一，其任务描述如下：

您要从 [forkopolis.com](http://forkopolis.com) 网站购买一套路易十四餐叉。请从网站首页开始，去找一套路易十四餐叉。

最后，根据难易程度，按照从最简单到最困难的顺序排列任务。从执行简单任务入手，人们会对产品和任务过程有好感。

### 3. 物色合适的人选

现在，找一些符合步骤 1 所创建的特征的人。找五六个人，这些人要与你预期对产品有兴趣的人相似，像这样快速运用可用性测试，就可以更充分地认识真实用户使用产品时发生的问题和误解。从身边的人中物色合适人选是最快的方法。如果身处大公司，可以从与产品没有任何关系的部门找些同事。如果是在小公司，可以找朋友、家人以及同事的朋友和家人。可以找在办公室大厅的人。可以找街头的行人。还可以找一些不熟悉产品的人，无论喜欢者或不喜欢，对产品都没有偏见的人(所以不要找溺爱的祖母、公司最大竞争对手的首席执行官)，只要这些人和你期望访问网站的人有点相似就行。除非产品专为开发人员而设计，否则就不要找靠开发网站谋生的人，因为他们懂的东西太多。



联系这些人，告诉他们想请他们帮你评估一下正在开发的产品效果。除了任务清单上方的产品简介外，不要透露其他内容。告诉他们不需要做任何准备，只要能来参加测试就行。提前一两天给他们安排好半小时单独访谈的日程，访谈间隔时间为15分钟。

### 4. 观察人们执行任务的过程

首先，写一个剧本，你和受邀的评估人会根据剧本进行评估。把网站简介写在剧本页顶部。评估人所能看到的所有产品介绍就这么多。不要告诉他们其他情况。在现实世界中，人们往往只能看到简介加链接。重新拿纸和笔写下任务，一页写一项。如果用户以前使用过该项服务，则不要写用户不知道的任何信息。

接下来弄一台电脑，找个安静的房间，你和评估人可以在那里不受干扰地讨论产品。一般情况下，找间僻静的小会议室即可。确保房间里没有任何与产品有关的东西，以免分散评估人的注意力，也不能出现任何可能混淆评估者的信息。所以，房间里不能有笔记、公司宣传海报、没擦干净的白板和展会鼠标垫。

根据任务设置好计算机。尽可能按照最通用配置设置浏览器，去掉自定义工具栏、自定义颜色、显示选项和无关书签。人们需要从起始页进入每个编写好的剧本，把这些起始页用书签保存起来。

评估人来了后，请他们为要进行的事情做好准备。确保他们感觉舒适。可以像这样介绍测试过程。

- 他们受邀到此，是为了帮助你了解产品哪些地方有用，哪些地方让他们感到困惑。
- 虽然称为测试，但不是要对他们进行测试，而是邀请他们来评估产品，谈谈他们对产品的看法，所以他们怎样做都无所谓对错。要强调一点：如果不能完成任务，也不是他们的错；如果他们说出对产品的负面看法，也不会伤害任何人的感情。
- 他们要大声说出所有想法，这一点很重要。建议他们详细叙述在做什么，以及为什么这么做。





- 你会留在同一房间里，一边倾听他们发言，一边做笔记。但他们要当你不存在，而专心执行任务和详细叙述。

(很可能需要提前把这些要点写成具体文字，放在产品说明开头的剧本中。)

如果参加测试的人感觉良好，你也给了他们初步指导，接下来即可开始阅读产品介绍和任务说明页。请他们以最佳方式执行任务。但如果几分钟后他们还不能完成某项任务，应该请他们执行下一项任务。要强调他们随时都可以大声说出自己的想法。

接下来，请他们发言。你就坐着观察，安静地记笔记。如果他们被问题卡住了，不要告诉他们应该单击哪里或者看什么。不管什么情况，都不要告诉他们应该怎么做。如果他们看上去特别沮丧，就告诉他们有些事情无法完成不是他们的错，请他们执行下一任务即可。

如果所有任务都完成了，或者半小时时间已到，就可以请他们停止。请评估人向自己谈谈他们的总体印象，他们是否会“在现实生活中”使用该网站。然后送份礼物感谢他们为测试所花的时间(当地餐馆或者书店的礼券、公司餐厅午餐券、一桶汽油……只要适合受众就行)，向他们表示谢意，并送他们离开公司。

最后，重启电脑，清除缓存和浏览历史，将主页设为空白页，为下一位评估人进行评估做好准备。

## 你学到了什么？

可用性测试一结束，马上问自己下面几个问题。

- 哪些功能起作用了？
- 有没有用户总是误解的东西？如果有，是什么？
- 有没有错误总是发生？如果有，是什么错误？
- 他们有没有按照你期望的方式执行任务？如果没有，他们又是怎么做的？
- 他们执行的顺序是否与你期望的一样？如果不一样，又是什么





顺序？

- 他们发现了什么有趣的东西？
- 你期望他们会发现有趣的东西，但事与愿违，他们并不觉得有趣。
- 他们明白网站用途吗？他们有没有错过好想法？
- 他们完成了多少任务？哪些任务(麻烦)困难最大？
- 他们在什么时候感到受挫？当时他们在做什么？
- 网站符合他们的期望吗？如果没有，哪些地方让他们感到失望？
- 你了解他们的期望吗？
- 他们感到困惑了吗？当时他们在做什么？

到这里，你应该知道产品哪些地方存在问题。你可能看到有几件事情一再，再而三地发生。也许人们并不理解你为某项功能指定的名称。也许他们没看到关键功能。也许他们对所提供的功能不感兴趣。也许，他们喜欢这个产品，它能满足他们想要的一切。知道所有这些情况是好事，因为能揭示问题出在哪里，(而且同样重要的)能说明哪些地方没有问题。

忠告：

邀请朋友和家人进行可用性测试快速、简单而且方便，但它属于应急方法，朋友和家人能告诉你产品存在的问题，但(多半情况下)他们无法代表实际用户受众。只要有可能，都应该邀请实际受众代表来进行可用性测试。

## 下一步

完成了第一个可用性测试，你会认识到这种方法适用于哪些地方，对你有什么用。读完本书第I部分的余下各章(一直到第4章结束)，读者应该能整合产品研究计划、廉价可用性测试(比如邀请朋友和家人进行可用性测试)及其他多种方法。本书第II部分(第7章~第13章)将深入探讨这些方法，解释如何超越基本测试来了解产品开发中各方面的用户体验。第III部分(第14章~第19章)将帮助你展现研究结果，以确保开发团队对产品进行的改变真正有益于受众，并引导产品一直朝着用户需要的方向健康发展。



## 第3章 迭代开发与平衡用户需求

在完美、人人平等、无忧无虑的世界中，产品开发过程的唯一宗旨就是让用户开心。开发过程完全以用户为中心，关注那些不计任何代价创造理想用户体验的想法。所有软件(硬件、录像机、汽车以及任何有人机界面的东西)都被标准化、优化、一致且透明。一切都关注于帮助用户执行任务。

但世界远远不够理想。任何任务都不可能只有一种完美执行方法。即使是理想的用户解决方案，也无法总能产出理想的产品。而且，产品一般都不是考虑用户利益而单独创建的：它们是由具有赢利目的那些公司创建的。赚钱和满足人们的需求是两个完全不同的目标，它们可以结合在一起，但永远是截然不同的两个目标。

此外，现代软件——尤其是网站——所处竞争环境独特。它不仅要满足用户和开发方，而且常常还要满足另一类利益相关者：广告合作伙伴。广告不只给用户体验带来新要素，而且在此基础上创造了一种贯穿开发过程的全新业务关系。过去，用户体验质量和产品利润是驱动软件设计的两大主要要素，现在有了第三个要素——广告效果(图 3.1)。三类利益相关者在进行一场持续不断的拔河比赛。如果任何一方拉得太厉害，那么其他两方都会受苦。



图 3.1 影响产品优先级别的拉力



本书侧重于了解用户体验，不会深入讨论广告效果或公司赢利能力。当然无论如何，开发时记住此类拔河比赛和考虑用户体验开发环境是至关重要的。

第4章将关注用户体验。在本章后面的内容中，我将展示解决这些拉力的一种方法。但首先从这三者角度出发，看看促成产品成功的要素。

提示：

唐·诺曼 (Don Norman) 在其著作《设计心理学》(The Design of Everyday Things) 中列举一些精彩例子来说明无法使用的产品。这些例子说明，我们虽然不一定知道什么东西能使产品可用，但我们一定知道产品什么时候无法使用。

## 最终用户眼中的成功

产品的最终用户体验是其成功基石。出色的用户体验虽然不是确保产品成功的唯一因素，但糟糕的用户体验绝对是导致产品失败的“快车道”。但体验质量不是二进制——产品完全无法使用的情况非常罕见——但令人惊讶的是，用户体验平庸实际上比彻底失败更糟糕。如果东西完全无法使用，至少能清楚说明哪有问题，而即使其他地方设计完美，但某些东西会间歇性发生问题——购物车放弃率高或者只有 40% 的人能用搜索引擎找到东西——这却会造成产品表现不佳。虽然平庸的用户体验不会引人注意，但是甚至可能会成为整个商业风险中最严重的问题。

出色的用户体验由什么构成？答案是因人而异，因产品而异，因任务而异，但如果东西有用，有效，而且是预期受众想要的，那么把这个东西定义为“可用”就是一个很好的通用定义。

### 功能性

如果拟使用该产品的人认为产品或者产品的一部分对他们来说有用，那么产品或者产品的一部分就是有用的。用户期望每个产品都能做一系列事情，它只有做到这些事情，人们才会认为它有用。这个想法很简单，但值得注意的是，人们常常会忘记这个想法。例如，有台高端低音炮，自带放大器和遥控器，这台低音炮从表面上看就是台自带控制非常少的扬声器——有些功能，厂商只放在遥控器上。其中一项是均衡器模式切换控制（“剧场”、“音乐厅”和“电台”等）。不幸的是，如果遥控器丢了，放





大器就永远处于最后一次使用的均衡器模式中。因此,如果低音炮设为“电台”模式,遥控器丢了或者坏了,其主人听所有音乐都被放大了,好像是场玎玲脱口秀,声音轰轰作响,还几乎没高音。这种做法首先就破坏了购买昂贵低音炮的目的:让音乐好听。

还有种现象更为普遍,即界面复杂或者难以理解而造成关键特性被深度隐藏。可视化编程问世之前,编辑录像的过程需要用编程来控制就是个经典例子:这项功能难以理解,所以录像机中可能都没有包括此功能。

### 有效性

人们基本上都注重效率,而产品的速度和易用性取决于用户界面的效率。产品的用户界面能让任务一步完成,也能让任务花很多步才能完成。这些步骤可能很直观,也可能被隐藏。也许界面需要用户同时跟踪许多事情。从传统角度看,这些因素都可以归结成速度,人们在给定条件下能以多快的速度执行任务,同时出错最少。

### 愿望

虽然可用性专家和工业设计师(当然也包括市场人员)早就认识到这一点,但在优秀用户体验中,这是最难明确的。用起来很好的东西,却很难捕捉到让人惊奇和满意的地方,但它绝对属于最佳产品设计的不可或缺的一部分。它与产品包装(美学)和营销都截然不同,但对两者都会产生情感反应。

#### 可用性与设计

归根结底,可用性就是优秀设计。这并不是说,所有优秀设计都能用,因为有些精心设计的东西来源于一个无法使用的方面(形象、技术和价值等)。例如,很多昂贵的名牌家具很漂亮,一眼就能识别出来,但出自这些名家的书架常常无法放很多书,椅子坐着也不舒服。同样,Unix 操作系统令人难以置信的优雅和强大,但需要多年实践和记忆才能使用其最基本功能。可口可乐公司的 OK Cola 包装炫目,但一点儿也不好喝。另一



方面，雅虎重视最终用户体验，结果大受欢迎，但其视觉形象却缺乏容易识别的特性。最终，尽管做了那么多营销努力，很难用的产品最终也不会被大量使用。人们使用不能用的产品，要么无法完成需要做的事情，要么不能很快完成，要么用得也不开心。

当然也有例外情况：产品碰巧很独特，大家都会容忍其可用性问题，因为它功能独特（内容管理系统是个很好的例子：系统超级难用，但如果不用，就几乎无法管理海量内容）。但这种情况很少。几乎所有其他情况下，产品可用性都是产品成功的关键。

## 广告商眼中的成功

否认广告——及广告商——是全部或者一部分体验是一种天真的行为。但在互联网出现之前，就软件体验而言，广告一直微不足道。20 世纪 80 年代，字处理软件没有广告。电子表格工具没有联合赞助商。在 Zork 游戏中杀死怪兽也不会从联合航空公司获得飞行常客积分。但在网上，广告无处不在，甚至广告里面还有广告（图 3.2 是一家典型的广告驱动型网站，请数数有多少联合品牌协议、广告和产品放置标记）。

而现在，即使是非网络产品，如 Intuit 公司的 QuickBooks 软件，也在界面中加入了营销消息。页面上有赞助商链接、联合品牌广告内容、插入式广告、背景式广告（pop-under）、流量共享协议、联盟计划、奖励措施以及其他形式的赚钱方法。页面有 20 个不同内容区域，基本全是广告。甚至能赚钱的订阅服务也建立在与另一家公司关系的基础上。

所有这些东西都属于用户体验，但它们都只是广告商想要的。用户要能看到这些广告，采取行动，并记住这些广告。最终，广告商主要对自己所选市场的广告效果感兴趣。他们愿意等待，并提高广告受欢迎程度，只要用户体验能从正面影响广告销售就行，他们也会对最终用户体验感兴趣，但他们的终极目的是想知道某个广告交易能带来多少收入。在互联网上，广告商用“流量”（traffic）和“知名度”（awareness）来衡量结果。



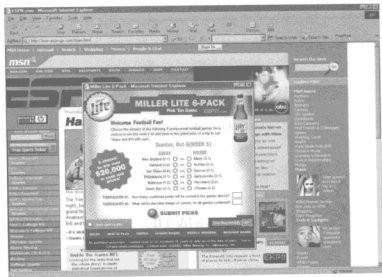


图 3.2 2002 年 ESPN.com 网站上的广告单元

## 流量

提示：

不管采用哪种方式，传统互联网广告都是到其他网站的链接，广告商的基本目标时常与用户或者公司的目标背道而驰。简而言之，广告商在一定程度上想要让用户离开你的公司网站去他们指向的网站。评估用户体验中的平衡的时候，这种关系往往最难解决。

互联网上有三个主要指标来衡量广告表现：展示次数(impression)，点击次数(click-through)和销售量(sell-through)。

“展示次数”衡量广告插入到被下载网页的次数。因此，它可以衡量出广告应该被看到的次数。

不幸的是，只把广告放到网页上并无法保证有人会看到广告，就像整个公告牌并不代表每个开车经过的人都能看到一样。因此网上广告行业开发了第二个衡量标准，“点击次数”。“点击次数”衡量有多少人点击了某个广告，人们对某个广告有多浓厚的兴趣。当然，广告的目的不只是产生兴趣，所以点击次数并不是完美的衡量方法。

次数对于其他类型的广告或伙伴关系，如联盟计划或流量共享协议，这一根本分歧不会这么深刻或者明显，但它依然是影响当事方之间关系的潜在主要因素。

广告效果最直接衡量标准是“销售量”，特定用于衡量出售东西的广



告。它能衡量实际有多少人通过某个广告购买了某一产品。但这种衡量标准也有问题：人们可能是在离开刊登那则广告的网站很久之后才购买产品的。

## 知名度

广告效果属于大型研究的主题，但网站研究中有三个共同指标，品牌知名度(Brand Awareness)、品牌亲和力(Brand Affinity)和产品销售(Product Sales)。

“品牌知名度”衡量有多少人知道产品。“品牌亲和力”衡量人们喜欢产品的程度，说得更通俗一些，人们和产品之间的关系。“产品销售”是底线：在广告期间或者在广告后不久，有多少人购买了产品，无论是什么产品在做广告。

这三项指标不一定需要互相依赖，有时候也会彼此产生矛盾。例如，奥兹莫比尔汽车(Oldsmobile)推出“这不是你老爸开的奥兹莫比尔”广告，最初相当成功，大大提升了品牌知名度。人们记住了它，品牌亲和力高：广告目标市场的很多人重新考虑这家汽车公司的形象，认为它比以前认为的更年轻，更富有冒险精神。不幸的是，广告的影响令人失望，因为人们去展厅看实物时发现，它还是老爸开的奥兹莫比尔，因此产品销售无法体现广告的成功。原因何在？因为汽车实物带来的用户体验与广告带来的用户体验不匹配。

最后一点，即使是最佳传统营销衡量系统——印刷广告斯塔奇得分(Starch Scores)和尼尔森电视节目收视率——也只能是估计广告对提高知名度和销售的效果。互联网更容易得到(日志文件比比皆是)，但广告战成功的决定因素往往仍然是一门由媒体买家的胆量来决定的艺术。

## 公司定义的成功

除了极少例外，公司投资于产品开发的目的是为了赚钱或者一直寂

提示：

可以访问互联网广告委员会网站 [www.iab.net](http://www.iab.net)，找到互联网广告标准。





寂无名。公司衡量产品成功有两种主要方式：利润或者产品如何从整体提升公司的品牌。

### 利润

虽然在网络热潮时代，人们曾经一度把利润抛之脑后，但网站也好，餐叉也好，创建产品的根本目的就是为了赚钱。也许不会马上赚钱，或者不能赚很多钱，但没有什么事情是为了赔钱而去做的。非营利性组织的产品例外，赚钱是公司所有活动的首要原因(就非营利性组织而言，赚钱也是首要原因，只不过它并不想成为最赚钱的组织)。

### 推广

在某种程度上，所有东西自身就是广告，但时装或者极端情况除外——我租过车，车上贴满赫兹公司的广告——用户体验中的企业自我宣传通常相当微妙。互联网是个例外。互联网上，市场巨大，发展神速，竞争激烈，充满了类似产品，更为重要的是，网站本身就在积极为自己打广告。网站需要做几件事情来推销自己：它需要不同，可识别，令人难忘，它需要传递功能。只为这些事情而设计创造出来的产品可能表面看起来闪闪发光，但缺乏实质内容，在开发过程中忽略这些情况同样也很危险。

无需用赞美之词去描述产品所存在的差异，大家都知道个性化对于产品是很重要的。产品要与所有竞争对手区分开来。例如，曾有一段时间在搜索引擎界，最大搜索服务很少有差异。例如，Infoseek 和 Excite 的界面和服务几乎一模一样，雅虎和 Lycos 也是这样(图 3.3 和图 3.4)。为 HotBot 做竞争性研究时，我们发现，这些界面在用户眼里混在一起，趋于一致和互换。因此，没有一个界面能传达出去这个网站而不去其他网站的理由，也没有任何迹象表明这个网站比那个网站更有价值。所以人们并不关心在用哪个网站，也不想知道为什么这个网站比那个网站好。





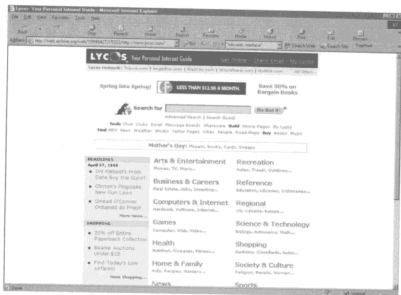


图 3.3 1999 年 5 月的 Lycos 网站

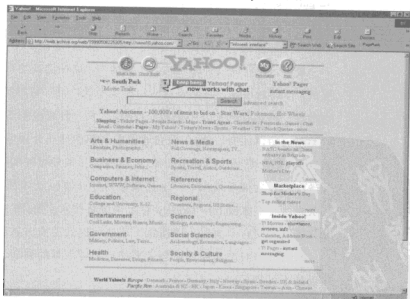


图 3.4 1999 年 5 月的雅虎网站





要发现产品和竞争对手的不同之处,产品必须要有能使自己鹤立鸡群的特征:布局、颜色、编辑语气或者其他重要东西(甚至可用性! )。CNET 网站编辑的任务广泛,同时面向大众,不能变换语气和内容,因此他们集中从视觉设计上显示身份。他们做得很成功,很长时期内,人们看到黄色阴影就能识别出他们。同样, Netscape Netcenter 采用蓝绿色,而雅虎采用“渐变灰色”,在顶部放置圆形按钮建立标识。其他网站采用了不同的战术。HotWired 每天都在网站上放一个动画首页。ICQ 首页上的文字比有些网站的全部文字都多。亚马逊上几乎所有书都可以通宵发货。即使会干扰网站功能,但所有这些都能让人们很容易记住并识别出这些网站。

认知心理学和广告中,有“识别”这个概念,识别和仅仅记住截然不同。识别意味着独特辨识,并且更快更好的辨识。虽然识别主要和视觉相关,但还涉及网站名称、网址、标识及态度。它可以包括所有这些能让网站从所有竞争对手和其他网站中脱颖而出的方面。埃里克·艾迪嘉德(Erik Adigard)是顶尖网站设计公司 M.A.D.的负责人之一,他说,界面应该做到这样:把它放到邮票大小的地方,或者和其他另外一百个界面放在一张纸上,人们都能把它一眼识别出来。

最后,还有语气,即网站的“精神”。这种交流的语气不同于程序对程序、市场对市场,它们要适应每种类型的产品、每种类型的用户。《纽约时报》网站看起来并不像报纸(没那么大,没有油墨污点,标题放置位置不同,角落还有广告),但它通过布局和字体充分说明了它的“报纸特性”,人们几乎马上就能看出来(虽然不会花很多时间阅读报纸的13岁滑板少年不一定能看出来——这也就是为什么滑板新闻网站看起来更像滑板杂志,而不像报纸)。Google 看起来像是搜索引擎(图 3.5)。为什么? 因为其首页中间有个大框,“搜索”这个词也做了突出显示。

自我推销对产品和公司会带来很小的可衡量的利润,但它对于产品成功仍是不可忽视的关键因素。



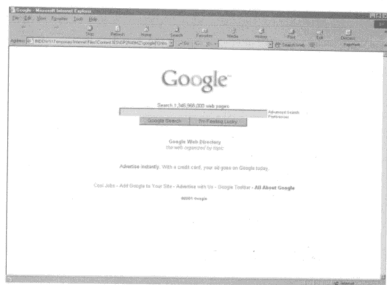


图 3.5 Google 在 2001 年左右的界面

## 平衡系统：迭代开发

如果只关注这些成功标准的某一个而忽略其他标准，势必会影响产品的长期成功。例如，电影放映前看几个预告片不错，但周五晚上看两小时的预告片不太可能会受欢迎。就像我们怀疑一个人只会说自己多么令人难忘但实际是否受欢迎一样，从长远看，只特别强调形象而忽视可用性和广告效果的设计难以取得成功(尽管能获得一些设计奖项)。就像人们很少能记得住街道限制牌一样，与平衡可用性、良好广告效果和强烈形象三者的产品相比，只注重可用性的产品在外观上会很平庸，不受人欢迎。

这些因素之间的平衡随着产品的成熟而变化，从而导致问题更复杂。在产品生命周期早期，人们第一次看到产品时，独特和功能也许是最重要的。此时，想让广告在产品上发挥最大效力可能会导致首次用户看不懂界面，这样你的品牌就与粗鲁的公司商业形象联系在一起(一般情况下，没人希望有这种联系)。但到后来，品牌已经建立，用户明白产品用途，也





学会如何使用之后，就可以增加广告数量，提升广告关注程度。

如果列出这些因素之间的交互复杂程度，会让人望而生畏。与第1章提到的快速简便可用性测试相比，就好像是要从做冰块突然变成了捕捉暴风雪中的所有雪花。需要有系统方法来整合发现问题和创造解决方案的过程，关注单个要素，而同时又不忽视整体。接下来轮到迭代开发上场了。

迭代开发的中心思想是通过试验和发现错误而不断改进。迭代开发关注目标、改进重点并完善产品，一直达到目标，而不是从一开始就要创造完美梦想。每个周期都包括相同基本步骤，每个周期都利用更丰富的信息来激发过程。不断创建、检查、继而重建解决方案，一直到通过一致、定期和可预见的方式同时满足商业需求和用户需求。

### 迭代开发怎么不起作用

在有计划的迭代开发逐渐受到欢迎之前，流行的开发方式是“公司命令”(Corporate Edict)和“瀑布方法”(仍有大量公司在采用这些方法，而且数量惊人)。

“公司命令”指有人或者小组决定要做什么，公司其他人开工，不准提问题。因为签发命令的人缺乏对问题的整体认识，从而造成这种方法存在缺陷。如果总设计师(或者不管是谁签发的命令)不了解商业环境、用户、商业合作伙伴的需求、技术能力，产品很可能会失败，有时甚至是惨败。

有一个真实例子。有一家颇受欢迎的数字贺卡公司，该公司首席执行官想给朋友发一张他最喜欢的贺卡。网站经过彻底重新设计后，他很难找到自己想要的那张贺卡。他觉得很多网站用户也会碰到相同问题，就坚持让开发团队为网站开发一个搜索引擎。经过几个月的开发，公司开发团队开发出了一个全功能网站搜索引擎。这位首席执行官利用此搜索引擎找到了他最喜欢的那张贺卡，速度很快而且方式很多。不幸的是，搜索引擎发布后，几乎没人用。经过一番研究，他们发现很多人都对寻找一张特定贺卡的概念很陌生。他们喜欢把类似贺卡分在一起，但没有兴趣发送——或者找到——一模一样的贺卡。信息架构重构后更接近了人们的期望，此时的网站比之前带搜索引擎的时候在使用和满意度(及广告收入)方面大幅



增加,消耗的资源和时间都更少。该公司通过命令创建了搜索引擎特性,错过了一条重要信息,误解了自己的核心竞争力,浪费了开发人员数月时间。

“瀑布方法”(图3.6)虽然不那么武断,但比公司命令也好不了多少。按照流水线方式工作,工作人员首先编写出确定最终产品每个细节的广泛需求文档。也许需求是建立在目标受众之实际需求和能力基础上的,但也很可能是文档作者大胆猜测、闭门辩论的结果。如果假设都是错的怎么办?如果公司需要变更怎么办?即使有内置的反馈机制,僵化的瀑布方法也几乎无法回溯,就像瀑布不可能倒流一样。如果需要回溯——实际上总是会回溯——这种模式很僵化,回溯会很昂贵。

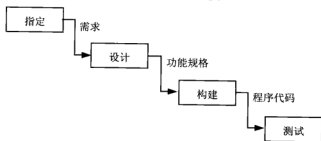


图 3.6 瀑布方法

两种方法都有同样的致命弱点:两者都缺乏内部的理智检查步骤来修正假设以符合现实环境。两种方法都要依赖于假设一开始就正确,初始数据也完整。如果最初设想有问题,那么最终产品就会有错误时间向错误人员提供错误解决方案的风险。

### 迭代螺旋

迭代开发方法应用于大型软件和制造部门已有多年的历史。它们有多种名称,仅举几例:快速应用开发、统一软件过程(RUP)、全面质量管理、联合应用开发和进化生命周期。虽然这些方法的具体细节有所不同,但都抱有通过周期性数据开发而逐步完善的内在思想,虽然需要五步或者更多步骤来描述迭代,但其背后的核心思想都可以归纳成三个基本阶段(图3.7)。



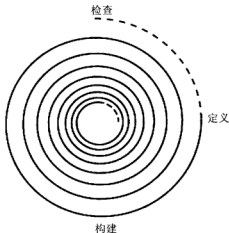


图 3.7 迭代开发：以最终产品为中心进行开发，边开发边调整

(1) 检查。本步骤试图定义问题以及受到影响的人。提出问题，分析要求，收集信息，进行研究，同时评估可能的解决方案。列出长处和短处，并分出优先级别。研究客户要求和能力，评估现有产品或者原型。举例来说，也许公司外联网能带来新客户，但支持邮箱总塞满了邮件，发件人看起来找不到要找的内容。也许是因为界面有可用性问题，但也可能是因为缺少一项基本服务，或者因为用户群体不符合预期。

(2) 定义。指定解决方案。也许外部支持邮件说明产品缺少一项基本功能特性。在该阶段，随着不断发现目标受众的真实需要及能力，产品变化的更详细细节逐渐绘制出来。

(3) 构建。执行解决方案计划。该阶段最花钱、最费时(要占用多达一半的开发时间)，如果没有检查阶段收集到的数据支持，又没有定义阶段的仔细规划，创建阶段所完成的大部分工作都是浪费。

目前为止，迭代方法类似于瀑布法，而造成迭代开发不同于流水线开发的是创建后面紧接着又从检查开始的下一轮。每个周期——从初步检查到发布之间有很多周期——并不会产生出完整产品，而是提高理解质量，从而充实功能特性集。因此，每次迭代都会对项目进行调整，从而开发过程能贯穿和顺应新信息以及商业环境变化。从理论上最大限度地减少不必

提示：

我第一次提出这些图表和原则的时候并不知道这些思想归功于巴里·鲍依姆所提出的“螺旋开发”方法。他在1988年5月的《计算机》杂志介绍了这些概念。



要开发，同时让产品更符合人们需要。

### 迭代开发的优点

#### 弹性

一开始无法知道项目所有制约因素。无论是什么开发过程，随着开发不断进行，受众以及公司的需求不断显山露水，技术限制也不断凸显。公司命令和瀑布方法都依赖于初始条件，它们很容易受到后期信息的破坏。

例如，1997年有几个月，推送技术成了自“叉匙”(Spork)以来最伟大的东西，大量公司围绕这个想法开发出整套产品和商业模式。他们没有去大量研究用户需要什么，喜欢什么，就这么理所当然地认为既然PointCast这么受欢迎，那么自己的产品也会受欢迎。

迭代方法可以在需要的时候保持灵活，项目开始的时候不需要灵活，只需要锁定与合理解决方案有关的事情。华盛顿大学信息学院的助理教授戴夫·亨德利(Dave Hendry)，他将这一过程称为“从低保真度及高度暂时性到高保真度及低度暂时性”。最初，产品很粗糙，也有很多地方要改变，还有许多基本问题需要回答。随着开发过程的继续，这些问题陆续有了答案，细节陆续得到填补，原型看起来更像是成品，灵活性降低。良好应用迭代开发能逐步减少灵活性，同时收集适当信息以确保只着眼于正确的决策。

#### 适应性

每种设计都是权衡的产物。或者说得更准确一点，任何复杂产品的设计和生都涉及很多权衡。车子又大又快，肯定不会像又小又慢的车子一样省油。如果本书采用较大字体，就势必会多印几页。每件产品的每个部分都是权衡的结果，而权衡是在产品创建过程中做出来的。几乎每次权衡都会改变产品的基本特点。有些权衡会导致产品发展成为供某一群人更多使用，有些权衡会导致产品朝着更盈利的方向发展，而有些权衡会诱使人们更渴望使用产品。理想的权衡是引导产品同时在这三个方向发展。

我们往往很难明白如何做出正确权衡。就像岛屿上不断进化的新有机





体一样，公司里的孤立想法只暴露在某些环境条件下。它必须要面对一些天敌，对付某种气候，适应某些类型的疾病。让它暴露在最终生存环境下才能知道这些想法脱离庇护所还是能否生存。然而，当把最终产品推到糟糕的大世界，让产品自立时——不像瀑布方法带来了震撼——迭代开发为了产品能在野外能蓬勃发展，会努力了解环境并预测如何调整想法。

### 共享远景

提示：

整本书中，我用的是开发团队。在这一点上，我是指负责创建和维护产品的那群人。根据公司结构不同，这群人会包括来自许多不同学科的代表。有些开发团队实际上是产品团队，负责产品的所有方面。这些团队包括视觉设计师、商业策略师、市场研究人员、质量保证工程师等。从我的角度来看，虽然团队大多数人员不写代码，不设计屏幕布局，但这些团队也都是开发团队。

除了创建良好产品外，迭代开发理念还关注整个公司不断改善用户体验和公司盈利能力。其重点已不再是创建一系列单一的一次性产品，而是不断进化工具和技术，以满足客户和公司的需求，而不管这些需求是什么，如何变化。

迭代很容易用于营销产品，设计标识，开发商业目标，或者建立支持系统。为了达到最大效果，每个负责产品的人——项目、市场营销、信息架构、设计、业务开发、客户支持——都是单一迭代开发过程的一部分。人人都需要和核心开发团队一起参与迭代过程，共享信息并改进产品。例如，确定好最初市场，营销队伍可以根据开发团队对工作属性和受众工作习惯的研究，进而研究市场细分的规模及构成。这两套信息可以合并用于产生一套想要的特性，业务开发团队能用这些特性来寻找能提供或者加强这些特性的潜在合作伙伴，或者是寻找那些对这些特性特别感兴趣的大客户。

不仅整个公司可以一起打造产品形象，而且使用相同研究方法能考虑到公司各个方面不断变化的需求，从而减少了产品发布后计划变更的需要。把产品作为系统解决方案来开发，而不是单一发布，企业能做到从战略角度利用资源，拟订长远规划，同时又能应对短期开发。

### 迭代开发问题

尽管有优点，但迭代开发并不完美。迭代开发给整个开发过程带来大量不确定因素，会使想深入特性开发的开发团队感到沮丧。迭代开发要求纪律和专门项目管理，因为迭代开发过程复杂，要求每次迭代都关注产品的一面，而此时其他突出问题也需要关注。迭代开发要求回溯检查早先假





设,可能会导致开发时间延长。虽然大多数情况下,实施迭代开发的最大困难是打造企业文化——从首席执行官到营销再到客户服务——了解迭代过程并集成到公司工作方式中的文化。

## 迭代开发与用户研究

迭代开发特别适合基于互联网的产品,因为可以快速生成原型,并进行评估。从雕刻角度看,就像先与石膏、陶土或者蜡打交道,然后再刻成石头或铜牌一样。变化可以迅速付诸实施,而发布周期可紧可松。为了达到快速反应,搜索引擎每周都可以进行一次充分迭代。

与传统软件产品不同,网站产生的直接用户体验至关重要。可用性一直都很重要,但对于传统软件,不会有人使用产品会随时不再使用的危险。此外,由于用户(或者公司)已支付了软件费用,并计划长时间使用,让软件发挥作用能获得奖励,因此即使界面支离破碎也能容忍,直到找到替代品。这种稳定性和忠诚度,几乎没有几个网站有福享受。对其他网站而言,用户体验一开始就要做对。

用户体验研究不必作为迭代开发的一部分。即使不做任何用户研究,也可以进行迭代开发(例如,极限编程作为一种高度迭代开发过程,就没有明确的研究过程),而用户体验研究方法可用于软件开发方法的任意部分。但两者结合起来,效果相当好,一个研究项目结果能回答之前项目提出的问题,同时还能指导之后的项目。

用户研究所提供的检查用户观点的方法前后一致、快速、可控且彻底。用户体验出现在螺旋开发的每个旋转处,随着问题出现而提供回答方式。在早期,收集用户背景资料,研究用户所做的工作,描述他们的问题。然后,按照人们愿望或者需要的程度排除特性的优先级别。确定优先级别后,什么类型的人群会要产品,产品能为他们做什么,他们应该了解什么,记住什么,这些都一清二楚了。在对细节进行设计和测试时,即可知道唯一需要关注的事情是如何展现,而不用关注人们的需要或者产品功能,因为这些都已经彻底研究过。如果用图来说明,那么用户研究会贯穿整个开发





螺旋，在产品开发阶段在不同时间采用合适的方法。图 3.8 把用户研究计划案例描绘成迭代开发螺旋的一部分(本书后文将讨论这些具体方法)。

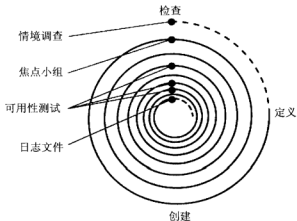


图 3.8 研究计划案例

## 案例：日程安排服务

以下例子经过简化和理想化，旨在说明迭代开发过程能对假想产品产生哪些作用。大多数情况下，事情不可能这么顺利，但本例足以解释事情的发展情况。

假设因为有了某种易于适应的后端技术，你的公司想做一个基于互联网的约会日程安排产品。

### 第一轮

#### 检查

最初将目标受众假设为大忙人，他们长期出差，需要容易访问的高级日程安排工具包。产品收入来自服务所带来的广告以及高级特性的订阅费。



在第一轮研究中，你拜访了很多大忙人，观察他们如何管理日程（一种情境调查形式，详见第8章）。你发现大忙人用现有技术（daytimer 和 Palm Pilot 等）安排工作日程非常顺畅，他们通常不愿意使用新技术，除非新技术比正在使用的技术更有用。他们宁愿不早使用新技术，除非他们知道值得这么做，而且同事也会用新技术。他们似乎担心服务和互联网整体的可靠性，比如说在很忙的一天却因为断线而无法上网，这将是一场灾难。

说句老实话，这说明目标市场对你的产品没兴趣，除非你的产品能打败现有产品。这会导致你的产品只会对一部分市场有吸引力，无法带来足够收入，开发成本无法消化。有种方法是产品寻找更大市场（也许是学生市场），但假设你决定仍继续开发原有市场，但会采用不同策略。与你谈过的有几个人表示有兴趣将日程安排解决方案用于个人生活，而不是用于工作（这部分市场似乎是已经被覆盖）。结果说明这类受众有以下特点。

- 他们的个人日程表几乎和工作日程表一样复杂。
- 他们需要与朋友和家人共享个人日程表。
- 他们无法使用办公软件，因为公司防火墙外无法使用，而且家人和朋友也不可能访问得到。
- 现有日程安排软件看起来都完全关注商务日程安排任务。

#### 定义

意识到这些情况后，你决定将目标受众锁定为繁忙的执行官，但想法有所改变以更好适应他们的生活方式。功能重点转为个人日程表共享，产品描述得以重写，目标定为帮助人们以明显优于现有方法的方式来共享日程表。产品描述详细定义了软件需要解决的问题，并明确列出了目标（或者范围、采用项目管理术语）之外的问题。同时，重新定位市场营销和产品形象工作，将精力集中于该服务的个人特质上。

#### 创建

采用新问题定义，重写产品描述，以反映日程安排应用工具的新用途，以及对受众需求的重新认识。该阶段大部分时间都花在创建产品所提供的特性以及优点的详细清单。和工程团队一起检查清单，确保所提出的特性





在软件能力范围内。此外，还需要创建初步研究计划，列出需要回答的问题，需要调查的市场，以及下一轮研究需要关注的地方。

## 第二轮

### 检查

把产品描述带给由繁忙执行官组成的几个焦点小组(参见第9章)后，你发现他们虽然很欣赏通过互联网共享日程表的想法，但他们担心安全问题。此外，他们认为这类系统最重要的部分是能快速输入信息，共享也需要便捷。有个人说，每天用共享日程表花五分钟就能搞定所有事情。他们还提到了其他功能，包括朋友、同事的日程表跟家人的日程表能单独分开，能自动获得特别活动(如体育比赛)的日程表。

### 定义

虽然核心想法很牢固，但需要解决若干关键功能需求，才能确保系统成功。软件目标中增加了安全性、输入速度提高和日程组织这三个需求，并传达给产品营销团队。

### 创建

根据这些想法，重新定义解决方案，日程安排系统采用“层”的方式，人们可以在常规日程表上增加层。有些层是家人日程表，有些是共享的业务日程表，还有一些是电视节目和体育比赛广告内容，后者不仅利用了日程表的个人性质，而且能带来潜在收入来源。你修改系统描述，包括此功能，并解决焦点小组提到的问题。

## 第三轮

### 检查

你担心“每天五分钟”的要求。非正规可用性测试(参见第10章)表明，很难做到每天5分钟就搞定一天的日程表，但如果人们渴望这种感觉，就应该要满足他们。你决定进一步展开研究，了解人们在个人日程表上真



正会花多少时间，日程安排管理中是否存在共同的趋势。(使用情境调查法)观察六个人管理日程表，你发现他们平均每天要花费20分钟——而不是5分钟——来处理个人日程表，他们最头痛的地方是不知道是整个家庭的日程表，也没法从被邀请人那得到是否参加活动的确认。你还了解到，不管在什么地方，他们平均每天要检查3到10次日程表，这便为每天可能的广告展示次数提供了参数。通过几位用户的意见，你还发现产品还有另外两个潜在市场：青少年和医生。青少年日程表复杂，涉及多人(大部分都能上网)，而医生办公室要花很多时间来安排日程、确定日程安排以及提醒预约病人。

你还对主要目标受众进行实地调查(参见第11章)，发现他们确切的技术能力和愿望，他们还使用哪些相关产品和媒体。结果发现，他们在家通常有几台新型电脑，整个家庭共用这些电脑，但每次只有一位家庭成员上网。你认为这说明所有家庭成员需要一种简单方式来使用日程安排服务，同时又不会看到其他人的日程表(并保护隐私)。

#### 定义及创建

和前两轮一样，你需要完善产品目标。目标中加入了日程表共享、家庭日程表及确认。最后，你创建了系统概况，该系统能实现所有这些目标，然后再写一份系统详细描述。

### 第四轮

#### 检查

调查还表明，你所感兴趣的家庭中每人至少有一部手机，而且使用频繁，他们对体育比赛感兴趣，也看许多电视节目。因此你决定再进行一轮焦点小组，以调查人们是否会对服务的手机界面版本有兴趣，他们需要不需要用“层”来显示运动比赛和电视时间表(两者都是吸引广告市场的潜在目标)。

这组焦点小组说，日程表共享及确认是目标受众最想要的东西，而家庭日程安排和特别活动特性虽然很想要，很酷，但不是那么重要。手机界





面版本很有趣，但有上网手机的人多从未用过手机上网，他们担心会发生尴尬和困惑的事情。你发现青少年认为共享个人日程表很不错，特别是他们可以安排即时消息提醒和聊天、可以用手机访问日程表的时候。医生——上一轮研究建议的受众，业务发展和广告人员也希望得到他们，因为他们有购买力——但医生看起来对日程安排服务没兴趣。虽然理论上有用，但他们觉得不会有足够多的病人使用该系统，因而无法抵消职员培训费用。

### 定义

利用新资料，定义两类完全不同的受众：忙碌的执行官(初始受众)和高度社会化的青少年。两者对产品展现形式的需要截然不同，即使两者都会使用基本日程安排功能。虽然意识到可能没有足够资源能满足两者，但你还是把产品一分为二，分别定义每组的受众需要和产品目标。

### 创建

你为每组新受众创建新的产品描述。纵然对青少年需求的研究不如对商人需求的研究那么完善，但你觉得自己充分理解了青少年小组的问题及解决方案，所以开始用纸原型来表达解决方案描述。根据描述，你对基于互联网和电话的界面都做了纸原型。与此同时，根据你的指引，营销团队针对主要市场，把即将投放的广告重点放在共享和邀请能力上，而针对青少年市场，他们把广告重点放在易用的电话界面和电视节目层上。

## 第五、六、七轮

产品目标受众、他们需要的功能、想要的功能、想要的顺序以及大致如何展示，这些都确定以后，就可以开始动工了。经过多轮可用性测试，而且还用焦点小组测试了营销情况。每过一轮，就越能了解如何将日程安排系统和其他产品保持一致，同时让它更容易被理解、更有效，以及保证广告内容被看到但又不显得唐突。

此外，还创建了管理系统，以便职员和赞助商可以添加和管理内容，对消费者进行测试的时候同时对职员和赞助商进行测试。



提示:

本例是产品周期高度简化和理想化后的版本。其主要重点是说明用户研究与产品开发的交互关系,并不是会持续6个月的开发周期的详尽说明。同时,这里正在发生的情况,通常项目管理任务也会发生。安排资源、建立预算、软件编写和测试,等等。研究构成开发过程的一部分,但它并不一定耗费时间,也不一定是主要驱动组成部分。本段描述强调持续研究产品用户如何能发现新市场、需要、愿望和能力,不断使其成为人们生活中更好、更强的组成部分。

第七轮结束后,产品即可发布。

## 第八轮

立即开始用户群调查,即一系列定期调查观察用户群如何变化以及变化方向中首先要做的调查。这样能让你确定赞助并利用新用户群出现后引发的特别需要。

还要着手进行广泛的日志文件分析和客户反馈分析(参见第13章),以便于了解人们是否按照预期使用该系统,他们遇到了哪些问题。

此外,继续进行实地调查研究,看看人们的其他相关需求。账单管理是一项常见日程任务,开始考虑到产品不但可以进行日程安排,还可以进行支付家庭账单,也许未来还要开发一套完整的家庭管理工具。



## 第4章 用户体验

从用户角度看，体验是呈连续性的。网站、浏览器、电脑、所处环境以及生活相互交错并彼此渗透。用户了解的东西不仅会影响他们能完成的事情，还会影响到吸引他们关注产品的东西，而吸引他们关注产品的那些东西直接会影响他们了解产品的愿意程度。如果网站在视觉上具有吸引力，也许会更促使用户花更多精力来了解和使用网站。如果他们觉得网站用起来简单，也许会促使他们更经常地使用。

如此一来，便很难定义“用户体验”，因为从搜索按钮上的文字，到配色方案，到引发的关联，到描述它的语言语气，再到客户支持，用户体验可以扩展到人们与产品交互时的几乎所有事情。了解这些要素之间的关系需要不同类型的研究，而不是只是计算能多快完成任务，也不仅仅是测试公司 Logo 有多么好记。

但是，试图一次就观察到全部用户体验会让人云里雾里，可以把它分成若干容易管理块来进行理解。就网站(及其他信息管理产品)而言，创建用户体验一般包含三类工作。

- **信息架构** 针对产品试图传达的信息而创建基本组织系统的过程。
- **交互设计** 向用户呈现组织系统结构的方式。
- **形象设计** 彰显产品的个性和吸引力。

本章将通过关注研究需要、工具以及参与这些事情的相关人员的头衔(Title)来介绍这三个方面。

### 信息架构

**信息架构**(Information Architecture)是人们体验网站最抽象的层次。人们创建的所有信息都有某种形式的基本结构，也总要有某种组织思想来定





义在一定产品中的所有信息如何互相适应。这种结构往往非常明确，例如电话簿、国会图书馆或雅虎分级目录。在这些例子中，如何安排信息基本都没有问题。

但有些时候，不太容易看清楚创建者的整个想法，这种结构是隐性的。如果信息架构不明确，它会采用另一个组织结构的比喻方式隐藏起来，用暗喻方式把一种信息和另一种信息对应起来。想浏览的人利用暗喻访问其“真实”信息架构。隐性架构有时能让事情更清楚，有时会更混淆事情。Macintosh 和 Windows 系统采用的桌面暗喻针对文件及其操作，清晰说明而且组织了大量概念，而雅虎的 Geocities 网站托管服务采用的地理暗喻几乎与网站上的内容无关。Macintosh 系统中，把文件图标从一个“文件夹”移到另一个“文件夹”，视觉运动较少，系统要求的信息量较少，而比如在 DOS 中，要边敲键盘边理解整个文件系统布局才能进行文件操作。但在 Geocities 网站上，地理暗喻要依赖于文化暗示（“硅谷”几乎没有传达用“硅谷”来命名的板块的含义），而如果从字面理解又可以分解。圣克拉拉和山景城都是位于现实硅谷中的两座城市——而如果 Geocities 在其硅谷中也有包含这两个名字的板块，它们能传达那些板块中的什么内容呢？很少。实际上，雅虎意识到了这一点，很早就放弃使用地理暗喻，而采用更明确的组织结构；现在没有了硅谷，取而代之的是电脑和互联网。

无论显性还是隐性，信息都需要某种结构，但建立结构的人甚至常常可能都没有意识到他们正在创建信息架构。他们的组织方式可能深埋在他们对某一主题的知识或者对某一任务的理解之中，而对这主题或任务，他们无法想象出其他思考方式。例如，采用公司组织图来组织企业信息网站，其信息架构对网站创建人看起来合理、明确，但从用户角度看，这种信息架构可能与用户要如何看待公司无关。用户在此类网站中寻找支持信息，不会关心会在一家台湾公司的北美子公司运营部门找到技术支持人员；他对购买的产品有问题——从他的角度看——所有其他信息都完全无关，既不是向他提出问题，也不是回答他的问题。





## 信息架构师

信息架构师的工作是让隐性架构显性，让它符合用户需要、期望和理解的东西。架构师让用户能浏览信息、并理解所看到的东西。受众即使不到达也能预测接下来要发生的事情，并能从 A 点到 B 点也不会迷路，这是最终目标。

如果基本模型不明显，人们会创造一个明显的模型。人类一直都在寻找数据中的模式，即那些能用来简化我们与世界的交互、能让我们更容易理解世界的东西。即使我们不知道怎样称呼，也无法人工创造，我们仍然会创造出图画、模式或者故事来理解我们所处的信息空间。人们总觉得他们发现了股市市场或者轮盘赌局中的模式，而绝大多数时间并不存在任何模式。因此，虽然开发团队中可能并没有特定的信息架构职位，但总有人在创造信息模式。尽管可能自己也不知道，但总有人扮演着信息架构师角色。

### 提示：

这些需要很多都与交互和形象设计师的要求相重叠，下节提到的大部分研究对它们同样有用。为了把这些方法大致按照在开发过程所处的顺序展现出来，我将介绍信息架构关注的几件事情，它们真正关注的是构成整个产品基础的东西，无论目前正在研究或者设计哪方面。

## 信息架构师的信息需求

开发过程通常一开始就会产生(或者应该产生)信息架构，所以解释它的各种研究都十分重要。具体来说，信息架构是与几个方面(受众是谁，他们如何考虑任务，他们使用什么语言，现有信息架构是否起作用)相关的信息。

准确知道哪些人会使用产品往往是创建信息架构的关键。不同人群会有不同情境来理解将向他们展现的信息，对如何使用信息也有不同假设。描述目标受众越精细，信息架构师就越能准确地使网站按照受众心目中的方式工作。“人口统计数据”是衡量任何团体的基本方法，它能描述人们的身体和就业特点。典型人口特点包括年龄、教育程度、收入、工作头衔等。

对网站而言，也要建立互联网使用特征，以总结团体的互联网体验。典型互联网使用特点包括网龄多长，花在网上的时间有多少，以及用互联网做了哪些事情。

采用恰当的术语是成功交互的最重要要素之一。大部分界面基本上都



由文字构成，而文字意思会含糊不清，容易被误解，所以人们的理解特别关键。检查人们用来描述任务的文字以及他们如何组织这些文字是产品能满足人们期望和需求的一个关键。

受众的心智模型(mental model，也称心理模型)构成了创建信息架构的第三块主要内容。心智模型可以说明人们目前如何理解主题，对如何完成或者组织一定任务他们建立了什么画面，他们使用什么术语名称和关系。例如，虽然超市中的食物主要根据食物类型或者储藏要求进行摆放(瓶装的放一起，冷藏的放一起等)，但英迪·杨(Indi Young)<sup>①</sup>开展的研究表明，人们购物时很少用这些术语来考虑食物。人们常常根据膳食计划或者食物搭配列出购物清单。他们的心智模型建立在膳食基础上，而不是建立在外形或者储藏方式上。因为网站不必跟超市一样，它可以组织成更接近人们计划膳食的方式，而不是接近食物储藏所受到的限制方式。

创建服务之前要进行心智模型研究——可以发现人们如何执行任务，以用软件界面进行模拟，而界面设计之后也要进行同样研究——可以了解用户的想法与设计师的想法的吻合程度。

## 有用工具及方法

对信息架构而言，最有用的工具指既能深入了解目标受众是谁，又能深入理解他们如何考虑信息的工具：他们如何组织信息，如何确定信息优先级别，以及他们如何命名东西。

- 对于全新服务，潜在用户特征能说明什么人会使用该服务、如何使用以及为什么使用。详见第7章。
- 对于现有服务，调查能说明谁已经在使用服务。详见第11章。
- 如果确定了目标受众，情境调查和任务分析便是创建心智模型的基本工具。情境调查能揭示(或者没有揭示、但应该揭示)典型用户使用产品的情况。任务分析能准确确定用户如何思考他们正在

① 英迪·杨的著作《心智模型》(Mental Models)对此有更详细、完整的描述。清华大学出版社，2010年出版。——编注





做的事情、做了什么假设、使用了什么文字来描述任务，以及这些任务如何与他们要做的其他事情进行交互。关于这些方法，详见第8章。

- 卡片分类，指人们把想法分成他们认为合理的类别的过程，卡片分类有助于揭示人们如何组织主题，如何命名这些分类。详见第8章。
- 分析用户填写的日记有助于揭示心智模型随时间而变化的情况。日记分析能深入了解用户期望，并理解用户更熟练使用工具时的变化。详见第12章。

这些工具产生的知识能用来确定良好信息架构，并为回答产品的其他基本问题提供情境。例如，知道哪些特性最有吸引力能马上用到产品营销上。同样，研究网站心智模型能发现目标受众未得到满足的需要，说明构思中的产品需要更多特性——偶尔——甚至也会产生其他新产品。例如，如何使用搜索引擎的心智模型显示出人们期望能用搜索引擎来搜索最新消息。但这并不属于搜索引擎应该关注的事情，但这个想法看起来仍然很宝贵，所以可以剥离出来成为完全独立的产品。

## 交互设计

传统意义上，所有用来操作机器(无论真实机器还是软件模拟)的突出物、按钮以及显示都被当成是“用户界面”来考虑。传统机器和静态信息世界中的确是这样。但动态信息世界中并没有单一界面。互联网上每一页都有不同界面。进入用户当前体验的一切东西都能被当成界面：用户看到什么、听到什么、读到什么和操作什么。界面体验不仅和功能相关，还有可读性、导航以及(害群之马)广告有关。总之，界面包含了交互的方方面面。

### 交互设计师

交互设计师控制着直接用户体验。他们决定如何根据信息架构进行浏



览，如何安排用户看到需要看到的内容，并如何保证用最清晰的方式加上适当重点来展现合适的数据。交互设计不同于信息架构，就像设计和放置路标不同于道路铺设过程一样——信息架构师决定地形最佳路径，而交互设计师放置路标，并画出地图。

## 交互设计师的信息需求

与为信息架构收集的信息相比，交互设计要求的更具体、更有针对性。确定了心智模型和目标市场之后，工作重点就会转移到交互细节。根据开发过程所处阶段，设计师需要知道的一般资料要么关于其设计是否正确，要么关于人们实际是否能够按照应该的方式做事情。

- 任务流(task flow)指发生进行有意义的事情所必需的一连串动作。研究任务流包括了解人们按照什么顺序查看要素；对下一步有什么期望，他们需要什么反馈，结果是否符合他们的预期。
- 界面的可预见性(predictability)和一致性(consistency)很关键。研究能确定人们需要多少可预见性才能顺畅执行任务流，以及不同任务流需要多一致才能让人们感到熟悉。
- 网站特性与特定界面要素重点(emphasis)之间的关系也同样重要；例如，屏幕右侧的大幅图解是否会导致注意力从左侧核心功能发生转移，在界面不同位置重复出现某个特性是否会影响人们的使用频率。
- 不同受众(audiences)。与熟练用户相比，初次使用的用户需要不同特性，使用方法也不同。青少年理解术语与40岁的人理解的不同。如果产品要服务于不同目标市场，应该要知道市场想要什么以及他们能用什么产品。

## 有用工具

任务分析是发现交互顺序的最有用工具。它还有助于建立术语，而且在一定程度上能建立特性重点。

目前，焦点小组是能确定人们优先级别事项的最简单方法。焦点小组





提示：

形象构成了产品品牌的很大部分，但不是整个部分。品牌构成的用户体验部分令人难以置信的强大，而且能把用户期望渲染到用户几乎忘记其他所有因素的地步。不幸的是，围绕哪些东西构成和影响品牌，有很多混乱认识、争论和权威发言。尽管很重要，但品牌建设以及研究——老实说我解脱了——不在本书讨论范围。

可以准确知道目标市场认为服务最有价值的地方，而且有助于确定人们对特定特性形成的交互关系(但无法知道人们的实际行为，也无法知道系统是否可用)。有关焦点小组，详见第9章。

基于任务的可用性测试常常是获取能立即用于交互设计信息的最有用工具。虽然不是发现用户需求和满足用户需求过程的最佳工具，但它是发现人们是否理解已创造的东西的最佳工具。可用性测试也是最灵活的方法之一。几乎在开发过程每一点都可以进行可用性测试，并能自始至终提供宝贵见解。第10章将介绍如何进行可用性测试以及如何分析。

日志分析能对人们去的地方及频率给出直接确凿数字。如果分析包括点击轨迹，它还能说明人们如何到达那些地方。第13章将介绍日志分析问题及方法。

## 形象设计

产品形象传达产品价值，并贯穿但又独立于产品信息架构和产品交互设计。产品形象代表着网站风格、感觉及共鸣。它代表着令人难忘的东西、独特的东西。有些情况下，它会超越产品功能的重要性(但很多营销活动都证明产品形象无法克服糟糕功能，此时产品形象并没有那么重要)。产品形象结合了网站功能、外表、联想、编辑语气以及特定特性的重点。

- 有些要素对决定网站形象发挥作用，包括网站文字的写作风格。文字能传达大量网站价值。网站有趣吗？和技术有关吗？不摆架子吗？这些因素构成网站编辑语气。
- 一致视觉主题能为一个或者一组网站创建识别要素。例如，几乎所有 CNET 资产都采用黄色背景、绿色高亮的圆形红色标志。使用过某个 CNET 网站一段时间之后，大多数人都能识别出这种组合。即使他们最后去从未见过的 CNET 网站，他们也能将它关联在一起。还可以采用更基本的视觉一致来维护形象。
- 网站强调的特性也能告诉客户网站出自哪里。两个网站可能都有购物车特性和公告栏，但把公告栏放在正中的网站，和把购物车



作为主要焦点的网站，两者给人们造成的印象完全不同。

- 通过关联现有品牌，耐克网站看起来与已有形象及品牌明显相关。在线品牌关联和定义线下品牌的关联一样：标识、颜色、口号以及共鸣等。

## 形象设计师

形象设计师的工作是传达区别于竞争对手且同时和公司其他产品保持一致的形象。虽然形象设计师的工作与网络营销密切相关，但两者有一个关键区别：形象设计师的目标是让网站体验愉快、独特，人们不用的时候也能记住网站，而不是说服他们尝试使用。

## 形象设计师的信息需求

形象设计师的信息需求类似与市场营销研究部门的信息需求。但有一些关键区别，因为设计师更关心产品的直接体验，而不是市场对产品品牌的认知和接受程度。因此形象设计师需要的信息和人们的直接情绪反应以及他们以后能记得多半有关。假设已经定义了产品目标受众，形象设计师需要知道以下信息。

- 与其他同类产品相比，产品的竞争优势。形象设计师能知道需要强调哪些特性，因为产品在这些方面优于竞争对手。
- 用户注意力方向。这部分信息包括人们看着界面时看到的东西、忽视的东西及他们认为重要的交互部分。
- 当前用户是谁。这部分信息包括他们与目标市场的人口统计情况和互联网使用构成比较，他们喜欢什么，不喜欢什么以及对产品的印象。
- 他们喜欢和理解的参照物以及联想。例如，年龄超过 50 岁的人不会像青少年一样喜欢明亮的颜色。

## 有用的工具

无论从当前看，还是从长期看，形象设计的目标都是要创造令人满意、





令人难忘的用户体验。因此，形象设计工具类似于市场研究工具，其结果能和市场营销者共享，(有时)也能从市场营销者那里获得结果。

- 焦点小组(focus group)是发现产品哪些东西最具吸引力的最快方法。焦点小组提供的信息既与人们的关注点有关，也和人们的偏好有关。
- 调查(survey)能确定现有受众的人口统计情况以及互联网使用构成。调查难以确定潜在受众，但善于找出谁在使用产品，以及他们是否正在使用任何竞争性或者互补性产品。访谈时可以用这些信息来确定用户为什么选择正在使用的产品。
- 竞争性分析(competitive analysis)能用来创建竞争性产品的特性清单，清单能用于找出最有价值特性的研究。关于竞争性研究，详见第 14 章。

## 用户体验研究

在所有人中，用户体验研究员的工作范围最广泛。用户体验的各个方面给创造优秀产品的人带来不同要求和限制。所有人都有不同需要，不同词汇，不同限制，而且常常根据不同日程表进行操作。但是他们对信息有相似又相互关联的需要，而常常自己并没意识到。深入了解产品用户、用户观点及其能力，并把这些情况在合适的时间提供给合适的人，这就是用户体验研究员的工作。研究员地位独特，可以集中所有信息——所有信息要求——并让信息具有意义，让整个开发过程更加合理化，效率更高。

把所有这些人召集在一起，同时把开发过程各方面整合成单一连贯的开发文化，这是第 5 章的主题。第 5 章将涉及如何在关键点集成研究，并同时把用户作为产品创建团队的连贯、不可分割的一部分。







## 第 II 部分 用户体验研究方法



好设计,更有爱,关注微信



idesign123



## 第5章 研究计划

注意：

只要关系到用户体验目标，这些指导便能展现从空白石板开始的理想化情况。但是许多项目情况并非如此，这些项目已经有预先假定的目标和过程。这类项目中列举目标比较容易，但如何把研究整合到现有过程来创建统一理念就比较困难。德博拉·J. 梅休的著作《可用性工程生命周期》及休·拜尔和卡伦·霍尔兹布拉特的著作《情境设计》描述了一种能彻底整合体验研究的完全迭代开发环境。

不要凭空搞研究。每部分用户研究都是一项不断发展的研究项目，即使项目不正规。研究项目正规化优点众多：能提供可使有限用户研究资源发挥更大作用的一套目标日程表及在每个阶段所需提供的结果；避免多余、冗长或者仓促研究。一旦决定进行研究，就应该马上着手制定研究计划，即使只是微不足道的可用性测试或者客户拜访。结构良好的研究计划也是一种沟通工具，可以让公司其他人能根据你的日程表开展工作。有了研究计划，同事便能了解用户研究的好处，并为他们提供向用户提问的平台，提出对研究过程的期望研究结果。但即使不把研究计划告诉任何人，研究计划仍然能证明其是有宝贵价值的，它能帮助你决定何时进行何种研究。

研究计划包括三个主要部分：为什么需要研究(目标)；什么时候进行研究(日程表)；需要多少费用(预算)。这三部分反过来可以分为实际的几块，如报告格式和时间表。

### 目标

在开始编写研究计划之前，需要知道两点：为什么需要研究？如何得出研究结果？为了对产品产生最大影响，可以用这两件事情来发现需要回答哪些问题。

前者涉及决定公司优先事项以及设置能满足优先事项的研究目标。而后者涉及了解开发过程，以便研究能对最终产品产生最大影响。

每个产品都与公司每个部门产生交互，而每个部门都有一套不同方法来衡量成功。对开发而言，成功可以是符合日程表或者符合每千行代码的



好设计，更有爱，关注微信



idesign123



缺陷数量。对营销而言，成功可以是正面评论数量和网站流量增加。对形象设计而言，成功可以是产品与企业品牌的紧密关联程度。对客户支持而言，成功可以是他们需要解决的问题数量变少了。对销售而言，成功可以是它所带来的利润。这些成功中，每种都为最终产品带来不同的角度，每种都要求有不同的用户体验。

研究可以在任何方向进行。为了获得最大收益，需要把重点放在产品最重要的特性上。但产品“特性”不仅仅指屏幕布局。产品最重要的特性指那些会影响公司商业目标的特性。因此，第一步要针对产品用户体验如何影响公司目标列出问题清单。每一个问题都代表研究计划的一个目标；它们都关注研究计划，并从公司收益最大化的角度发现产品可改善之处。从公司角度收集并组织问题，这些问题不会(也不应该)只关注有效使用产品的用户能力。广告效果、客户支持负载、品牌认知度等都是目标。

换句话说，因为用户体验影响到产品方方面面，所以方方面面都需要考虑。这个过程包括以下三个步骤。

- (1) 收集问题，并将其表现为目标形式。
- (2) 排列出目标的优先顺序。
- (3) 把目标重写成需要回答的问题。

### 1. 收集问题，并以目标形式写下来

理想情况下，任何项目开始之前，公司每个人都应该知道为什么有此项目，项目能为客户做什么，项目如何为公司业务的开展提供帮助。项目要有明确、具体及可衡量的目标，而且每个人都清楚目标。

不幸的是，生活从来都不理想。因此，确定重要问题的过程本身就是研究项目。

从找出利益相关者入手。每个部门都拥有每个产品的一部分，但有些部门比其他部门所占比重高(或者更关心)。产品经理——至少很可能是最重要的交谈对象——能帮助确定利益相关者清单，因为产品经理知道哪些团队和个人在项目占有最大利益。这份清单当然肯定包括项目、设计、营销和广告销售，但也会包含对产品成功有利的任何其他部门。如果产品



是主要赚钱的东西(或者是不赚钱),或者是全新产品,还可能会出现重要管理人员。

SPORT-i.COM 利益相关者清单样本

- 艾里森, 产品开发副总裁
- 埃里克, 交互设计
- 米歇尔, 市场营销
- 孙, 前端开发
- 珍妮特, 销售
- 埃德, 客户支持
- 利夫, 质量保证
- 琼, 形象设计
- 马亚, 罗勃, 频繁用户

如果某个部门不止一个人负责产品,应该找到最近处理产品的人。此人有可能经常处理该产品,或者能告诉你谁在处理。有了利益相关者清单,应该发现他们认为的最重要问题是哪些。可以把利益相关人全部召集到一起,花一下午时间给产品列出公司范围的优先事项,或者和每个人单独谈(对管理人员和其他忙人必须要这么做)。应该向每个人(或者部门)提出下列关键问题。

- (1) 对于每天的日常活动,产品目标是什么?
  - (2) 有没有找出产品不符合目标的方法?如果有,具体又有哪些方法?
  - (3) 你对产品有没有想回答的问题?如果有,具体又有哪些问题?
- 和所有部门代表谈完之后,列出目标及问题清单。

SPORT-i.COM 用户体验目标及问题

人物	目标及问题
艾里森, 产品开发副总裁	访问者到购物者的转化率更高。 黏性更高: 人们更频繁地回来光顾。
埃里克, 交互设计	帮助人们更好、更经常使用搜索引擎。 为什么这么多人会使用购物车,但然后又放弃使用?

#### 提示:

如果利益相关者的目标相互冲突该怎么办? 广告销售经理想通过引入额外广告单元来提高收入,而同时内容副总裁要在首页增加更多新闻报道。由于界面空间有限,这些目标可能会互相不一致。想在此阶段解决这些问题还为时过早,但调查它们之间的关系是项目的重要近期目标。





续表

人物	目标及问题
米歇尔, 市场营销	让人们知道我们会全年提供给他们最喜欢的当地体育比赛最新信息, 不管他们住在哪里。
孙, 前端开发	是否值得为了动态地图而等待 Java 插件载入?
珍妮特, 销售	到本财年结束收入提高 30%。
埃德, 客户支持	减少对过期促销的支持电话量。 把更多支持工作从电话支持转移到电子邮件支持。

和内部每个人谈过之后, 应该跟几个用户谈谈来了解他们的观点。听起来像第 22 条军规: 为什么还没有研究用户需求计划, 就开始研究用户需求? 因为把用户的观点纳入研究计划能保证研究会尽早关注用户, 而且能在公司内部建立这样一个先例, 证明利用研究结果的重要性。请阅读第 6 章, 了解如何找到用户进行交谈的提示。把这些信息加到清单中。

人物	目标及问题
马亚, 罗布, 定期用户	Sport-e.com 能不能记住他们对每项体育比赛的偏好品牌? 想知道在这些类别中哪一个对他最有价值。

作为此过程的一部分, 应该去收集公司中四处传播的其他用户体验知识, 不用开展任何原始研究就能回答研究问题。这些知识包括市场营销所做的调查, 客户支持反馈摘要, 开发小组开展的访谈以及可用性顾问编写的未使用报告。用户体验研究员能充当信息采集人和集成人的角色, 并存储所有用户相关信息, 传播用户体验研究价值信息, 并在这一过程中获得利益相关者的信任。第 15 章将描述更多信息源。

## 2. 建立目标优先级别

根据访谈情况, 你对于既定目标的相关优先级别有了一些想法。有些目标之所以重要, 是因为它们被认为会阻止人们使用产品的一个关键特性。有些目标也重要, 只是因为它们能使产品区别于竞争对手产品。还有一些目标不太重要, 是因为它们会造成资源耗费或者因为公司内部在激烈辩论某话题。



要了解研究的顺序，应该先建立问题的优先顺序。如果优先顺序不清楚，可以尝试以下练习(这些练习采用了第10章中的相同方法，此方法适用于选择具体特性)：

#### 优先级别练习

在问题清单旁，增加一栏“重要程度”。采用“1”~“5”分制顺着清单逐项打分，“5”分表示特性“必须有”，是产品成功关键，“1”分表示功能“不错”，但并非至关重要。

接下来，再增加一栏“严重程度”。本栏反映问题有多糟糕。同样采用“1”~“5”分制。“5分”表示严重问题(通常指会立即影响底线的问题)，“1分”指恼人问题，或者指应该知道的信息。

两栏分数相乘，相乘结果写入第三栏“优先顺序”。这种方法能将合并放大两个因素，从而确定什么调查什么问题。根据“优先顺序”栏排列清单，就能得到调查产品用户体验的顺序。

目 标	重要程度	严重程度	优先顺序
帮助人们更好、更频繁使用搜索引擎	3	4	12
本财年结束时收入提高 30%	4	5	20
访问者到购物者转化率更高	5	3	15

现在应该有了一张公司从整体考虑的重要用户体验问题优先级别清单。清单中不应该有超过6个左右的“大”问题，而应该12个左右的问题都是较小、更具体的问题。这些问题加在一起，再加上日程表、预算及一些更充实的目标，就构成研究计划核心。

### 3. 将目标重写成需要回答的问题

手上有了产品目标，下面开始将访谈过程中产生的目标重写成用户具体问题或者重写成需要收集的信息。把狭义问题扩展成一般主题，以便找到问题根源。例如，如果有人说到了一个具体特性，就可能需要把关注点拓宽到包括此特性存在的根本原因。

问题应该简单化，以便能产生最大效果。每个问题都会带来新的问题





和事项，其中许多都很复杂、而且互相关联。但现在是在“轻松挑选时间”。给每个目标挑选一两个问题，所选问题如果能得到回答，就能解决目标。以后研究和修订计划，就能挑选出更微妙的问题及重要交互。

重写成研究问题的 SPORT-i.COM 目标	
事 项	研究问题
访问者到购物者转化率更高	为什么有些访问者没有变成购物者？
提高粘性：人们更频繁回来光顾	人们回来光顾的原因是什么？
	用户人群的人口统计构成是什么？他们如何使用互联网？
帮助人们更好、更频繁使用搜索引擎	人们如何浏览网站，特别在寻找具体东西的时候？
为什么很多人使用购物车，而最后又放弃使用？	人们期望购物车有什么功能？购物车的什么地方让人们感到失望？
让人们知道我们会全年提供给他们最喜欢的当地体育比赛最新信息，不管他们住在哪里。	人们认为哪些服务有价值？他们关心哪些服务？他们关注哪些屏幕显示信息(他们关注屏幕上的特色促销区吗)？
是否值得为了动态地图而等待 Java 插件载入？	人们如何浏览网站，特别是需要横向移动的时候(就像使用动态地图时)？
Sport-e.com 能不能记住他们对每项感兴趣的体育比赛品牌？	个性化有多重要？哪些东西对个性化最有用？
想知道这些类别中哪些对人们更有价值？	人们在 Sport-e.com 如何购物？是根据比赛、品牌、工具、价格还是其他东西？

## 用具体问题扩充一般问题

最后一步，把较大、较抽象的问题转化为研究能回答的具体问题。

当然，单一研究项目或者单一方法都不能回答所有这些问题，但有一张能从中挑选问题的清单将有助于组合研究日常表。此外，列清单过程常常有助于定义较大事项的界限，揭示较大事项之间的关系，有时候





还能发现新问题以及假设。

一般问题	具体问题
为什么这么多人弃用购物车?	弃用购物车的人与完成交易的人比例是多少? 人们在什么页面弃用? 人们在什么页面最频繁打开购物车? 人们是否理解购物车页面上的说明? 他们知道自己弃用了购物车吗? 他们知道购物车有什么用吗? 什么情况下他们会打开购物车? 他们如何使用购物车? 他们如何在网站上购物?

## 提示

说明:

确定所提问题的符咒很简单。测试能测试的东西。如果问题答案无法变成行动,就不要问这些问题。如果测试的东西无法改变,也不要测试这些东西。例如,如果特性还没确定,就询问有关产品形象设计问题一般没什么用,因为特性会极大影响了产品形象。同样,确定交互设计之后再研究特性渴望性也是无用功,因为设计假设了人们想要产品展示给他们看的东西。

- 不要利用用户研究来证明观点,也不要为了判断立场或者强调观点而创造目标。研究过程的目标应该是发现人们真正想要的东西,以及这些发现的真实程度有多高,而不是发现观点(无论你的观点还是利益共享者的观点)是否正确。
- 透彻了解产品。如果能了解软件当前如何发挥作用,就能更准确制定研究目标。尽一切可能变成用户:阅读文档,参加培训课程,与技术支持人员讨论。
- 做好处理产品基本问题的准备。如果研究过程中出现问题——即使问题是“我们应该介入这块业务吗?”——就应该有处理问题的方法,并建立能回答问题的研究(或者至少应该把问题呈报给能解决问题的高层利益相关者)。

## 日程表

有了要回答的问题,研究计划就即将接近尾声了,但尚未完成。在它变成研究计划之前,我们需要一个日程表,而这个日程表需要整合到现在





的开发系统中。

如果运气好，已经明确定义好开发过程，而且有良好的日程表，以及具体的可交付物。但如果运气不好，研究计划可能会变成开发日程表(如果你能说服开发团队他们值得花时间这么做)。

经常和产品直接负责人进行讨论是确定实际开发过程的最快方法。产品直接负责人可以是项目经理、高级产品人员或者营销经理。把问题优先级别清单给产品直接负责人看，并充分利用开发日程表来重新安排优先级别，去掉清单中会影响短期开发效果的研究。要关注大问题，但应该在方法和日程允许范围内随时加入次级事项。

第 18 章将介绍把用户体验研究集成到企业文化的更多方法。

每个研究项目的产出不仅提供了如何改进产品的信息，而且还能延续到下一个研究项目。调查描述当前受众，便于知道应该招募哪些人。情境调查总结受众问题。焦点小组能告诉你人们对哪些问题感受最深，等等。研究所提供的每条信息都有助于了解用户是谁或者他们想要什么。这些信息可以用于关注随后的研究。

因此，虽然受到当前优先级别和开发日程表的限制，但仍应该组织项目日程表顺序，便于一个项目能让其他人了解以便带来随后的更多项目。在实践中，这意味着先收集一般信息，再收集具体偏好以及行为的信息。

### 从开始阶段开始研究

如果现有产品处于重新设计的早期阶段，以下顺序是行得通的(详见第 4 章，了解把研究整合到开发过程的其他想法)。

#### 早期设计与需求收集

- 内部发现，以确定商业需求及制约因素。(详见第 18 章。)
- 调查，以确定现有用户基础的人口统计细分和互联网使用情况。(详见第 11 章。)
- 用户目前行为的日志文件分析，如果有日志文件的话。(详见第 13 章。)



- 理想用户的个人特征记录和分析, 基于现有用户或者竞争对手产品的用户的知识。(详见第7章。)
- 现有产品可用性测试, 找出现在所存在的交互问题。(详见第10章。)
- 情境调查用来揭示用户碰到的问题(包括对产品以及辅助性的任务), 任务分析用来说明用户如何解决这些问题。(详见第8章。)
- 2~3 轮焦点小组, 以确定人们对所提解决方案是否觉得能实际帮助他们, 同时哪个功能点的价值最高。竞争性焦点小组用来确定竞争对手的产品用户认为最有价值的地方和竞争对手的产品让他们失望的地方。

实际上, 如果给定发布日程表, 一般只能采用几项方法, 但假设有 4~6 个月时间能专门用于开发过程的这一阶段, 那么对产品进行大幅改动, 这就是一张理想的任务清单。

### 开发与设计

- 用原型进行 4~5 次连续可用性测试, 来实施解决方案, 并测试原型效果。竞争性可用性测试用来确定竞争对手产品的强弱点。

### 发布后

- 调查和日志文件分析, 用来对比过去行为与现在的产品变化。
- 日记用来追踪长期行为变化。
- 情境调查用来研究人们实际如何使用产品。

## 从中间阶段开始研究

用户研究计划常常从开发周期中间开始。此时, 已经确定了这些问题: 用户是谁? 他们有什么问题? 采取哪些解决方案? 这些不能修改——至少要等到下一个开发周期才能修改。这种情况下, 研究计划应该从步骤着手, 这些步骤能在发布前给产品带来直接好处, 并能为下一开发周期计划更多基础研究。这类项目可采用下列顺序。





### 设计及开发

- 快速迭代可用性测试及竞争性可用性测试，直到发布为止。（关于竞争性研究方法，详见第 14 章。）

### 发布后

- 发布前后都展开日志文件分析，以了解客户行为变化情况。
- 调查用来确定用户人口构成。

### 需求收集

- 对现有人口开展情境调查，以确定下次发布需要解决的突出问题。

上述步骤做完之后，研究计划就作为第一个计划继续下去。

## 将研究问题组织成项目

回到前面的例子。研究问题清单中，看起来有一些关于人们当前行为的定量问题(放弃交易的人和完成交易的人的比例，发生放弃购物的网页等)，还有一些问题与人们动机和对网站的理解有关。通过日志分析可以回答第一组问题，而第二组问题主要通过可用性测试获得解答。有这么一个问题“他们在网站上如何购物？”，这个问题很可能过于抽象，难以通过可用性测试得到答案。这个问题属于需要较长时间、采用几种不同方法研究才能解决的基本问题。可以从一轮情境调查开始。

### 方法选择

挑选正确方法并进行分类有困难。使用方法经验越多，就越能更好地了解哪种方法最适合解决哪些问题。如果没有任何使用这些方法的经验，那就从本书说的内容入手，选一个看起来正确的方法。然后试试看。如果这个方法无助于回答问题，先记下来它擅长收集的信息，再试试另一种方法。

下表表示平衡后的方法。该表提供了这些方法的基本概况，当然还不够全面。



名 称	所处开发阶段	持续时间	周期时间
特征 (第7章)	开发过程起始阶段	两周以上的2~5天工作	每次主要设计开展一次,或者在定义新用户市场的时候
			<b>描述:</b> 开发者把受众描述转化成不同的虚拟角色,以理解用户需求的相关性。
			<b>优点:</b> 低成本方法,能为产品团队创建良好沟通工具。关注特定用户的需求;而不是“所有用户”的需求。
			<b>缺点:</b> 主要基于团队对用户理解,而不是基于外部研究。
情境调查及 任务分析 (第8章)	初始问题定义	2~4周,不包括招募时间	每组主要特性开展一次
			<b>描述:</b> 观察人们如何解决问题,以创建定义用户当前理解及行为的心智模型
			<b>优点:</b> 针对正在解决的问题建立广泛理解
			<b>缺点:</b> 劳动密集型
焦点小组 (第9章)	早期开发特性定义	2~4周,不包括招募时间	每组主要规范开展一次,然后在每个特性群组后开展一次
			<b>描述:</b> 对6~12位受众代表进行结构化小组访谈。
			<b>优点:</b> 发现人们的优先级别和愿望,收集人们的言行,调查群体对于这个概念的反馈。
			<b>缺点:</b> 取决于参加者的集体想法,愿望容易被误解成需求。
可用性测试 (第10章)	贯穿设计及开发	1~2周,不包括招募时间	频繁开展
			<b>描述:</b> 结构化一对一用户访谈,用户同时用产品原型执行特定任务
			<b>优点:</b> 成本低,能发现交互问题。
			<b>缺点:</b> 无法发现潜在需要,仅仅发现在操作过程中的问题。
调查 (第11章)	开发起始阶段、发布后以及重新设计前阶段	2~6周	主要重新设计之前开展一次,之后定期开展
			<b>描述:</b> 邀请随机所选受众代表填写问卷,将回答的定量摘要制成表格。
			<b>优点:</b> 定量描述受众,把受众细分成子人口,调查他们的看法以及优先级别。
			<b>缺点:</b> 无法解决人们为什么有这些看法的问题,也无法发现人们实际需要。受到选择偏见的影响。





续表

名 称	所处开发阶段	持续时间	周期时间
不间断研究 (第 12 章)	贯穿产品生命 周期	不断进行	发布后定期开展
	<b>描述:</b> 对用户进行长期研究; 通过日志和顾问委员会完成研究。		
	<b>优点:</b> 调查用户观点和使用模式随时间和经验的变化情况。		
	<b>缺点:</b> 劳动密集型。需要长期参与。		
使用日志及 客户支持 (第 13 章)	开发初始阶段, 发布后以及重 新设计前	不定	发布后定期开展
	<b>描述:</b> 定量分析服务器日志及客户支持评论。		
	<b>优点:</b> 不需要额外数据收集。揭示实际行为及看到的问题。		
	<b>缺点:</b> 无法提供行为或者问题的原因。		

如下表所示, 根据了解到的公司优先事项级别, 采用一些方法把问题进行分类, 并制定出粗略日程表。日志分析可以立即开始, 而可用性测试需要几周的准备及招募时间。情境调查可以随时开始, 但在本假设情况中, 因为没有足够资源能立即进行完整研究, 所以可以立即开始小轮研究, 同时假定以后可以进行更多研究。

内 容	时 间	问 题
日志分析	立即开始	放弃交易的人与完成交易的人比例是多少? 人们在哪些页面放弃交易? 人们最频繁从哪些页面选择购物车?
可用性测试	立即开始(招募现在开始, 2~4 周后进行测试)	人们是否理解购物车页面上的说明? 他们知道会选择购物车的吗? 如果放弃交易, 他们知道放弃交易了吗? 哪些情况下他们会选择购物车? 他们如何使用购物车?
情境调查	立即(招募现在开始, 1~2 周后进行访谈), 并持续 进行	他们在网站上如何购物? 在网站之外, 他们如何购物? 他们知道购物车是什么吗?



根据事项优先级清单,继续按上述方式扩展所有项目。扩展的时候,要寻找问题相似之处,还要寻找能把研究结合起来自动解决多个问题的地方。此外,要寻找竞争性分析会产生有趣观点的地方。

也可以从方法清单开始,并用方法清单生成进一步的问题和研究想法。例如,可以这样说:“我们用情境调查能知道什么?能解决任何研究目标吗?”,然后通过其他方法进行工作。

最终清单中的单条记录与下表所示的内容类似。

内容	可用性测试
时间	立即计划, 2~4 周后开始测试
问题	<p><b>购物车问题</b></p> <p>人们是否理解购物车页面上的说明?</p> <p>他们知道会选择购物车吗?</p> <p>如果放弃交易,他们会知道放弃交易了吗?</p> <p>哪些情况下他们会选择购物车?</p> <p>他们如何使用购物车?</p> <p>完成购物需要多长时间?</p> <p><b>导航问题</b></p> <p>人们如何找到特定物品?(他们使用什么工具?怎么用的?)</p> <p>人们如何从一个主要板块进入另一个主要板块?(使用了哪些工具?怎么用的?)</p> <p><b>促销板块问题</b></p> <p>人们在首页按照什么顺序看东西?</p> <p>人们理解促销板块的信息吗?</p>

从时间安排目的出发,可以把这个考虑成单一单元。可以把它放入到时间或者项目管理项目中。测试内容规格会根据测试计划发生变化,但这么做,你能知道什么东西可以安排时间,还能知道什么时候在测试什么东西。

## 跨多个项目进行提问

多数互联网开发日程表节奏都很快,研究计划应该集中在短期,但不能忽略长期计划。关注下一个版本的目标,而把产品以及用户的根本问题





留给长期项目，看起来总是很诱人。但从长远看，更深层次的答案恰恰最终决定着产品成功或者失败。这些问题不能推迟，因为这些问题具有普遍性或者很难回答。事实上，应该尽早开展调查，因为越早产生结果，就能越快制定出产品核心变化的计划、设计以及实施。

但如果产品即将发布，无论从产品角度，还是公司研究员角色来看，只重点关注深层问题几乎并不是一个明智的计划。因此，研究计划应该由一组平行项目构成，而长期目标贯穿若干项目。每个项目都能解决有待回答的短期问题，而且还能每次都提出几个推动认识基本问题的关键问题。

可以采用网格来表示，栏代表每个基本问题，行代表研究项目。这样就可以跟踪拟解决的项目基本目标所提出的问题。这样能防止研究过程忽视长期目标。下表代表了此类表格，黑暗阴影单元格代表哪个项目解决哪个目标收集的资料(项目 4 将对项目 2 和 3 研究的问题进行深入调查)。

	搜索引擎结果	导航理解	放弃使用购物车
可用性测试 1			
焦点小组 1			
日志分析			
可用性测试 1			
等等			

## 注意：

应该经常更新研究计划：每轮研究之间、软件每次重大更新、每次增加了用户体验知识以及公司事项优先等级变化。每周或者每两周更新一次很常见。记录每次更新有助于跟踪所有更改。

当然，并非所有项目都能给每个目标提供数据，但在脑中记住这种结构，就可以准确地满足长期的需求，而同时又能完成短期工作。

## 计划格式

谈到这点之前，我都有意回避提出研究计划应该有特定结构或者表现形式，因为这两者会根据需要和资源而变化。两者应该具有灵活性，而且适应环境。如果使用项目管理和日程安排软件，就可以表现很多计划。如果要定期向管理层呈报计划，可以采用能并入产品开发计划、更正规的书面形式。如果企业文化更喜欢使用即时贴，可以把计划贴到墙上。无论哪种形式都可以。报告应该采用符合对你有帮助且让你觉得舒服的风格，但报





告要能共享,并且能整合到更大的产品计划中。计划可以是一个文档,可以用这个文档来表达研究结构,向产品团队以及整个公司推销研究价值。

每个研究计划都应该包括以下内容。

- 设定期望。进行研究的人和接收结果的人都应该知道正在进行什么研究、如何进行以及他们预期的结果。不要对任何单轮测试设定过多期望。一轮测试无法验证整个项目有效,也无法说明整个项目不恰当,也不应该期望测试能达到这种效果。研究计划还要能表达出它具有灵活性,可以随时更改。可以用版本号甚至是过日期来说明这一点(“本研究计划有效期至2003年12月2日”)。
- 设定日程表和责任。谁该做什么?什么时候做?研究日程表如何整合到更大的开发过程?从短期看,这些内容要具体,但从长期看,会具有普遍性,用于直接与更大日程表绑定的研究除外。
- 指定目标。从整体看,每个研究项目以及研究计划应该有与其关联的具体目标。在过程起始阶段收集到的目标能驱动研究的具体化。应该明确研究具体内容。
- 指定产出。根据利益相关者的需要,每个研究项目都应该有产出,指定要呈报的信息。理想情况下,要说明实际交付的东西(报告、演示等)。

## 预算

预算基于可用资源花费,但很可能包括下面三大块。

- 人们的时间(包括你的时间以及研究团队的时间)
- 招募和奖励费用
- 设备费用

根据我的经验,要计算定性用户研究项目所需的时间(包括项目管理和时间和典型低效),例如可用性测试,大致可如下表所示进行。





任 务	时 间
单一研究项目准备工作(所有事情, 除了重复研究)	10 小时
招募以及日程安排	每位被招募人 2~3 小时
开展研究 <ul style="list-style-type: none"> <li>● 情境调查/任务分析</li> <li>● 焦点小组</li> <li>● 可用性测试</li> </ul>	每人 5 小时 每组 3 小时 每位参加者 3 小时
分析结果 <ul style="list-style-type: none"> <li>● 情境调查/任务分析</li> <li>● 焦点小组</li> <li>● 可用性测试</li> </ul>	每人 5 小时 每组 4 小时 每人 2 小时
准备通过邮件发送的报告	12 小时
根据报告准备 1 小时介绍演讲	6 小时

定量研究, 比如调查和日志分析, 根据任务复杂程度以及工具和专门知识花费的精力会有很大不同。训练有素的优秀统计员几小时完成的事情, 另一个培训较少的人要花上几天才能完成。

奖励费用详见第 6 章, 但大多数研究的奖励费用(如 2003 年春季在美国旧金山)会下降到约每人 90 分钟 100 美元。

同样, 设备费用也取决于研究野心及文档要求的空间。可以一天花 5000 美元租一间可用性实验室进行一系列可编写成文档的可用性测试, 也可以不花钱, 用借来的摄像机和上一年节日聚会剩下的磁带在会议室开展焦点小组。

## X 公司研究计划

以下摘自用户体验研究顾问英迪·扬(Indi Young)准备提交给一家消费类产品比价网站开发团队的研究计划。该计划提出广泛研究计划, 目的在于在重大重新设计之前能够广泛了解问题和用户。因为要用于演示, 所以该计划不像内部文档, 而是包括较多过程说明, 同时还包含较少用于内



部传递的细节(包括类似之前提到的所有表格),但总体而言,它清楚地勾画了此类研究计划充分扩展后会是什么样子。

### 摘要

为了对 X 公司及其竞争对手的产品进行快速用户研究,本研究计划叙述了 X 公司的需要和目标的要点,同时提出了达到目标的日程表。计划包括 5 轮可用性测试、4 个焦点小组以及开始持续进行的情境调查过程。还包括 7 月 8 号当周进行的所有研究日程表,同时提交了估计的预算。

该研究计划有效时间:从 2003 年 5 月 22 号到 2003 年 6 月 26 号,6 月 26 号同时提交更新计划。

### 研究事项

根据与设计、信息架构、产品开发、营销以及客户服务代表的交谈,我们确定了研究要解决的五大问题。

- 许多人都使用核心产品比较服务,但不到 1%的人(基于对追踪 cookie 的分析)从网上商店买过东西。
- 顶层页面获得大量访问,而较深层内容,尤其是产品详情板块访问量不大。
- 竞争对手的设计较粗糙、更混乱,但他们获得了两倍流量,而广告费用的情况也相似。
- 除了知道用户比较产品外,基本不了解用户会在什么情况下使用比较服务。
- 人们经常抱怨无法再次找到曾经发现的特定产品。

### 研究结构

研究将分成平行的两块:交互研究和现有用户人口的特征。

### 直接用户体验

为了给下次版本发布提供可操作的结果,我们将马上开始测试过程,以评估现有网站界面。这样可确定哪些设计要素最能发挥作用,哪些最有用,哪些特性最引人注目,同时能发现哪些地方无法发挥作用。还包括竞





竞争性分析，用来发现竞争对手的产品在用户体验上存在的优缺点。

采用的方法将包括4轮可用性测试以及潜在的焦点小组研究。

- 可用性测试：我们会对 X 公司每组主要目标受众的用户进行四组一对一结构化的、面向任务的访谈，每组48位用户，共20~32次访谈。每次访谈约一小时，关注人们对界面要素的了解程度，他们对结构和功能的期望，他们如何执行关键任务。然后分析访谈录像带，发现特性使用趋势及特性偏好。6月5号~6月2号每周都有一轮访谈。每轮访谈结束后2~4个工作日内完成发现总结报告，并提交给 X 公司相关部门。每轮访谈均次用当时的最新原型，并集中于当时公司利益共享者以及之前研究所确定的最紧迫的用户体验问题。
- 第五组测试任务采用相同格式和相同任务，但会针对 Y 和 Z 公司提供的服务分别进行。
- 焦点小组：如果发布前不需要额外可用性测试，我们将进行一系列3个焦点小组，每组6~8位用户，分别来自两个关键用户群的用户，会员中的研究人员以及购物者(根据营销研究所得到的市场细分研究所定义)。这些小组将关注发现用户认为最有价值服务有哪些部分，以及服务表现离他们的需要和期望差多少。
- 另外，还会进行竞争性焦点小组，通过讨论 Y 公司产品来发现用户对 Y 公司产品的熟悉程度。

### 现有用户个人特征的记录和分析

此外，我们将进行创建现有用户角色特征的计划，以及更好了解他们如何比较购物。这样能(我们希望)发现使这项服务拓展并进入人们生活的机会，并能更适合人们生活以进一步鼓励用户使用。

所用方法是1~2人的情境调查。

情境调查：我们将拜访3~5位用户的家庭或者办公室，这些用户要能代表主要目标受众。我们把拜访时间安排在他们预期会针对特定物品进行比较购物时，我们会观察并记录(视频录制和书面笔记)他们如何执行这项任务。我们会分析视频和笔记，并建立他们用来比较购物的过程模型，列举他们采用的工具和方法，他们面临的问题，他们如何解决这些问题。



### 日程表

以下日程表安排了已制定的研究。为了能在可用时间内最大程度利用研究，大部分工作都会在不同任务之间平行进行。这些可用性测试都会涉及相同数量的准备以及招募工作，两者可能同时发生在一次测试期间，而此时正在进行和分析下一次测试。

焦点小组需要大量准备工作，但因为小组本身时间相对较短(每组 2 小时)，所以可以在同一周进行(虽然日程安排允许延续一周)。还有竞争焦点小组，但这两种类型有各自形式的交付物。

情境调查项目并不是要在这段时间内完成，因为已经确定了虽然了解产品使用情境非常重要，但了解最紧急的可用性需求的优先级别更高。所以，虽然日程表上没有列出交付产出，但日程表上列出情境调查的准备工作。

### 研究日期

如下表所示。浅色阴影表示准备工作周，深色阴影表示测试及分析周。

	5/20	6/05	6/12	6/19	6/26	7/03	7/10	7/17	7/24	7/31	8/07
可用性测试 1											
可用性测试 2											
可用性测试 3											
可用性测试 4											
竞争性可用性测试											
焦点小组 1											
焦点小组 2											
焦点小组 3											
竞争性焦点小组											
情境调查											

### 预算

这些是已安排的预算，包括总估计时间和总估计花费。这些估计数字





根据经验得出，未来研究计划中将进行调整，以反映随研究进展的实际数额。

## 5 个可用性测试

准备工作	10 小时
招募及安排日程(假设有 40 名参加者——32 人参加 80 小时 常规测试, 8 人参加竞争性测试)	
开展测试	120 小时
分析测试	80 小时
写报告并呈报结果	15 小时
与开发整合(会议, 演示等)	10 小时
总时间	315 小时
招募奖励(25~40 人)	2500~4000 美元
供给(食物, 录像带等)	500 美元
总费用(不包括工资)	3000~4500 美元

## 焦点小组

准备工作	10 小时
招募及安排日程	40 小时
进行并分析小组	20 小时
写报告并呈报结果	15 小时
与开发整合	5 小时
总时间	90 小时
招募奖励	2400 美元
供给(食物, 录像等)	400 美元
总费用	2800 美元

## 可交付的成果

每次完成可用性测试的结果之后，都通过电子邮件将它们发送出去。每封邮件包括测试程序要点、参与人简介、行为中观察到的趋势摘要(适用于初始研究目标)、他们遇到的问题以及一系列支持性简介。每次测试的结果都会安排时间演示给所有受影响的人。有了演示，分析师就可以针



对结果回答问题，并对测试过程做出进一步解释。

所有常规焦点小组的结果会被收集并整理成一份报告，报告一完成就通过电子邮件发送出去。除了列出所用程序的要点和提供适用于研究目标的趋势摘要之外，该报告还会分析不同市场细分之间的观察到的不同之处。最后，焦点小组有一份单独的报告，该小组会把 Y 公司服务的用户的价值观及反应与 X 公司的用户进行对比。

## 维护

如果每次都有新知识，就应该修订研究计划，这是很重要的。随着团队和公司用户体验的认识不断增加，一切都会随之而变化。尤其是研究目标，应该重新评估、提炼和重写，以便将新增的知识都考虑在内。

因为所做的每次研究都可能会影响到对很多不同研究目标的理解，所以要尽可能整理和巩固所有用户体验相关知识。有个好办法能达到目的，建立一个微型内部网站，并把所有问题输入开发团队的缺陷跟踪软件，该网站包含所有报告和目标，还能把每个目标链接到相关的适用信息。

最终，应该就能得到一组相互关联的文档，大概构成了用户的全图。



## 第6章 通用工具：招募与访谈

无论什么研究，都要做两件事：寻找能提供最佳反馈的人，并向他们提问。招募和访谈，这两个要素构成了每个成功研究项目的基础，虽然其他章节也会提及，但应该有专门章节来讨论这两者。

### 招募

如果找错了讨论产品的人，即使其他一切都很完美，研究结果仍会是错误的，这比没有做研究还要糟糕。因为通过这样的研究，你会认为研究结果无法代表真实用户的观点和行为。从烤面包机到导弹制导系统，每个产品都有目标受众。你需要了解实际使用、理解并购买产品的人的体验。其他任何人的体验都只有边际用途价值，甚至还会具有欺骗性。因此，如果要做导弹制导系统，却邀请北美烤面包机爱好者来讨论应该如何改进制导系统，得到的反馈就几乎不可能有助于改进为军队生产的产品(生产出来的导弹很可能带有30秒钟弹出定时器☹)。

为研究而去寻找、邀请并给合适的人安排日程，这一过程就称为招募，包括三个基本步骤：确定目标受众，找到受众代表，说服他们参加研究。

被招募者会随着项目发生变化，但招募过程一般需要两周时间，典型的招募日程表大致如表6.1所示。

表 6-1

时 间	活 动
t-3 周~t-2 周	确定一个或者多个目标受众。
t-2 周	招募初始群体，或者如果有合格的参加者的数据库，可以从中预筛选出初始群体。





续表

时 间	活 动
$t-2$ 周~ $t-1$ 周	筛选出最终候选人。
$t-2$ 周~ $t-1$ 周	发送邀请给主要合格候选人。
$t-1$ 周	发送邀请给次要合格候选人。
$t-3$ 天	为所有候选人创建日程，同时创建备用候选人联系名单。
$t+1$ 天	跟进参加者和研究员。
$t$ = 测试时间(一天或者多天)	

忠告：

要请人进行全职招募，而不要在其他工作之间进行招募。招募本身就属于全职工作，而且工作量总是超乎预期。为了降低成本，培训一下临时工或者实习生，他们就可以胜任全职招募工作。

招募到一位参加者并安排好相应的日程至少需要准备两小时的时间，如果从未招募过这类参加者，需要的时间会更多。

### 挑选受众

任何用户体验研究开始之前，要充分了解谁会使用产品。如果受众的背景资料不完整，产品也缺乏成熟目标，项目就会处于没有人想要、也没有人需要的危险境地。

招募开始之前，要确定受众的背景资料，并在研究开始后与被招募者进行更新。从受众的人口统计、互联网使用和技术背景资料入手，并缩小范围。如第4章和第11章所述，人口统计背景资料用来描述人的身体及就业特征；互联网使用背景资料用来描述人们的互联网经验；而技术背景资料用来介绍他们普遍的计算机技术经验。你需要这些高素质的人成为你的研究参与者。例如，普通用户是否访问过竞争对手网站并没有关系，但如果要招募竞争性分析焦点小组成员，就需要招募有使用竞争对手产品经验的人。如果要对网站新版本进行可用性测试，可能就要招募使用过老界面的人。或者可以反过来，挑选从未见过产品的人，看看他们对“新鲜事物”有何反应。这些因素一般不会从整体上对确定产品目标受众发挥很大作用，但对于确定如何组成优秀用户研究受众很重要。

招募步骤开始前，要针对具体研究明确理想研究受众。可以从建立产品目标受众的人口统计和互联网使用的背景资料入手。如果所开发的产品能帮助人们安排活动日程，主要受众就会是要参加很多聚会并且有一定互





联网经验的年轻专业人士。

### 事件日程安排的背景资料

#### 人口统计

年龄：25~35 岁

性别：男性或女性

受过大学教育

年收入：60000 美元

考虑产品营销，在目标市场中，用户的典型因素可能包括用户的可支配收入，但如果发现这是用户在整个目标用户群中所占比重很大，就把这个作为研究因素之一，也许不会为你的研究带来什么结果和好的反馈。问问你自己，研究受众与整体目标市场之间的区别。什么样的人能对研究目标给出最佳反馈？

- 哪些市场细分最受研究影响？
- 只有一个群体还是有多个群体？
- 哪些因素对研究的影响最大？
- 哪些是愿望特征？哪些是非愿望特征？

探讨这些问题的答案，并编写相应的背景资料。删除不影响人们使用或者看待产品的因素，并增加信息。要关注让他们成为理想研究对象的因素，并分离出来。

### 事件日程安排研究的背景资料

#### 人口统计

年龄：25~35 岁

性别：男性或者女性

受过大学教育

#### 互联网使用

在家或者上班时必须要有个人电脑

必须使用互联网



续表

## 互联网使用

必须有一年或者更长时间的互联网经验

必须每周使用互联网 3~5 小时用于个人任务(购物、研究等)

要有使用互联网效率应用工具的经验(日历、电子邮件、文档管理等)。

## 行为

要定期举办社交活动,至少有另外三人参加,而且最近举办过一次

社交活动必须会明确涉及邀请他人

小心,不要过分确定目标受众。如果发现自己要对不清楚的地方定义多个条件,可以物色能见面的人,或者如果某些限制条件相互排斥,可以考虑将研究分成若干群体进行。因此,如果要给卡车司机做网站,那么长途司机和短途司机的看法也需要考虑,不能把两个群体混在一起研究,也不能只找开两种车都有经验的司机,可以考虑对两个群体进行相同研究。

## 找到受众

提示:

请商业招募公司一起招募还是自行招募,很大程度上取决于现有资源。招募过程非常耗时,尤其是要招募比较罕见的群体或者大忙人。但如果组织得当,加上足够的时间,招募过程基本还是相当简单的。

有了研究背景资料,接下来就是物色符合背景资料的人。有两个步骤:先多找一些人,再找到合适的人。

商业招募师有各种行业从业人员的数据库。对他们来说,找到人的任务相当简单:搜索数据库,进行基本人口匹配,然后根据已确定的互联网使用、技术和行为标准进行筛选。如果要采用商业服务进行招募,接下来的任务就是建立适当剧本,招募师可以根据研究的特定标准来筛选人口。这种剧本被称为筛选文档。不过,如果是自己招募,就需要建立自己的数据库。

## 个人数据库

建立自己的数据库并不困难,但很费时间。如果要为下个月的研究进行招募,最好还是要借助专业服务。但数据库一旦建成,这种宝贵资源可以使用多年,并能用于所有不同类型的研究。

数据库中的数据越多越好,但前提是这些人愿意使用、了解并欣赏你





的产品。如果数据库中都是些不合适的人，使用数据库所带来的麻烦远远多于它能够产生的价值。

可以使用各种方法来丰富数据库，但采用电子邮件进行招募是目前最有效的方法。首先，设置一个用于招募工作的专用电子邮件帐户。要把所有招募相关信息收集到一个地方，专用电子邮件帐户能迅速完成这项工作。

### 从家人和朋友入手

首轮招募中，给公司每个人发邮件，或者在公司新闻邮件组中发通知。邀请同事的朋友、室友、重要人物和家人通过参加研究来帮助公司(当然有酬金)。问哪些人与被测试产品没有直接关系。请他们发送电子邮件到专门设置的电子邮箱。

回复这些电子邮件，解释整个研究过程，同时附带一张问卷，询问人们的基本人口统计和互联网使用情况。完成问卷的时间不要超过5分钟，内容要尽可能具体，以便于比较答复。问卷中的问题要有助于了解与需要进行研究的产品相关的用户行为，但问题不要太具体，否则你会发现用户答案存在相当大的差异。因此，在问卷中设置问题询问受访人在很长一个阶段经常购买的产品类别比询问他们上次购买的产品会更有用。

感谢您有兴趣帮助\_\_\_\_\_做出最佳产品。我们不断努力以创造有创意、引人注目、有用而且可用的服务和产品。作为一部分工作，我们正在实施一项制度，邀请一般公众评价正在开发的新产品，以确定这些产品哪些地方设计成功，哪些需要改进。

因此我们邀请您来这里。

在向您展示我们的更新、更好的想法之前，我们需要了解一些信息，帮助我们确定哪些人更适合哪些产品。因此我们编写了以下问卷。请填写并尽快发回以下地址。我们会尽快与您联系，安排评估日期和具体时间。根据产品的准备就绪情况，可能需要等待几周到几个月。

问卷完成后，请发送至：<mikek@adaptivepath.com>。



## 问 卷

说明：多项选择题，请在最佳选项前的中括号内填 X，如[X]。

1. 您每周有多少小时的上网时间用于工作？

- ☐ 没有
- ☐ 少于 5 小时
- ☐ 5 ~ 10 小时
- ☐ 10 ~ 20 小时
- ☐ 20 ~ 40 小时
- ☐ 超过 40 小时

2. 您每周有多少小时的上网时间用于娱乐？

- ☐ 没有
- ☐ 少于 5 小时
- ☐ 5 ~ 10 小时
- ☐ 10 ~ 20 小时
- ☐ 20 ~ 40 小时
- ☐ 超过 40 小时

3. 您最经常看哪些信息？

4. 您主要从哪里获取新闻和信息？

5. 您最熟悉哪种计算机平台？

- ☐ 苹果 Mac
- ☐ 微软 Windows
- ☐ Linux/Unix

其他(请说明)

6. 您最熟悉什么软件？

7. 您有时间来我们在旧金山的办公室参加评估吗？

个人信息

姓名：

请用[X]选择您喜欢的联系方式。

☐ 电子邮件地址：





[ ]电话:

[ ]传真:

十分感谢。

—迈克·库涅夫斯基<[mikek@adaptivepath.com](mailto:mikek@adaptivepath.com)>

Adaptive Path 公司

提示:

要找的人数应该是预期参加第一轮研究的人数的10倍。

有了答复,接下来需要把答复收集到数据库中。从数据库标准来看,这种数据库不会是大型数据库,所以并不需要采用特别重量级的软件。微软 Access 或者 FileMaker Pro 都是优秀的通用数据库程序,实际上表格处理软件都可以,只要能对所有字段进行全文检索即可(但如果有了 100 或者 200 条记录,基于表格处理软件的数据库会变得很难用,因此从长远看,专门数据库软件更容易维护)。发出请求邮件后,要安排足够时间——如果在一家大公司,每天至少要几小时——录入数据。如果邮件列表很大,估计会有很多答复,可以考虑采用互联网表单自动化输入某些数据。通常情况下,只要电子邮件脚本能把输入内容按照一致形式导出,就足够大大提高数据录入工作效率。

### 拓宽视野

招募到第一批朋友和家人,完成第一轮研究后,接下来需要继续建立数据库。如果每月要进行的研究方案包含 4~5 个研究项目,大约需要 500 人左右的数据库。如果要进行定期研究,应该把建立数据库作为定期“基础设施维护”的一部分例行工作,总是从低层次开始。

通过这些方法招募到了最初的用户,现在还需要一些方法继续招募更多、更大的用户群体。但找到这些群体具有一定挑战性。没什么简单方法。对于某些研究受众而言,比较容易找到愿意参与研究的人,例如,“想买杂货的人”。而对于另外一些而言,就很难找到愿意参与研究的人,例如“财富 500 强的首席执行官”。下列方法在过去的实践中发挥过作用。

- **社区电子邮件列表以及在线公告栏。**有些社区有自己的电子邮件列表。无论是生活在一定地区的人,或是使用某种产品的人,还是在同一行业中的人,世界上都有大量邮件列表。电子邮件一



般不会发送给不可能参加研究的人，因此这些覆盖特定城市或者地区的邮件列表可能最有价值。应该确保了解社区规则，而且只有获得了组织者的许可，才能发送邮件。永远不要发垃圾邮件(发送未经许可的邮件)。垃圾邮件没有任何作用，而且人们收到时会很生气。邀请信要突出个性化，并明确说明为什么要发送邮件和参加研究有何好处。

- **邻居。**邻居公司是很好的研究资源。他们可能很熟悉你的公司，很方便在午餐休息时间去你的办公室。给邻居公司的办公室经理写张条子，请他们帮忙在公共场所张贴邀请海报。
- **用户。**如果网站已经收集到了用户信息(无论是通过销售还是通过市场调研)，就可以根据清单邀请用户。筛选出和你在同一个地理区域的人，并发消息给愿意接收公司邮件的人。不要发送垃圾邮件，并清楚说明选择他们参加研究的原因和研究性质。
- **新员工。**新员工会带有偏见，但他们对内部人士和项目不够熟悉，因此他们能受到激发，而且时间宽裕。我知道有家公司会邀请每位新员工，不管是接待员还是副总裁，刚进公司就参加研究项目。但要合理把握员工参加研究的道德和法律问题的敏感性，他们可能认为是对他们能力的评估，所以要在邀请中明确说明研究不属于工作要求范畴，而且评估结果不会用来评判他们的工作业绩。
- **参加过研究的人。**符合招募标准的人可能会认识与其相似的人。如果对相同目标受众进行多轮研究，可以请被招募者向其朋友讲清楚研究情况及如何加入(甚至还可以做一些宣传页加以说明，请他们带回去)。
- **广告。**可以在自己的网站和其他网站上投放广告。确保广告明确说明招募的性质，并说明参加研究有奖。确保广告具有吸引力(毕竟是广告，免不了与其他广告竞争)。例如，动画横幅广告可以写成“住在旧金山？想赚 60 美元吗？点击这里！”
- **传统方法。**需要能来办公室的人，因此也可以采用传统广告方





式。可以在当地报纸刊登分类广告，但把广告放在求助栏可能会被误认为是工作机会。如果在城区，可能会有传单分发服务，广告能分发到城镇每个角落(大学校园布告栏是特别好的资源，而且如果大学生也属于部分受众，这是招募大学生的好途径)。如果你雄心勃勃，甚至可以买块广告牌(如果你真这么做了，请告诉我一声，因为我还从未听说有人真这么做，我想知道广告牌能发挥多大作用⑤)。

如果你有些困惑，不知道如何定位目标受众，可以先找到符合研究受众背景资料的一个人。然后跟这个人坐下来进行头脑风暴，一起想想其他方法以找到与他类似的人。谈谈这个人出入的地方——线上线下都可以——还有习惯和爱好。向他介绍前面清单提到的一些方法，并询问如何调整清单才能符合他们这类人的习惯。

但不要招募具有相同特质的人，也不要招募用同样方式招募过的人。背景资料的每个属性都有一组值能满足，要尽量招募值范围内的人。例如，如果参加者年龄界定为 25~35 岁，那么被招募者最好能覆盖整个年龄段，而不要全是 35 岁的人。

### 更多招募技巧

- 除了向人们发送电子邮件阐述要点外，还可以用网页来阐述研究方案，而无需在电子邮件里进行大段解释。这些网页可以由多个相同页面组成的产品原型或者是网页，当然也可以是纯粹的信息页。网页上清楚说明招募目的、奖励机制、研究的临时性，人们如何获得更多信息。随着招募进程渐趋成熟，可以在网页上添加一些常见问题。
- 在所有邀请中清楚说明研究地点以及希望招募的地理区域。由于无法预知谁会看到邀请，因此一般情况下，不要邀请住在地球另一端的人(甚至是 100 英里外的人)，除非你愿意支付交通费用。
- 跟踪人们何时以及如何发现你的方案。这样能知道哪些方法最有用，此外，参与者的招募方法还能使你对参与者有一个预判。如





果从你的公司买过东西，这些人会更了解你的产品，并且比通过邻居招募到的人具有更强的观点，而后者又比从报纸看到广告的人具有更强的观点。

- 每6~8个月就要更新一次每个人的信息，以了解其信息是否已发生变化，以及他是否仍然对研究有兴趣。要留意数据库最近的一次更新，并定期发送电子邮件给最后更新日期超过6个月的人，请他们确认他们在数据库中的信息是否与实际相符。
- 尽量减少重复招募的数量。大多数情况下，如果近6个月内没有参加过任何形式的可用性或者市场调研，一般都可以招募。但在某些情况下，必须要降低要求——比如，时间紧张，被招募者的背景资料罕见，或者被招募者属于经常被招募的群体(例如IT经理的数量很少，但同时又是很多市场研究的目标)。
- 如果有人的反馈特别好，务必将它记录到数据库中(可以通过“优秀反馈”的复选框形式)。如果短期内需要一两个人，就可以找到表达能力强而且又有见解的人。
- 如果时间允许，可以进行小型招募测试。先招募一两个人，并通过他们的表现进行研究。如果能提供良好反馈，他们就属于合适的目标对象，就可以用同样的方式招募其他人。

### 预筛选

有了潜在参加者的数据库，接下来需要开始找到合适的人，邀请他们参加研究。在研究开始前的两三周，应该预先筛选数据库，筛出打算邀请的人。通过数据库搜索工具，筛选出你感兴趣的那部分用户普遍互联网使用参数模型。尽量找到能代表目标受众范围，同时又能符合参数的人。尽可能利用收集到的资料，不断将名单缩小到最有希望的人。例如，如果要测试服务的聊天室组成部分，可能就需要寻找把社交工具作为最喜欢网站的人。

从数据库中选好，形成初始名单，再把名单分为主名单和次名单，标记出最有希望的人。之后会用此名单来确定安排参加者和时间的优先





### 提示:

不要把邀请或者筛选文档发送给整个名单。否则,对不适合的答复进行排序,还会惹恼不够标准或者没有兴趣参加的人。

级别。

确定名单后,接下来把筛选文档发送给每个人。

### 筛选文档

筛选文档可能是招募过程中最重要的部分。筛选文档是个脚本,可以筛选出能提供良好反馈的人,并去除掉仅仅符合之前我们确定的参数模型的人。招募师可以通过电话或者电子邮件阅读筛选文档。

合理使用筛选文档可以找到对产品感兴趣的人,可以巧妙优雅地谈论其经验。如果筛选文档使用错误,情况最好的时候只能找到对产品勉强有兴趣的人,而情况最坏的时候,找到的人对产品既没兴趣,表达也不清楚。

筛选文档会随着项目不同、招募师不同而变化,但大多数筛选文档都有一些通用规则。

- (1) 问题数量控制在 20 个内。游戏存在总有其原因。约 20 个问题就可以了解一个人的所有情况。对于大多数目标受众, 10~15 个问题就够了,如果通过数据库进行预筛选,不到 5 个问题就够了。
- (2) 保持简短。整个筛选过程要控制在 5~10 分钟。
- (3) 清晰具体。要让受访者明确知道应该提供什么答案。
- (4) 不要使用术语。语言要简单、直接、清晰。
- (5) 询问准确日期、数量和次数。这样能消除一个人的“偶尔时间”会是另一个人的“所有时间”的问题。
- (6) 每个问题都要有目的。每个问题都要有助于确定某人是否属于受众。不要使用偶然问题,或者“知道了会很好”的问题,因为这些问题的答案无助于招募,而且还会浪费大家的时间。可以在测试开始的时候就提出“知道了会很好”的问题。
- (7) 从一般到具体排列问题,特别是进行电话筛选的时候。问题在筛选文档中出现得越早,被去掉的潜在参加者就越多。这样既节省了招募师的时间,又节省了参加者的时间,因为较先提出的问题可以避免之后出现一些不相关问题。例如,如果年龄是确定目标受众的重要因素,而互联网经验的多少可以非常灵活,就要把年



龄问题放到互联网经验问题之前。

- (8) 问题不要有引导性。问题中不要带有任何价值判断或者答案提示。“您受到过互联网延迟时间的困扰吗？”这个问题就具有暗示性，即暗示人们明白延迟，并应该受到过困扰。相反，要用更普通的方式进行提问，再寻找具体答复。“网络上有没有事情经常打扰您？如果有，是哪些事情？”这类问题可以和前一个问题一样得到答案，而且不带偏见。
- (9) 清楚说明研究形式。要说明研究目的、方式、花费时间、奖励、和参加者是否需要提前做准备(或者参加者是否什么也不需要做)。
- (10) 保持弹性。要让招募师知道每个问题答案的可接受范围，以便不会漏掉会提供有价值反馈的人。

#### 电话筛选文档样本

以下电话筛选文档是为一家在线日历网站所准备的。目标对象多为网站现有用户，但筛选文档做了扩展，包括了一些潜在用户和高级用户(以便了解不同类型的用户体验)。

筛选文档分为三部分：为招募师介绍哪种类型的受众为理想目标受众；招募师为参与者所做的介绍；询问参加者的主要问题和每个问题的筛选条件。筛选文档较为复杂，对于首次用户测试而言可能不太容易懂。

##### 目标受众

总共 6~9 人：4~6 位当前用户，1 位高级用户，1~2 位非用户(但可能是潜在用户)

##### 当前 eCalendar 用户

定期使用和最近使用过 eCalendar 的人

男性或者女性

任意年龄，但最好是 35~50 岁之间

个人收入 6 万美元

家中或者上班时能上网

每周使用互联网时间的不少于 5 小时

#### 忠告：

虽然很容易想到招募多个目标受众，但受众描述要非常接近，同时被招募人的主体要来自单一目标受众，这两点是关键。否则就会出现数据分散的局面，由于缺乏足够信息而无法识别模式。先确定主要受众的优先级别，同时招募足够多的人(通常为 4~6 人)，然后再为次要受众进行招募，这种方法能避免以上情况的发生。





有一两年互联网使用经验

上个月至少创建了 5 个 eCalendar 事件

至少使用了两个月的 eCalendar

#### **eCalendar 高级用户**

频繁使用 eCalendar 和定期使用其高级特性的人

男性或者女性，但最好是男性

任意年龄，但最好超过 40 岁

每月至少创建 20 个 eCalendar 事件

使用过提醒和日历叠加功能

[其他背景资料与当前用户相同]

#### **潜在用户**

从未使用过 eCalendar，但对 eCalendar 有所耳闻

每周至少有 5 个约会

[其他背景资料与当前用户相同]

请注意，有些标准很具体，比如上网时间、记录进日历的事件数量，而有些标准只是一个范围(比如性别和年龄)。

#### **时间地点**

1 月 20 号或者 21 号，上午 8 点到下午 6 点

可以来旧金山市中心的办公室

和 eCalendar 无直接附属关系，也不是竞争对手(但可以听说过 eCalendar)

最近 6 个月未参加过可用性或者市场调研

#### **当前用户筛选文档**

您好，我是来自\_\_\_\_\_的\_\_\_\_\_。我们想找几个有兴趣参加付费产品评估的人，您可能会对此产品有兴趣。这不是推销电话，并且任何时候都不会有任何形式的推销或者销售引诱。

评估是 1 月 20 号或者 21 号在旧金山市中心进行的一对一访谈。评估会在工作时间进行，约持续 1 小时。如果参加，您会收到 60 美元现金。



访谈内容将严格用于研究，您提出的所有意见都会保密。如果您有兴趣参加，我需要问您几个问题，看看您是否具备我们要找的使用经验。

#### 提示：

还可以采用更通用的介绍（“我们正在进行一项研究，想问您几个问题”），但有可能在完成整个介绍过程后，又发现某人在某天没空，或者无意参加这类研究。

介绍既能设定人们的期望，还可以作为第一轮提问，因为人们会立即说明在指定日期是否有空，或者他们是否愿意参加。要仔细描述研究的一般性质，同时不要说出会影响人们答复的描述。

问 题	回 答	说 明
1. 您或者您的家人是否从事以下行业？	<ul style="list-style-type: none"> <li>● 市场研究</li> <li>● 广告或者媒体销售</li> <li>● 公共关系</li> <li>● 可用性或者质量保证</li> <li>● 网站设计或者开发</li> </ul>	如果答案均为肯定，则终止。

去掉所从事行业可能会产生利益冲突的人。去掉从事广告、可用性研究、网站设计和市场调研工作的人，因为他们会特别留意研究要发现的各种问题，所以他们给出的观点不大可能不带偏见（即使他们想不带偏见都不可能）。

“终止”用来告诉招募师停止招募过程并结束访谈。筛选文档末尾有一些终止措辞。

问 题	回 答	说 明
2. 我们需要不同年龄的人。您属于哪个年龄阶段？	<ul style="list-style-type: none"> <li>● 小于 30</li> <li>● 30~34</li> <li>● 35~39</li> <li>● 40~45</li> <li>● 46~50</li> <li>● 大于 50</li> </ul>	终止，跳至问题 3。 终止
3. 我们还需要不同收入的人。您的年收入大概有多少？		如果小于 6 万美元，则终止。否则，跳至问题 4。
4. 您在家或者工作时是否有个人电脑？	<ul style="list-style-type: none"> <li>● 有</li> <li>● 没有</li> </ul>	跳至问题 5。 终止
5. 您在家或者工作时是否能上网？	<ul style="list-style-type: none"> <li>● 能</li> <li>● 不能</li> </ul>	跳至问题 6。 终止





续表

问 题	回 答	说 明
6. 您什么时候开始使用互联网?		如果少于1年,则终止。否则跳至问题7。
7. 您平均每周使用多长时间互联网?		如果少于5小时,则终止。否则跳至问题8。
8. 您每周要跟踪多少安排好的会议或者事件?		如果少于5个,则终止。否则跳至问题9。
9. 您上周用过在线服务跟踪安排好的会议吗?	<ul style="list-style-type: none"> <li>● 用过</li> <li>● 没用过</li> </ul>	跳至问题10。 终止
10. 跟踪了多少个?		如果少于5个,则终止。如果多于20个,跳至问题11。否则跳至问题13。
11. 您使用过哪些在线服务?		如果提到eCalendar,跳至问题12。否则终止。
12. 您使用eCalendar多长时间了?		如果2个月或者更长,跳至问题13,否则终止。
13. 您使用过eCalendar的哪些特性?	<ul style="list-style-type: none"> <li>● 提醒</li> <li>● 地址簿</li> <li>● 日历叠加</li> <li>● 好友名单</li> </ul>	如果是日历叠加和提醒,考虑安排高级用户日程,跳至问题14。
14. 您正在使用eCalendar.com或者其公司提供的在线日历来管理项目吗?	<ul style="list-style-type: none"> <li>● 使用</li> <li>● 不使用</li> </ul>	终止 跳至问题15。
15. 您是否参加过市场研究访谈或者讨论组?	<ul style="list-style-type: none"> <li>● 参加过</li> <li>● 没参加过</li> </ul>	跳至问题16。 安排日程
16. 上次参加是什么时候?		如果是6个月前,则终止。否则,跳至问题17。
17. 请用几句话描述一下您最近最喜欢的网站及原因。		[记下来]如果表达不清楚就终止。否则就安排日程。
终止		问题结束。感谢您参与。我们现在无法安排您参加本次项目,但我们将来可能还会再次邀请您参加其他研究项目。

提示:

潜在用户筛选文档一样时,去掉问题12、13和14。



虽然如果是从数据库挑选出了很多名字，也已经滤掉了许多问题，但还是要确保信息的准确性，所以要再次核实信息。此外，有关年龄和收入问题可能会冒犯一些人。实际上在很多情况下，这些信息并不会影响到人们的行为，可以去掉这些问题，如果是次要信息，可以把问题挪到最后面。

#### 提示：

如果时间会促，或者只能采访几个人，有时可以从访谈过的人中间进行筛选，而不要去掉这些人。找到能提供良好、可信、明确反馈意见的人是获得有用信息的快捷方式。这种情况下，要告诉分析师参加者的背景，因为这可能会影响分析师解释结果的方式。

如果最近参加过用户研究，人们可能会不自觉地给出带有偏见的答复，因为他们熟悉研究形式，能预测到“合适”答案。而我们希望的是没有经过事先准备、没有偏见的答复，因此要从研究中去掉最近参加过用户研究的人，除非没有其他选择。此外，有些人觉得奖金是增加收入的好方法，会为此而愿意参加任何形式的市场调研或者可用性研究项目。要完全避免邀请这些人，因为他们不太可能会提供自然或者真实答复。

类似上表中的开放式问题有两个目的。招募师能了解潜在参加者的表达能力，还可以收集到不太容易转化成多项选择问题的信息。要保存这些信息，而且由于很费时，因此每个筛选文档中不要超过一个开放式问题，但也不能过滤掉许多人。（也就是说，有些招募师喜欢一开始就提出开放式问题，因为不会把人吓跑，还能及早发现表达不清的人。）

问 题	回 答	说 明
结语		我的问题就这么多。非常感谢您的参与。尽管我们不会马上安排合适的人选，但我们会再次致电你参与我们的另一个不同的项目。

#### 提示：

如果采用更一般的介绍，也可以通过其他不太具体的方式代替终止说明，比如“这就是我要问的问题。非常感谢您参加我们的研究。”

### 电子邮件筛选文档

就所提的问题种类而言，电子邮件筛选文档和电话筛选文档相类似，但前者要求参加者做出回答。电子邮件筛选文档在形式上更接近调查问卷。电子邮件筛选文档需要清楚说明所需答案范围。问题要更精炼、更简单。参加者会觉得比较容易填写，因为他们不会觉得必须要写大段说明，因此更愿意回复（对多数人而言，回答 20 个问题非常枯燥而且费时）。类似下列问题：

您每周要跟踪多少个安排好的会议或者事件？





在电子邮件中要采用如下措辞：

正常一周时间中，您大概要跟踪多少个安排好的会议或者事件？(请选一项)

- ☐ 没有
- ☐ 1~5 个
- ☐ 6~10 个
- ☐ 11~15 个
- ☐ 16~20 个
- ☐ 超过 20 个

以符合电子邮件程序的方式书写邮件，并格式化，然后把筛选文档通过电子邮件发送给邀请名单上的人(千万记住将“回复：”地址设置成研究专用电子邮箱！)。避免使用 HTML 格式的电子邮件，因为许多电子邮件程序无法正确显示 HTML 格式的邮件。此外，如果有快速创建在线调查的系统，也可以建立网页，把所有问题放到快速填写表单中(但要提供人们能通过邮件答复的方法，防止无法上网或者访问在线表单出现问题的情况)。

### 日程安排

有了潜在参加者名单，接下来需要创建日程表。首先要定义日程安排表。什么时间对研究人员而言合适？有没有时间不适合目标受众？(例如，医生和网络管理员在工作日几乎不可能有时间参加研究。)在某些时候关键观察员是否需要在场？公司会议？假日？

有了日程安排表，接下来就开始执行日程。有许多方法可以安排人们的日程，你可以采用适合你的且自己使用方便的安排日程的顺序。下列安排顺序曾成功用于 Wired Digital 公司，与其他招募师采用的顺序相似。

- 写邀请函。
- 邀请主要候选人。
- 收到答复，并安排主要候选人日程。





- 邀请次要候选人。
- 收到答复，并安排次要候选人日程。
- 确定主要候选人。
- 确定次要候选人。
- 发送感谢信给未安排日程的候选人。
- 日程到期前一天，发送电子邮件和打电话给所有参加者。
- 创建并分发所有参加者的日程安排。
- 创建并张贴指引标识，引导参加者来到测试地点。

如果事件有固定日程安排，比如焦点小组，或者关键观察者日程受限，可以直接进行邀请。

如果情况比较灵活，可以根据参加者的偏好建立日程安排。询问候选人他们有空的时间段，或者给候选人应当出现的时间段，两者效果一样。可以允许参加者驱动日程安排，并预先直接告知他们，他们所决定的时间对于日程安排是非常有价值的。请他们提供一个主要时间和几个备用时间，这样就可以消除很多时间上的冲突。

### 邀请

无论采用电话还是电子邮件，邀请都要重申研究内容、参与的重要性、参加者报酬、研究进行的地点以及参加者应当出现的时间(或者就网站访问而言，即期待参加者访问网站的时间)。

又到了 Wired 界面测试小组的可用性研究时间，我们从数据库中选中您接收邀请。下期系列时间为 2003 年 12 月 8 号周三和 2003 年 12 月 9 号周四。如果您愿意参加，请务必于 2003 年 12 月 3 号本周五之前答复以下筛选文档。

详细介绍：

-日期：12 月 8 号周三，12 月 9 号周四

-时长：1 小时

-地点：Wired Digital 旧金山办公室

测试结束后两周内，我们将支付给您 60 美元的支票作为酬金。研究开始前，我们会请您签署一份保密协议。





请于 2003 年 12 月 3 号周五之前答复以下陈述或者问题。

\* 如您无法参加本期测试,也没关系。但为了帮助我们保持资料更新,请从下列陈述中选择一项:

☐ 我无法参加本期测试,但请在 WITG 数据库中保存我的记录以便我能参加将来的测试。

☐ 我无法再参加 WITG 的研究计划。请从数据库中去掉我的姓名和地址。

\* 如果您能参加本期测试,请回答以下问题。

对于下列每家网站,请列出您平均每周的访问次数。0 代表以前访问过,但不是经常访问;1 代表每周访问一次;2 代表每周访问两次,以此类推。如果从未访问过,留白即可:

\_\_\_\_\_《纽约时报》网站

\_\_\_\_\_《连线新闻》网站

\_\_\_\_\_CNN

\_\_\_\_\_雅虎

\_\_\_\_\_ZDNet

\_\_\_\_\_Slashdot

\_\_\_\_\_Salon

\_\_\_\_\_其他: \_\_\_\_\_(请说明)

请选择日期和上午 11 点到下午 6 点的具体时间,确定第一约定选择和第二约定选择:

1. 约定日期和时间——第一选择

☐ 12/8/03 周三, \_\_\_\_\_(仅限 11:00 AM - 6:00 PM)

☐ 12/9/03 周四, \_\_\_\_\_

☐ 任意一天, \_\_\_\_\_

2. 约定日期和时间——第二选择

☐ 12/8/03 周三, \_\_\_\_\_(仅限 11:00 AM - 6:00 PM)

☐ 12/9/03 周四, \_\_\_\_\_

☐ 任意一天, \_\_\_\_\_



好，就这么多问题。收到您的答复，协调好日程表后，我们就会告诉您最新测试日程表。感谢您的参与！

——迈克和凯文

### 确认和再确认

日程安排可能会涉及大量日期和时间表，因此我们在《连线》(Wired)发现应该最后确认参加者同意参加的时间。可以采用简短清晰的信息，内容如下所示。

亲爱的特里：

感谢您答复下周可用性研究的邀请。我们想约您在以下日期和时间进行研究：

日期：2003年11月15号周一

时间：下午5点

**\* 非常重要——请务必确认约定：**为了确认本次约定时间有效，请答复本邮件或者拨打电话(415)235-3468！！如果您不答复，我们会认为您将不参加本次研究，同时会安排另一位参加者在此时间段参加本次研究。

所有确认都应该包括一行文字，以起到“因为日程紧张，请务必在约定时间准时出席，或者能提前出席”的效果，同时还要有具体的交通路线和停车说明。

### 地址

我们在蒙哥马利市场大街582号，霍巴特大厦602室。大厦前就是湾区快速交通和市区车站，所有列车都会在那里停靠。路面公交有2路、7路、9路和14路。如果您自己开车，最近的收费地下停车场位于米什大街2号，从米什大街往东即到。

### 其他事宜

本期研究持续约1小时，参加后两周内您将收到60美元支票作为酬劳。我们为本次测试安排的日程表很紧，如果您可能会迟到，请拨打电话(415)235-3468通知我们。如果条件允许的话，请提前十分钟到达，我们





将万分感激。

再次感谢。下周见！

——迈克和凯文

此外，在邮件开头用一两句话总结邮件内容，所有邮件采用标准格式，并在邮件结尾加上联系人名字、电子邮件地址以及电话号码，我们发现所有邮件都采用这些标准化格式能发挥作用。如果参加者无法赴约或者有任何疑问，这样能鼓励他们尽快与联系人联系。保持文案简洁，但又不失幽默和个性，过程看起来会少一点点冷冰冰的感觉，人们会更愿意阅读。

最后一点，在已安排好的约定日期到期前一天要同时发电子邮件并打电话提醒，这是防止用户失约最有效的办法。这样既能提醒人们，又能巧妙强调他们参与的重要性，同时一旦他们真的不能来，你也有时间重新安排其他人来参加。

### 选择奖励

奖励是对人们帮助的鼓励与回报。因此，奖励要起到两个作用：说服人们分享时间和经验，并向公司传达价值。永远不要在奖励上欺骗人们。

虽然能有机会对喜欢的产品发表看法是强大动力(而且要始终强调这点)，但对于大多数人来说，最佳奖励还是现金。人人都知道钱怎么用，而且现金马上就能用。2003年在旧金山湾区，普通消费者研究的基本现金奖励公式为每分钟0.75美元，而一对一访谈以及专业人士访谈为每分钟1美元。因此，每小时访谈要支付60美元奖励，两小时的焦点小组需要支付每位参加者70~120美元(根据背景资料的复杂程度而定)。在波士顿和芝加哥差不多也是这个标准。

如果参加研究的人观点独特或者愿意在困难时间来，要给予额外奖励。额外奖励的金额根据实际条件设定。如果需要人们第二天就来，20~40美元的额外奖励比较合适。如果低于200美元，繁忙的经理可能就不会愿意打乱一天的时间安排。通常要预算一个范围，然后让招募师利用奖励招募到优秀候选人。

对某些人而言，即使给再多的钱作为奖励，他们也不会参加研究。对



这些人，需要采用其他的奖励方式。执行官可能比较看重能和同行业的其他执行官会面的机会，因此把研究作为能和对手交流的机会就会吸引他们。有时人们会重视慈善捐款，而送豪华餐馆礼券或剧院门票效果也很好。有家豪华邮轮航线网站把喜欢旅行的已退休百万富翁作为客户。这家网站支付了飞往纽约的全部费用，招募到了这些百万富翁。虽然为每名参加者花费了数千美元，但能从全国各地招募到准确目标受众也相当值得。而仅提供等额的现金就无法体现对参加者的理解与尊重，从而无法说服他们参加研究。

有时有些人愿意参加研究，索取的比付出的少，甚至是免费付出，但不要利用这类受众。愿意低价参加研究的群体都属于自我选择的人：他们要么有大把时间，要么真的就是渴望与你交谈，要么是因为需要钱。符合这些标准的人不太可能代表目标受众。

## 容易犯的错

### 错误人选

忠告：

对按照约定按时参加的人，即使是错误招募了他们，也务必要支付酬劳。他们满带诚意而来，并不知道自己是错误受众。同样，如果招募师给错误的人安排了日程，而问题是出在筛选文档上，也要全额支付招募师。

有时候会招募到错误群体。也许是因为有个关键筛选问题的措辞不对，所有人的回答都与你的期望相反。也许是因为确定目标受众时忘记了一个关键要素。也许是因为招募师总是采用最小标准，而你只想挑选若干符合描述的人。不管是什么情况，有时就是会错误邀请了错误的人。进入房间，却发现一屋子的人根本就不该在这，每位研究员都有这种经历。

准备工作，优秀筛选文档，精心挑选的目标受众，加上在研究开始时询问一些关键问题，这些都能尽量减少招募到错误的人，但无法避免。如果你发现自己正处于这种状况之中，可以做两件事：取消访谈，然后重新招募，或者尽量从小群体中获得信息，希望一些反馈能起作用。人选错误有时也有启发作用，因为可以拿来和从合适人选那获得的回答进行比较，以更好了解整体用户人口(也许目标受众并非那么独特?)。但如果时间紧张，而结果准确又是关键，完全可以直接取消访谈，而把精力放到获得合适人选上。





决定重新招募(或者要指责招募师)之前,必须要搞清楚招募为什么会失败。有些情况下,原因显而易见——要的是新手,但不是从未见过电脑的人,有时原因就不是很清晰,为什么这组用户与你预期的用户不符。找到导致招募错误群体的主要原因:是因为年龄?因为经验?或者可能是因为还没有确定关键因素(例如,“工厂工人”和“在车间工作的工厂工人”)。或者因为你对用户行为的预期不符合你目标用户真正所想的。而在有些情况下,有可能没有办法招募到合适群体,因为这个群体根本就不存在。

### 失约

任何受邀群体中,总会有约四分之一的人会因为这样或者那样的原因而失约,这似乎是人类的天性使然。从数据库中邀请额外的人,并去掉毫无理由就失约的人,可以补充缺少的人。

如果要确保不会有人失约,可以在每个时间段做双重安排。这样会导致招募和日程安排的工作量翻倍,奖励费用也会翻倍(因为不管最终是否需要他们参加研究,都必须支付报酬给每个到场的人),这样就太麻烦了,所以不值得。但在一定条件下还是有用的,例如给一位极其重要的人物安排日程来观察研究。如果资源充足,并且还要进行次要研究项目,也可以安排额外人员参与这一类型研究。还可以把可用性测试作为主要研究项目,而把问卷或者卡片分类任务作为次要项目,因为这些项目不会像可用性测试需要和研究人员产生那么多互动。甚至可以请不需要观察可用性测试的开发团队成员对参加者进行访谈(如果进行访谈的开发团队成员有一些用户研究经验)。

同样,可以在某时间段内安排一位非固定人选,对该候选人说等待3个小时,万一有人失约,就可以请他来做替补。但很难找到一天中能抽出2~3个小时的合格人选,所以这种方法不会像之前所说的那么有吸引力。支付给非固定人选的奖励要和其等待时间成正比,是安排在单一时间段内的人所得奖励的2到3倍,另外,如果他们不需要作为候补人选,也可以参加次要研究。

如果日程安排和设施租借预算允许,还可以在初始研究开始一两天后创建“虚拟”时间段。无法按照原定时间参加研究的人也许会愿意接受额



外时间段中的时间。

### “下雪天”

施耐德咨询公司用户体验顾问及负责人卡罗琳·施耐德建议在经常下雪的地区要提前确定备选日期，以防止路况糟糕的情况。如果有可能发生天气恶劣的情况，她建议在测试场所附近挑一所学校，并请人们检查学校的停课报告。如果停课，有可能就是因为道路无法行驶，人们就能自动知道他们被重新安排到备选日期。

务必记录失约人员名单。除非人们提前联系了你或者招募师说他们不能来，否则就不要支付奖励给他们，并把他们从数据库中去掉，或者请招募师这么做。

### 偏见

招募方法会造成找到的群体带有偏见。如果要从日间交易讨论名单中招募年收入3万美元的20多岁的人，他们给出的观点可能会和从劳工关系邮件名单招募年收入3万美元的20多岁的人的观点有差异。如果要谈论的是音乐，差异可能没什么影响，但如果要谈论的是政治，差异就会产生一定影响。

几乎所有招募方法都会对用户人口产生偏见。不可能真正做到随机抽样，因此存在一定数量的偏见在所难免。偏见是否会影响项目，这是要回答的问题。一般情况下，偏见不会影响项目，但招募时要考虑到潜在偏见，以及偏见将如何影响研究。可以在数据库中设置一个字段，用来填写每个人被招募的情况，并且要评估会造成偏见结果的来源。例如，如果有一半参加者出自同一来源，就可能要考虑招募方法本身是否会影响结果。

### 匿名

公司形象和声誉也会造成参加测试者的态度产生偏见。知名公司会让人感觉很亲切，而无名公司或者产品会比实际看起来更加微不足道。先入为主会改变人们的行为方式，同时还会影响研究结果。

在招募、安排日程的时候，如果不想被招募者知道公司名称，可以使





用替代身份。不要欺骗人，但可以采用替代电子邮件地址(甚至可以采用免费邮箱)，描述公司和研究的同时不要透露具体细节。一般情况下，人们并不介意你不透露公司或者产品的名称。

### 青少年日程表

青少年很难招募，即使他们是最抢手的用户人口之一。找到、吸引、联系到他们都很困难，他们需要获得父母同意才能参加，他们的日程表很复杂，还有联邦和各州法律限制他们被雇佣。要尽可能多地获取联系信息(移动电话号码、呼机号码和父母电话号码)才能确认他们的日程表没有变化，并且还要在研究前一天与他们联系，即使这样也要准备好最后一分钟变卦，所以要比成人预约更多。

### 创建以及空间准备

虽然从技术上讲，并不属于招募缺陷，但安排参加者日程可能会有个共同问题，即公司没有准备好迎接他们。如果他们到了办公室，他们碰见的每个人都应该在期待他们到来。门卫应该指引他们去往哪里，接待员应该知道他们的到来。还要有清晰的指示牌为他们指明正确方向。

## 与专业招募师合作

专业招募师能解决招募中的很多头痛问题。专业人士能获得广泛市场的大量群体资料，并快速给他们安排日程。如果时间紧张，或者不想自己维护数据库，也无法确定如何招募特定受众，专业人士就能帮上大忙。

但和专业招募师合作需要更多参与，不能仅仅是拿起电话说你需要这样或者那样背景资料的人。招募师并不了解你的业务或者研究要求。研究团队要认真准备并合作，以便招募师能在你需要的时候找到你需要的人。

### 上哪里物色招募师

招募师从事市场研究行业，因此他们常常和提供广泛市场研究服务的公司会有联系，这些公司提供的服务涵盖特殊会议室到主持焦点小组，到设计运行整个研究项目。如果你身处大都市，找招募师应该不难。一般来





讲，查询电话簿中的“市场研究”类别，就能找到半打。也可以联系美国市场营销协会([www.ama.org](http://www.ama.org))、营销研究协会([www.mra-net.org](http://www.mra-net.org))或者意见和市场调研欧洲协会([www.esomar.nl](http://www.esomar.nl))，他们会告诉你其他资源。

选择招募公司时也要做些调查。和其他任何服务一样，要向多家不同公司询问报价，并要求他们提供上一批用户体验研究客户作为参考。如果是为可用性测试进行招募，要询问他们能否为“一对一访谈”(或者，最好能为可用性测试)进行招募，以及他们能否为异地场所招募。跟进他们提供的参考资料，并询问招募的准确性，招募师如何应对变化，招募到多少人，招募标准具体到什么程度。询问招募公司是否为网站或者软件进行过可用性招募。多数大公司都从事过一些，但许多公司只是涉及过，对如何解释可用性研究中人们的反馈没什么经验。

这些公司如何获得用户名单，他们的数据库有多大，很难得知这些情况，因为这属于他们招募业务的核心内容，但如果能知道的话，可用于进行信息对比。

招募可能只是招募公司的部分业务，因此业务齐全的市场研究公司可能不愿意只进行招募，并可能会坚持出租其研究空间(会很昂贵)或者提供其他服务。这种情况下，可以考虑使用其服务，但如果你只关心招募本身，就要找一家只从事招募的公司。

对于小型、快速的用戶研究项目，最好能找到只从事招募和日程安排的独立招募师。这类公司比一般市场研究服务公司少见，但大多数城市都有这类公司。如果找不到，可以去问问同事或者查询可用性或者人机因素组织，如可用性专业人员协会(UPA)或者电子计算机协会人机交互技术联盟(ACM SIG-CHI)。

### 招募师能提供什么

找到合适人选，给这些人安排日程，回答问题，提醒他们赴约，每位职业招募师都要能处理这些过程中的方方面面。此外，有经验的招募师还可以把目标受众缩小到你感兴趣的人，从而有助于你关注研究。如果不能确定如何指定研究所需的合适市场，可以请招募师帮助建立背景资料。





有些招募师会为你编写筛选文档,而且他们常常会在实践中修改筛选文档,以更好适应自己的风格和策略。要参与这一过程,在招募开始之前要一份备份,以确保适当强调对你最重要的要素。

除了提供参加者的背景资料外,许多招募师都愿意做笔记,记录下没有被安排日程的人的答复。这种记录方法还达不到调查要求,但可以揭示目标受众中的未预期模式。

很多招募师还会管理奖励支付,在合适的时候支付支票或者现金。另外,较大型招募公司能为你在不同地理市场进行招募。

### 招募师需要你做什么

向招募师提供完整的受众描述是最重要的。他们很乐意根据所提供的描述进行招募,但由于不了解你的业务,因此如果你没有说出对受众的假设,他们就无法根据这些假设进行筛选。如果描述不够具体,可能会招募到不想要的人。如果你说要找互联网经验不超过一年的人,但又没说希望那些人稍微熟悉互联网,招募师就可能给你找来一帮从未见过浏览器的人。另外,描述要尽可能具体。如果正在找白领工人,要明确定义“白领工人”的意思。

列出不需要的人员清单。哪些行业会有潜在利益冲突?以往研究中多少参加者算过多?哪些公司开发竞争性产品?

要提供足够时间(至少整整一周),在招募开始后要避免改变招募参数。参数变化会惹烦招募师,他们可能会把麻烦转嫁到花费中。如果你取消招募,要准备支付至少一部分奖励费用(如果在当天取消,要支付所有奖励费用)和所有招募花费。

告诉招募师去哪里物色合格候选人。如果你已经有了客户名单,比如订阅邮件列表的人,可以根据你所处的地理区域进行筛选,把名单提供给招募师(要注意名单上的人会知道你的产品,预先对产品有正面概念,否则他们不会订阅邮件列表)。如果你不知道去哪里物色候选人,要尽量多提供一些信息给招募师,让他们知道去哪寻找目标受众以及你期望的代表占整个用户人口的比例。如果你不能提供任何信息,那就只能请招募师进



行电话调查，这可能就需要较长时间，而且花费也不低。

要提供如何处理边际候选人的指导。明确说明哪些品质可以有灵活性以及灵活程度。（“我们想要招募年龄在25~35岁之间的人，但如果其他所有参数都非常吻合，也可放宽到20~40岁。”）

向招募师描述研究。这有助于招募师了解该如何回答问题，并可能会为结构化目标市场描述提出更多想法。是一系列小组？是一系列一对一访谈？需要在特殊设施中进行吗？是在参加者的家中进行吗？重点是放在他们的态度、经验上，还是放在他们使用原型的顺利程度上？要告诉招募师提供给参加者多少这类信息才算合适。

最后，解释任何必要的术语，以便招募师能合理解释人们的答复。如果要找的人是定期购买“热插拔、能容错、低频、军用规格、窄U型架”的IT经理，可能需要告诉招募师这些词都是什么意思。招募师在交谈中不一定需要使用这些词，但了解这些词的意思有助于理解提出的问题。

### 他们要花多少钱

2002年我在旧金山写这本书时，每招募并安排好一位参加者的日程，一般要花费60~200美元，100美元是比较典型的标准。消费者处在此区间的较下端，专业人士位于较高端。对于几乎没有限制的常见群体，比如至少使用过一次互联网的杂货店购物者，甚至可能会更便宜。对于其他群体，比如用企业知识管理系统运营组织的人力资源副总裁，花费可能会显著增高。

其他服务，如编写筛选文档或者汇总答复，可以计入总花费，或者按照小时收费，每小时费用50~100美元。

### 如何应对招募问题

不要接受糟糕招募。如果确定没有严格执行筛选文档中的优先级别和问题，应该要求退款或者进行更佳招募。但首先要确保受众描述不会让招募师有悖于你的意图来解释受众。虽然不愿受这些约束，但大多数招募师还是会重新招募不符合目标描述的参加者。

最后，出于对招募师的礼貌，如果有特别好或者差的参加者，应该告





诉招募师。这有助于招募师将来提高招募效果。

## 访谈

本书提到的大多数研究都可以归结为一种方法：访谈。观察至关重要，但要真正了解用户的体验，必须要进行提问，即访谈。可用性访谈——几乎是所有用户体验研究的基本组成的另一工具——不同于调查记者的采访，也不同于未来雇主进行的面试。相比而言，可用性访谈更正规，也更标准，而且作为一种非引导性访谈，可完全去掉访谈中提问者的观点。

### 访谈结构

无论单人午饭聊天还是 10 人焦点小组，几乎所有用户体验访谈的基本结构都类似。该结构呈沙漏形状，从最普遍信息开始，逐步深入到越来越多的具体问题，然后再回到较大的观点，并通过摘要和总结得出结论。以下方法把标准访谈过程分为六个阶段来进行。

- (1) **介绍。**所有参加者进行自我介绍。小组成员要互相认识，这样他们会感觉舒服，因此小组自我介绍能强调参加者、包括访谈人之间的相似之处。相比而言，一对一访谈中的介绍，可以为访谈人建立中立、富有同情心的形象。
- (2) **暖场。**回答问题或者参与讨论的过程需要每个人都有适当想法。访谈的暖场阶段是为了让人们脱离常规生活，把注意力集中到对产品的思考和问题回答上来。
- (3) **一般问题。**初始产品细节问题应该以产品和人们的使用方式为主题。重点关注态度、预期、假设、方法以及经验。提出此类问题能尽早防止开发团队对产品的假设会扭曲人们的想法。通常情况下，产品在此阶段甚至尚未命名。
- (4) **深入关注。**介绍产品或者产品想法，而人们关注细节，如产品能做什么、如何操作、是否可以使用以及直接体验。对于可用性测



## 忠告：

每个新访谈剧本都需要演练。找位同事或者样本参加者进行演练，加上录制设备和原型，并进行适当修改。

试而言，该阶段构成了访谈主体，但对于要发现新问题的情境调查而言，可能因此无法进入讨论。

- (5) **回顾。**在此阶段，人们可以从更广的范围评估产品或者想法。该阶段讨论可以和“一般问题”阶段相对比，但其讨论重点是深入关注阶段引入的想法如何影响早先所讨论的问题。
- (6) **总结。**该阶段通常时间最短。总结是访谈的正式终结，参加者回答完最后一个问题后不会感觉被晾在那里，并且能把讨论带回到最一般的管理主题。

## 非引导性访谈

曾经有位著名科学家在调查时问过如下问题：

您的雇主或者其代表有没有为了欺骗您的部分收入而要手段？<sup>①</sup>

这是引导性问题。往下阅读之前，请先思考为什么是引导性问题。什么东西在暗示“正确”答案？作者想获得的是什么实际信息？对于这个问题，怎么问才会成为非引导性问题？

提出此问题的科学家是卡尔·马克思，他很清楚他想要的答案，答案不是“没有”。

引导性问题是所有社会研究的祸根，因为它将提问者的个人偏见带入了本应该只能是回答者的观点中。但想避免引导性问题，说起来容易，做起来难。提问者要一直保持警惕，并且要坚信必须无条件了解人们的想法。

非引导性访谈是不会引导答案或者导致答案有偏见的过程。它是获得用户思想、感情及经验，而不会用访谈人筛选想法的过程。

## 中立访谈人

作为编写和提出非引导性访谈问题的人，你的任务是去掉你对产品的了解和感觉。要忘记所有辛苦工作和创造力。要抛弃所有成功希望以及失败恐惧。忽略一切听说或者想过的关于该产品的事情。要完全中立，就好

① [T.B. Bottomore and Maximilien Rubel, eds., Karl Marx: *Selected Writings in Sociology and Social Philosophy* (New York: McGraw-Hill, 1956), p. 208; as cited in Earl Babbie, *Survey Research Methods* (Belmont, California: Wadsworth, 1990), p. 37]





像根本不是你要访谈一样。只是你对一件事、一件你毫不关心的事情进行提问。

听起来很苛刻，但必须这么做，才能理解人们给你的正面和反面反馈，并且把开发产品的过程与他们想要或者需要的东西关联起来，而不是你认为他们想要或者需要的东西。否则，永远都只能看到一线天空或者一片云。实际上两者都需要兼顾。

要运用禅的智慧，要让提出的问题不会导致受访人的答案带有偏见，会涉及很多自我保持距离以及严格检查关键假设。如果你对产品很熟悉或者有很大兴趣，研究起来就会特别困难。首先，你会觉得花费大量精力而没有提出明确问题，或者是问题本身显得矫揉造作。有了一些经验之后，哪些问题会引导人们、如何组织问题以便能得到最自然反馈，这些会变得清晰起来。最后——当你获得非引导性问题的智慧之后——问题将会变得自然，分析起来也会更容易，而且得到的无偏见答案能给结果带来更多把握。

### 编写非引导性问题

最重要的是，每个问题的重点应该是回答问题的人。应该关注体验，而不是推测。我们对自己行为的理解很少和我们的真正行为相对应。如果要把自己放到别人的情况中，我们会把事情理想化和简单化。这样有益于理解人们的想法，但几乎无益于理解他们的行为。类似“这项特性有用吗？”的问题很容易被误解成“在所有情况下，您认为某个地方某个人会觉得这个特性有用吗？”即使大多数人都能理解字面意思，但潜在误解会造成所有回答都值得怀疑。“这项特性对您现在的工作有价值吗”就阐明了观点。

同样，问题要关注直接体验(immediate experience)。人们的当前行为比他们的预测能更好预测其未来行为。如果要问“您觉得这有趣吗？”他们会想象只要在某种程度上觉得有趣，就会回答有趣。但理论上有趣的东西和人们能记住并能回忆起来的东西相差很大。如果他们现在找到了令人信服的事情，他们可能会继续发现它令人信服。因此，用“如果现在能用，



您会使用吗？为什么？”来提问，答复会更有用。

问题要客观(nonjudgmental)。不要让回答问题的人觉得你在期待某个特定答案，也不要让人觉得没有正确答案。可以明确说明这点，但如果问题提出时就重申这一观点，效果会更好。“难道您不认为如果 PDA 上也可以使用会更好吗？”这个问题在暗示提问者认为这是个好主意，回答问题的人除非没听明白，否则不会不同意。“如果将来 PDA 上有此特性，您会使用吗？”这种说法并没有暗示提问者有预期答案(尽管它受到了稍后提到的二元问题困扰)。更好的提问方法是问：“您有没有想过使用此特性的其他方式？”回答人说出初步想法后，就可以迅速开始讨论 PDA。

问题要关注单一主题(focused on a single topic)。采用“和”和“或者”连接两个想法的问题会导致含义不清，因为常常会分不清要回答哪部分问题。“这个产品怎么适用于学生或者上班族？”实际上是两个问题。而问题的答案可能不足以区分出两者。

保持问题的开放性(open-ended)。如果给出的选项有限，即使人们的观点不在选项内，或者如果他们觉得应该有一个以上的选项，他们也会选择其中一项。他们会调整自己对选项的定义，挑选出跟感觉最接近的一个选项。但这并不是他们的真实感觉。要保持提供封闭式问题之外的选项，除非能绝对肯定选项覆盖了所有可能性。但这种情况极少发生，因为大多数情况都是在辨别细微差别。“下面列表中，哪个特性对您最重要？”这个问题就假定了有重要特性，并且其中有一个比其他更重要。更好的办法是采用下面的说法，“从 1 分到 5 分表示出下列特性对您的重要性，1 代表最不重要，5 代表最重要。如果某项完全不重要，请用 0 表示。请写出我们没有列出的重要特性”，或者也可以采用完全忽略特性命名的机制，比如“该产品有没有对您特别有用的地方？如果有，是哪些？什么东西让该产品有用？”

避免二元选择问题(binary question)。二元选择问题是一种特别狡猾的封闭式问题。其形式有“是/否”，“真/假”，或者“这个/那个”，人们被迫做出黑或者白的选择，而他们的态度并不接近这种极端情况。“这是好产品吗？”这个问题漏掉了许多人们态度中的微妙之处。虽然能迅速知





道人们的即兴意见样本，但实际上了解他们认为哪些地方好、哪些地方不好更有价值，而不仅仅是他们认为整个事情是好还是坏。“如果有的话，您喜欢这个产品哪些地方？”

### 进行非引导性访谈

非引导性访谈和其他访谈一样，但必须要更细心倾听你和交谈人的对话含义，注意会产生偏见的信号。可以做一些事情来提高答复质量。

定义术语(define terms)。词语往往含义不清，容易误用。“那个东西”可以指按钮、特性，或者整个网站。词语的个人定义会不同于字典定义或者开发小组的定义。有人会用“模块”(module)来指一个简单功能，而开发团队可能会把功能很复杂的簇群称为“模块”。使用技术术语时，要确保首先明确其定义。只要有可能，应该使用答复者的词语定义(即使你通常并不会这么用)，但要确保首先了解定义(可能要请答复者进行定义)。在群体互动中这一点尤其重要，因为人人都有着与自己不同的定义。

不要强求观点(don't force opinions)。对于某些事情，我们就是没有什么观点。我们可能从未定性想过某个问题，也可能因为信息不够而无法形成观点。征求观点时，大多数人都能给出观点，但这种观点通常不会经过认真考虑或者深入研究。如果提出的问题需要意见，要确保人们的回答能包含观点。“如果能自动完成，这个东西会不会更好？”这种说法对没有使用“这个东西”经验的人没有意义。

重述答案(restate answers)。利用问题解决疑问有个最佳方法，即采用不同词语重述应答者的答案。这样能澄清术语的许多微妙之处，并验证你理解了答案，答复者也理解了问题。答复者做出回答后，你可以马上说诸如这样的话“我听到您说……”，用你理解的方式把听到的内容说出来，但要使用不同词语。但要避免不要替换答复者使用的“正确”词汇。首先要调查他对术语的理解。如果有人说“订单信息”，而实际指的是“确认页”，要请其详细解释希望在“订单信息”网页上看到的内容，然后再使用“确认页”术语来重述答案。

跟进例子，但首先要等待无引导性答案(follow up with examples, but





always wait for an undirected answer first)。有时人们理解了问题，但可能不知道如何回答。如果问题措辞准确，回答起来就不会有困难。虽然有时你会故意想问一个含义广泛的问题，来看看人们如何理解概念或者发现他们最普遍的想法。可以在你认为需要的地方，准备一个(或者两个)例子。参加者给出最初答案后，可以使用例子再来调整他们的想法。比如，正在进行焦点小组，可以对新特性进行头脑风暴。如果人们定义的特性过于狭窄，似乎陷入了僵局，你可以说：“如果现在要把您喜欢的促销商品条目通过电子邮件发给您，会怎么样？”看看参加者顺着这条思路会不会产生其他想法。例子不要超过两个，因为例子本身往往会强烈干预人们的认识。

利用人工物品保持人们的关注点并触发想法。人工物品是人们工作的物质产品：笔记、纸张、工具等。把参加者带回到即时环境中，向参加者提出定期必须处理实物(或者软件)方面的问题。如果有人谈到“购物车”，就可以问：“这个购物车怎么样。”如果是实地访谈，请他们把实物拿出来看看。人们想象和讨论的理想条件常常和他们生活的实际条件并不一样，而他们使用的物品能提醒他们在理想条件下所缺少的繁杂细节。

要留意自己的期望。注意观察让你惊讶的事情，并且预测受访者下一步会出现的情况。尽管本节开始就提出忠告，但对访谈情况不可能毫无准备。要从访谈交互过程中做一些假设或者预计，而这些事情会影响如何操作访谈。如果能注意到这些假设，就更容易避免产生这些假设。

不要说参加者错了。即使某人对产品工作方式或者用途的了解完全不同于产品本身的意图，也不要说其观点错误。研究其观点，努力了解观点的出处，以及为什么会有这种想法。可能是这个人的理解与他人或者你的观点不符，但不是错误观点。

仔细倾听用户对你的提问。这些问题能揭示人们如何理解产品或者情况的很多东西，对了解人们的经验和期望也非常重要。探讨人们为什么要问这个问题。例如，如果有人问：“它用起来就是这个样子吗？是这么用的吗？”，可以用提问的方式进行回答，以揭示提问者心智模型的更多情况：“您认为它就是这么用的吗？”或者“您期望它就是这么用的吗？”

从语言和意图上保持问题简单。利用问题来发现新的假设和认识，而

#### 提示：

访谈期间，观察员可以在场。观察员在场会造成访谈不太舒服，但观察员可以记笔记，也可以增加访谈角度。观察员的参与程度要由主持人决定，但一般有一位室内观察员就足够了，而且应该介绍其身份。我发现，在特定时间允许观察员提问效果很好。





不是来证明观点或者评判行为。好的问题能把引出观点或者看法所花费的努力降到最低。分析答案是很有意义的，它是对问题的证明和评判。问题要关注获得最清楚的原始信息。

随时回顾磁带。靠记忆力和笔记很容易会漏掉关键陈述或者微妙区别，应该随时安排时间回顾磁带——无论是录音还是录像——以证讨论的观点是否准确地代表了访谈过程，以及如何才能更好地进行今后的研究。

### 常见问题

- 应该是开放式问题的封闭式问题。“这三个徽标，您最喜欢哪个？”如果人们一个都不喜欢，这样问就并不是特别有用。“这些徽标，有您喜欢的地方吗？”如果人们有喜欢的地方，这样问就能告诉你人们喜欢哪些内在特征。你可以根据人们的喜好定制徽标，而不是任意选择一个。
- 有复杂答案的二元问题。“每日更新对您而言是重要特性吗？”这个问题忽视了是或者不是的全部原因。也许人们并不打算每天检查网站，但每周一次更新会很有用，也许根本就没有必要更新。“关于每日更新，您觉得哪些地方有意思？”能说明哪些地方有趣。
- 词汇过多或者词汇有多重含义，用词要准确。“如果在网站上找东西时迷路了，您因此而绝望，您会怎么办？”“绝望”并不准确。它可以解释成“非常迷茫”，也可以解释成“完全迷失，不可能找到任何东西”。可以把问题重写为“在网站上找东西时，您发现不知道如何返回到之前访问的地方，您会怎么办？”
- 要求人们预测未来。如前所述，如果人们计划自己的未来行动，他们往往会过于简单化和理想化，造成预测和实际会做的事情几乎不相干。人们更擅长解释其正在进行的行动，而不擅长解释预测的行动。如果有兴趣了解特定情况下人们的行为，可以将其放到这种情况下(或者是合适的仿真条件下)。
- 施加强权威或者同伴压力。例如，“大多数人都说用这个工具找信息很容易。您也这么觉得吗？”或者“设计师拥有导航工具设计



的丰富经验，他们想出了这个设计。这个导航工具对您能起多大作用？”这些问题几乎都能简化成实际问题：“请描述一下您使用该工具的体验。”

- 假设知道答案。我发现自己会一边听问题答复，一边假设答复是我已经听过的另一种版本，而只仔细听出乎我意料的答案。人们有时甚至会使用许多和你预期的相同词汇，但关键性的否定或者倾向性的陈述可能会推翻或者从根本上改变他们所说内容的含义，因此要仔细听每一个字。
- 假设他们能回答问题。并不是每个人都明白他们知道什么、不知道什么。如果问某人某件东西在同级别中是否属于一流，其实就已经假设了他熟悉同一级别的所有产品，他们可以在所有产品中做出平衡的、有见地的评估。

并非只有制定问题时才会出现麻烦，对答案的解释也取决于提问方式。提问时，需要注意几种行为，这样能迅速发现这些行为并采取相应的后续行动，减少之后分析的歧义。

- 人们并不总是愿意说出他们的真实想法。有时他们说是，是为了避免说不是时会产生冲突。注意观察能说明其真正意思的线索。回答时犹豫不决，或者回答不符合之前的陈述，这些都可能是线索的表现形式。可能还有更微妙的线索，比如有人边摇头边说是，或者突然不说话了。要能在此时发现情况，并请对方解释清楚。通常情况下，给他发言权就是给他信心说清楚真正意思。
- 人们有时回答的问题和所提的问题不是一回事。在人们对主题思考困难的情况下——也许因为他们正在执行任务，或者要记住情况——他们很容易会听错问题细节。参加者有时有自己的时间安排，他们真正想讨论的是那些你没有问的问题。要认真倾听他们说的内容，判断与你所问的内容是否直接相关。如果明显偏离，要打断他们，并再次向他们提出问题，但要采用略有不同的措词和重点，不要害怕坚持问一个问题。





## 何时打破规则

如果遵守以上全部规则和建议，对话显然会很枯燥，而且可能比遵守规则消除的偏见更糟糕。人们与你交谈，如实回答提问，应该感到很自在。与他们交谈，你也应该觉得很轻松。

因此，构建问题时，要把这些规则当成建议，尽可能全面遵守。但如果问题看起来很难回答，可以随时即兴访谈，人性化访谈，要提供范例或者让参加者“摆脱困境”。访谈既是非引导性，又让人感觉舒服。最终，最佳访谈就是在你需要的时候能够给你提供所需信息的访谈。要做到这一点，每个访谈都要采取不同方法，这些规则和指南有助于尽可能获得最佳信息，但必须要知道如何恰当实施这些规则和指南。

### 录制访谈

只要有可能，每次访谈和互动都应该进行录像。许多人把视频文件看成是录音的奇妙形式。有时候的确是这样，但视频可以显示互动过程中音频无法捕捉到的关键时刻。有人说“是”的时候耸了耸肩，他们其实想表达“否”，这对于准确理解观点非常重要。鼠标在一个按钮短暂停顿，然后点击了另一个按钮，说明特性的核心很混乱，而且通过录像，主持人不必一边做笔记，一边思考如何主持。

如果仔细引入和放置，录像并不昂贵，而且对大多数人来说会迅速融入背景，因此相对来说，录像更加不引人注目。可以在访谈一开始就说明有摄像机，把摄像机放在位置不显眼的三脚架上，访谈可以继续正常进行。磁带会成为永久记录，可用来挖掘关键细节以及准确引言(包括语言和动作)。

拍照使用的设备较少，并且可以收集访谈具体项目和安排收集特写，但它会破坏进程，因为研究人员需要停止会话进程以便拍摄照片。但在某些情况下，比如在有安全意识的组织中进行情境调查访谈，拍照是进行记录的唯一办法。这类情况下，应该同时对访谈过程进行录音。



## 第7章 用户背景资料

“谁会使用你的产品？”

“每个人都会用！”

“他们用它能做什么？”

“什么事情都能做！”

随着技术和带宽发展，通过互联网可以实现实时语音通信，因此很多企业进入实时语音通信市场，来促进和开发实时语音通信。许多公司都目标远大，照此推理，目标面向所有被问及谁会使用他们的产品时，有家公司的一位执行官答道：“每个使用电话的人！”

尽管当时雄心勃勃，也有好故事能打动投资者，但在1998年，本来可能会使用互联网电话的人远远少于“每个使用电话的人”。受众定义过于广泛，这些公司实际上就没有定义受众。当然，如果能抓住大市场的一小部分，这些公司就能宣布自己活得很好，但这种推理一样有缺陷。对这家公司而言，瞄准大市场的小份额也是不可能完成的任务。即使要获得大市场的一小块，他们也需要了解整个市场的动态。世界上有很多电话用户。要了解如此轻率定义的市场，这些互联网电话公司需要了解一大块世界人口的需求、能力和愿望——而且还是1.0产品。

这家有问题的公司最终意识到必须集中精力。想开发完全通用的产品遭到失败后，他们确定了更集中的受众定义，但在此之前已经花费了很长时间去开发大多数用户(用户本身就很少)并不关心的特性，而且这些特性跟产品其他部分也无法配合。

这就是用户背景资料宝贵的地方。如果创建了用户背景资料(或者如库珀交互设计公司的创始人艾伦·库珀所称的角色(persona)，该公司是这种方法的主要创新者和倡导者)，就能创建一个或者多个虚拟用户，并模





型化他们及其生活。从一般特点开始，模型化购买和使用产品的特定个人。每份最终背景资料都是一个工具，有助于定义理想、比较理想和现实。根据用户背景资料进行角色扮演，可以显示采用其他方法难以发现的功能和需求，而同时能在产品特性头脑风暴环节增加一份真实感(减少“漂亮因素”，雅虎资深交互设计师克里斯蒂娜·沃德柯这么说)。有了用户背景资料，就可以关注用户的具体形象，团队每个人都可以进行参考并把它作为开发想法的共鸣板。跟“每个人”相比，工作要简单得多。

本书主体内容是了解用户是谁，他们能做什么，他们想要什么。这些信息能从长远指导产品，但即使在开始收集确凿数据之前，就要把重点放在用户及其体验上。记录和分析用户背景资料，根据直觉、判断及手头信息模型化用户。这样可以迅速了解良好用户体验由什么构成，可以让你的产品跳出“一切为了所有人”的陷阱。背景资料能提供工作的概念范围以及关注研究其他部分的基准。如果要跟进调查、情境调查或者用户测试，都有可以用来比较结果的模型。

除虚构这一事实外，用户背景资料就代表了你要认识的人，要花时间在一起的人，要喜欢的人。这些资料不仅仅是人口统计数据清单上加张脸，而是为了做出更好产品，要与你一起工作的实体。

## 何时建立用户背景资料

开始开发产品之前，要考虑用户背景资料。受众概念要和产品概念同时形成。别忘了，你在为某人做东西，那么这个人是谁？

进行详细、繁琐的产品开发工作之前，产品要最大限度保持灵活。可以调整产品能做什么以及其为谁服务，然后再全身心投入开发。这些早期决定能决定整个开发过程的方向，而且需要花费大量精力和资源。早期错误会困扰很长时间。你想要更有价值的特性，或者想要的用户群体需要大量支持而回报很少，这些情况都值得怀疑。精心建立一套用户背景资料，有助于理解为什么要做正在做的东西、人们为什么会买这个东西。

研究计划的首要任务始终应该是建立一套用户背景资料日程表(表 7.1)。



建立过程会需要几天，只要有白板、即时贴，再加上开发团队的参与就够了。它还有另外的作用，有助于建立新团队，并提供他们沟通的共同基础。

表 7.1

时 间	行 为
$t-2$ 周	组织场所，并邀请参加者(一般是内部的利益相关者)。开始初步研究。
$t-1$ 周	完成参加者日程安排。继续初步研究。
$t-2$ 天	完成初步研究并总结。和参加者进行确认。
$t-1$ 天	准备场所，确认日程表。
$t$	背景资料练习(通常为 1 到 2 天)。
$t+1$ 天	开始编写背景资料。
$t+3$ 天	完成编写背景资料。提交给整个开发团队。分发文档。

## 如何建立用户背景资料

创建用户背景资料是一项集体活动。它并不是一个人或者一个群体的思想。它必须经过整个开发团队的决定、充实与商定(协商一致或者妥协)。产品开发各方面的关键成员都需要参与进来，因为每个人都会给背景资料带来不同信息、不同角度。

有两组人需要参与背景资料的建立过程，一组人已经知道用户特征，而另一组人需要知道这些特征。

需要参与记录和分析背景资料的小组样例	
已知用户特征的小组	需要知道用户特征的小组
市场调研	工程
支持	形象设计
销售	交互设计
业务开发	信息架构

每个人都会对他们希望成为用户的人带来独特看法。信息架构师会讨





论他们期望人们如何组织信息。营销和销售人员会关心用户关注的东西，还会关心与其他产品相比这个产品吸引用户的原因。交互设计和信息架构会关注人们要完成的具体事情和人们与他人的互动情况。每个人都会揭示用户的不同一面。所有这些情况放在一起看，就能丰富背景资料，就能揭示出用其他办法难以确定的限制和机会。

要邀请 5~8 人参与记录和分析背景资料的过程。如果人太多，会议会变得臃肿；如果太少，用户背景资料会无法用于整个团队。

### “这不是吹嘘” Quackery

如果你的公司或者团队是第一次进行这类练习，可能必须打消一些人的顾虑。如果从未参加记录和分析背景资料练习，很容易会搞成营销乱吹或者空洞的团队建设实践(我们都明白我们多喜欢这些练习)。要克服这一点，就需要把记录和分析个人背景资料变成合理但充满乐趣的一部分开发过程。

通过这种方式，人人都能对同样事情进行讨论，也有了有效方法来了解受众所受的限制和愿望，这些都是需要传达的最重要事情。通过建立有名字的客户模型，可以建立有效的简略表达方式。如果要把一项特性描述给“财富 1000 强大型企业中不经常使用 SAP 的采购人员”，可以说“这是为里昂纳多开发的”，营销、工程和设计都能知道受众特质，知道受众会如何使用该特性。“里昂纳多”代表着对一类用户体验问题的共同理解，而这些问题对产品成功很重要。

建立人物背景资料过程的其他好处——能理解微妙的交互问题，把人们的愿望与他们的理解模型耦合在一起，以及团队建设——都是这种沟通方式优点的副作用。

## 初步研究

要先做一些调查，然后再将所有人召集到一起。在建立受众背景资料之前，先对你的受众有一些直接的了解，通过这些了解快速建立受众的简





要描述，然后，再一步步地通过解决问题的方法完善它。

### 从内部研究着手

首先可以从周围的人着手。对于目标受众，公司内部有些员工有自己的想法。如果这个人不是你，也许就是业务开发部门的人。如果不在业务开发部门，也许就是产品设计人员。如果不是产品人员，也许就是公司首席执行官。如果不是首席执行官，也许就是首席执行官的理发师。反正某个地方有某人对受众是谁有想法。(如果没人有想法，问题就大了，就不是用户体验研究可以解决的了。)也许所有人都有想法。

从询问本地专家开始——接触过受众的人——请他们向你描述一下受众。如果产品有既定客户群，可以和与客户群直接打交道的人进行讨论——销售人员、支持人员、市场研究人员、技术销售顾问、培训人员等。他们对真正受众是谁有直接经验。这些群体常常会有详细的人口统计背景资料，还能给出受众大致分类的市场研究报告。支持人员有时会有“前40强”的用户疑问和问题清单。

#### 忠告：

客户不总是用户！确定、研究并购买产品的人可能永远都不会使用产品。大型信息技术组织中，经理可以决定购买产品，并在全公司部署，却永远都不会使用该产品。父母经常购买自己从来不用教育软件。要请本地专家找出负责选择产品的人和最终使用产品的人。

假设要为一家已设立的银行建立小企业结算服务。该银行的营销团队已经向小型公司销售结算服务多年，自然知道向谁销售、那些人的价值观和需求。如果有一个电话号码专门用于小企业客户，接听电话的人就需要经过专门培训，并且要有和这些客户打交道的宝贵个人经验。

访谈这些本地专家，请他们谈谈与用户及客户的个人体验。可能要招募一些开发团队成员帮忙访谈，也许可以给每位参加者安排一次访谈。每次访谈不要超过一小时左右。收集专家所了解的受众特征。用户是什么样的人？靠什么谋生？使用产品的频率有多高？是否有常见行为？有没有让他们觉得特别高兴的事情？常见问题有哪些？请专家用小故事说明典型客户的情况，特别是有兴趣的典型情况。

针对以上银行项目，你可能想要知道几位经理的名字，这些经理有和小公司客户打交道的经验，打电话给他们或者去拜访他们。

你可能会看到自相矛盾的背景资料：销售人员的一组销售对象会完全不同于联系支持部门的人；市场研究部门建立的背景资料会不同于业务开





发部门正在使用的背景资料。这种情况很典型，也很正常。不管是否矛盾，都要把所有背景资料收集起来。不要在建立背景资料开始之前就想解决矛盾(不要陷入哪份背景资料更“正确”的争论)。

### 与用户交谈

了解目标用户(或者应该是谁)之后，接下来要安排访谈大致像用户的人。如果能接触到产品的实际用户(银行分行经理可能会有一些可以共享的客户联系信息)会很好，但和受众关系密切的朋友和亲戚也是很好的访谈对象。和3~10人进行交谈。请他们谈谈现在的产品在哪些方面对他们有帮助，请他们描述碰到的问题和正面体验。要记下有用的引言、难题和故事。

创建背景资料前一天，要尝试从收集到的信息中提炼出共同线索。这些人有什么共同点？有没有频繁发生的问题？有没有受到欢迎的认知？把不同描述中矛盾的地方记下来，因为人们认识问题方式的不同之处和相似之处同样重要。

会议开始之前或者开始时，要总结并呈报发现结果。

### 列出属性

要立刻完成建立用户背景资料，最好能在一天之内。通过这种方式，参加者能专心关注信息，并创建一致模型。连续两天也可以，但是如果时间过长，建立过程被打散，时间会冲淡人们的记忆、降低人们的热情。通过一次完成(有时时间长)，会产生生动画面的背景资料。

第一步，列出受众的属性清单。给房间里所有人一叠即时贴，请他们把对受众了解或者怀疑的地方写下来。内容可以尽可能广泛。例如，有人列出了小企业结算业务网站的如下内容：

40多岁  
店铺经理  
女性  
固执



懂计算机  
白天没有太多时间  
晚上上网  
管理两家店铺  
每店每周收一次货  
有30名雇员  
有几个孩子  
不拥有店铺,但像所有者一样进行管理  
等等

请参加者列出所有内容,直到他们再也想不出来为止(大约需要半小时)。他们每人应该会写出(典型有)50条,如果能写出更多条也可以。

告诉参加者,列出属性时不仅要利用他们各自领域中初步研究获得的知识,同样重要的是还要利用直觉。实际上,人们在自己内心往往有更加丰富的模型,但一般不愿拿出来分享,因为,人们自己都还不能完全信任它。要鼓励他们分享怀疑、预感和希望,并且要让他们感到舒服,因为没有谁比其他人更“正确”。所有属性都不需要有明确信息,因此总会出现一些经验猜测。这些想法早点产生最好,只要标记为猜测,就可以根据出现的新信息修订背景资料,这样就没问题。此外,如果建立背景资料的小组想解决一些特定问题,也可以开发背景资料的某些特定方面(比如目标或者特定任务)。

属性列表不是核查清单,也不是属性的无穷列表,它是思考如何描述用户的种子。每项属性都带来很多问题,可以用来充实背景资料。

### 人口统计

市场调研报告和业务开发文档中都能找到人口统计描述。人口统计描述采用年龄、头衔、行业、收入以及邮政编码等传统衡量值来划分受众。由于要解决人们的体验问题,而体验超越普查数据(其中许多核心人口制约因素来源于此)能够进行衡量的程度,这些信息仅仅是开端。附录中有一套相对完整的人口统计问题,一些关键问题如下所示。





- 年龄。用户多大年龄？如果他们代表的是公司，公司有多少年历史？
- 性别。根据性别如何划分用户人口？就他们如何使用产品而言，性别是否有关系？
- 收入和购买力。如果销售的东西是个人用途，用户的家庭收入是多少？作为个人他们受到哪些经济制约？如果东西是商业用途，他们的购买力有多大？他们所处的行业受到哪些限制？
- 位置。用户在城区，郊区还是农村？他们来自特定地区、国家，还是分布在世界各地？
- 文化。他们的主要语言是什么？他们的文化是否会影响购买或者使用产品的方式？
- 头衔。他们的工作头衔是什么？
- 公司规模。如果产品是给在公司上班的人使用，公司有多大？公司赚多少钱？公司利润高还是低？

不要过细划分人口。虽然应该知道受众的人口统计构成，但不要沉迷于获得用户收入或者地理位置的每条细节(当然除非这些信息对产品很重要)。

### 技术能力

用户所使用的硬件和技能是产品使用道路上的主要制约因素。电子商务服装商店 boo.com 遭到惨败有许多原因，但最直接和最严重的原因就是他们的第一个网站——花费了数百万美元进行宣传——对其一大部分用户而言无法使用。网站要求客户要有更大的带宽，计算机要有更强的计算能力。许多潜在客户因为无法访问，甚至从未看过商店网站。

- 计算机。典型用户使用哪类计算机？使用了多长时间？
- 显示器。用户使用哪种显示器？他们使用最新的超高分辨率显示器还是廉价的低分辨率显示器？
- 网络连接。他们上网的速度多快？
- 经验。他们硬件用得怎么样？他们用硬件做什么？一般情况下，



他们使用计算机有多顺畅？哪些经验形成了用户对计算机和互联网的态度？

- 浏览器品牌和版本。他们使用哪些浏览器？什么版本？浏览器平台决定了用户的很多体验，因为不同浏览器代表了不同环境(有时候是非常不同的环境)。
- 操作系统。与前面类似，浏览器所处的操作系统会对人们认识信息产生制约。

### 互联网使用

就人们要使用产品的情境以及他们对产品的期望而言，人们如何使用互联网非常重要。

- 经验。他们使用互联网多长时间了？使用频率？用于个人目的和工作任务的频率分别是多高？
- 典型任务。他们在网上做什么？上网频率？

### 环境

人们使用产品的环境会影响他们认识产品的方式。在工厂车间出价竞拍一吨钢材，在办公室座位或者起居室出价竞拍，两者会不同。描述的复杂程度取决于产品性质，但对于人们使用产品的情境总是要有模型。

- 使用地点。是在家用这项服务吗？是在办公室用吗？如果是在办公室，是哪类办公室？如果是在家，会不断被孩子和宠物打断吗？他们共用一台计算机吗？
- 使用时间。他们是在工作期间还是上班时间使用产品？他们会在早上首先使用，还是在看完电影回来后使用，还是在一天最后一次挤奶之后使用？
- 工具情境。他们在使用你的软件的同时还会使用什么其他工具？这些工具有多重要？你的产品如何符合这些工具？
- 竞争。哪些产品或者服务在争夺用户注意力？是什么竞争性质？竞争性产品有哪些优势？





### 生活方式/消费心态

除了使用产品的方式外，人们还有利益观、思想观和价值观。这些可能不会直接影响产品的直接使用，但对思考用户如何体验产品仍然很重要。创建背景资料时，要考虑使用直接范围之外的用户生活，并且从人的角度定义他们是谁，而不仅仅是从顾客或者客户的角度。

- 价值观和态度。客户的价值观是什么？在他们的生活中，什么东西重要？节约？速度？轻松？有趣？舒服？他们对类似服务有何体验？
- 媒体。他们如何知道新产品、服务和想法？他们阅读什么杂志？为什么读那些杂志？看什么电视节目？最经常访问什么网站？
- 行为。除了使用你的网站之外，他们还做什么事情？还喜欢什么网站？在生活中还做哪些事情？有哪些娱乐活动？

### 角色

人们在生活中扮演着不同角色，他们与其他扮演角色的人产生互动。公设辩护律师和检察官在法庭上激战，但周末会一起打壁球。人们在生活中扮演的角色经常会比其头衔或者职责更重要。例如有家针对塑料厂商的物资交换网站。一开始这家网站认为，网站的目标受众是销售和采购经理。然而，受众研究表明，采购经理很少参与日常采购管理，而且从群体整体看，他们的电脑技能都很薄弱。他们的工作职责和实际做的事情并不相同。进行大部分物资交换的人都是他们的助手。

- 头衔。涉及哪些群体？如何称呼他们？
- 职责。每个群体的工作职责是什么？他们能得到什么回报(正式回报和非正式回报)？怎么得到回报？
- 培训。他们工作时接受过哪些培训？多长时间他们会获得针对新特性的再培训，或者是培训？
- 权力。他们对产品有哪些职责？能选择产品还是被指派使用产品？付钱还是有其他人买单？如果出了问题怎么办？
- 关系。他们和谁一起工作？做决定时他们向谁咨询？使用产品

#### 提示：

如果目标市场和用户市场产生分歧，角色就会显得特别重要。目标市场由做出购买决定的人构成，而用户市场由必须要和选择一起生活的人构成。理想情况下，这两组人会发生重叠或者相通，但并不总是这样。如果为两组人分别定义背景资料，并且指明两者之间的关系，就能澄清用户体验的需要。



后，他们会和谁一起工作？那些人需要什么？

- 互动。他们与其他人如何互动？那些关系将来会发生变化吗？是什么性质的互动？信息型？规则型？商业型？

### 目标

人们总会有原因才使用你的产品。他们想要比现在能更容易、更快、更便宜、更有趣地完成事情。应该是哪个原因？确定最终结果(包括短期和长期)有助于解释人们的行为和态度。

- 短期。用户想解决什么问题？能解决他们的什么业务需求？这些需求随着时间会如何变化？
- 长期。重复使用这项服务对整体业务会产生什么影响？
- 动机。他们为什么要使用该产品？什么东西在驱动他们使用？是个人目标？还是职业目标？目标与用户工作或者生活有什么关系？他们是自愿使用你的产品，还是其雇主或者学校要求他们使用？
- 结果。每个目标成功完成意味着什么？如何定义成功？有没有最终目标？最终目标是什么？
- 痛苦。什么东西让他们无法使用现有工具来解决问题？有多痛苦？是否有多个痛苦点？这些痛苦只是小困扰还是造成停止的严重问题？

### 需求

需求有两种不同类型：直接关系到手头任务的需求和偶然型、非功能型但仍重要的需求。前者是目标表达——已完成的目标就是已满足的需求，但后者与产品功能没有直接关系，而更短暂、更情感化。有些用户可能需要获得放心感。其他用户可能需要觉得有控制权。还有一些用户可能需要觉得占了便宜。这些需求都不会直接关系到产品功能，但仍然重要，因为它们含有把功能性产品转变成良好用户体验的秘密配方。

- 功能性。如果要解决问题，要达到目标，必须要做什么？区别合格和完美的因素是什么？





- 情感化。要欣赏产品，他们需要做什么？有没有用户没有注意到的需求？用户不烦恼和用户很着迷，哪个更重要？他们害怕什么？
- 理由。他们为什么要用你的东西？为什么停留在你的网站？什么东西引发他们使用？竞争对手提供的原因是什么？

### 愿望

愿望不等于需求。愿望指人们觉得需要但实际并不真正需要的东西。从用户角度看，可能很难区别两者，但创建背景资料时必须区分两者。愿望得到满足，可能并不会让产品更有效或者效率更高，但能让用户体验更好，因此要加以考虑。

- 已表明的心愿。用户说他们想要什么？
- 未表明的心愿。他们实际想要什么？

### 知识

正如黑色侦探小说中所说，“她知道什么？什么时候知道的？”背后的故事能告诉你人们如何来到你的网站，他们为什么来。闷热的8月下午，她本来有很多事情可以做，而她走进了你的办公室，呃，点击到了你的网站。因此，你发现“她”来到你家门口是什么情境？

- 领域知识。受众有多了解他们要做的事情？他们是想找到有效工具的专家，还是想学习些正在发生的情况的新手？
- 产品知识。客户有多了解你的产品或者服务？他们有多了解自己的需求？了解多少现有选择？
- 竞争关注程度。他们对不同品牌的竞争产品有多留意？有多关注？

### 使用趋势

人们使用产品的模式会决定他们对产品有什么要求和期望。

- 频率。他们只会使用你的服务一次，还是会在工作中定期使用？
- 考虑。他们是否要进行大量研究才决定最终采用首次使用的服





务？他们每次都会进行研究还是只是这一次研究？

- 忠诚。他们对所选产品有多忠诚？只是选择最便宜或者最方便的产品，还是固定一个产品使用一段时间？

### 任务

最终，人们会使用你的产品来做些事情。是什么事情？任务指用户为了更接近目标所采取的行动。对于用户背景资料，要从用户角度看待任务。应该从用户思考的粒度来考虑任务。用户很少会考虑数据库行锁定(而软件工程师会考虑)，他们想的是把钱转到支票账户或者是进行酒店预订。背景资料要包含用户通过系统想要做的最重要任务清单(不论现有技术实施是否能执行这些任务)。

- 理由。他们为什么想做这件事情？
- 持续时间。这件事他们会做多长时间？
- 顺序。各个任务的先后顺序如何？用户认为哪些地方顺序很重要？顺序在哪些地方不重要？任务如何连在一起？
- 关键程度。任务有多重要？这些任务不常见但绝对重要吗？例如，数据恢复不会频繁发生，但发生的时候就非常重要。
- 方法。他们想如何完成任务？
- 模型。是否已经有其他方式能完成任务？他们在用什么工具？怎么用的？

创建任务时，你可能会注意到相同背景资料下完全独立的簇群。例如，卖房人和买房人执行的几乎所有任务都不同，虽然这些任务在主题上有关联，而且卖房人和买房人很相似。要能识别出这种重要差别。两个簇群不仅仅定义两种不同背景资料，更定义了两个不同产品。

### 分簇属性

有了收集好了的属性大清单，接下来要把属性归入背景资料中。可以采用类似于拜尔和霍尔兹布拉特在情境调查中采用的亲和图(Affinity Diagram)方法，但要简化一下。





### 忠告：

如果参加者有安排或者状态不合适(比如说小组主要由初级工作人员构成,但其中有一位高级副总裁),而又强调特定用户观点,要小心背景资料会带入偏见。如果背景资料中有明显矛盾的地方,要进行讨论,要强调用户市场不同于目标市场,除非有确凿数据,否则就是猜想。

- (1) 挑选某个人,按照他觉得能建立有用的用户背景资料的方式,请他把即时贴分成若干族群。针对在线结算业务网站,他可能会建立一个族群,背景资料包括“40多岁”,“男性”,“副总裁”,“互联网新手”以及“负责公司预算”,他还想建立另一个族群,背景资料包括“40多岁”,“女性”,“店铺经理”以及“每晚记账”。每个人都会有自己的受众画面,但这种方式会产生多份背景资料。
- (2) 请另一位参加者把她的即时贴放到上述相关的族群周围,一次放一张。不用管即时贴如何关联,只要她觉得它们的确有关联就行。如果有必要,请人们复制即时贴,这样属性可以归入多份背景资料中。
- (3) 如果参加者觉得有即时贴不适合归入现有族群中,就应该把那些即时贴放到新的族群中。
- (4) 每个人都分好了自己的即时贴以后,开始进行小组讨论。即时贴分得合理吗?它们为什么要分在那里?可以移动即时贴,直到每个人都对族群达成一致。
- (5) 请组中的每位成员重复这一过程。

属性很多,因此可能要花费数小时才能分完,而且讨论会发现对用户需求和对他们要做的事情的假设。

可能有些族群定义了主要用户市场,还有一些较小的族群定义了潜在市场或者次要用户的背景资料。这些族群会类似于业务开发或者销售部门已建立的目标市场细分,但在受众与产品的关系方面更精确。

属性都放在白板上后,就停下来,接下来开始讨论从族群中看到的趋势。有明显趋势吗?趋势符合参加者预期的用户类型吗?哪些地方让人惊讶?

如果有许多不同族群,可以减少背景资料的族群数量。主要族群的数量要保持在3~8个。超过8个会很臃肿,少于3个将无法对有用背景资料提供足够的多样性。

这个过程很灵活。一旦发现属性会如何“自然”成族,就可以随时组



合或者修改代表主要用户市场的族群。用户市场的切实可行的背景资料比墨守陈规更重要。

忠告：

有一种错误概念认为，每个网站只有一种用户类型。这并不正确，没有任何单一背景资料能描述用户。有些用户市场可能会比另一些更重要，但没有一个用户市场能代表所有用户。把用户当成同一群体进行归纳或者讨论是粗暴的过分简化行为，应该要避免。

目前为止，你定义的族群——以及随后在此基础上建立的背景资料——应该互不相同。如果有几份用户背景资料有一些共同属性也没关系，但在所有主要方面的用户背景资料不能都相似。最终目标是要建立一套用户背景资料，不仅能显示潜在用户的广度，还能提示最重要的产品功能。大多数人会使用大多数相同特性，这可能是事实，但还要创建使用较少共同特性的用户背景资料。这可能会要用你会觉得对某人很重要的属性来扩展或者超载某个“自然”属性族群。例如，如果要检查人们如何从多台机器上使用网上结算业务系统是关键，就可以把“在家里和工作时积极使用互联网”增加到合理族群中。

同样，如果有大量受众没有被表现出来，而你也了解了一部分用户市场，就应该继续与这部分群体进行讨论，直到你觉得已经识别出了主要用户群体。请再次记住，这些不是销售市场，而是用户市场，这些受众不一定符合销售和市场营销团队销售产品时所认为的受众。

### 游击方法

从属性清单列表创建用户背景资料，有种较少的劳动密集型方法，即缩短或者消除即时贴分簇过程，并关注较少背景资料。试试通过属性清单得到三份优秀用户背景资料，而不要请每个人把所有情况都写下。这种方法需要有优秀小组主持人参与，因为参加讨论的每个人都要有机会贡献自己的观点。小组要首先拿出几份主要用户背景资料，进行充实，直到看起来比较完整为止。利用尚未使用的属性来增加其他用户背景资料，直至每个人都觉得所有重要用户群体都得到了描述。

### 围绕族群创建人

目前为止，创建过程都与定义目标市场类似，但接下来就会产生不同。目标市场定义了物品或者服务的市场，此市场既要大到足够产生价值，也要均匀到可以忽略不计个体之间的差异。要做到这点，就需要抽取某些





个体之间的细节。而另一方面，用户背景资料引入了个人详细资料来创建现实用户肖像以关注特定用户需求，而不是关注市场趋势。

也就是说要虚构东西。

现在要通过在属性之间创造足够“叙事胶水”，让簇群富有生命力，让它们看起来在描述个人。你会想把每份人物背景资料分配给组中的某人；这个人能充实背景资料，并带回小组进行完善。所加细节不能是主要定义因素(主要定义因素都已经确定了)，而只是能创造丰满性格的细节。

很多情况下，增加细节和创建属于所定义属性范围内的特定“事实”一样都很直接。例如：

40 多岁  
女性  
有孩子  
郊区  
店铺经理

变为

多丽丝·华盛顿  
45 岁  
丈夫史蒂夫  
住在密歇根州理芳尼亚  
有两个孩子：克瑞阁 17 岁，安迪 13 岁  
管理密歇根州特洛伊市的三家 7-11 便利店

增加足够多的细节，让人们看起来真实可信，但尽量不要引入会干扰基本问题的特性。如果生活在密歇根州利沃尼亚市还是伊利诺伊州威拉米特市并不重要，就不要太关注多丽丝的住所，选个地方就行了。不要创建癖好。强大特性往往是从重要属性中提取而来。因此，如果基本受众由典型青少年构成，就不要突然把背景资料搞成 45 岁的职业登山者和爵士乐喇叭手。

现在对属性的剩余部分进行充实。你已经确定了多丽丝为别人工作，



她对其领域“知识有自信”，并对业务有“直接(但不是最终)责任”。现在，可能要写个小故事，把这些特征变成更广泛的人物背景资料。

便利店都归萨米所有，多丽丝从25岁多就开始为萨米工作。多年来，她从职员做起，一直成为了萨米最信任的员工，对于这一点，他们俩都认为多丽丝比萨米更知道如何运营便利店。

你可能想加入一些具体的结算业务信息。

她在一家店的后场办公室工作。每天傍晚，她开车去另外两家店，检查库房，并检查当天的收入和收款机磁带。她把三个店的所有钱都存入Comerica银行(在银行的上班时间内)，然后回去做账。她常常要工作到很晚，晚饭后继续做账。

上述内容可以和她的计算机以及软件使用情况合到一起。

她家里有一台新计算机——安迪不聊天的时候，她就用来上网——她工作时用的计算机至少已经用了五年了。一直没有升级，因为做账占用了她大量时间，她不想因为设置新系统而耽误工作。

创造故事，把所有属性绑在一起，突出多丽丝会定期碰到的问题和产生的忧虑，从而完成人物背景资料。可以想象成自己要模仿别人的生活，把所有要素集中起来，并产生意义，这些要素能讲述一个真实的人、一个真实的角色如何生活的连贯故事。继续充实属性簇群的其余部分。可以把讨论群体分成小组，请每个小组花一小时给每份背景资料编个故事，然后把故事提交给整个团体。

做完这一步，会得到五份人物背景资料(就本例的目的而言)。

多丽丝，三家7-11便利店的经理

萨米，拥有多丽丝管理的7-11便利店

克拉伦斯，古董及工艺品商店店主

基斯和杰夫，建筑承包商，KJ建筑公司的共同所有者

詹妮弗，自由作家

做完后，讨论所有人物背景资料。这些背景资料可信吗？他们看起来





像要使用该服务的人吗？

### 确定优先级别

并非所有人物背景资料都同等重要，因此工作组要确定优先级别。业务需求可能决定优先级别，此时优先级别和目标市场顺序相类似。设计需求也可以决定优先级别。某个受众可能对于产品成功很重要，但该受众的需求附属于另一受众群体的需求。关注更具挑战性群体的需求，就在于满足他们的需求，又在满足从财务上考虑更主要用户的需求。例如，保时捷汽车的设计理念是超越豪华性能汽车迷的期望。我们大多数人都不是豪华性能汽车迷，因此这种理念就能确保如果购买了保时捷，我们大多数人都不会印象非常深刻，也会觉得满意，而那些观点对于保时捷形象非常重要的车迷而言，他们仍会推荐保时捷汽车。

在结算业务的例子中，这种理念意味着，如果有人像多丽丝一样要运营三家繁忙的便利店，这些人的需求很可能包含了只运营一家店铺的克拉伦斯的需求。克拉伦斯可能有其他专门需求，但能满足多丽丝的大部分相同功能都能满足他的一般需求。也许像克拉伦斯的人会比像多丽丝的人多，因此克拉伦斯可能是销售和营销的目标市场，但服务了多丽丝的需求就能满足两者的需求。

从开发角度看，确定优先级别还能创建确定竞争解决方案之间的连贯方式。如果有两个办法能解决问题，应该总是选择一个有利于主要人物背景资料的办法，以确保连续性和一致性。

### 讲故事

有了背景资料，就可以由此来创建情节。情节是描述一个人如何行为或者如何思考任务或者情况的故事。它们包含对人们根据动机、期望以及态度所做事物的叙述。根据背景资料进行角色扮演，考虑所创建的人的情况，并通过他们的眼睛看待他们的问题以及你的解决方案，这些都可以创建情节。

从某种程度上说，故事是用户个人特征记录和分析练习会产生东



提示:

如果你上大学时晚上表演或者玩角色扮演游戏,现在就是把这些知识用到工作中的大好机会。把那些即兴舞台表演方法,《龙与地下城》的技能重新拿出来,给同事们看看是怎么完成的!(可能要把剧作家塞缪尔·巴克利·贝克特作品集和20面骰子留在家里。)

西。用户个人特征记录和分析会对故事模型产生影响。我们通过故事连在一起,进行思考和沟通,以叙述方式从人际关系了解世界。有了模型,就能更容易理解产品的微妙之处,更容易看到产品关系和问题。统计表格或者市场抽象描述基本无法像故事那样揭示关系。

从每份用户背景资料角度,复述初步研究时写下的一些故事。这些步骤可以作为一组、个体,或者成对完成。多丽丝会如何处理你知道已经实际发生的情况?她会做什么?她如何思考?而建筑承包商杰夫会如何处理?

下一步,要为每份人物背景资料建立理想情节。如果在天堂,每件东西对他们该怎么起作用?在什么情况下软件能预计他们的愿望?软件怎么帮助他们做出决定?什么是完美情况?

多丽丝下午5点回家。她和孩子说说话,和史蒂夫吃晚饭,到了晚上七点坐下来做账。她相信存款袋已安全抵达。她登录到网站,检查当天收款总数。把一天的总数拖到电子表格中,可以看到每家店铺的运营情况,还能计算一周的业绩,并和去年同期进行比较。7点20分,工作结束。

一次引入一个制约因素,看看故事会怎么变化。假设多丽丝无法将一天总收入自动输入网站。她会怎么办?如果无法下载到电子表格,会发生什么?如果产品过于繁琐或者复杂,她什么时候会停止使用该产品?根据背景资料进行角色扮演来确定哪些价值观驱动多丽丝的决策,什么价值观对她最重要。

现在想想你为多丽丝和其他用户背景资料所列出的问题。什么价值观和制约因素会影响角色看待那些问题的方式?面对任务X或者问题Y时,这些用户会做什么?理解并传达不同人如何处理情况的方法有助于创建适用于所有人的解决方案。

## 利用背景资料

创建背景资料的过程有益于背景资料本身。它创建了一套能传达用户、其观点及其问题的共同理解文档。但要想从这一过程充分受益,则需





要持续使用。每天产品开发过程中都要记录、分享、更新并使用这些背景资料。

### 文档

向整个团队传授背景资料以及如何使用背景资料,这是确保背景资料长期有用的最佳方法。并非所有使用背景资料的人都是开发小组成员,即使开发小组成员也需要提醒。

所创建的文档应该是对创建内容提炼后的结果。即使已经为背景资料建立过程进行了大型的角色参与研究,最终文档也要简短易读并关注项目的特定需求。

用户背景资料是一种简洁沟通工具,因此首先要为背景资料确定受众,以便能以适当方式展现背景资料。不同群体有不同需求:形象设计与品牌有关,工程与技术制约有关,业务开发与合作伙伴及竞争对手有关,交互设计与功能及特性组织有关。所创建的文档要能反映受众的需求和关心。

团队中每个人都需要知道人们要完成的事情、碰到的问题、愿望、重要行为以及扮演的角色。基本背景资料会包含以下要素:

- 背景资料的名字
- 人口统计描述
- 被提取特征者的目标
- 他或者她的需求
- 他或者她的能力
- 他或者她对产品和任务的观点

另外,其他小组会因为他们使用背景资料的方式产生特定需求(表 7.2)。

表 7.2

受 众	信息需求
形象设计	用户对品牌的关系和情感反应的模型
业务发展	潜在免费以及竞争性产品和服务





续表

受众	信息需求
交互设计	使用目标、特定任务、使用模式
信息架构	熟悉领域任务以及已知命名机制
工程	主要互联网使用以及技术能力，还有使用目标、特定任务以及使用模式

肖像是最终要素。能代表真人的图片能给背景资料增添人性色彩，背景资料看上去会比一段文字描述更“真实”。此外，发现合适照片的过程能让许多已做出的决定具体化。有些产品团队会在工作区域附近把主要用户背景资料的照片挂在相框里，而那里本来要挂一张一百年前的“总裁”照片。许多产品团队会创建海报，里面包含了用户照片和关键属性(他们把海报设计成老西部风格的“通缉”海报)。还有些团队会拿出自己的座位，把所有与背景资料相关(因此与项目相关)的东西都展示出来。这样会进一步强调每天有真实的人要依靠你。此外，挑选照片很有趣，也是用来结束长时间头脑风暴环节的好办法。

找到库存照片来源(设计部门也许有 CD，或者可以去免版税商业图片网站找找，比如 PhotoDisc)。照片来源会包含许多团队都不认识的不同的人的照片。翻翻照片，一直到找到一张合适的“多丽丝”为止，然后就把这张照片当成多丽丝。

## 共享

共享用户背景资料最具挑战性的部分是以简明而有用的方式传达背景资料。最佳传播方式取决于开发团队的规模和构成。较小的团队自身可以成为参与背景资料具体化过程的一部分，也可以和某位参加者一起工作。而较大的团队则需要更多正式文档编写方法。

但对于任何规模的团队而言，介绍会议都是有用的步骤。要向整个团队说明创建过程，并介绍背景资料。介绍特定背景资料的创建过程，并且走查每份背景资料的主要特点，能给背景资料打下基础，并加强功能的有用抽象性。

用简单好用的表格总结背景资料。总部位于旧金山的设计公司 Hot

### 忠告：

创建背景资料时不要用认识的人的照片或者名字。也不要根据你认识的人创建背景资料。扩展或者改变背景资料的时候会分散注意力和导致限制(“但查理没车!”)。





### 提示：

本章着重介绍创建用户背景资料，并不涉及把背景资料融入到开发过程的具体规范。如何优秀处理背景资料与软件设计和业务发展相吻合。请阅读艾伦·库珀的著作《软件创新之路》。

Studio 创造了一种背景资料的单页“文摘”版，(根据库珀的建议)做成薄板，并分发给开发团队每个人。这样，每个人都有了可用的背景资料。薄板更耐用，更容易从一叠文件中找到，并且不会被意外扔掉。

虽然可以共享大量信息，但是信息总是比每个人能记住的信息要多。管理员(或者保管员)有助于集中信息。管理员在创建过程中负责更新人物。也可以作为向开发团队进行解释的资源(“杰夫对 Palm Pilot 下载选项有何看法?”)。

### 利用背景资料进行开发

很容易走一遍开发背景资料的过程而再也不使用背景资料。要让背景资料有用，就需要积极使用背景资料。最直接的，它们可以作为招募筛选人的基础(尽管他们太具体了无法完全使用：一般你需要不同类型，而不是满屋子都是多丽丝)。更重要的是，应该鼓励开发人员开会时通过背景资料进行交谈，并且思考特性如何能对不同背景资料发挥作用。

有种最简单的方法能让人们以背景资料进行思考，即在文档编写和规格中使用背景资料的名字。例如，描述一项新特性时，可以讨论它能如何帮助杰夫或者萨米，以及你觉得他们会如何使用。

利用背景资料来评估竞争对手。去竞争对手的网站，尝试确定他们所创建的背景资料。尝试识别出符合你的背景资料的人会如何使用竞争对手的网站。网站对谁有吸引力？竞争对手哪里成功？哪里失败？

当然，并不是背景资料的所有方面都有用或者可用。有时候，你可能花了很多时间开发特定背景资料或者一组情节，却发现有些小方面比整个情况都有用。这属于典型的正常情况。不要后悔你做的其他工作，把它作为背景资料就行了。

### 定期更新

应该定期检查并更新背景资料。

例如，你首先确定了多丽丝 45 岁，是两个孩子的母亲。但是调查后发现，用户基群中没有很多 40 多岁的妇女符合多丽丝背景资料的其他部



分。用户基群主要由 50 多岁、有孩子或者根本没孩子的女性构成。多丽丝的年龄以及孩子的数量并不是产品关键，因此就可以改变她的背景资料。改变后，她 55 岁，有两个孩子。反过来，如果在该银行现有客户中发现几乎没有像多丽丝所属的企业，可能要转移重点，并将精力集中到更具代表性的背景资料上去。

进行其他研究时，应该对符合背景资料一般描述的人进行比较。如果发现很多地方不符合想法，应该调整背景资料，以反映真实用户群。

如果有时间，还可以通过访谈来检查背景资料的健康性。招募匹配用户市场的人，并进行访谈，关注之前对他们行为和价值观的假设。他们如何适合你对市场的想法？如果差异很大，应该调整背景资料(但大幅调整之前，应该访谈一些那个市场的人，以验证差异的一致性和实质性)。

可能还要考虑用户体验随时间会如何变化。多丽丝定期使用该服务 6 个月后会发生什么？一年后会发生什么？如果有天她发现网站重新设计了，会发生什么？原来的组可以定期举行会议(每 6 个月左右一次)，来考虑新信息会如何改变你对背景资料的理解，背景资料的目标、动作和需求如何随时间变化。

## 实例

杰夫是在线小商业结算业务服务的潜在客户。以下是对其简短描述。

**杰夫·门罗**

建筑承包商，和基斯·格里默德共同拥有 KJ 建筑公司(KJ Build)。

**个人描述**

35 岁

成年后从事木匠工作，主要为类似他现在拥有的小型建筑公司工作。

每周工作 50~70 小时，在密歇根州安阿伯市和戴克斯特市为单亲家庭修建住房。

和基斯共同拥有 KJ 建筑公司，5 年。





## 肖像



## 技术方面

杰夫有一台使用了3年的戴尔电脑，和妻子珍妮特共用。他们通过当地互联网服务提供商 MichNet 公司有一个上网帐号。他用这台电脑和朋友互发电子邮件，看看红翼棒球队的情况，偶尔会看一些木工和房屋建造网站。

去年，他从 sears.com 网站买了一些工具，大部分作为实验。

他几年前购买并安装了小型企业财务软件 MYOB，但从未成功使用过，他还是用办公桌上的纸和计算器记账。

## 角色

KJ 公司雇了一位兼职会计，杰夫和基斯共担财务职责，但杰夫承担了日常的大部分记账工作。

将来，基斯和他会扩张 KJ 建筑公司，招聘几位永久雇员，一次可以同时建造两栋房子，每人管理一栋。

## 任务

杰夫和 KJ 建筑公司都在大湖银行公司有帐号，杰夫使用该公司的网络服务来检查个人账户、支付账单。

他执行的典型任务如下：

- 支付承包商的薪水
- 支付材料费
- 向客户收款，通常是分期形式



- 交税期间, 收集发票和费用单, 交给会计
- 追踪应收账款
- 支付多种小企业费用(牌照, 证券等等)
- 支付营业税(密歇根州安阿伯市和戴克斯特市)

他常常用信用卡解决短期资金问题。

### 当前问题

回款有时候很好, 有时候很差, 偶尔会造成现金流问题, 因为在 KJ 公司收到最后一笔合同金额之前, 就必须要支付材料费和按照小时收钱的分承包商。

有时杰夫不得不堆积账单, 因为他太忙了, 而进一步导致了支付过程复杂。

所有文件需要组织好, 交税时要用。

如果客户付款推迟, 有些信用卡和设备费用就无法支付。

### 愿望

希望支付材料费和分承包商能更轻松, 能追踪到哪些合同已过期。

想减少所追踪的文件(和数字)数量。

希望和基斯更轻松共享财务信息。

想能更好地预测将来的现金短缺。

### 价值观

杰夫喜欢为像自己现在这样的家族企业工作, 而不是为大型企业工作。

最后, 背景资料仍然只是开发产品时需要考虑的这些因素的一些便于理解的抽象内容。但是, 思想和价值观拟人化是一种古老方法, 以故事形式强化人类的记忆本能, 可以促进对复杂思想的交流。这可以追溯到《伊索寓言》和篝火故事所表达的一些寓意。通过创建角色, 并让他们参与到一些具有代表性的故事中, 我们不只创建了一种有利于沟通受众理解(认知)水平的高效工具, 还拥有了一个有助于第一时间洞悉受众情况的高效工具。



## 第8章 情境调查、任务分析和卡片分类

尽管网站包含很多信息——娱乐、广告、信息——但其实每个网站也是工具。网站是很高效的工具，相较于其他方法，它能帮助人们更容易、省时、便宜地完成事情，即使网站只是一些更有趣、更好玩的捷径，它仍然可以作为一种工具。网站能让人们的生活变得更轻松。网站能解决问题。

工具用于解决问题，因此，要建立合适的工具，就需要先定义问题。有许多方法可以提出问题。可以了解目标受众及其要做的事情，从而猜测出来。这种方法快，但充满危险：如果你不属于目标受众成员（作为开发者，你基本不属于目标受众成员），你对用户问题性质及严重程度的理解和用户的理解不会一样。你可能会决定某人需要大锤，但实际上那人只需要一颗小钉子。

另一种方法是请目标受众代表描述问题，但这也会把你引入迷途。人们往往会理想化需求和愿望，他们的陈述常常不符合实际需求和行为。许多网站都非常困惑为什么没有人使用个性化特性，而调查报告却说每个人都想要个性化特性。原因很简单：理想情况下，人们希望能有与其偏好和需求完美吻合的一切东西，但实际吻合过程比他们起初想象的难得多。因此，他们会喜欢符合自己想象的个性化服务，但如果他们自己面对必须采用个性化方式才能完成的任务，他们就会宁愿使用普通服务。

此外，明显的问题是并非总是真正的问题。有人想要新锤子和较小的钉子？也许她真正需要的只是廉价鸟笼，因此她要做一个。鸟笼做好了，也许她永远不再需要锤子和钉子了。要发现人们的真正问题和需求，了解他们实际做事情的情况，最佳办法就是你自己来发现。本章提到的方法用来揭示目标受众如何生活，如何思考，以及他们都碰到了什么问题。



最好在创建具体解决方案之前完成这些步骤,而且通常要在开发过程前就完成。理想情况下,甚至可以先完成这些步骤,再提出开发方法或者建议解决方案。然后,可以把所有开发精力都投入去解决目标受众的问题,而不是要调整现有方法以满足其需求。这种情况实际很少发生,但仍然要尽早进行这类研究。

## 情境调查

“情境调查”的权威著作——本章正受惠于此书——是休·拜尔和卡伦·霍尔茨布拉特的著作《情境设计:定义以客户为中心的系统》(Contextual Design: Defining Customer-Centered Systems)。他们如下定义这一步骤:

情境调查是一种实地数据收集方法,深入研究精心挑选的个人,以达到更全面理解所有客户工作实践情况的目的。通过询问和解释,情境调查能系统地揭示客户群的共同点。

换句话说,情境调查是一种有助于了解人们生活、工作所处真实环境的方法,能揭示在这种环境下人们的需求。情境调查会揭示人们实际做什么以及他们如何定义对他们实际有价值的东西。情境调查能揭示出意想不到的竞争对手(只有其他网站,还是你还要和真实世界现象进行竞争?)和人们的实际价值观。

情境调查以人类学和民族学为基础,其研究方法涉及拜访人们并观察他们的工作情况。通过仔细观察人们,并仔细研究他们使用的工具,可以了解人们面临的问题,并了解产品如何才能符合他们的生活。

### 什么时候适用情境调查

如前所述,理想情况下,每个开发周期都会从情境调查步骤着手,而不是从技术、解决方案、问题陈述或者开发团队着手。研究人员可以挑选目标受众,研究他们,发现他们的需求。他们的问题也许最主要是数据交





换和传输方面的原因，但也可能根本不是。也许用户最大的问题是把咖啡撒到打印稿上。在纯理论开发周期中，任何问题都可以接受，都可以发现并解决。第一个需求产生某种软件解决方案。而第二个则产生杯座和层压纸。

由于本书的重点与信息有关，我们将忽略杯座解决方案，因为软件设计者可以解决的问题有实际限制，但要记住人们遇到的问题和其最终目标不仅仅是有关信息的。考虑用户的体验不能局限于他们对产品的行为。他们的目标与其生活有关，也可能是任何事情。因此，从用户角度看，有关滑板技巧的网站不是要共享或者归类滑板技巧信息——这是次要目的，而是要踩上滑板，做出新花样。信息和网站都是让查找技巧指导效率更高的工具，但其真正目标在于滑板，而不是网站。

①大多数项目都源于需要解决哪些问题的想法和如何解决这些问题的粗略想法。②情境调查可以更明确、有重点地关注这些想法，发现在实际情况中这些问题是如何发生的，这些问题实际到底是什么，人们是如何解决的。因此，最好在创建解决方案之前完成情境调查，通常是在开发周期最开始的时候。

情境调查也可以在开发周期之间或者作为重新设计的一部分来完成。这些情况下，情境调查能告诉你人们如何使用产品、何时使用产品以及用产品来做什么事情。情境调查可以作为对最初假设的检查，也可以作为发现产品进一步发展的用户需求挖掘的方法(表 8.1)。

表 8.1

时间安排	行 动
$t-2$ 周	组织并安排参加者时间。
$t$	开始访谈。为开发团队开始分析时间安排步骤。
$t+1$ 周	完成访谈。回顾录像带，记笔记。完成分析时间安排。
$t+2$ 周	准备小组分析房间。核实参加者时间。
$t+2$ 周	分析亲和图(一天时间)。
$t+2$ 周+1 天	开始结果分析。写报告，并呈报给利益相关者。





### 情境调查过程

因为要走出办公室，进入客户的工作场所和家庭，所以进行全面准备工作特别重要。不可能有机会能回去拿忘带的东西，你是给客户带来公司第一印象的人(不是指等候室里的又酷又新的真皮家具)。

#### 目标受众

本书第6章详细叙述了如何选择合适目标受众，但目前是要挑出你觉得想用产品的人。也许他们已经使用过你的产品。也许他们在使用竞争对手的产品。不管是什么情况，这些人都应该和最终使用你所提供的东西的目标受众有相同背景资料。

要尽可能详细地指定目标受众，并专注其行为。

- 人口构成情况？
- 互联网使用概况？
- 他们经常执行什么任务？
- 他们经常使用什么工具？
- 他们是否必须偶尔使用某些工具来解决特定问题？
- 他们如何使用这些工具？

要关注最重要的客户。你的产品也许会吸引不同群体的人，但只能有几个关键目标市场。事实上可能只有一个。要把所有研究都集中到那个市场，一直到你自己觉得自己明白他们的行为情况，然后再关注次要市场。

#### 招募

有了背景资料，接下来就需要找到符合资料的人。第6章对招募进行了完整叙述，但为情境调查进行招募时需要考虑以下情况。

首先，应当决定要拜访多少人。要拜访的人数取决于分配给研究的时间和现有资源。拜尔和霍尔茨布拉特建议观察15~20人，但由于受到访谈数量和观察所需分析时间的限制，拜访15~20人无法实现。对于第一轮调查，拜访5~8人应该足以说明目标受众的主体如何工作(或者娱乐、购物或者产品正好要关注的重点)。如果发现没有达到研究目标，或者觉





得第一轮研究结果有问题，则应安排第二轮调查。

### 安排日程

确定候选人群后，就需要为他们安排日程。情境调查环节所需要的时间会从数小时到一整个工作日，时间长短取决于任务长短以及要收集的辅助信息数量。在观察人们的同时，人们需要进行你要研究的各类行动，这是安排日程最重要的标准。可能需要与他们进行商谈，以便等你出现后他们再执行相关任务，或者请他们等你到达后再执行任务。

访谈会持续 1~3 或 4 个小时。可以每天安排一两次访谈，还要有时间回顾观察。如果有多位观察员，可以一次安排多个访谈(后勤工作会更容易，但需要额外设备)。

由于要在现场进行研究，因此可以让参加者了解一下将要发生的事情。告诉他们研究的一般目标，研究要多长时间，会使用什么设备，你会看什么东西。不必具体说明(事实上，刨除一些细节，人们能产生更多自然答复)，但在你到达之前，参加者应当明白日程。告诉他们完全不用为你要来做准备，告诉他们重要的是你要看看他们如何工作，无论好坏都没关系。

如果要研究办公环境中的人，常常需要得到许可，并签署保密协议。有时也可以承诺匿名进行隐形研究，但视频设备会非常明显，很难忽视有个陌生人整个上午都在盯着显示器并记笔记。如果有任何疑问，请已排好时间的人告知目标研究人员，并尽早安排好需要的各种表格。

奖励支付要能反映观察的时间长度，大部分研究要支付 100~200 美元。但公司可能有政策限制支付给其雇员，这要提前决定(被研究的公司和委托研究的公司是同一家公司尤其如此——企业内部网或者内部软件项目往往就属于这种情况)。如果不允许支付现金，小礼物就可以了(但不适用于政府机构或者受管制行业)。要同等对待后续访谈，除非你和参加者提前达成一致只支付一次费用(这种情况下，奖励支付要反映总的花费时间)。



### 了解领域

为了了解人们正在做的事情并正确分析数据，需要熟悉他们做的事情。也就是说，要熟悉人们在工作中使用的术语、工具和方法。不必知道人们工作的所有细节，但要熟悉他们工作的领域。

如果对任务一无所知，可以在开始研究之前找位熟悉任务的人带你走查一遍。如果时间够，也可以采用“体育节目解说员方法”，请一位专家以评论方式解释另一个人在做什么。专家不必解释复杂技术，但要足以使你熟悉工作。

如有可能，可以自己尝试一下任务。如果是用软件做事情，可以借一份拷贝(或者培训手册)，用上几小时。如果需要用力体完成，而且试试也没有风险，就可以试试看看(做披萨和录入数据都可以试试，但像核化学就不能试了)。如果要研究的是技术环境，可以请技术支持或者质量保证成员带你走查一遍典型任务，看看他们作为内部用户专家是如何执行任务的。

一般而言，对受众要做的任务了解越多，就越能在观察时更好地解释其行为。

### 开发情境

作为一部分准备工作，要清楚知道你的期望。对于被观察者，你期望他们如何及什么时候做对产品重要的事情，你期望他们如何看待某些要素，把这些期望都写下来。可以和开发团队的其他成员来一起完成，请他们采用第7章描绘受众一般行为的同样方式来描绘你期望人们采取的特定行动。这种方法能给你一个平台，比较观察到的行为。

访谈时，一边观察，一边在脑中记住这些情节。对看到的情况保持开放心态，如果观察到的行为不符合期望，就记下来，以引发更深层调查。

### 实际操作

除了与情境调查有关的前述所有准备工作外，还要做几件事情，因为在外边不如在办公室方便。

#### 忠告：

注意，不要让情境导致观察偏见。把情境当成框架来建立访谈，不要用情境对人们一般怎么做或者不怎么做来制定期望。





### 提示：

有时候，现实和预计的情况截然不同。你可能期望能发现一种情况：比如，使用典型工具的典型一天情况，而你发现的东西完全不同，主要系统发生故障危机或者人们争先恐后赶时间。这种情况下，要注意突发状况如何解决，并对突发情况与预期情况及他人经验进行比较。如果这种情况非常偶然——每五年才发生一次或者被访谈人已经在从事的工作和你要研究的任务无关——应该请他们描述他们的正常工作，也许是请他们说与他们现在工作相反的事情。如果这种情况和你想要尝试完成的事情相差很远，要重新安排访谈时间，那时他们的体验可能更符合研究目标。

- 列出所有要带的东西——每支铅笔、每盒录像带和每本笔记本。第一个访谈前一周就开始列出清单。不管什么时候发现还有东西要带，都要加入清单。访谈前一天要确定清单上的一切东西都准备就绪(我会在已有的东西旁边打勾)，缺的东西务必补全。到访谈当天，把东西放入背包或者车上，并从清单中划掉。
- 如果你被允许参与(无论是允许你观察或者作为参与者参加)，要确认已准备好备份。
- 带上支付奖金费用需要的所有东西。可能要从财务部拿付款单、支票或者装现金的信封。空白收据单页也有用(但有些会计部门会把参加者许可作为收款证明)。
- 了解如何操作设备。要提前搭建试验场所，并模拟用户工作环境。提前一两天，设置好所有设备，就像要在现场使用一样，插入所有电线，延长所有三脚架，打开所有相机，启动所有笔记本电脑。然后分解，并再次组装。找副好耳机，检查一下音频质量。好的音频质量是录像过程最重要的部分。
- 确保供应充分。带上延长线，多带两盒录像带、两叠纸和几支笔。如果搅带了，或者因为访谈令人兴奋用完了两本笔记本，这些情况都无法预料。如果打算用笔记本电脑来记笔记，务必携带额外电池或者交流适配器(一个就够了)。
- 制定用餐休息计划。如果密切观察某人几个小时，会非常消耗体力，你也不想满办公室找饮水机，而又担心漏掉关键时刻。可以带一些瓶装水，并且计划好在访谈间隙吃点东西，或者和参加者一起吃午饭，这是用不太正式的方法了解他们工作背景的好机会。

### 实施情境调查

在情境调查条件下，获得良好反馈内容的关键之一是和参加者建立融洽的关系。既然想尽可能“自然”地观察他们的工作情况，就要设定对每个人角色的期望。拜尔和霍尔茨布拉特定义了可以尝试建立的几种关系。



- 师傅/学徒模式，你是学徒，被观察人是师傅。学徒通过观察学习手艺。学徒偶尔可以提问，或者师傅会解释要点，但师傅的主要角色是工作，边工作边叙述在做的事情(不用思考，也不用解释)。这样，“师傅”就能关注细节，避免因总结而掩盖成功设计的关键细节。
- 伙伴关系是师傅/学徒模式的扩展形式，访谈人和参加者结成伙伴关系，一起尝试提取工作细节。参加者会更多关注他/她的工作中不可视的元素，而访谈人会讨论如何把这些工作变为假设的功能，并且努力提问题，讨论这些功能变为界面后的操作方式。偶尔请参加者倒退几步，针对某一行动或者陈述发表评论，或者思考一下行为的原因。这会潜在改变参加者的行为，但也能在关键点提供关键信息。

拜耳和霍尔茨布拉特还指出要避免以下几种关系。

- 访谈人/受访人。一般情况下，受访人通过回答访谈人提出的问题来揭示信息。除非特别要求，受访人不会透露细节。你不希望情境调查时发生这种情况。你需要了解参加者的工作和思想来驱动访谈。如果你发现自己在像记者一样鼓励参加者说些什么，就需要重新把重点放到受访人的工作上。
- 专家/新手。你可能是软件专家，但他们是领域的专家。正如拜耳和霍尔茨布拉特所建议的：“一开始务必清楚说明你在场的目的是观察他们的工作，因为只有他们才了解自己的工作实践，这是正确设定客户期望的前提。你在这场并不是要帮他们解决或者回答问题。”他们建议，要明白访谈目标不是解决问题，而是了解问题是什么，参加者自己如何解决问题。如果让你像专家一样，就要采用非引导性访谈方法来转移话题，“您觉得它应该怎么打印？”
- 不要成为客人。你是否自在并不重要。你在场的目的是了解他们的工作，而不是获得他们的热情接待。要对现场的礼仪保持敏感。如果在拜访的最初几分钟，有礼貌能主导行为，那就保持礼貌，





### 忠告：

有时候，管理层想让你报告特定员工及其表现情况，而不会提前告诉他们你会进行报告。这种情况下，要解释一下这样做的道德问题——它违背了受访对象对你的信任——可能会违反劳动法。

### 提示：

我采用行动一词来指代任务中的操作。大多数情况下，就是指花几秒钟就能完成的事情。还可以用单一简单想法进行描述。行动可以组成任务，任务可以满足高层次目标。任务粒度可以遍及全局。任务可以简单如填写表格，也可以复杂到类似于挑选汽车。

以博得参加者的好感，但要很快鼓励他们继续工作，并进入以伙伴关系为基础的对话。

- 另一个要避免的角色是“老大”。你在场的目的不是评估或者批判被观察人的表现，务必明确这点。如果他们有这种感觉，就可能无法以典型方式做事情。此外，如果是他们的经理要求他们参加研究，难免会被误认为是一种考查他们工作的狡猾方式。要强调你不会评价他们的表现。如果有可能，一旦从管理层获得许可，就要自己联系人，并为他们安排好时间，而不要当成来自上头的命令。

### 情境调查结构

情境调查结构类似于大多数访谈结构，但其主体是由受访人的工作驱动，而不是由访谈人的问题来驱动。其结构遵循第6章所描述的一般访谈：介绍，暖场，一般性问题，深入重点，回顾和总结。

介绍和暖场应该是参加者和访谈人互相适应、对观察如何进行设定期望的时间。内容包括签署所有保密表格，说明项目的主要细节，并设置设备。检查图像和声音记录良好，到访谈结束之前都不要再调整设备，因为调整设备会分散人们对工作的注意力。描述师傅/学徒模式，并强调你的角色是观察员和学习者。提醒参加者讲述在做的事情，但不用深入解释。

确定了定位，接下来要提出一些一般性问题以了解受访人是谁，从事什么工作，要执行什么任务。请参加者描述典型一天：参加者经常做什么事情？偶尔任务是什么？今天的任务属于典型一天的哪些地方？不要过于深入钻研参加者执行任务的原因，而要重点关注行动以及次序。

接下来是主要观察阶段。这一阶段至少要占访谈三分之二的的时间。大部分时间要用于观察参加者在做什么事情、用什么工具以及如何使用。从请他们描述一下要做的事情着手，对学徒而言——已经足够告诉学徒发生了什么事情，但没有详细到会打断工作流程——然后请他们开始工作。作为学徒，你可能偶尔会要求解释、澄清或者走查行动，但不要让这些要求来驱动讨论。我发现，偶尔记记笔记，而大部分精力关注他们的言论和行



**忠告:**

保持诚实是观察的关键部分。如果你觉得被观察人执行任务的方式和没有被观察时候的执行方式不一样,就要问清楚情况,问问他们有没有按照应有方式在执行任务,或者问问应该怎么执行任务。如果没有按照应有方式执行任务,即使“真的一团糟”,也要请他们说说应该怎么执行。

**提示:**

如果人们担心隐私,就应该提供隐私策略。请被观察人告诉你电话或者会议是否需要保密——或者告诉你所讨论的信息是否属于秘密——告诉他们你会停止观察,直到他们告诉你可以重新开始。这种情况下,找一个能去的地方(可以是附近的会议室或者餐厅),并让他们完成后再来告诉你。

动这种方法很好,但我会看录像带来获得丰富引证以及捕捉交互的微妙之处。有些人建议在现场做大量笔记,而录像带作为备用。不管是什么情况,都要有明确方法来突出要跟进的问题。我把要跟进的问题写在单独的地方,和笔记其他部分区分开。

如果完成了任务或者时间到了,主要访谈阶段就结束了。接下来要立即提出深入问题跟进访谈,这可以澄清很多情况。有些情况不适宜打断(如果是观察外科医生或者股市交易员,整个观察期间都不适宜打断),提出问题会导致任务流中断。要尽可能在参加者记忆清晰的时候多提问题。为了减少人们遗忘,甚至还可以倒录像带,给他们看看取景器中你要请他们描述的顺序(但只有很快能找到这些内容的时候才能这么做,因为花時間提问要比花時間操作摄像机更好)。正如施乐帕洛阿尔托研究中心的资深科学家维多利亚·贝洛蒂所说,“你永远都无法明白真正发生的事情,除非和人们谈谈他们在做的事情。[跟进]访谈……能让你理解看起来奇怪或者无意义的事情。”如果在分配时间段内要提太多问题,或者如果太复杂,应该安排另一次会议来澄清问题(要马上安排时间,一般要在初次访谈后1~2天内就安排好,因为人们的记忆会迅速消失)。

请受访人从其角度谈谈情境调查体验来总结访谈。有没有什么地方让他们感到焦虑?有没有什么事情受访人想以不同方式进行?作为学徒,你有没有什么东西可以做得不同?

拜尔和霍尔茨布拉特总结了以下的访谈过程精髓:

卓有成效的访谈并不是指循规蹈矩,而是指在访谈期间要关注某个人。学徒模式是进行访谈的最佳起点。情境调查的以下四项原则能修正调查行为,以更好地获得设计数据:1)情境,前往工作地点并观察工作;2)伙伴关系,在工作期间讨论工作;3)解释,发现客户措辞及行动背后的含义;4)焦点,挑战你的假设。

因此,你最需要的是摒弃任何偏见,睁大眼睛、竖起耳朵,千方百计了解工作过程,并洞察幕后的工作机制。

**收集什么**

观察人们工作时,要留意四种信息。每种要素可以是临时构成或者正





式，可以共享或者单独使用，可以具体或者有弹性。

- 他们所用的工具。可以是正式工具，比如专业软件，也可以是非正式工具，比如笔记。要记录下他们是否按照工具设计的方式来使用工具，或者是否用于其他目的。工具之间如何交互？是什么品牌？是显示器上的即时贴还是 Palm Pilot(掌上电脑)背面的即时贴？
- 行动发生的次序。行动顺序对了解参加者如何思考任务非常重要。是不是存在工具或者受到办公室文化支配的固定顺序？顺序什么时候会有影响？是否有可以平行完成的事情？是连续完成，还是和另一项任务同时进行？中断会如何影响次序？
- 组织方法。人们会为了方便整理信息，也会由于需要而整理信息。信息整理可以在人们之间进行共享，也可能专属于被观察人。目标受众如何组织他们使用的信息要素？是根据重要性吗？如果是，如何定义重要性？是根据方便性吗？顺序有灵活性吗？
- 产生的互动。在知识传递过程中，谁是重要当事方？是人？是步骤？共享了什么信息(输入和输出是什么)？互动属于什么性质(信息性质、技术性质、社会性质等)？

当然，这四个方面相互交织在一起，有时很难发现线索。参加者可能选择一个顺序来处理数据，或者数据的组织采用了某个特定顺序。要留意行为会涉及许多制约因素的情况。可以通过选择一个合适的时机提问或者在后读的采访时采弄清这些问题。

### 人工物品

人工物品——就情境调查目的而言——是人们用来帮助自己完成任务的非数字工具。记录和收集人们的人工物品非常具有启发性。例如，如果关心人们如何安排时间，就可以拍下他们的日历，看看他们做了什么注释，或者把他们使用作为团队日历的办公室白板的情况记录下来。如果关心他们购买食品的方式，可以收集他们的购物清单，并录下他们在超市挑选商品的过程。想收集手术后备外科手术医生的器械会有困难，但可以录制他们安排这些器械的过程。找一台大存储容量的数码相机就能真正有助于收





集人工物品，就可以把人工物品拍成数码照片，人们就可以在报告和演讲中用来制作图解。

总是要确保获得许可才进行复制或者收集人工物品。

以下笔记摘自对一位健康保险交易员使用已有在线系统创建需求方案说明书(RFP)的观察：

查看纸质表格中的保险覆盖范围(需要摘要)。用笔划出保单覆盖范围章节。

进入 Plan Search(保险方案搜索)页。

打开 New Search(新搜索)窗口。

“我想要 90/70 加 5/10 医药费保险方案，但不管什么情况，我想看到所有 90/70 保险方案。”

没有看表格就填入计划细节。

查看搜索结果页。

指着最上方的保险方案说：“Aenta 公司有份包含糖尿病的 90/70 保险方案，所以我把这份保险方案作为基点，我就能明白需求方案说明书中要有什么内容。”

点击 Aenta 保险方案查看详情。

用浏览器的打印按钮，在大厅打印机(大概三个座位远)上打印出保险方案详细情况。拿到打印出来的纸，并放到需求摘要总结表单最上面。

想详细了解类似保险方案的最近报价。

返回到搜索结果页。翻滚结果，点击 Blue Shield 保险。

### 视频录制

我建议用磁带录制每次情境调查(也就是你每次进行的访谈)，因为相较于其他方法，录像带提供的廉价记录方式能进行更细致的分析。有些关键事件只会发生一次。笔记会分散访谈人/主持人/观察员对参加者的注意力，而且还可能无法捕捉到所有会影响事件的细节。两人一组会感觉亲切，并且要有两位熟练的人参加所有访谈。录音无法记录肢体语言或者面





部表情，而这些情况对了解人们的实际想法和感受非常重要。只有录像能记录所有现场体验。即使在最糟糕的环境条件下，录像也胜于录音；而如果环境条件非常好，录像的好处远远胜于录音。

但是，理想的访谈录制成视频需要一些练习，特别是实地录制时。有两点会导致录像不成功：光线不足和噪声。

- 光线。大多数摄像机都不如人眼敏感，因此人眼看起来完美的房间，对于普通摄像机太暗。如果光线不足，就捕捉不到环境细节。幸运的是，最新几代的消费类摄像机对低光照足够敏感，甚至在最暗的环境下都可以拍摄。如果估计到必须要在黑暗环境下进行摄像，尽量要找到在2勒克斯(照明单位——译者注)或者更小的光线强度下能拍摄的摄像机。这样敏感度会足够高，即使局部阴影的东西也可以进行摄像。要留意不同光线强度以及摄像机补偿光线的方式。明亮显示器放在黑暗房间里，画面其他地方会太暗。对准晴天空光取景，房间内所有东西都会在阴影之中。许多摄像机可以弥补这些情况，但需要适当设置。
  - 声音。计算机风扇的嗡嗡声，空调的啾啾声，隔壁餐厅传来的笑声，最终录制下来的不想要的东西会令人大吃一惊。从技术上而言，声音效果好可能是整个录制过程最重要的组成部分。录得好需要技巧。大多数摄像机都带同向麦克风，能录下摄像机前方(有时是后面)的广泛声音。如果是在迪斯尼世界录制家庭活动，这种方法会很好，但要想从背景噪声中独立出参加者的声音，这种方法就不行了。找台摄像机，带内置麦克风，外置有麦克风插孔和耳机插孔。弄副好耳机，插入耳机插孔，拿着摄像机转一圈，听听内置麦克风如何选取声音。达到现场后，要调整摄像机和耳机，尽量减少会录下来的外部噪声。可能必须要牺牲画面质量，或者必须要外接麦克风并单独安装。如果情况极为糟糕，可以使用外部的长线夹领麦克风。整个访谈过程中都要保持耳机连接到摄像机，并且每隔一段时间就小心检查一下音频质量。
- 即使是最好的设备有时也不够用或者显得多余。有些情况下不可能或



者很困难进行录制。还有些情况无法获得录像许可。这种情况下，常常可以采用录音方式，但录音只能获取部分信息，而且会需要更多密集劳动来进行分析。在可以的情况下进行录音，但要准备好用纸和笔进行记录。

## 如何分析情境调查数据

情境调查的产出并没有整齐的层级结构，而是对成功失败的描述、列举使用情境细节、以及收集大量人工制品。

——大卫·亨德里，助理教授，华盛顿大学信息学院，个人电子邮件

如何解释数据会随着要检查的任务而不同。情境调查有助于了解人们解决问题的方式，人们眼中的重要性，什么对他们是有意义的，还有哪些需求未被满足。主要有以下六种分析方法。

- **了解人们建立的心智模型。**人们不喜欢黑盒。他们不喜欢未知的东西。他们想了解工作机制，以便预测其用途。如果工作或者工具操作并不明显，人们会创建自己的模型。这种模型有助于解释看到的结果，并设定对未来操作方式的期望。这种模型可能与工具的实际功能毫无关系，但如果符合人们自己的经验，他们就会使用。了解被使用的心智模型可以让你充分利用这种理解，并满足人们的期望。
- **了解人们所用的工具。**你要创建的工具应该能取代人们正在使用的工具，因此要知道他们在使用什么工具以及如何使用这些工具。他们可能会查看特殊网页，并在线比价引擎找到最便宜的价格，而不是如你所期望的慢慢浏览目录。他们可能会把地址保存到精心组织的通讯录软件 Palm Pilot，或者放在满口袋的纸片上。跟你竞争的也许不是 Outlook，也不是 Daytimer，而是餐巾纸！
- **了解人们用哪些术语来描述所做事情。**言语能揭示模式和思维过程。如果要选购食品，人们会根据膳食、热量或者食物类型展开讨论。他们可能会使用一个词（“面包”）来代表另一个词（“碳水化合物”），也会误用技术术语（比如用“动力”来谈论汽车悬





挂系统)。注意人们使用的词汇和使用方式,能揭示思维模式方面的  
的大量情况。

- **了解人们所用的方法。**工作流程对于了解人们的需求、现有工具的不足非常重要。了解人们解决任务所采用的办法能揭示他们所用工具的优缺点。如果有人用字处理软件写邮件,然后复制粘贴到电子邮件程序中,这就说明了他/她对于这两个产品的延伸功能的期待。如果有人边看搜索引擎边在便笺簿上匆匆忙忙记网址,则说明搜索引擎和浏览器在这个方面的弱点和强项在于满足他这方面需求的长处。
- **了解人们的目标。**每个行为都有理由。了解人们为什么要执行特定操作能揭示人们工作的内在结构,而他们可能自己都不知道这种结构。目标看似直接(“我需要找到今晚的电视节目单,还要打印出来”),但其背后的原因会揭示人们正在使用的系统的许多情况(“我想在东部时间看《辛普森一家人》,这样电视正常播放时我可以打保龄球,但我不知道哪颗卫星会转播,没办法像搜索网站一样搜索收视指南”)。
- **了解人们的价值观。**价值观系统是心智模型的一部分。我们常常会受到自己的社会和文化背景驱动做决定,就如同我们爱理性驱动所做决定一样。在什么背景下使用这些工具?选择这些工具完全是因为功能,还是有其他因素?品牌重要吗?(如果重要,为什么?人们把什么品质和工具联系在一起?他们有没有注意到其他品牌能做同样事情?)他们的工作环境和相同工具的其他人的工作环境相同吗?如果相同,为什么这群人会采用这种方式?

全面分析收集到的数据有几种方法。拜尔和霍尔茨布拉特推荐他们所称的亲图法(建立在次郎川北和发明的同名方法模式上,工业质量管理界称为KJ图法)。此方法创建出所有观察的层级关系,并集群出趋势。以下是对这种方法的解释(第9章情境设计完整叙述了两人的方法)。

- (1) 观看录像带,每2小时访谈创建50~100条笔记(访谈时间越长,



笔记越多,但笔记数量和访谈长度不一定成正比)。笔记是对工具、次序、交互、心智模式——任何东西的单个观察。给笔记编号,并标识出引发这些行为的用户(拜尔和霍尔茨布拉特建议使用编号,例如:U1, U2, U3 等,而不要采用名字)。编号可随机。

- (2) 召集一组人到房间来,房间要有空白的墙、大窗户或者大白板。拜尔和霍尔茨布拉特建议每人 100 条笔记。最好都是开发团队的成员。请开发团队做分析,能形成对客户需求的集体理解,建立对比,并且每个人速度相同。请这些人封闭工作一整天。要在一天内完成全部亲和分析工作。
- (3) 将小组分成若干对分析人员。给每对分发相同数量的笔记(理想情况下,每对 100~200 条笔记)。
- (4) 在黄色即时贴上写个笔记(是的,的确就是黄色,对于即时贴的颜色,拜尔和霍尔茨布拉特说得非常明确),并贴到墙上/窗户上/白板上。
- (5) 请小组成员把笔记放到黄色即时贴周围,一次一张。这些笔记之间有什么关系不重要,只要小组成员觉得有关就行。
- (6) 如果对于给定的笔记簇群,再也没有可关联的笔记,把一张蓝色即时贴放到这组旁边。在蓝色即时贴上写个标签,总结并命名这组笔记簇群。拜尔和霍尔茨布拉特建议在标签中避免使用技术术语,而应该采用简单词组来说明“把所有单个笔记集中在一起的工作问题”。
- (7) 针对其他笔记,重复此过程,同时用蓝色即时贴命名组。
- (8) 尽量保证每组蓝色即时贴中黄色即时贴的数量在 2~4 张之间。一条笔记不能做一组,超过 4 条笔记的组常常可以分成更小的簇群。但如果没有明显划分方式,多少条笔记能构成一组并没有上限。
- (9) 随着小组数量增加,他们建议使用粉色即时贴对蓝色即时贴进行分类,使用绿色即时贴对粉色即时贴进行分类。

最后,分光所有黄色即时贴,小组分析人员对哪条笔记属于哪一组,





以及如何给蓝色、粉色和绿色笔记加上标签都达成一致(图 8.1)。此时,就能得到一个层级图,引用拜尔和霍尔茨布拉特的话,这张图能显示“到目前为止,小组了解到的工作中每个问题以及每件事情,而且都同实际例子绑定在了一起。”

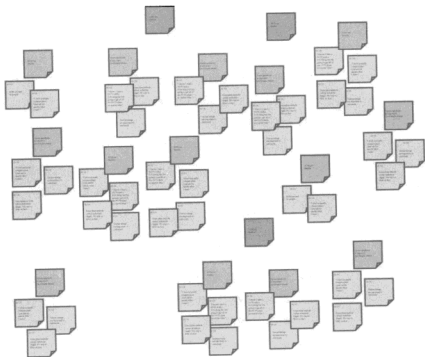


图 8.1 亲和图的一部分

观察保险经纪人可能会产生含有以下片段的亲和图。

需求方案说明书是能收集到大部分信息的工具。

需求方案说明书具有灵活性。

人们会阅读需求方案说明书。(笔记 35, U2)

“我一般不会比较保险方案。我就想要需要的东西。”(笔记 20, U1)  
比较保险方案能提供关键信息。



检索规格非常重要(大部分都缺乏)。

“我想要包含 5/10 药品的 90/70 保险方案，但不管怎么样，我都要找到所有 90/70 保险方案。”(笔记 55, U2)

有些保险方案不覆盖一定行业(合法)。无法过滤出这些方案。(笔记 43, U3)

承保人等级对客户很重要。(笔记 74, U3)

检索输出需要大量过滤。

“大部分时间我都在找信息。”(笔记 26, U1)

除了亲和图法外，还可能采用基于专家阅读数据的更传统的分析方法。观察员都可以碰头，讨论他们所观察到的现象，并猜测这些现象之间有何关联，有何含义。这种方法不是很严格，但可以大量减少数据(即时贴常常会产生大量数据)，并建立足以充分描述人们行为和态度的可行模型。如果开发团队没有时间进行研究，也可以请一名研究员来完成此过程。但只要有可能，都推荐使用亲和图步骤，因为亲和图所带来的好处有利于向开发团队传达人们的需求，并最大限度地利用收集到的数据。

### 建立模型

许多情况下，要符合人们现有工作实践所需要的条件，建立模型来满足未满足需求和创造解决方案足矣。根据从亲和图提取的信息，拜尔和霍尔茨布拉特定义了五种模型：流程模型、次序模型、人工物品模型、物理模型和文化模型。

- **流程模型。**代表“人们为完成工作而进行的沟通”。流程模型展示了在开发团队成员之间传递的信息、知识和人工物品。其要素正式或者非正式，书面或者口头都行。流程模型力求获取在产品生命周期沟通过程中发生的交互、战略、角色和非正式结构信息。
- **次序模型。**代表“工作完成的步骤，每一个步骤是由什么引发的和正在完成的意图”。次序模型显示事情完成的顺序、什么(“触发器”)引发了这些步骤、每一步骤的目的(“意图”)以及这些步骤如何相互依存。次序模型要详细到开发团队可以一步一步地了





### 提示：

频率不等于重要性。因为对于系统设计而言，常常发生的情况并不比罕见的情况更重要。例如，大多数被观察人会用纸记录特定交易，每天可能要记几次，这个问题普遍存在，但并不比由于搜索引擎没有最新信息，而使其要花半小时浏览新文档一样重要。纸质笔记可能会经常用到，明显可以补偿该搜索引擎的缺陷，但也可能代表用户可以忍受的解决方案。照此推论，用纸质笔记没有搜索系统那么重要，而针对此问题，该搜索系统根本没有好的解决方案。

解任务的完成过程。

- **人工物品模型。**代表人们如何使用真实世界的工具来实现目标。从简单照片、绘画或者对象影印入手，人工物品模型“扩展人工物品的信息以显示结构、战略和意图”。人工物品模型更深入地了解人们使用的工具、使用过程、碰到的问题以及(最重要的)为什么这些工具是必要的。
- **物理模型。**代表用户工作的实际物理环境。通过物理模型可以了解工作布局、工作环境中的物品、人们对环境有哪些控制(没有哪些控制)以及他们如何利用环境来完成工作。
- **文化模型。**代表对用户价值观及其自我认识的理解。文化模型把产品放入用户生活和他们生存的现实世界环境中。文化模型包括他们体验的正式组织——他们的其他职责、竞争环境——也包括非正式组织——与工作相关的情感、工作环境、用户审美价值观和风格，等等。

### 产生结果

许多情况下，“啊哈！”这灵机一动的时刻要么来自实际观察过程，要么来自亲和图创建阶段。啊哈！可能完全能转化成产品概念，并提供足够信息，因此接下来可以开始研究产品规格(而不是问题规格)，例如焦点小组和纸原型。

过早下结论从来都不是好事，但即使时间有限，研究结果也要形成文档，用来说明对受众的了解，并和开始时做出的假设和情节进行比较。文档中包含对用户心智模型、工具、术语及目标的思考，可以作为“了解陈述”，以推动其他研究，并作为整体汇入研究目标。每获得一条额外用户信息，就输入文档，并要根据陈述了解新信息。如果新信息有矛盾或者让人困惑，可能需要找一组新人来重复研究。

如果有更多时间和资源来看收集到的信息，就可以看看这些信息。通过数据，可以进一步洞悉人们工作时使用的心智模型和任务流程。拜尔和霍尔茨布拉特的著作有很大部分篇幅用来讨论具体做法。





## 任务分析

任务分析的主要目的是比较系统需求与操作者的操作能力的需求,如果有必要,改变这些要求,从而减少错误,实现成功的完成操作。

——B. 柯万, L. K. 安斯沃思, 《任务分析指南》

很多方法都属于任务分析的定义范围。这些方法包括从结构化标准问卷到可用性专家进行的非结构性访谈。对于本书,我采用了自认为最常见的定义,即任务分析是一种分层次分析单个任务的结构化方法,用于了解任务组成部分中的相互交互。

也就是说,任务分析与情境调查密切相关,而且事实上,可以在情境调查过程中为任务分析收集数据。就重点程度而言,任务分析不同于情境调查。情境调查试图了解围绕和解释任务的整个情境,而任务分析的重点在于任务本身。事情的确切顺序是什么?涉及哪些工具?过程中哪里有灵活性?在任务不同部分,人们需要和使用哪些信息?

### 何时适用任务分析

如果已经知道要解决什么问题,并且想知道人们如何马上解决问题,此时就可以采用任务分析。此时通常属于螺旋开发过程的检查阶段或者瀑布过程的需求收集阶段。任务分析要求你知道任务是什么,并大致明白受众是谁。如果已经在某种假设下,已有了某个解决方案,也可以完成任务分析,但理想情况下,最好在投入精力研究特性和技术之前完成任务分析,因为之后完成可能会对如何实施解决方案的基本假设造成大量修改。

因此,如果已经决定在线开设超市风格的杂货店,但订购系统尚未建成,那么现在就是发现杂货购物者如何挑选杂货的好时机。你要再一次寻找情境调查中的很多相同东西:工具、次序、组织和交互。但是,情境调查过程调查的是人们决定吃什么、如何吃和在哪里吃等决策过程,而任务分析则只关注在超市里购买杂货的任务流程。

此外,虽然任务分析能揭示广告或者促销出现的位置,但其目标是优





化这个交互过程，以便发现提高可用性的机会，而不是发现宣传或者广告的机会(表 8.2)。也就是说，任务分析方法仍然具有很大未开发潜力，如果应用到整个用户体验领域，而不仅仅是其立即务实目标，这种方法可以产生有趣结果。

表 8.2

时间安排	行 动
$t-2$ 周	组织参加者并安排时间。
$t$	开始访谈。
$t+1$ 周	完成访谈。回顾录像带，并记笔记。
$t+2$ 周	开始结果分析。
$t+4$ 周	完成结果分析。呈报给开发团队。分发文档。

## 如何进行任务分析

这里描述的任务分析形式是两种方法的松散组合，即柯万和安斯沃思在其著作《任务分析指南》中提到的分解方法和层级任务分析(hierarchical task analysis, HTA)。

### 准备工作

任务分析准备工作和情境调查准备工作几乎相同。需要挑选目标受众，进行招募并安排日程。需要学习领域知识，开发任务如何开始的情节，并计划访谈。最后，根据进行访谈的方式形成条理，以便能有足够用品、合适的设备和知识来进行分析。

### 收集信息

柯万和安斯沃思在其著作中所发表的成果是在网络兴盛之前完成的，并集中于已安装软件及其周围问题。因此，他们虽然把公开资料和专家作为来源来讨论发现构成任务流的行动，但对于互联网而言，往往无法找到这种来源。此外，重要的并不是人们应该如何工作，而是人们实际如何工作。柯万和安斯沃思所处的任务领域是核能和核化学，其安全要求人们遵



守既定规则，要比其他领域严格得多。现代软件环境中，直接观察再次成为数据收集的基础。

访谈形式和情境调查大致相同：访谈人作为任务执行人的学徒，观察参加者执行任务，并听其说明关键要素。学徒观察并了解任务的微妙之处，偶尔会需要参加者解释动机和选择。

访谈与情境调查的不同之处在于关注点。每件事情都是为了理解参与者如何执行手边的任务。

- 在特定点，他们认为哪些东西是选项？
- 有哪些工具可用？
- 他们如何选择这一项，而不选择另一项？
- 他们在哪里改变了想法？
- 过程如何变化？
- 他们在哪里发生错误？哪些是常见错误？
- 什么造成了错误？如何纠正错误？
- 过程中有哪些输入？有哪些输出？
- 他们所执行任务的频率和重要性是什么？
- 失败的风险是什么？

收集到的信息在信息结构和关注点上与情境调查一样，但更强调次序信息和使用的准确工具。

以下笔记摘自(一个在线家具购买工具)对室内设计师挑选沙发的观察。

“我拿出目录。首先看看[制造商 A]，如果[制造商 A]的沙发是用于候车室，我就去看看[制造商 B]。我有一大书架的目录。最近，我一直在互联网上找目录，但看纸质目录能获得更好的结果。很难找到东西。每个网站都非常不同，要花时间才能在网站找到东西。”

有四本目录：A，B，C 和 D。拿出 B 目录。在索引页中找到“沙发”，然后跳到沙发部分。

“如果以前看过类似沙发，我通常会找几个沙发进行选择，权衡成本，





库存情况以及发货时间。根据照片无法判断出实际颜色，图片常常无法显示面料颜色。我会打电话给销售代表，尽量弄到完整图片。有些公司和我业务量不大，他们也会出现向我销售其他东西。”

用小便签给几个沙发做上记号，准备给厂商打电话问问面料/颜色库存情况。

“我知道很快就能买到，因为以前就买过，但我不敢肯定他们有我们想要的面料。我会记下来问问他们。他们有时候会用不同面料来做，即使他们并不正式发行，但还是要看情况。”

“我要全部算一下。我们需要什么，要花多少钱，假定他们有面料库存，多长时间能拿到货。我要填写采购申请单，翻翻目录并挑出我想要的。我会写段说明给客户。内容包括项目解释、购买清单、日程表等。我可能没时间把一切都解释给客户听，所以必须确保说明要充分。”

“不同厂商提供不同质保。但我要买的不是质保。我买的是匹配和生产时间。成本也很重要，但如果客户已经决定了预算，成本马上就会限制选择。”

准备打电话给厂商。

### 如何分析结果

任务分解和层次任务分析是互补方法。前者描述输入、输出以及任务中的行动动机，而后者把行动组成连贯流程。这两种方法使用的顺序取决于主要需求。如果主要关心任务组成部分如何组合(例如，如果要寻找工具可以优化任务速度的地方)，就可以从层次任务分析方法着手，充实分解关键细节。如果关心的是创建符合现有实践的工具，就可以从分解着手，并以层级形式重新组合部件。

#### 任务分解

任务分解是打碎任务、分成组成部分行动的过程。说起来容易做起来难。如果有人操作控制面板的带按钮和刻度盘，每次按下按钮或者读刻度盘就是一个行动。但如果还是这个人买新车，任务分解就会比较困难，因



为所采取的步骤——选择、权衡、比较——几乎都不明显。这种情况下，要鼓励参加者说出所有想法，并请其解释具体行动。

要在脑中记住最终目标，有助于挑选出每个行动最相关的方面。如果脑中有具体工具或者解决方案，可以在任务分解时强调一下。有时针对人们面对的问题，可以根据脑中已有的解决方案倒过来工作，看看解决方案是否符合问题。

描述每个行动。柯万和安斯沃思列出了几十个已用来描述行动不同方面的类型。

- 目标。为什么这里有这个行动？这个行动如何把任务向目标推进？
- 线索。什么事情让参加者觉得要执行该行动？
- 对象。行动对什么进行操作？
- 方法。行动内容是什么？
- 选择。此时有没有其他行动？怎么选择了这项行动？

描述行动的时候，要做一些错误预测。问问如果行动没有完成或者有错误，会发生什么。如果有足够多的行动产生了有趣的错误情况，可以做一个单独的错误分类，用它来描述每项行动。

对于每项行动，都要尽可能多的提供答案种类。描述行动的过程经常会产生问题，并引发新解释，因此要走查分解几次，以确保一致性和彻底性。

以下是购买家具任务的分解片段。

准备分解，提前确定可能会有用的类别，然后创建表格，边观察边填写。这样一来，就能知道每项行动有什么信息，还需要什么信息。如果每项任务包括很多行动，或者行动来得很快，任务分析会非常艰巨。这种情况下，更好的做法是密切观察任务，并将任务细分。

行动名称	目的	线索	对象	方法	选择
列出需求	明确选择类型		Word 模板：尺寸（长宽高）、颜色、预算、风格	和客户交谈	





续表

行动名称	目的	线索	对象	方法	选择
拿到现有目录	列出现有选项		目录, 需求清单	根据需求, 比较目录中的选项	
设定顺序, 仔细阅读目录	从最有名/最有可能的厂商开始	了解厂商选择	目录, 需求清单	翻阅目录, 根据需求比较选择	首先看 A, 如果是用于候车室的沙发, 再看 B
标记目录中的项目	找到主要候选厂商		目录, 需求清单	视觉检查, 并与需求清单相比较	
标记出需要跟进的项目	分离出需要进一步调查的项目	如果不清楚特定项目是否有选择都在	目录, 已标记项目	视觉检查, 并与需求清单相比较	
调查已标记项目	根据需求, 完成现有选项清单	需要进一步调查的项目清单	目录, 需要跟进的项目清单	致电销售代表	
全部选项	为客户做出最终清单	完成所有跟进工作	已完成的选择模板, 预算模板, 需求清单	根据选择及成本填写预算模板	

## 游击任务分解法

由于资源或者时间有限, 在许多情况下实际都无法进行这种级别的任务分解。此时, 可以采用非正式任务分解。利用任务分解的基本思想, 但不要了解任务所有细节, 填写所有方面, 而要关注较大的想法, 只利用自认为最重要的方面。如果经过几次访谈, 觉得可以写出基本任务, 就写在即时贴上, 创建一些初步组织, 然后组成流程。随着访谈人数增加, 可以调整数量以适应新信息。虽然过程不彻底, 但要快得多。

## 综合在一起(层级任务分析)

层级任务分析由三者构成: 提取形成任务的个别行动, 将其放入层级关系中, 并建立在层级关系中如何移动的规则。最终产品是一张流程图,



图中包含导致每项行动的次序，做出的选择以及行动后果。

层级任务分析不必紧随任务分解之后完成。在访谈结束和被观察用户执行任务之后，就可以坐下来进行层级任务分析。该程序是种自上而下的分析方法，基本上是交互开发人员、信息架构师和技术文档工程师站在白板旁边，努力理解事情是如何完成的过程组合。

- (1) 从目标着手。可以是参加者的最终目标(“装修房子”)，但往往需要分析抽象的最终目标，而这种分析对于简单任务分析太复杂了。可能要从更直接的情况着手(“获取现有沙发清单”)，并由此开始工作。
- (2) 确定目标的子目标。子目标指达到主目标之前发生的所有事情。例如，如果目标是获得现有沙发清单，就需要从目录获得清单，也可以从家具公司代表那里获得一张目录中未列出(或者没有库存)的沙发清单。每个目标的子目标很少会超过三四个。如果的确超过三四个，就可能要创建几个包含若干中间任务的中间目标。
- (3) 确定如何安排行动，并创建如何行动之间的流程计划。内容包括确定哪些行动必须跟在哪些行动后，哪些可选，哪些可以互换等等。如果目标没有自然流程，可以挑选一个看起来见效的流程，如果发现有困难以后再进行调整。
- (4) 对每个目标及子目标重复分解过程，一直到得到单个行动为止。

最终结果是大家熟悉的线框图，单个目标变成一棵目标树，下方是每个目标的子组成部分，并链接到目标树，从而创建出完成目标的程序(图8.2)。

如果要正式分解任务，除了不是在宣传单上建立目标及关系清单，而是采用从访谈中提取到的目标及关系之外，程序几乎相同。这种程序是自下而上的方法，类似于情境调查中创建亲和图的方式。

- (1) 从把行动进行排列组合成与相同结果关联的组着手。平行发生的行动可选，或者可以互相替换，并要打上相应标签。
- (2) 根据结果给问题类别打上标签。这些标签就是子目标。
- (3) 根据其完成的目标，排列组合子目标。





- (4) 给这些类别打上标签，这些标签就是目标。
- (5) 重复这一过程，直到所有行动都进行了分类，所有问题类别都进行排序并打上了标签。
- (6) 把子目标类别组织到目标类别下方，用箭头表示目标与子目标如何关联，创建图表。

该程序产生的图和图 8.2 一样。

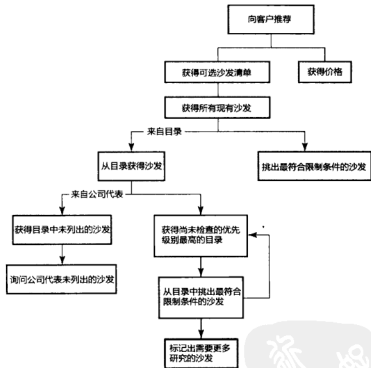


图 8.2 购买沙发分析图(部分)

### 检查图表的有效性

有了图，还需要利用一种从功能分析系统方法(柯万和安斯沃思在其著作称为 FAST)演化而来的方法来验证图的有效性和完整性。

- 层级图中，指定方框上方的方框能回答“我为什么\_\_\_\_\_?”





的问题，\_\_\_\_\_就是方框中的文字。根据以上的例子，如果要回答“我为什么要和公司代表谈谈未列出选项？”从图中可以知道，答案为“因为你需要知道哪些选项未在目录中列出，这样就能列出所有现有沙发的完整清单，可以向客户列出推荐可选清单。”

- 指定框下的方框能回答“我怎么\_\_\_\_\_？”的问题。如果要回答“我怎么从目录得到沙发？”，看图就可以知道，答案为“找到优先级最高的未检查目录，从中挑选出沙发，并同时记下需要更多研究的沙发。”
- 带箭头的项目只能往一个方向(通常代表“怎么”)。如果每个方框的两个问题都能回答，就代表图是完整的。

完成任务分析图之后，可以请几位知道任务的人走查一遍，以确保分析图符合他们的经验，从他们的体验看，次序、任务及目标是他们熟悉的。

当然，这种详细方法费时，也无法捕捉某些任务的微妙之处，但它会迫使对揭示基本程序真正性质的任务进行深入思考，并且理清复杂的交互关系。理解人们的行为方式之后，就能轻松创建符合他们方式的产品。任务分析提供了符合人们现有行为并满足其迫切需求的具体框架。此外，因为是在分析之前进行了描述，任务分析还可以显示出产品的弱点和冗余：可能人们要完成的任务并不是产品原先设想的，也有可能是人们目前执行任务的低效方式，而产品可以使其简化。

如果任务分析图足够准确，开发人员能形成一致说明书、产品任务完成检查清单(目标清单)以及产品完成顺序(任务流)。

## 卡片分类

卡片分类<sup>①</sup>是用来揭示人们如何组织信息、如何分类和如何关联概念的方法。

① 有关卡片分类方法的更多实践案例，请参见清华大学出版社2010年出版的《卡片分类》，作者 Donna Spencer，译者柴林艺。——编注





## 卡片分类何时进行

知道需要组织哪些信息时，最适合进行卡片分类，但要在实施组织解决方案之前进行。卡片分类通常是在产品目的、受众以及特性都已经确定之后，但在开发信息架构或者设计之前再进行，把卡片分类放到设计过程的某个中间位置。这样可以把卡片分类和情境调查及任务分析区别开来，后两者都处在开发过程的开始阶段。

另外，卡片分类是一种快速简便的方法，因此无论何时改变信息结构或者要对现有结构加入要素，都可以采用卡片分类。

## 卡片分类过程

### 准备工作

卡片分类的目标是了解受众如何理解你提出的信息空间。因此，不同于其他方法，卡片分类对应该邀请的人没有任何额外限制，只要招募到符合目标受众的人即可。通常情况下，从受众中招募 4~10 人就能充分了解人们是如何组织信息的。

一次安排一个人，参加者不会感到有竞争压力。如果有地方能让参加者安静坐在那里，而且如果附近有人可以回答问题(但如果只有一个人作为监视员，每隔约 15 分钟就要安排一个人，以便监视员有时间向每位参加者介绍卡片分类)，也可以同时安排几个人。花一小时进行卡片分类研究绰绰有余。

### 整合卡片

卡片分类过程核心就是卡片，这并不奇怪。在一叠坚固的相同卡片上，为需要组织的信息写上名称(或者使用字处理软件的邮件合并功能，并打印到邮寄标签上，然后可以贴到索引卡片上)。可以是具体板块名称，可以是考虑采用的术语，可以是网站不同板块背后的概念，可以是要使用的图片，甚至还可以是个别网页的说明。除了文字，要使用相同卡片来减少干扰。



卡片数量可多可少,但一叠标准卡片(52张)的数量能很好平衡卡片数量不够而无法进行足够分类和卡片数量过多而造成困难的情况。如果要对几百个类别进行卡片分类,可以考虑分成更多可管理块,进行多次测试。

卡片上的文字要能反映测试内容。如果想发现人们如何组织概念,可以在卡片上用一两句话解释概念。但如果想了解人们怎么理解一套图标,而无需了解人们怎么理解你对图标的定义,在卡片上写出标题即可。

### 分类!

召集参加者,过一遍所有初始格式,向参加者介绍卡片分类的概念。可以说说以下内容。

这有一叠卡片。您在网站上可能看到的東西都写在卡片上。我想请您按照您觉得有意义的方式把卡片分成组。您自己决定要花多长时间。分组没有对错之分。请尽量把所有卡片都分成组,但不是所有卡片都一定要分成组。您不需要说明为什么卡片属于同一组,只要您觉得分组正确就行。关键是您觉得合理就行,不用关心其他人是否觉得合理。

提供一叠即时贴,几支笔,还有一些小夹子或者橡皮筋。分类完成后,如果行的话,请他们给每组写上标签,但并不是每组都需要标签(不要提前告诉他们要请他们写标签,因为会造成人们根据标签来分类,而不是根据自然感觉进行分类)。完成后,请他们用夹子或者橡皮筋把卡片绑在一起,并把标签放在每组上面。

最后,请他们把分类组成更大的类,但不要移动卡片,如果能想到名称的话,请他们给更大的类命名。然后,用更大的夹子或者橡皮筋,把源(Meta)类绑在一起,并总结这一环节。

### 卡片分类分析

卡片分类产出有两种分析方式,正式方式和非正式方式。

#### 非正式方式

从所有参加者那里拿到卡片类别,然后观察这些卡片。把类别复制到





白板上。看看类别趋势，就可以推断出人们如何直接理解各种要素之间的关系。例如，如果人们把“News”(新闻)，“About us”(关于我们)和“What we like”(我们喜欢的东西)放在一起，说明他们关心的是集中归类产自你的观点的所有信息。但是，如果他们给“News”(新闻)写上“Latest Deals”(最新交易)和“Holiday Gift Guide”(节日礼物指南)的标签，也许说明他们把有时效的信息联系在一起。两者之间的差异可能说明两种信息架构的差异，一种能符合用户期望，而另一种迫使他们寻找信息。

可以从整体观察这些族群，也可以一次跟踪一张卡片，看看这张卡片被分入的组。不要从字面上处理这些族群。人们现有的组织方式可能无法产生可扩展或者功能架构。相反，要看看把它们绑定在一起的内在主题。留意人们没有分类的卡片，还有分类不同的卡片。这些卡片给人们带来麻烦了吗？是因为名称吗？还是因为其基本概念？还是因为它和其他要素的关系？

看完所有参加者分的所有族群，也列出所有主题，然后要看看他们分配给各族群的名称。看看标签之间的关系、标签和标签下族群的关系，应该就能得到建立接近用户群预期的信息组织方式的架构基础(即使信息架构师不会采用这种架构，营销人员也能用这些术语向潜在客户解释产品)。

### 正式方式

族群分析属于统计学分支，用来衡量多变量环境下物品之间的“距离”，并试图找到变量空间下紧密联系的族群。以白板为基础的卡片分类分析就可以完成这样的工作，但需要数学方法才能完成彻底的一致分析。你可以跨组发现有些对象在许多维度都类似但在其他方面可能都无明显相似之处。由于人们很难可视化超过三维的东西，而判断相似性又常常需要超过三个变量，因此这种方法可以用来“发现”其他方法难以发现的族群。

族群分析的经典例子来自市场调研，即目标市场分类。人们在许多方面都不同：年龄、收入、种族、居住地、性别、购买的东西等。根据一个标准挑选类型相对很简单，但要根据所有标准挑选族群就要困难得多。如



何定义雅皮士？他们开路虎吗？不，有些开宝马，有些开路虎的也不是雅皮士。他们的年薪都是12万美元吗？不，有些人赚3万，但过着年薪12万的生活。计算社会中的族群很难，但族群分析经常可以提取出有用差别。

就卡片分类而言，它属于反向工作。卡片分类通过人们组合的族群发现潜在变量。是否经常有些东西比其他东西被更多分在一起？某些卡片之间是否有隐藏关系？光看卡片很难发现所有这些情况。

不幸的是，数学族群分析不简单，没有计算机很难完成。统计软件包，如SAS和Statistica，都包含族群分析模块，但很昂贵，而且还需要了解分析中采用的统计过程。

幸运的是，有一种软件专门用于卡片分类数据族群分析，而且截至2002年夏天，它都是免费的。1999年，IBM用户参与小组向可用性研究业界提供了EZSort软件，分析卡片分类过程变得更加容易。EZSort把参加者创建的组作为输入，产生树形图来显示卡片群组之间的关系，并揭示卡片族群以及族群的族群。这些图能更容易从相似处分离出强烈共同亲和关系，更容易看见较大主题，而通过层层卡片及分组很难发现这些情况。

截止到2002年秋天，可以在以下网址下载该软件：

[http://www-3.ibm.com/ibm/easy/eou\\_ext.nsf/Publish/410](http://www-3.ibm.com/ibm/easy/eou_ext.nsf/Publish/410)

软件使用过程很简单：把所有卡片名称和参加者输入程序，把卡片名称拖动成堆，软件就再次创建了每个人的族群，族群名称与虚拟堆关联在一起，然后运行软件。产生了一张图，采用树形图表示所有项目之间的关系(图8.3)。一棵“树枝”距离另一棵树枝越远，两者的相关性就越小。

图中显示出该公司有关信息的强族群(图顶部的树)，这可能说明人们期望能在同一个地方看到所有这些(或者至少能以同样方式处理这些)信息。同样，对一般产品信息，比如推荐、评论和“免费东西”，都用类似方式做了考虑，同时“免费东西”和同组中的其他东西略有不同。其他要素还形成了较小族群。两条垂直线代表会员资格“门槛”——要分在一起的项目需要有多紧密关联——每个项目下方的数字代表了每条线代表的计算得到的亲和关系。

#### 提示：

要想进一步了解EZSort，请查看董建明(音)，雷利·马丁和保罗·沃尔多的文档或者论文，网址为[http://www-3.ibm.com/ibm/easy/eou\\_ext.nsf/Publish/410/\\$File/EZSortPaper.pdf](http://www-3.ibm.com/ibm/easy/eou_ext.nsf/Publish/410/$File/EZSortPaper.pdf)。



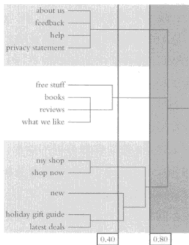


图 8.3 典型购物网站的 EZSort 图

### 利用卡片分类确定优先级

卡片分类主要是一种组织或者命名方法,但其变体也可以用于其他目的。信息架构师杰西·詹姆斯·加勒特是我在 Adaptive Path 公司的商业伙伴,我俩开发出一种方法,利用卡片分类来了解人们如何确定功能的优先级别。

召集一组(12 位参加者)来组织 50 张卡片,卡片上描述了在网站上展现当前事件现有以及潜在的特性(可以在研究人们如何阅读在线新闻时,以长期一对一访谈形式完成)。每张卡片上都用一句话描述网站上信息展现或者组织的方式。首先请参加者把卡片分成四堆,这四堆描述了他们使用网站时觉得功能对他们有多大的价值。四堆名称为“最有价值”、“有些价值”、“最没有价值”和“没有任何价值”。一边请参加者把卡片分入这四堆中,一边请他们讲讲对主题的想法。这样做的目的是了解为什么他们做出一定选择,他们如何思考这些特性。

参加者完成卡片组织后,请他们重复一次,但这次只针对“最有价值”堆中的卡片,同时采用一套不同放置标准。新任务是根据他们认为会使用这些特性的频率来进行组织。这样做是为了把吸引人们来到网站的特性和



他们觉得最立即有用的特性区分出来。

为了了解参加者认为的价值所在，可以采用 0~5 分，给每个分类打分，以表明其偏好强度：

- 0 - 没有价值
- 1 - 最小价值
- 2 - 有些价值
- 3 - 最有价值，很少用
- 4 - 最有价值，有时候用
- 5 - 最有价值，经常用

对参加者选择进行评级，计算出他们评级的中位值。有些选择的中位值相同，因此需要进一步进行组织。计算出选择的标准差，代表了给定类别所有评级的一致程度，低偏差代表更大一致程度。第 11 章对中位值和标准差进行了定义。这份清单首先根据偏好进行排列，再根据一致程度进行排列，开发团队能更清楚了解客户价值观，从而帮助他们安排自己开发精力的优先级别。

卡片分类过程揭示了人们的现有理解及偏好，可以显示只检查簇群清单无法发现的微妙关系。它还能了解概念之间的相互关系，因为偶尔检查时，看似牢固的关系经过仔细分析结果却可能较弱。

情境调查、任务分析和卡片分类可以带你深入发现目标受众如何理解周围环境及其需求。作为规范方法，三者可以尽早关注项目，消除许多错误转变，减少今后询问产品重大基本问题的需求，而此时开发应该关注产品体验，而不是其目的。



## 第9章 焦点小组

焦点小组属于结构化小组访谈，能迅速揭示目标受众的愿望、经验和优先级别，而且花费不高。虽然由于有时和欺诈营销有关而受到中伤，但焦点小组并不像人人常常所认为的那样声名狼藉。焦点小组既不是治疗糟糕产品的灵丹妙药，也不是用来证明不合理决策的伪科学巫术。如果有好的主持人引导，经过认真分析，并合理展现，焦点小组会是一种优秀方法，能发现人们对某一主题的想法，尤其是他们的思维过程。焦点小组能揭示人们需要的东西，而这对于要决定哪些应该是体验的一部分以及如何展现是至关重要的。

焦点小组最初被称为“有重点的访谈”(focused interview)，人们在20世纪30年代将其发展成一种社会研究方法，后来在二战期间改进成为改善士兵生活的一种方法，上世纪50年代市场营销采用了这种方法。因此，焦点小组很可能是最古老的用户体验研究方法。焦点小组系列属于严格主持的小组讨论，而参加讨论的人来自产品目标受众的一小部分。讨论目的是让人们舒适地说出想法和感受。

对于软件或者网站开发，焦点小组用于开发周期早期，而此时想法产生、特性优先级别以及了解目标受众的需求至高无上。焦点小组能说明人们认为哪些特性最有价值，人们为什么这么认为。作为一种竞争性研究工具，焦点小组能发现人们认为竞争对手的产品或者服务最有价值的地方以及失败的地方。焦点小组有时甚至能完全发现产品或者服务的新竞争者或者应用。

焦点小组提供了一种短时间内听取大量第一手体验的方法，可以给开发团队打下早期坚实基础，用来分析产品及其用户需求。作为一种可观察、可触摸、无术语的方法，焦点小组能吸引公司产品开发成员，而他们通常





不会有机会、专业知识或者时间参加用户体验研究的过程。

简而言之，焦点小组提供了从用户角度快速、低成本以及(通过认真准备)轻松看到现实的独特机会。两位兼职工作人员就能设置和运行一系列焦点小组，并在三周内分析结果。但如果人数相同，一位专职研究人员开展情境调查所花费的时间几乎是开展焦点小组的两倍。此外，调查需要抽样的人数大得多——大大增加后勤和分析的复杂性——同时对受访者的动机和态度的了解更少。

这并不是说，焦点小组提供的信息和情境调查或者调查提供的一样。这三种方法能发现的信息不同，但同样有用，但如果在很短时间内需要大量可靠信息，焦点小组是一种廉价方法。

## 何时适用焦点小组

焦点小组的目的不是推断而是理解，不是概括而是确定范围，不是陈述，而是对人们如何认知提供观点。

——理查德·A. 克鲁格，《焦点小组》，第87页

明白何时采用是成功使用焦点小组的关键之一。这种方法很直接，有弹性，但并不适用于所有情况，也不适用于产品开发的所有阶段。

### 焦点小组的长处

焦点小组善于发现愿望、动机、价值观和第一手体验，换句话说，它是获取人们态度和看法的工具。焦点小组是一种环境，人们(在理想情况下)感觉很舒服，并说出想法和感受。人们可以分享问题观点和假设，这些都是体验核心，并与现实世界情况相关。

焦点小组结合情境调查和任务分析，就能在写代码或者在白板上画图之前建立人们当前做法及其动机全景。

作为竞争性分析工具，焦点小组用于产品开发的早期阶段，但也可以在重新设计或者更新周期完成。在产品开发周期开始时，了解基本问题和





### 提示：

焦点小组能发现人们对他们的需求和价值观的认知。这不能说明焦点小组能发现人们的实际需求，或者真正对他们有价值的东西(从某种客观意义上来说)，只是说明他们认为需要的东西、他们声称有价值的东西。但把相对论(还有斜体字)放在一边，知道对需求的认知和知道这些需求一样重要。最终，人们会根据他们认为产品能做什么和产品符合他们需求的程度来判断产品。它不会取代准确功能或者需求评估，但产品对其服务展现越靠近人们对需求的认识，人们就越有可能使用产品。

认知通常是最需要的东西。此时开发团队也在努力寻找产品要解决的问题，应该如何解决，与其他方案相比，使用他们的解决方案为什么对消费者有价值。同样，在早期过程引入竞争产品的用户，很可能会发现为什么人们认为竞争产品有价值，他们觉得最关键的特性是什么，什么问题经常困扰他们，他们觉得竞争产品还有哪些不足。除了从明显的市场营销角度，这些信息可以立即驱动特性和交互开发，并在配置资源之前定义出更接近目标受众愿望的用户体验。

后面的开发周期中，焦点小组有助于识别和安排特性的优先级别。知道人们为什么看重某些特性有助于确定开发哪些特性、按照什么顺序开发。此外，焦点小组可以作为头脑风暴环节，可以达到协同作用，因此相较于个人，参加者聚在一起可能会产生更多想法。

我为一家网站实践过一次焦点小组，网站内容集中于为互联网新手准备互联网开发相关新闻和信息——哪些公司在做什么，什么软件何时发布，等等。我们根据目标受众进行了一系列焦点小组，结果发现大多数互联网新手一般也是计算机新手，他们不关心互联网行业，只关心软件工作的基本情况以及何处寻求帮助。此外，几乎所有人使用互联网的过程都有困难，他们厌倦了使用互联网，也没兴趣安装或者学习新软件。他们想知道能用它做什么，怎么用它来改善自己的生活。根据这些信息，这家网站决定不强调互联网新闻和信息，而强调互联网的自助性质，在一个地方收集流行软件产品信息。

### 焦点小组的不足

首先，焦点小组无法获得可用性方面的信息。一个小组的人没有好方法能告诉你他们是否能使用某个界面或者某个特性。他们只能表达自己是否喜欢你的想法，但不能告诉你他们实际是否会使用。

其次，焦点小组的结果无法量化推广到更大的用户人口，因此它不能取代调查。具体来说，焦点小组所创建的广义模型是基于分析师的观察结果。这些模型假定适用于类似的受访群体，但这些模型不具有统计意义，因此无法保证某个小组中的反应比例符合更大用户群体的反应比例。这是



焦点小组和调查的重要区别。调查有统计学意义(或者应该有),调查所观察到的比例可以推广到更大群体,而焦点小组无法做出这种保证。

因此,对于调查整体目标受众中存在的普遍现象,焦点小组不是一种充分的方法,但能告诉你受众为什么会这样表现。一旦知道了“为什么”,就可以通过统计学研究进行验证,例如调查。但确定趋势常常就足以采取行动,不再需要特定的量化。正如了解商业街区并不需要知道具体地址一样,了解某种普遍现象也不用知道确切规模。

焦点小组创建出的情况会对参加者和分析师都具有欺骗性,分析师通过字面解释焦点小组中的陈述,而不是提取出所暗示的用户态度。由于错误理解焦点结果所造成众多产品(例如许多故事片)做得更差,而不是更好,就反映了这种欺骗性。漫画书有个有趣例子,可能有些未经证实:为了能给超人迷想要的东西,人们进行了一个漫画消费者(10~12岁的男孩)焦点小组,问他们崇拜什么样的人。他们的答复经过字面解释,在20世纪60年代有段时间,不管焦点团体认为崇拜什么人物,人们都崇拜超人,这就产生了一系列超现实故事,钢铁侠穿着印第安人的服装成了警察局长或者跟乔治·华盛顿见面(而温顺的配角吉米·奥尔森变成了巨大的宇宙海龟)。它导致创造性破产,故事情节错综复杂到不可思议,最终不得不完全取消,漫画要一切从头开始,似乎这些故事没发生过一样。

最后,如果要证实观点或者判断立场,就不能用焦点小组。数据收集过程和分析是以多层次的人类判断为基础。因此,如果作为证据依据,结果是值得怀疑的,实际也正是如此。焦点小组是极好的探索过程,能深刻了解动机和思考过程,但它并不会(也不能)用来明确证明或者推翻东西。

## 四种焦点小组

软件或者互联网开发圈有四种常见焦点小组。所选小组类型取决于想要回答的问题类型,而反过来又要依据产品开发所处的阶段。不要受到这些类别的限制;它们只是提供粗略指南。

**1. 探索性。**这类小组就某一主题给出普遍态度,帮助开发人员搞清楚产品的最终用户怎么理解产品,用什么词语谈论产品,用什么标准判断





产品。例如，有一家家具公司就很想知道人们购买家具时使用什么标准，人们在线如何购买类似物品。在开发的开始阶段，他们进行了焦点小组，最初发现人们坚持要“实际”看到家具才会购买(因此就否定了在线销售家具的整个业务方案)。进一步的讨论表明只有某些类别的产品(比如沙发和床)，才会因为无法直接体验而受到怀疑。而其他家具(桌椅)，大多数人都会仅根据图片和说明就购买，而且事实上，人们喜欢这样做。

**2. 特性优先级别。**这类小组关注哪些特性对小组最有吸引力，为什么最有吸引力。一般情况下，这类小组一般在接近开发周期的开始阶段进行，而此时已经明确了产品的一般轮廓。这类小组会假定参加者对某种产品有兴趣，讨论以“参加者希望产品能为他们做什么”为中心。例如，有一个网页制作服务焦点小组，而参加者对社区服务不感兴趣，对能帮助他们建立和营销自己主页的工具也不感兴趣，该网站想传达的“社区感觉”和自己宣传的语气基本无法发挥作用。对于参加者而言，网站的价值在于其工具和免费磁盘空间。

**3. 竞争性分析。**应该知道人们看重产品提供的哪些价值，同样也应该知道竞争对手网站有哪些地方吸引和排斥用户。这类小组常常是匿名进行(不提及委托客户的名字)，想了解人们和竞争对手有什么关系，人们认为竞争对手的用户体验哪些方面有价值，以及哪里不能满足他们的需求和愿望。例如，有一家在线技术新闻网站的竞争性焦点小组，小组表明人们觉得不是明确新闻的大部分内容都纯属多余，大多数人一天只访问一个新闻网站几分钟，他们看重的是每日更新和到其他网站内容的链接。人们觉得观点、深度背景新闻和旧新闻专题收藏没有价值。

**4. 趋势解释。**如果发现了行为趋势，无论是通过调查答复和客户服务反馈意见，还是日志分析，常常都很难确定众多潜在原因中哪些是主要原因。焦点小组有助于通过调查用户的动机和期望来解释行为。这类焦点小组一般作为重新设计周期的一部分或者针对特定问题。例如，有一项调查显示使用搜索服务时，人们会使用两三个不同的服务，而没有特定模式，但无论使用过什么其他搜索服务，也无论使用过多少其他搜索服务，大多数人都使用过雅虎搜索。参加过调查的人组成了焦点小组，结果发现他们



首先会使用雅虎，因为他们觉得雅虎没有包含很多内容，而且即使雅虎包含很多内容，其结果质量明显高于竞争对手。因此，这就是为什么他们总是首先使用雅虎，其目的就是为了以防万一。

## 如何实施焦点小组

焦点小组开始之前，应该要确定以下几件事情。

- 日程表。如果有足够时间来审查执行情况，会产生最佳结果。好的日程表(表 9.1)能为一切事情提供足够时间，特别是招募和指南写作，同时也有足够灵活性允许犯一两个错。

时间选择	行 动
t-2 周	确定受众及范围；马上开始招募。
t-2 周	确定要调查的广泛主题；开始编写指南。
t-1 周	编写初版讨论指南；与开发团队讨论准确的主题措辞；检查招募工作。
t-3 天	选择时间编写第二版讨论指南；与开发团队讨论；应该完成招募。
t-2 天	完成编写指南；安排时间进行排练；设置并检查所有设备。
t-1 天	上午进行排练；检查时间，并根据需要调整指南问题；做最后招募检查。
t	实施焦点小组(通常为 1~3 天内，取决于日程表)；与观察者进行讨论；收集所有笔记的原件。
t+1 天	放松一下，做点其他事情。
t+3 天	观看所有录像带；做笔记。
t+1 周	整合笔记；编写分析。

- 目标受众。即受邀参加测试的人。具体而言，你需要知道能给你最佳反馈的目标受众子集合。
- 研究范围。焦点小组系列可以有几组，每组只有几个人，也可以有十几组，每组有十个人或者更多人。小组和人的数量取决于问题的复杂性、要探索的答案深度以及要知道的答案的确定性。





- 研究主题。并非所有小组都对讨论所有主题有好感，也并非主题都适合小组讨论。仔细选择主题、深思熟虑编写讨论指南有助于从小组讨论过程中获得最大信息，而不会牺牲研究深度和结果清楚性。

## 确定日程表

焦点小组系列的典型日程表从头到尾需要花约三周时间，应该要提供足够时间用于招募和编写讨论指南。

## 挑选目标受众

参加者是否愿意一起讨论主题的原因不是他们之间的实际区别，而是他们是否认为每个人都不相同。

——大卫·L. 摩根，《定性研究焦点小组》，第47页

对焦点小组而言，也许和本书提到的其他任何方法相比较而言，选择合适的受众更是关键。要让人们感觉舒适，谈论他们的经验和价值观，他们需要知道他们不会受到同组中其他人的评价。因此，与大多数其他类型研究不一样，焦点小组集中在同质受众。

定义“同质”取决于研究和小组所处的情境。参加者不必在所有方面都一样，只要在对于研究中需要的方面一样就可以了。例如，几代人之间的差异会导致小组无法讨论自己的音乐品味。

应该从理想目标受众中选择可能提供最有价值的反馈的一个或者数个子集合。合适的小组会根据不同情况而变化：首先，需要目标受众的确定背景资料，资料要透彻到能了解目标受众的人口、互联网使用和构成情况。例如，如果只是要了解现有用户认为你的服务哪些地方有价值，就要选择代表实际受众的最大子集合，但如果要了解为什么有那么多人用户发生转移，就必须更具体：必须只关注符合标准的，且只使用过网站一次而再没有使用的人。

通常情况下，可以选择几组来广泛地了解体验，并且可以作为关注特



定群体体验的方式。例如，如果医生应该使用你的系统，但你发现只有医学院学生在使用，很可能就需要跟这两个群体进行交谈。医生会告诉你他们为什么不用，而医学院学生会告诉你他们为什么用，也许你需要两个单独的网站。

从子小组成员的相似角度出发来定义小组，如果觉得有些小组的成员彼此都没有好感，就不要把他们放在一起。从人口和互联网使用术语着手，对于人们在小组中如何互动，如何应对某一特定用户体验而言，收入、种族、性别、阶层、年龄、工作和计算机经验都会产生影响。为了能建立可以顺畅讨论主题的小组，必须考虑这些差异。定义基本背景资料之后，要进一步根据行为划分群体：谁做了什么？他们还做了什么？有些区别会影响一个群体，而不会影响另一个群体。

以下是一些例子。

- 选择子小组研究汽车购买体验，结果发现在混合小组中男性比女性谈论得多。研究人员对两个群体的观点都感兴趣，因此决定单独创建男性组，女性组和混合组。
- 研究如何扩大技术新闻服务的客户群时，研究人员决定除了符合基本人口统计和互联网使用情况外，还要知道受众应该每天花几小时阅读新闻。
- 有一家医疗服务网站确定尽管医生和医生管理者可能会使用相同服务，但由于两者在医院等级制度中的角色，两者在一起可能会互相觉得不舒服。因此，根据医生在管理工作上花的时间分离了小组。
- 有一家为老年计算机用户提供服务的网站认为，与居住在美国其他地方的相同背景资料的人相比，居住在旧金山湾区的人在互联网上会透露过多信息。调查是在华盛顿特区和密苏里州圣路易斯市进行的。

应该用自己觉得舒服的方式定义子群体。但不要根据子群体的观点或者偏好来定义子群体。根据所述价值观预筛选参加者可能会给小组带来不必要的偏见，还可能会搞砸焦点小组的核心目的，即确定价值观。例如，





不要筛选出“喜欢网上购物”的人，而关注网上购物的人，然后再找出这些人购物的全部原因。

### 招募

选好目标受众(或者受众)，接下来便是找到他们并发出邀请。为焦点小组招募，就像为其他研究招募一样，要尽早进行，并且一旦挑选好受众就要尽快开始。

有几件事能把第 6 章所叙述的招募和招募一组人区分开来，而第 6 章关注的是招募个人。

- 参加者的准确背景资料非常重要。有这么多研究，可能招募到接近目标受众、但并非理想目标受众的人。对于焦点小组，某一小组中的所有成员都应该符合受众的背景资料，因为只要有一个人不符合受众的背景资料就会破坏整个讨论。例如，小商店业主与大型厂商或者连锁店业主在一起讨论问题可能会觉得不舒服。虽然优秀主持人有时能克服这些问题，但最好首先避免这种情况。
- 不要招募互相认识的人。如果有两个人或者更多人互相认识——有时即使他们只是知道其他人是谁——都会造成小组讨论产生动态不平衡。两个人互相认识，抬高眉毛或者一声叹气就能代表很多东西，会破坏讨论，并无意中隐瞒重要信息。因此，如果招募时采用朋友和家人的方法，应该避免招募通过一个人推荐的人，如果采用数据库进行招募，应该选择最初在不同时间或者通过不同方法招募到的人。
- 避免招募了解焦点小组如何工作的人。焦点小组的进度取决于坦率回答和讨论小组中的舒适程度。焦点小组老手(常常是下意识)会给出他们觉得应该给出的答案，或者他们可能会预测讨论的下一个主题。虽然不会有致命效果，但这可能会扭曲讨论，导致由此产生的额外主持和分析工作。如果不得不招募有焦点小组经验的人，必须确保这种经验既不是频繁发生也不是最近发生，出于同样原因，永远不要招募来自接触过大量焦点小组的公司的人，即使间接接触也不行。因此，不要招募来自营销公司和广告机构





等的人。

- 筛选掉对于任意主题明显比其他人有更多了解的人。如果小组中有一位“专家”，这位专家的知识会吓到其他参加者，其意见会左右整个小组的观点。

选好目标受众，接下来为招募对象编写一份背景资料。

以下内容可以作为家居装饰网站的潜在用户背景资料。背景资料定义了积极思考家居装饰并且愿意在网上查找信息的一组人。

#### 焦点小组招募背景资料样本

##### 人口统计

年龄 20~55 岁。

如果达到行为标准，收入不重要，如果未达到，家庭收入需要超过 10 万美元。

性别不重要，只要是主要决定人就可以。

##### 互联网使用

在家或者上班时使用个人电脑。

在家或者上班时使用互联网。

有超过一年的互联网经验。

每周花 5~10 小时使用互联网用于个人任务：

- 购物
- 比较产品或者服务
- 收集主题信息(比如 Sidewalk, BabyCenter, ImproveNet 和 CarPoint)

##### 行为

在过去 9~12 个月中进行了家庭装修，或者要在将来 3 个月进行装修。

完全装修至少要花 2 万美元(可以是未来要装修的估计值)。

## 定义范围

焦点小组必须要小到人人都有机会能分享见识，但也要大到能提供多样化的认识。

——理查德·A. 克鲁格，《焦点小组》，第 17 页





如果要定义所进行的焦点小组的范围，需要确定两件事情：需要多少小组，每组需要多少人。

不要因为同样原因只进行一个焦点小组，也不要只根据一个人的话就对一类人做出结论：也许有些因素对小组很重要，但无法代表受众整体的意见。也就是说，一般情况下，四个焦点小组就足够了。第一组本质上属于彩排。到第三组，应该就会看到对前两组的意见和陈述的确认（有时是逐字逐句）。第四组应该用来确认前三组的趋势。如果到第四组还是有許多新信息（尤其是意见分歧），可能说明需要进一步的焦点小组，或者说明小组构成过于宽泛。

同样，多数营销教科书都建议每组 8~12 人，但如果要检查用户体验和态度，一般情况下，人数少一点会更好。用户体验研究所需细节的详细程度常常大于典型的营销焦点小组。缩小小组规模可以更深入地了解每个人。6~8 人的规模很好，可以平衡细节收集与观点的范围。如果想要深入了解每位参加者，或者如果单个焦点小组的时间很短，可能需要把数量降到 4 人。如果少于 4 人，讨论会更像访谈，无法产生更多人时的动态情况。

### 选择主题

对普通焦点小组而言，要调查 3~5 个主要主题。这些主题需要从项目整体准备措辞。“了解人们在研究保险时使用的心智模型”可能是保险经纪网站的目标，而推荐家庭建筑承包商的服务会关心“修理住房时，人们什么时候会寻找外包服务”。竞争性分析时，高端在线拍卖网站会有两个目标，即“了解什么原因会导致高档轿车卖家不在 eBay 上出售汽车”，或者“揭开哪些因素能帮助我们的产品创意能被苏富比(Sotheby's)<sup>①</sup>一类的公司看见”。

要足够关注这些目标，小组可以花约 10 分钟充分讨论每个问题。这些目标不要太关注于你会向参加者提出的实际问题，也不要变成其他方法

#### 提示：

并非所有人谈论所有主题都愿意。要选择目标受众愿意讨论的问题。同一产品的不同受众对同一主题可能有不同反应。

① 全球历史最悠久的艺术品和收藏品拍卖行。



(例如调查)能获得更好答案的问题。“列出竞争对手的清单”一般都太宽泛，很可能用调查得到更好的答案，而如果把目标设为“发现哪些因素会导致竞争对手的用户体验比我们的更有吸引力”可能更为合适。

## 编写指南

讨论指南是主持人需要遵循的剧本。它为焦点小组创建了一致框架和日程表。根据大致相同情境下的相同顺序，向小组提问，确保所有主题都得到充分讨论。这样讨论就会挖掘参加者的微妙意见，而同时也不会偏离任何主题。

先思考核心和问题，再编写完整指南。对于焦点小组的问题，应该做到以下几点。

- **仔细安排顺序。**问题会导致参加者形成一定的思维框架来思考一些问题，并会记住某些事件。仔细安排问题顺序能利用这种思维框架，小组讨论流程能变得更加“自然”，而反过来又能帮助参加者保持创造性思维，产生更好的见解。一般情况下，问题应该从最一般到最具体，每个问题都要能缩小一点讨论范围，并集中到之前讨论的子问题集。除非要引入全新主题，否则应从一般问题开始讨论，主题之间应该过渡自然。
- **无引导性。**正如第6章所述，问题不要暗示答案或者价值判断。应该把问题侧重于让参加者表达自己的想法和价值观。例如，“您认为谁的搜索服务更好，谷歌还是莱科思(Lycos)”就假定了参加者会认为一项服务比另一项服务有优势。相反，应该中立地提出此问题，“您使用谷歌搜索服务，有没有喜欢的地方？使用莱科思呢？都喜欢哪些地方？您有没有比较方法？如何比较？”
- **保持开放性。**问题不应该将答案限制成固定答复。问题应该鼓励人们保持开放心态并分享经验。回答较长，能说出更长的故事，而且比短的回答会少一些模糊。因此，不要问“您最喜爱哪个食谱网站？”，而可以问“如果要搜索在线食谱，你会找什么？”
- **关注细节。**问题应该鼓励参加者尽量具体回答问题。理查德·A.





克鲁格在其著作《焦点小组》中建议把“为什么”的问题分解成多个“什么”的问题，明确询问参与者做决定的原因以及影响因素。例如，“哪些因素使你决定在网上购买餐叉？”和“哪些因素让您选择这家网站？”相对于“您为什么选择这家网站？”能提供更好的看法。

- **个性化。**人们常常会试图向广大公众或者某些假定的受众推广他们的经验，而他们并不是那部分受众。因为想知道个人意见、价值观和经验，就应该强调个人经验，应该格式化问题，让问题关注人们当前的行为和观点，而不要提出选项来设计他们的经验。因此，“如果现在不得不重新装修厨房，您会根据哪些特性来挑一家住房承包商？”胜于“您认为哪些特性有用？”

- **清晰。**意思应该尽可能的清楚，特别是向参加者介绍新术语时。

当然，所有问题都要符合所有以上标准常常很困难(编写的问题既要具体又同时要保持开放性特别具有挑战性)，但要在脑中把它们当作指南记住，而且应该尽量遵守。

## 讨论指南模板

指南分成三个主要部分：介绍、主要讨论和总结。

以下指南摘自一家在线新闻网站的焦点小组，该网站(主要)想了解其现有用户挑选新闻网站的标准。

暖场及介绍(10 分钟)

介绍阶段为讨论定下基调，打破参加者之间的冷场，并解释焦点小组过程。

大家好。欢迎大家来参加讨论。

我是\_\_\_\_\_ (名字)。我是\_\_\_\_\_ (公司)的研究员，他们请我帮他们了解一下大家对他们的产品和概念有什么想法和观点。今天讨论的任何产品开发和我不相关，我对这些产品也不会有感情上的关系，所以大家可以畅所欲言。



大家都阅读了大量在线新闻，花了很多时间上网，所以我们邀请大家来参加讨论。我们今天要做的事情就是谈谈上网体验，这样公司创建出的服务才能最符合像大家这样的用户的体验。

告诉人们如何选中了他们，他们会互相有好感。让他们知道研究的最终目标有助于他们关注自己的想法。

讨论方式是随意聊天，但我会请大家关注一些情况。

讨论的时候，请大家尽量坦诚讨论，不管您说什么，都不会伤害任何人的感情，所以请大家无所顾忌准确说出感受。

我们需要大家的观点。观点没有对错——特别是针对我们讨论的东西——只是观点，所以如果您不同意其他人的说法，我们也想听听。

但请大家一个一个发言。

提供明确的对话规则，明确说明可以有不同意见，这样能鼓励人们说出自己的想法。

今天会讨论很多问题，没人想待上几小时，我必须请大家进行总结，便于讨论能继续进行。

如果尽早建立期望，人们就不会觉得受到干扰。

大家能想得到，玻璃墙后面有几位公司的人，我们讨论的是他们的想法，\_\_\_\_\_（主持助理的名字）会和我在一起主持讨论。  
\_\_\_\_\_可能会进来给大家一些需要的东西，可不用理睬。

明确说明玻璃墙(假如屋里有的话)能了解参加者对玻璃墙的焦虑感觉，有助于缓解焦虑感。有时甚至可以在介绍阶段请参加者对镜子后面的人挥手或者做鬼脸。但只要提一次就够了。

前面说了，邀请大家的目的是想听听大家的想法，无论您说什么都不会伤害到任何人。如果大家想知道，这有一台录像机，所以\_\_\_\_\_  
和我就不用坐着紧张记笔记，就可以专心听大家说。产品开发团队成员可能会看录像，但不会公开，也不会用于推广或者传播目的。





如果要对焦点小组录像，应该说清楚。

现在我要给大家阅读一下保密协议。我对每位访谈人都要读这个。保密协议会说明大家参加此类研究的权利。

作为参加此次研究的人，拥有以下权利。

- 可以随时终止。
- 可以随时提问。
- 可以随时离开。
- 不会受到欺诈。
- 回答将会保密。

这里有张表，需要您允许我们对讨论进行录像并用于研究。

应该告知参加者的权利，并获得录像许可。不告知他们显得不道德，在某些情况下还会触犯法律。但介绍这些内容时不要搞得很严肃，太严肃会造成紧张气氛。我通常会对要说的内容开开玩笑，(指着点心盘子)说这是一场秘密点心口味测试。

大家有没有问题？下面就开始！

现在请大家介绍一下自己。请大家说下面四点。

- 姓名
- 居住的城市
- 每周必看的电视节目或者出版物
- 互联网上觉得讨厌的事

(\_\_\_\_\_首先介绍自己，其他每个人轮流介绍)。

自我介绍能让大家互相认识，并对个人问题打破沉默，比如最喜欢看的电视节目或者小毛病。这些问题可以很具体，直接，可以与要讨论的问题无关。

### 一般新闻阅读(20 分钟)

您怎么定期获取新闻？

探询：您有没有定期获取新闻的其他办法？



您定期阅读哪些新闻网站？(在白板上列出清单)

探询：与其他网站相比，哪些网站更经常阅读？

探询：您更常阅读的网站有哪些地方吸引您？(向个人询问具体网站)

您有没有从一个最喜欢的网站更换到另一个？有多频繁(只换过一次还是有几个最喜欢的网站)？

探询：更换的原因是什么？

探询：您如何选择其他网站？

探询是挖掘特定主题的跟进问题。

更换后的网站上有什么新闻？(列张清单)。

有没有好的体育网站？政治网站？科技网站？

即使主题是关于科技新闻，问问其他类型的新闻可以防止讨论过早关注科技新闻网站。

有没有网站专注一类新闻(商业、娱乐以及科技)？如果您定期阅读，能说出几个吗？

线索：能说个例子吗？

优先级别练习(20分钟)

分发纸张，进行下列练习：

根据重要性，请对下列情况打分。从1-5分，1分代表最不重要。

如果您能想到有个网站胜于其他网站，请在旁边写出名称或者网址。不是每个情况都要有一个不同网站，有些情况可能没有任何网站。

- 给定主题的不同新闻数量
- 覆盖不同主题的数量
- 页面下载的速度
- 报道新闻的人
- 每条新闻覆盖的程度
- 定期专栏
- 网站的搜索引擎质量
- 网站视觉外观





- 新闻发生后报道的速度
- 网站的导航质量

(四处问问)比如您认为哪个最重要。

探询：为什么它这么重要？

探询：有没有做得好的网站？有没有其他人也写下了那个网站？

有没有您觉得新闻网站重要的属性，但没有列在清单中？

此类练习关注构成新闻网站具体特性，用于开始讨论。

### 竞争对手网站检查(20 分钟)

我们已经讨论过了一般新闻阅读，接下来要谈谈几个具体新闻网站。

(打开幻灯机)

有没有谁熟悉\_\_\_\_\_ (竞争对手网站)？

探询(对熟悉的人)：您阅读过吗？频率？与其他新闻网站相比，您是在什么情况下阅读的？

请描述一下\_\_\_\_\_的个性化。

\_\_\_\_\_是否属于具体类型吗？

您期望在那里看到什么新闻？

探询：与其他网站相比，有没有情况您觉得会阅读？

探询：有没有科技新闻，您不会去那阅读？

探询：专栏呢？有没有专栏会让您定期访问？

(对另一家竞争对手网站提出相同问题，对每组给出不同的竞争对手网站顺序)

### 发散及总结(10 分钟)

讨论快结束了。我想请大家来一次快速头脑风暴练习。问题只有一个，但我请各位都先思考一下，再说想法。

不考虑实际情况，您希望新闻网站服务能做什么，而现在的网站都没提供？

(等 30 秒钟时间，四处问问进行讨论，然后把想法写在白板上)

好了，现在是关于新闻的最后几个问题。





看得出来，大家讨论网上新闻都很兴奋，但必须要结束了。如果在回家的路上您想到了什么就请告诉我们，随时可以发送到这个电子邮件地址（在白板上写出电子邮件地址）。

最后一个问题：有没有什么地方我们能做得更好，不管是时间安排上还是实施上？我们是不是应该提供不同食物和啤酒，还是其他任何东西？

确定网站内容，接下来的子焦点小组就可以关注对具体特性的愿望，以及人们在日常生活中如何使用新闻网站。

编写好指南之后，要进行检查并安排时间。产品开发团队成员是优秀的指南评审人，尤其是产品和项目经理。他们能指出技术上模糊的地方，能提供另一个角度的观点，还能为其他团队观察焦点小组做好准备。

应该对指南进行测试。测试指南有个简单方法，找几个没看过指南的人（但不一定要实际小组的人那么多），然后进行走查，注意他们如何回答问题以及时间选择的准确性。此外，把第一小组当场彩排，来审查指南的有效性，并适当调整。

## 实施小组

### 物理布局

实施焦点小组的地方要有一个舒适的大房间，通风好，室温宜人，减少会分散注意力的东西。典型设置看起来像是会议室或者客厅，取决于参与者觉得是在舒适的工作环境中还是家里。实际的会议室或者休息室很容易转变为焦点小组设施。如果要在会议室举行焦点小组，而会议室的玻璃又对着内部办公室，此时要把窗户遮起来，或者让焦点小组尽量远离窗户。窗户如果对着外面，没关系，除非是在一楼，在一楼人们会受到太多干扰。焦点小组开始后，除助理主持人之外的每个人都要进入房间（为达到这种效果应该在大门贴个标志）。我曾经观察过的一个焦点小组就受到了给观察员送比萨的打扰。经过一些混乱，大家知道了比萨是送给观察员的。虽然参加者只是开开玩笑，但造成了不好的异动。

饮食属于非正式活动，能打破小组刚刚形成时的紧张气氛，因此，在可能的情况下应该提供食物和饮料。对于晚间的焦点小组，人们可能吃过





了晚餐，提供高能量点心(如饼干)常常会造成困乏。确保要提供素食和大量水。录音时，干脆的食物吃起来声音很大，因此，最好不要提供芹菜和土豆片之类的食物，应该准备大量的非碳酸饮料，人们喝碳酸饮料会容易停止评论，因为他们害怕说话时会打嗝。

### 如何创建座位顺序

克鲁格推荐采用下列步骤来安排人们的座位。

- 把座位和参加者的名字用会议桌牌的形式打印出来。把桌牌放在桌子上。每位参加者的名字印在两侧，字要足够大，从观察室和通过摄影机都能看到(1.5 英寸、黑色字母就很好)。
- 去等候室跟人们谈谈，看看哪些人很安静，哪些人喜欢说话。
- 想好每个人的座位安排(啰嗦的人坐在主持人旁边，这样他们会打破眼神接触，而不会破坏小组动态；安静的人交叉安排，便于主持人容易眼神接触并征求他们的意见)。
- 根据顺序排放桌牌。
- 把人们带进房间，看似随机的把座位桌牌放到桌子上，但实际是根据预先设计好的顺序。

房间里应该有办法让观察员看到过程。可以采用传统的双向玻璃镜墙，镜子后面是一间隔音室，观察员可以坐在那里。观察员的房间应该自有入口，便于观察员不用通过讨论室就可以进出。如果找不到这类房间，闭路电视是一种廉价和容易的解决办法。一般可以在附近房间用长视频电缆加上电视，只要观察室和讨论室互相不干扰即可。

只要有可能，应该录制焦点小组。录像很容易看到谁在说什么，并能看到肢体语言，肢体语言有时会暴露出对概念的否定或者说明对主题感到不舒服。如果录制过程不引人注目，会很快消失在交谈背景中。

焦点小组只需要一个摄像头，放在主持人身后就行，声音要好。如果有好的广角镜头，就不需要有摄像操作员，从而降低过程花费，还能减少有个人不停从桌子一边到另一边摆动相机所带来的麻烦和惊吓。

一般情况下，两个心脏形麦克风足以捕捉到所有参加者的评论，同时



不会有大量多余噪声。心脏形麦克风指有方向的麦克风，能捕获前方广泛领域的声音以及较小面积的背后区域。因此，不同于连接到多数摄像机的全向麦克风，心脏形麦克风可以远离噪音来源，例如窗户外面和通风口(见图 9.1)。

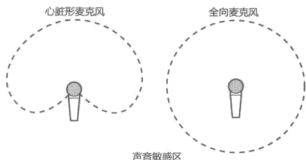


图 9.1 心脏形麦克风与全向麦克风的敏感度比较

如果是临时设置，可以在桌上放两个相对的麦克风，但容易受到振动干扰(铅笔敲打桌子，咖啡杯放到桌子上或者电脑的嗡嗡声听起来比在房间里更大)。在桌子和麦克风支架之间放块布或鼠标垫能降低音量。如果是永久设置，两个挂起来的麦克风可以最大程度地消除来自上方的环境噪声(例如空调和日光灯变压器)，同时还能很好地录制对话(见图 9.2 和图 9.3)。

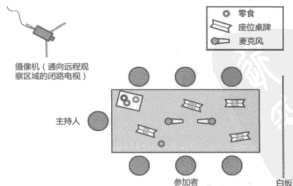


图 9.2 会议室风格的临时焦点小组布局



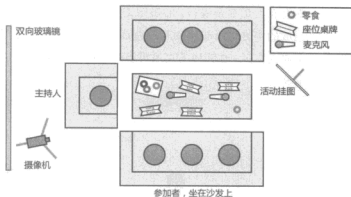


图 9.3 客厅风格的永久焦点小组布局

### 主持人

主持焦点小组是一种技能。主持人必须平衡小组中参加者的需求，让他们轻松自如地讨论和研究需求。要想不带任何偏见地主持焦点小组，需要实践、能力和正确心态。

主持人必须有的基本技能包括尊重参加者、仔细聆听以及快速思考的能力。通常情况下，主持人必须能预测下一个对话方向，要么能朝着那个方向驱动会话，要么能把对话引向更理想的方向，而参加者又不会觉得有人在进行主持。这会涉及主持人所说内容的许多微妙线索，主持人使用的语气，甚至是其肢体语言。

更具体来说，主持人必须做到以下几点。

- **始终保持控制。**大多数情况下，主持人可以通过身体语言和口头强调来保持控制，巧妙引导讨论围绕着某些参加者和主题。但如果讨论偏离到一个方向太久时，主持人要毫不犹豫地施加更多控制，并将讨论重新集中于必要的主题。例如，讨论新闻网站时，如果讨论偏离到当前头条新闻，主持人应该调整讨论来关注网站，而不是内容。
- **总是向前发展。**主持人应该监测会话进程，并在合适时间介绍主题，实现自然过渡，而不要停止讨论或者突然改变话题来控制进



程。这样讨论不会变成访谈，参加者能一个接一个回答主持人提出的问题。

- **不要判断。**主持人充当调停人帮助参加者表达意见，而不要扼杀他们的急切心情。因此，主持人不应该表达自己的意见，而应该想方设法让小组形成意见。有时这需要在维持文明交谈中学会的抑制习惯。例如，不管同意不同意，许多人会在其他人说话时点头以示鼓励。这种方式能表达同意，并认可表达的观点。由于主持人是“官员”，参加者可能会觉得这是官方对某种立场的认可，因此，可能会在表示反对意见觉得不舒服。
- **尊重他人。**在任何时候，主持人都必须充分尊重参加者，即使不同意他们的说法。焦点小组中每一位参加者都能有对开发团队有用的观点，即使不符合团队的实际情况。要让所有参加者都可以轻松说出想法，他们需要知道主持人对待他们的投入跟对待其他人的投入一样。采用不带判断的陈述以及严格控制身体语言能实现这一点，但最好通过诚实信念来沟通，即每个人的发言都必须被听到。
- **做好准备。**主持人要足够了解要讨论的主题，以便能够针对参加者的陈述提出具体问题。并不是说要获取和主题相关的专业知识，但主持人应该很好理解一般概念、术语和讨论所涉及问题的含义。主持人还应该对小组成员的习惯和环境有所了解。

#### 提示：

如果没有助理主持人，主持参加者、管理观察员以及分析最后结果的负担全落在主持人一个人身上。这种情况下，应该要让主持人可以离开参加者5~10分钟，回到观察区和观察员一起审查讨论，来帮助确定他们讨论的框架。

此外，主持人应该有良好的时间感(知道什么时候从一位参加者转移到另一位参加者，从一个主题转到另一个主题)、良好的短期记忆力(指针对人们的先前发言，用自己的语言来描述概念)以及随机应变的能力。

#### 助理主持人

没有助理主持人，也完全可以开展焦点小组，但如果有的话，能有助于过程的开展。

助理主持人是关键的分析人员，他们是联系焦点小组和外部世界的纽带。助理主持人照顾焦点小组参加者的需求，并收集信息，让主持人能保





持关注富有成果的讨论。

开始讨论之前，参加者第一次来时，助理主持人应该迎接他们，把他们带进讨论室，给他们看初始文件(如保密参加协议)，并给他们带茶点。随着讨论进行，助理主持人能带观察员的记录并处理他们的需求。

讨论过程中，助理主持人应该对有趣的讨论部分做大量笔记(如关键引言，参加者提出的问题，自己的观察)，并管理观察员讨论(也要记笔记)。焦点小组结束后，助理主持人可以使用这些笔记向观察员和主持人进行汇报。

### 主持讨论

主持讨论是一个在参加者的舒服程度和对研究产生有用信息之间取得平衡的过程。主持讨论很少有严格规则，因为每个组、每个话题都需要采用不同方法，不同主持人有不同风格，但都能发挥作用。

但对多数用户体验焦点小组而言，有如下共同指南，多数主持人也都会遵守。

- **提前花点时间和小组谈谈。**小组讨论前花5分钟和参加者随意聊聊，有助于了解参加者的个性。可以猜测谁很安静，谁会主宰讨论，谁很顽固，谁会搞讲座，等等。这有助于准备小组发展过程。助理主持人也可以和他们聊聊，并在小组开始讨论之前向主持人简要介绍一些参加者的个性。
- **遵守指南，但必要时应该抓住机会。**如果小组在讨论本应该过一会再讨论的问题，但参加者谈得都很好，那就谈这个问题。可以提出本来安排稍微晚会提出的问题(除非问题顺序绝对至关重要，否则可以请焦点小组等一会再讨论这个问题)，讨论完这个问题后，再把讨论拉回到之前的问题。
- **让每个人都参与讨论。**留意失去兴趣的人或者看起来发言不舒服的人。尽量找出造成他们厌烦或者不适的原因，并尽量缓解。有时只要喊某人的名字，并明确请其说出想法即可。但如果需要大量精力来让某人参加讨论对整体讨论效果有害，最好把精力花在



小组的其他成员上。

- **避免引入新术语和概念。**一旦引入新词和想法，往往会给讨论加上框架，并且会主导讨论。如果小组想用自己的话来描述概念，或者想建立自己对想法的模型，让他们这么做。即使这些术语应该完全可以听得懂，也不要使用行业术语，除非小组也这么使用。例如，“什么因素会影响您选择电子商务门户网站？”可能会产生混淆，因为大多数人对因素的理解不一样，“电子商务门户网站”也可能是个新词。
- **限制身体语言。**有很多身体动作会微妙表达出对某人观点或者某个特定角度的偏好，人们非常擅长发现这些动作，即使是不自觉发现。点头是向其他参加者发出观点“正确”的信号，退后是发出观点错误的信号，一直靠向某位参加者就说明那位参加者的评论更重要(甚至那位参加者说得很平静，你也会这么做)。你可能需要看看自己主持的录像带，看看自然身体语言会表达什么意思，尽量减少会表达偏好的行为。
- **澄清观点。**如果有人说得不清楚，或者和先前的陈述相违背，应该重申想法或者请那人说清楚，请那人举个例子来说明其使用的方法或者定义。
- **重申想法。**为了确保每个人(包括主持人和观察员)都能确切明白参加者说的意思，常常应该重申刚说的话，看看是否正确代表参加者的想法。如果听到不同的话，应该说清楚想法的核心内容。这种方法还强调主持人对小组的关系，巧妙向参加者表明他们虽然是专家，但主持人是控制人。
- **深挖其他观点。**如果有人说的话违背了先前的陈述，或者如果有人对正在讨论的内容做出一种新解释，应该彻底调查小组中的赞同意见和反对意见。让那个人说完他的想法，然后请小组其他人自愿说出想法或者体验，特别是同一话题的不同体验或者想法。
- **不要主导讨论。**尽可能让小组进行讨论。如果有参加者提出意见，可以给小组中的其他人一些时间(克鲁格建议五秒)来回应意见，





然后再继续提问。

- **提供思考时间。**不是每个人都可以立即阐明想法或者立即能理解概念。在讨论开始时，可以明确提供时间来思考或者记忆，来帮助参加者参加讨论。5~10秒可以满足一般需要，这些时间可以用在提问时，比如“您知道最近一次订购备件的时间吗？请想一下是因为什么情况而需要订购备件。”
- **适当时候来点幽默。**不要嘲笑参加者的话题，也不要嘲笑他们的陈述，但可以来点幽默。不要幽默过头——幽默过头看起来会很假，而且会造成人们认为你在娱乐他们的期望——但如果有些地方很好玩，笑出来也没关系，不要害怕说这件事情很好玩，要小心采用自谦的幽默方式。人们看重谦逊，但只有在不破坏主题或者情况的时候。
- **保持能量。**有时会发生小组闲聊或者对特定主题不感兴趣的情况。保持(和激励)高能量水平可以防止小组感到厌倦和沮丧。
- **让人们休息。**在房间里坐两个小时激烈讨论会很困难且枯燥。如果焦点小组进行超过90分钟，要让人们中途休息5分钟，好让他们收集想法，去交停车费，去洗手间，等等。

最后，放轻松，微笑一下，应该真诚地感受人们的体验。能有这样的机会分享人们的生活，好好享受一下。

### 提问

提问过程属于主持的一部分，但提问涉及的要素非常多，我觉得需要有专门章节来讨论。有几个方法有助于提问并澄清对问题的答复。

- **安排事项的优先级别。**如果有问题清单，就要请参加者明确优先级别，可以对最重要的事情进行投票(走查清单时随机投票！)，也可以给每个条目自发打分(每个人都对条目举手打分，主持人能迅速得到小组评级的平均值)。
- **写下意见，再进行讨论。**如果写下了，人们不太会改变想法，也不会在讨论时“随波逐流”。例如，在纸上首先给一系列主题评





分的参加者会更愿意参加其他组别的评级。

- **在可能的情况下，使用参加者的准确想法和言词。**如果已经向小组引入术语，或者定义某些概念，要修改跟进问题以使用相同术语和概念。
- **关键问题需要口头询问。**对于次要问题，使用和小组一样的术语感觉更自然，但主要概念应该用同样方式介绍给所有小组成员，以减少分析结果时的疑问。
- **准备好例子。**参加者并不总是明白问题，特别是对于新概念或者不熟悉的术语。应该准备一些好例子，在觉得需要的时候用来解释问题。首先要尽量不使用例子就获得反馈，但如果人们不知如何答复，就应该说出例子。例如，如果小组不知道从哪开始对特性进行头脑风暴，应该告诉他们什么例子可以考虑成新特性（“网站会把您喜欢的促销商品信息用电子邮件发给您”）。

#### 特性购买练习

本特利学院设计和可用性测试中心主任昌西·威尔逊对特性的优先级别提出了以下方法：给人们一张特性“菜单”，每个特性都有描述及根据估算每个特性开发需要花费的“价格”，以虚拟美元数额（100 美元的东西相对而言比较容易购买，而 500 美元的东西需要很多考虑）。请每个人检查特性列表，并“购买”特性，每个特性最多 1000 美元，把它们写在一张纸上。所有参加者都做出选择之后，可以四处问问，并讨论一下为什么人们做出这些选择。

#### 常见问题

许多焦点小组都会发生以下问题。

- **集体思维。**在小团体中，人们会受到限制（也称为易于，不同理论叫法不同），以避免与他人的冲突。避免冲突的最佳方式之一是根据小组的明显一致意见而趋于一致。因此，人们倾向于赞同房间里的人，即使他们实际并不同意，因为“归属”在感觉上比





“正确”更重要。前面提到的很多方法都是用来尽量减少集体思考，但仍会发生。要注意这种情况，并尽量减少其发生，要向参加者清楚说明不仅允许分歧，而且还鼓励分歧。

- **错误假设。**有些情况下，开发团队和主持人都会从根本上误解用户体验的性质，造成整个焦点小组都基于不正确的假设。避免这种情况的最佳方式是提前访谈一位目标受众代表。如果焦点小组已经开始，并发现了这种情况，要尽快改写指南，而且如果有必要，再安排新的参加者。
- **有人迟到。**除非当前仍在进行介绍，否则焦点小组一旦开始，就不要再加入，给他们全部酬金并道别。让迟到的人跟上进度会有破坏性作用，而且出现“陌生人”会让人们产生防备心理，会抵消之前的破冰效果。
- **群体沉默。**有时一群人就是不想说话。造成这种情况的原因会随组变化，但有时你会发现自己在在一个房间里，而一屋子的人都没兴趣讨论话题。根据情况不同，需要重新的讨论带入活力，并随时调整指南和时间表。但有时如果焦点小组根本没在讨论话题，就必须采取果断措施。此时需要主持人的创造性，主持人必须想办法让参加者参与讨论(例如，把讨论环节变成开放式的头脑风暴，或者转换到小组更感兴趣的问题上)。
- **参加者不发言。**人们不说话的原因多种多样。他们可能觉得对主题没好感，觉得他们的意见没有得到分享，或者因为他们没什么可补充的而感到尴尬。也可能只是害羞。如果知道每个人的观点至关重要，就需要让沉默寡言的参加者发言。让人们参与讨论的最简单方式是直接邀请他们发言，但如果邀请过了头就会变成了“呼叫和响应”式的动态讨论。请整个小组补充意见，但要直接看着不发言的参加者，这种信号表明需要他们发表评论，但并没有表达出他们是被挑出来的。非语言信号可以告诉我们谁有想法，即使他们觉得不舒服而没有表达。讲话开始前吸几口气、向桌子靠近、和主持人目光接触、另一位参加者说话时皱了皱眉，



所有这些都表明有想法在形成。

- **参加者过于健谈。**有些人不知道什么时候该闭嘴。这种人一般有两种：一种是有有趣想法但说出来有困难，另一种是没想法但又希望能进行另一个讨论。果断打断、重述参加者的职责可以克服前者的问题，但后者有点麻烦。如果人们谈论时没有明确目的，可以请他们总结一下自己的想法，然后继续往下进行，请另一个人发言或者提出不同问题。如果讨论还是漫无目的，应该礼貌提醒他们需要让房间里的其他人发言。
- **小组优势**(“Alphajerk”[阿尔法混蛋]效应)。只有一位参加者主导或者盛气凌人会毁掉焦点小组的作用。理想情况下，避免发生这种情况的最好方式是在焦点小组开始时请这种人离开。如果在等候室和参加者见面并观察他们，会明显发现这种行为模式。然而，一旦焦点小组开始进行，而且明显有人想“控制”小组，应该远离对其的注意力，而关注整个小组，并强调主持人驱动小组的概念。这种方法和对健谈的人的所采用的方法基本相同，但与针对其他只是啰嗦的人的方法相比，这种方法必须要更坚定使用(而且往往要更频繁使用)。此外，站起来从物理位置上占据讨论空间、走到白板前有助于恢复对讨论的控制。
- **参加者不合格。**人们有时会误解参加标准，或者会错误表达其体验。应该在早期排除这些人，因为他们很可能不能对讨论做出贡献，而且他们的存在往往使其他参加者感到不舒服。在环节早期根据关键筛选标准提出问题，并委婉请不符合标准的参加者离开房间，这样能让小组其他人的整体效率更高。
- **跑题。**小组会有螺旋跑题倾向。这类讨论可能是对价值和想法的讨论，如果对主题很重要，可以允许一段时间的跑题。如果是边际兴趣点，就要尽快结束跑题。主持人如果注意到人们看起来对某个跑题的话题有兴趣——不要打击任何主题——要鼓励小组结束这个主题或者推迟对此主题的讨论。
- **参加者有敌对情绪。**焦点小组中很少存在公开的敌对情绪，但某





些人(和某些想法)会造成敌对情绪。激烈分歧或者激进思想会导致人们增长敌对情绪。此时主持人要保持公正,并关注讨论想法,而不要关注情绪。因此,应该保持中立地对待充满感情内容的陈述。例如,如果有人说其他人的陈述都很愚蠢,应该询问那个人不同意什么,而不要去管“愚蠢”概念。

- **想法激进。**有时候有人说的东西会冒犯大多数其他参加者。可以利用这种情况引发所有参加者分享意见,也可以分化小组。为了指导讨论回到之前的方向,而不是让它使发言人沦为互相排斥的不同“党派”,应该问问是否有其他人分享这一观点。然后再问问是否有人有不同看法,并探讨他们的意见不同在那里。特别重要的是,主持人不能赞同这方或者那方——即使你坚决不同意所述看法——而应该关注事实陈述。
- **主题高度情绪化。**就像主题有争议性和参加者有敌对情绪一样,主题情绪化会很棘手。处理这些主题的最佳方法还是关注主题背后的想法,只要情绪释放,情绪化主题就不会成为唯一的主题。指南应该考虑此问题,如果预期到有情绪化反应,应该给参加者足够的时间来讨论自己的情绪。如果有主题意外变成了情感方向,应该尊重并讨论参加者的情绪,甚至需要破坏指南中的几个问题。但主持人要尽快摆脱这种讨论,同时要尽量了解导致情绪化状态的体验因素。

### 管理观察员

从一点可以往任意方向引出直线,但两点决定一线。

——卡罗琳·施耐德,施耐德咨询公司用户体验顾问及负责人

开发团队应该有尽可能多的人尽可能多的参加焦点小组。开发团队能够快速了解他们自己的想法,分析师能获得宝贵技术知识。出于同样原因,每个主题应该有一个以上的焦点小组参与讨论,如果打算进行观察的话,应该鼓励他们至少观察两个焦点小组。如果看过了几个小组讨论某个主题,就会更容易知道哪些现象属于特定群体,哪些是更普遍现象。



观察员会积极参加分析过程，因此他们应该知道如何处理观察过程。观察焦点小组有一个正确方法，也有很多错误方法。提前和新观察员见面，并请他们为观察焦点小组做好准备。以下指南有助于请观察员做好准备，并可以作为初始工作方向的一部分，也可以作为书面清单的一部分。

#### 焦点小组观察员指南

(1) **仔细听。**很容易立即讨论观察到的情况，但首先一定要听听人们真正说的内容。可以随时讨论你看到的东西，但别忘了要仔细听。

(2) **不要仓促下结论。**把人们的陈述当成指南，了解他们如何思考问题，他们的价值观是什么，但不要把陈述的具体情况当信条。如果小组中每个人都表示喜欢或者讨厌什么东西，并不意味着整个世界都会这么认为，但它清楚表明有足够多的人会这么认为，应该足够引起注意。

(3) **焦点小组没有统计意义。**如果五个人中有四个人说了一些情况，并不意味着 80% 的人都会这么认为。这说明有很多人可能会这么认为，但并不代表整个用户人口比例。什么也不是。只是数字而已。

(4) **焦点小组的参加者都是专家。**参加焦点小组的人知道他们要做什么，也知道他们当前是怎么做的。倾听他们的需求和体验，把他们当成顾问一样，他们能告诉你客户需要什么，而不要把他们当成产品消费者或者销售目标。

(5) **焦点小组不是灵丹妙药。**如果每个焦点小组都有几个好主意，焦点小组就足够有价值，但并不是参加者的每个陈述都应该详细研究。

(6) **随时把问题传给主持人，但不要过于频繁。**可以偶尔向小组提出问题，但一个环节中不要超过两个。应该简洁明了地写出问题，并且应该就好像要告诉主持人一样进行措辞。然后把问题给助理主持人，助理主持人会给主持人。在对话过程适当的时候，主持人会引入问题，但如果时间或者主题不合适，主持人可以决定不引入问题。

(7) **人是矛盾的。**听听人们如何思考主题，他们用什么标准来得出结论，不一定要听他们说出来的具体愿望。人们可能没有意识到两个愿望不可能同时实现，或者人们根本就并不在乎。准备接受偶尔的无聊或混淆。





人们的行为并不总是有意义或者有见地。

(8) **不要贬低人。**参加者有时候会说些情况表明他们没搞清楚。不要假设有人之所以没说什么重要的东西是因为他们从一开始就没兴趣或者没想法。了解为什么有人“没搞清楚”是了解那些“搞清楚”的人的观点的关键。

(9) 给主持人留点比萨。

### 提示

- 极少情况下，由于有人破坏了讨论，或者完全不合格，必须请其退出。助理主持人要做这份请人离开的苦活，但只有在主持人和助理都认为如果这个人存在，小组讨论就无法继续下去的前提下才能请此人离开。此时，有种好的处理办法，在小组中找到自然停顿的地方，并请人离开。开始的时候，人口统计表格或者保密协议是防止出现这种情况的好时机，可以和助理主持人讨论一下。之后可以采用“5分钟放松”或者类似借口来打破进程。在此期间，助理主持人要进入房间，礼貌地请人离开。应该感谢那个人的参与，支付承诺的酬金，并请他原谅。大家都明白发生了什么，但看到有人“被踢走”的兴奋感会很快消失。
- 有时候可以有多位主持人。如果有很多焦点小组，或者焦点小组时间特别长，或者需要用不同语言来完成，有时就需要有几位主持人。如果有若干人主持，他们应该过一遍每个问题，并确定怎么问，以及彻底调查哪些问题。
- 可以使用双向寻呼机或者无线聊天设备(如 Cybiko 手持设备)，不再需要助理主持人记笔记。主持人可以把设备放在桌子上，偶尔看看有没有来自观察室的新意见或者问题，在那里助理主持人(或者其中一位观察员)操作另一台设备。如果使用寻呼机，应该使用还没有很多条消息的寻呼机，以免分散主持人注意力，并关闭寻呼机的声音，打开振动方式来接收新消息，以免打断小组讨论。



### 聘用专家

焦点小组是一个复杂过程，可以选择聘用一家公司来专门组织焦点小组，而公司内部不用花费时间和精力。如果聘用公司来运行焦点小组，要考虑以下情况。

- 主持人是否了解主题？主持人需要能独立思考主题，并进行合适探讨。因此，所聘用的公司要么对特定主题有经验，要么对密切相关主题有些经验。如果没有这些经验，就要做好准备向主持人介绍主题、术语以及与主题相关的问题。
- 他们能招募到合适的人吗？给焦点小组找到合适的人是焦点小组成功的关键，因此，招募方法非常重要。问问该公司如何获得数据库，如何为焦点小组进行筛选。务必检查筛选文档。
- 周期有多长？每个公司的分析方法都不同。根据最后产出的不同，分析需要花费几天到几周。确保能知道他们计划要做的事情、需要多长时间。你可能要放弃漂亮的报告，放弃展示发现点清单的幻灯片演示，甚至与主持人和分析师的讨论。

和任何合同工作一样，应该要求公司提供最近完成的几个项目作为参考，并进行跟进。不要不好意思获得满足。如果发现不对劲的地方——招募、主持、分析——应该给他们打电话。如果有必要，应该要求进行额外焦点小组，费用由该公司负责。

## 焦点小组分析

“研究员必须保持警惕，避免陷入选择性认知。”

——理查德·A. 克鲁格，《焦点小组》，第130页

分析焦点小组信息有多少种方法，分析师就有多少种。从定义上而言，信息是有质量的，因此分析重点取决于焦点小组的目的。对于有些研究项目，发现参加者的心智模型是关键；对于其他研究项目，第一手体验最宝贵。还有一些项目会完全关注评估竞争对手。





焦点小组分析有两个基本过程，即收集数据和提取趋势。大量数据收集通常预示着有大量分析，但两者并不会严格一致。通常情况下，两者会相互交织，趋势需要重新考虑数据，而数据中的模式又会形成趋势。

### 收集数据

焦点小组会产生很多信息：剧本、引语、观察员的意见、模型和录像。组织并安排信息的优先级别是从中提取趋势的第一步，可以从捕捉瞬间信息、观察焦点小组的人所产生的直觉趋势开始，这些第一步假说有助于关注以后的数据收集。

#### 捕捉初始假设

主持人、助理主持人和观察员应当听取各组的看法。随着时间流逝，不同小组在记忆中会发生混淆，每场焦点小组结束后立即收集三者的看法能减少以后所需的创意和体验的纠结。应该记录每个人的笔记，并在访谈时收集他们的观察结果。使用指南是有效管理并引发记忆的方法。逐节走查指南，请主持人和观察员回顾他们的想法：什么是意外情况？什么是意料之中但并没有发生的情况？人们表现出来什么态度？他们主张什么价值观？他们提出什么有趣陈述(为什么有趣)？他们观察到什么趋势？参加者提供了什么有趣反馈信息？焦点小组有什么问题？这些意见往往是以后分析的基础。

简短汇报后，分析师要尽量完整写下对事件的记忆。随后，分析师要把观察从汇报笔记变成不同的主题或者问题，这些可以作为进行更正式分析的分类。

#### 誊写及编码

正式分析焦点小组过程应该从誊写开始。传统方法是聘请誊写服务。这项服务提供的文档会记下每个人说的每句话，可以很快收集到更多数量的引言。但不幸的是，誊写服务非常昂贵，而且需要很长时间，文档会很臃肿(两个小时焦点小组的文档会有 100 页，而且很常见)。较简单的方法是看录像带，只誊写分析师认为最重要的部分。





即使没有正式誊写,也需要回顾磁带。仅仅依靠记忆会漏掉很多微妙行为,会错误引用发言,观察员也会陷入集体思维方式。看看原始讨论可以澄清模糊的地方,并发现单凭记忆无法发现的隐藏含义。不必按照录制的顺序看录像;应该看你认为最有用那些录像。如果一个系列中有很多个焦点小组(比如说五个或者更多小组),但时间又很短,可以跳过那些“哑巴”小组(但看的录像如果少于四个,会漏掉一些关键问题或者有启发意义的用户描述)。

一边看录像带,一边把意见编成代码。编码是对答复进行分类以便跟踪趋势的过程。编码要有简短的描述性名称,每个编码都要包含一个想法或者需要跟踪的趋势。编码要能反映有兴趣研究的主题。如果要把人们在特定情况下的体验独立出来,可以对不同情况或者体验进行编码。如果想了解人们的优先事项,可以对偏好采用不同代码。可以从驱动指南编写的主题编码入手,在汇报情况时增加其他可能出现的代码。例如,如果原定目标是“了解人们研究保险时采用的心智模型”和“收集人们对现有保险的不足如何看待”,但你观察到人们被保险公司吓到了,初始编码可以如下所示。

#### 保险焦点小组顶层编码样例

模型:人们如何理解理赔过程和如何选择保险的过程,包括他们形成选择基础的参数以及评估参数的方法。

糟糕的故事:挑选保险的过程,或者填写理赔表格的过程中的一连串事件存在困难或者令人沮丧。包括错误的期望、失望,甚至保险公司完全无法兑现承诺。如果有正面的故事,可以编为“好故事”。

惊吓:如果参加者觉得受到保险公司的惊吓、受到过程惊吓或者无法控制过程。

当然,如果有需要,可以划分或者合并编码。如果决定需要区分人们受到保险商的胁迫和受到过程的胁迫,可以创建若干子类。但不要划分得太细。有一些社会调查研究可以用来对数百种不同活动和言论进行编码,但对于大多数用户体验研究而言,十几种分类已经足够了。

提示:

第13章将描述一种更正规的创建编码结构的方法。还有大量软件包可用于编码过程。





### 忠告：

焦点小组数据分析是一个充满争议的过程。提供良好可信分析的关键是保持分析师和产品之间的距离。即使你深入参与产品创造，现在都要尽可能保持客观。不要让你在录像或者进行焦点小组时产生的期望、希望或者任何结论影响你对参加者发言和行为的看法。要保持冷静，分析数据时，要装作从未见过这些焦点小组，而且要装作对网站或者主题一无所知。

### 提示：

如果时间紧张，可以跳过编码加快过程。主持人、助理主持人和观察员会根据直觉印象进行分析。这样能抓住最明显、最严重的问题，但容易造成评估团队的集体思维。从而对某些主题造成不必要强调的重视，而忽视讨论较少的主题。

看录像时拿上编码表。如果匹配编码或者看起来有趣而且相关，就记录下来，跟踪是谁说的，是在焦点小组进行到什么时候说的。如果有人说的东西真的能体现想法或者适合编码机制，就誊写出来。誊写的关键是捕捉人们有意义的发言，因此，目标虽然是准确誊写，但不要回避意译，要去掉无关的话、增加括号来提供上下文情境。但需要清楚标明哪些话是你写的，哪些是参加者说的。

完成对所选录像带的誊写和编码工作，回过头来检查一下誊写文档，检查编码的准确性，修改编码系统，如果有必要重新编码。

## 提取趋势

有了加上编码的引言清单，接下来就可以发现焦点小组的深层含义。本节介绍彻底、公正严谨的方法。时间和资源压力有时会造成这一过程很难全面实施，此时完全可以接受简化，但要小心简化，以免造成结果扭曲而不再有用。

焦点小组的分析方法与情境调查研究采用的方法类似。观察结果经过分簇和标记，为确定人们的行为和态度打下基础。这些趋势会被充实，并用假设进行解释，用数据来备份这些趋势。

从分类编码入手。代码代表原始意图以及检查数据时所观察到的趋势。充实这些编码，必要时进行修改以符合对数据新的理解。如果已经发现人们并没有在理赔时受到保险公司的胁迫，可以重写并添加一个编码用于描述人们感到困惑的情况(或者，如果要关注理赔体验，可以创建一个新编码来标记出对理赔的所有评论)。

接下来把主持人和观察员的观察结果，加上第一次分离、猜想出的分析，划分成为趋势和假设。尝试根据之前定义类别调整这些趋势，去掉重叠想法并澄清类别，并保存这些假设，之后可以用来解释趋势。

接下来根据修改过的编码重新整理数据引用和观察结果，根据所属编码，把所有情况组织起来(可以用字处理软件移动文字块或者写在即时贴上)。这样能有机会看看是否有办法来组织先前未编码的观察结果，并将观察结果和引用放到多个类别中。



族群需要能显示出可识别的趋势。有些趋势会被预料到而且很清楚。其他趋势会令人惊奇。可能有些情况你期望能找到支持想法的材料，但并没有找到。尽量试试根据相似性来组织趋势。甚至还可以给趋势组加上标签。

研究目标确定趋势很重要。关注心智模型的焦点小组系列能更加专注于人们使用的语言和隐喻，而不是分出特性的优先级别，后者更多和参加者的利益和自我观念相关。当然，人们的感受、价值观念和经验都相互关联，因此很少会出现观察结果完全符合趋势，或者趋势只影响到人们体验的单一方面的情况。在实践中，趋势之间的界限模糊不清，准确定义趋势的东西可能本身也不清楚。但在好的研究中，即使边界模糊，但中间线还是很清晰。

可以从数据提炼出以下内容清单。

- **心智模型。**心智模型和隐喻有关。心智模型是我们如何理解世界工作方式的心智表现。例如，乔治·莱可夫和马克·约翰逊在其经典著作《我们赖以生存的隐喻》(Metaphors We Live By)中描述了“时间就是金钱”的隐喻：谈到时间，说英语的人常常会采用类比方法把时间等同于金钱——“制造”、“浪费”、“花费”时间等等。协同软件生产商要了解这点，建立信息体系结构和命名界面要素会变得容易一些。有些人赋予软件个性，并从某种意义上把软件当成助手、朋友或者知己。当然，心智模型存在限制。除了在监狱，实际都没“获得”时间，因此不能过分从字面看待模型。在更现实的水平，某些人可能没有意识到他们可以挑战一些人的说法，在他们的心智模型中，他们只是去接受他人的意见，并不去对意见做更多的权衡，或是提出自己的新意见。
- **价值观。**人们喜欢或者不喜欢什么？他们采用什么标准来判断是否喜欢或者不喜欢什么？他们认为什么东西重要？他们采用什么过程来创建自己的价值观？他们的价值观如何相互关联？人们的价值观决定了很多对产品的体验方式。如果有人对网站内容很感兴趣，在某种程度上就可能会忽视各种交互问题。而采用另

编注：

有关心智模型，可参阅英迪·杨的著作 *Mental Models*，中文版由清华大学出版社2010年出版。





一种方式来展现相同内容可能会令人厌烦。人们的价值系统由喜欢、不喜欢、信仰以及他们对这些要素、物体、人和他们的生活关系构成。

- **故事。**故事是理解人们错综复杂的经验的强有力方式。故事能提供详细资料，说明人们的假设、做事情的顺序、如何解决问题(他们有什么问题)以及他们的观点。故事可以立即说明和澄清产品开发中的许多不确定性因素。不同人讲的故事之间的差异可以揭示目标受众有什么样的共同心智模型，以及开发团队可以预期到什么样的个人特质。
- **问题。**焦点小组头脑风暴能快速产生大量问题清单。即使没有正规的头脑风暴，在一群人中人们自然产生的同情感也能揭示很多问题。
- **竞争性分析。**人们不喜欢竞争对手产品的哪些地方？竞争对手产品有哪些？

上述清单还不完整。如果有一个焦点小组可能关心人们通常执行任务的顺序，而另一个焦点小组可能会想知道什么东西能让公司的徽标更出色。

以下是一些从数据中获取最有用信息的提示。

- 关注人们形成决定的方法。实际决定也很重要，但决定背后的原因更有启发性。
- 注意人们使用的术语。产品采用和用户一样的语言会更容易被接受。逐字剧本真的有助于解决这个问题。
- 留意矛盾的地方。人们说他们怎么行动和他们说想要什么并不一定和他们实际做的事情或者实际怎么使用产品相一致。
- 留意人们改变想法的情况。知道有人改变了想法能揭示出那人更看重的东西。
- 受欢迎程度并不一定代表重要性——人们认为重要的东西可能并不是他们谈论的东西——但受欢迎程度是强有力指标，应该记下来。同样，受欢迎程度不高并不意味着某个现象不重要，但如



果只有几个人提到了某个东西，那最好将其当成弱趋势。

可能也需要做一些数据定量分析。就整体而言，焦点小组数据的数值结果无法代表用户人群，但这些数值结果可以互相比较。如果在重新设计前后对相似小组分别进行焦点小组，而且讨论指南、主持人和分析师都相同，比较一些问题或者看法的数量可能是观察体验是否发生变化的有效方式。但这一过程需要严密控制，并且不能超出两组进行比较。

此外，分析任何焦点小组数据时，必须明确考虑潜在偏见。招募过程、问题措辞、集体思维以及主持人和分析师的个人经验都可能会影响答案，甚至外部事件也可能会影响人们的观点(例如，有一天新闻报道了交通事故，而前一天没有，人们对休闲旅游的看法前后可能就会不同)。密切注意可能已经引入的偏见，并在报告中明确说明什么时候会有偏见及其存在形式。

## 做出假设

解释趋势的原因是一个困难的过程。每个现象可能都有大量潜在原因，也有数据能支持矛盾的假设。分析师必须判断是否需要假设人们的行为和信念，或者是否就只说出这些信念而让开发团队形成自己的理论。业界一直在争论此问题，也没有硬性规定假设什么时候适用。如果可能，焦点小组中观察到的趋势可以指向困难的深层社会和心理问题(例如，“人们声称不喜欢广告”)，而这些问题需要解释或判断。知道问题层面常常就足能够解决问题，而并不需要知道确切原因。

但有时候，对问题潜在原因的分析以及他们之间的互比较更容易能找到解决方案。如果人们都受到保险公司的胁迫又感到愤怒，知道他们为什么受到胁迫和愤怒就有助于确定需要解决哪些更重要的问题。如果从收集到的数据中无法明确根本原因，可以开展更多研究。因此，如果从焦点小组中无法明确某人挑选保险公司的检查过程，但挑选过程中产品的功能又很重要，可能就需要一轮实地任务分析。同样，如果要知道目标受众中受到胁迫的人的确切规模，可以进行有统计意义的调查。

和所有用户体验研究一样，焦点小组产生的问题多于答案，但这些问题





题可能胜于开展焦点小组讨论之前提出的问题。

## 实例

以下报告总结了调查人们购买汽车保险时焦点小组系列的结果。  
AutoSafe.com 纯属虚构。

### 执行摘要

把 24 个人分成 6 个焦点小组，对他们最近的车辆体验进行访谈。他们一般都对汽车修理过程结果感到满意，但并不非常了解修理过程，他们需要修理步骤指导、每一步骤的选择，而不仅仅是具体建议。他们把保险公司当成可以信赖的保护者和向导来推动这一进程。修理过程开始的时候，如果人们要做出大量选择，而且对于之后发生的事情信息非常少，独立的第三方信息最有用。修理厂评估和“车祸解析”此时最有用，因为人们不信任修理厂，而且过程未知。

了解过程大致分为几个步骤：沟通(和家庭、工作和公司保险)，保险(为了指导过程开始)，临时交通，车身修理厂管理和善后管理(如果事情被延误或者发生错误)。如果发生受伤，有时候会牵涉到汽车保险，还会有平行的医疗过程。

除了帮助人们跟踪汽车修理情况，还能提供不带偏见的信息，把这种服务想法告诉人们时，人人都对其产生正面反应。人们觉得价值最大的地方是在澄清过程和不带偏见的观点。经过几个小时，人们认为电子邮件更新是有用的特性，但人们对实际的工作进程很好奇(也有些疑问)。

### 步骤

请参阅文档末尾的附录 A“讨论指南”。

### 评估人背景资料

完整报告包含参加者的人口统计情况图表、互联网使用、技术情况和汽车事故资料。报告接收人可以更好了解具体引用的上下文情境。



### 目的

- (1) 搞清楚人们如何了解修理/保险的过程。
- (2) 确定过程中的痛苦点。
- (3) 确定第三方在哪里提供帮助。

### 过程模型

这项研究的主要目标之一是了解人们如何考虑事故后的汽车修理过程。认识和心智模型分解成两部分：情感组成部分和步骤组成部分。情感组成部分包含人们对这一过程的一套感情和观念。步骤组成部分指人们对过程应该如何运作及对自己在此过程中充当何种角色的理解。

一般情况下，参加者看起来并没有深入到事故体验的维修方面。每组中，有很多人都对保险公司以及车身修理厂感到满意(但并不总是在相同时间)。许多人对最终结果都很高兴。最强烈的情绪反应看起来是在体验开始的时候——因为步骤的不确定性和预期费用(而且在有些情况下会涉及医疗问题)——和在最后，此时对汽车修理完成和索赔支付的焦虑程度最高。

此外，由于事故情况不同(有的受伤，有的推迟性受伤，有的车子能开，有的车子不能开，有的具有良好修理经验，有的没有)，在实际情况中，情感过程和步骤过程有很多不同。

因此，我试图尽可能根据答复进行概括，并提炼出情感和步骤组成部分，但会把它们放在彼此的情境内，以便能尽可能多地捕捉到参加者的想法。该模型有六个主要组成部分。

- 背景期望
- 事故后联系
- 保险公司物流
- 交通物流
- 车身修理厂物流
- 结果

略去模型和观察。用引言支持大多数观察。





多次提到人们的情感沟通问题，人们对过程中要发生的事情似乎了解不够——特别是谈到修理状态的时候。

罗伯特：“你必须自己跟进；好像没有跟进系统。你不知道发生了什么事情，到应该提车的前一天才知道他们会打电话给你，告诉你下周来提车。”

曼纽拉：“如果什么也不做，就什么也不知道。你要推动所有事情。”

### 结论

产品的维修跟踪部分虽然有价值，但其价值似乎并没有和客观推荐部分受到同样认可，因此，它是匹配人们需求的理想重点。

此外，由于人们对过程了解很少，而且很可能不会在事故发生前做任何研究来了解过程，过程解释就会是对他们体验的最大贡献。因此，匹配给他们现有最好的公式化模型，也许就没有给他们一个模型并解释其中的步骤一样有价值了。

因此，定义并解密过程，而同时让过程更加客观，可能比自动化更有价值。

## 附录 A：讨论指南

以下是上述研究采用的讨论指南。

### 暖场和介绍(10 分钟)

目的：介绍过程，并让参加者感到轻松。

大家好，欢迎大家来参加讨论。

迈克进行自我介绍。

- 市场研究员
- 和我们今天要讨论的产品没有关系
- 对这些事情不带感情因素，大家可以畅所欲言
- 帮助公司分析一个新想法

我们今天要做的事情，是要谈谈大家的体验，他们就能做出符合像大家一样的人的体验的服务。主要是随意聊聊，但有几个地方我会强调一些





事情。

讨论的时候，请大家尽量坦诚讨论。不管您说什么，都不会伤害任何人的感情，所以请大家准确说出感受。

我们需要**大家**的观点。观点没有对错——特别是针对我们会讨论到的东西——只是观点，所以如果您不同意其他人的说法，我们也想听到。

但请大家一个一个地发言。

今天会讨论很多问题，没人想待上几小时，所以只好请大家长话短说，便于讨论正常进行。

大家能想得到，玻璃墙后面有几位公司的人，我们要讨论的就是他们的想法，\_\_\_\_\_（主持助理的名字）会和我在一起主持讨论。和我在一起的是玛丽。玛丽会偶尔进来给大家一些需要的东西。不用管。前面说了，邀请大家来是想听听大家的想法，无论您说什么都不会伤害到任何人。如果大家想知道，这有台录像机，玛丽和我就不用坐着紧张记笔记，就可以专心听大家说了。录像带纯粹用于研究目的。产品开发团队成员可能会看录像，但不会给公众看，也不会用于推广或者传播目的。

### 保密协议

现在我要给大家阅读一下保密协议。我对每位访谈人都要读这个。保密协议会说明大家参加此类研究的权利。

作为参加此次研究的人

- 可以随时终止。
- 可以随时提问。
- 可以随时离开。
- 不会受到欺诈。
- 回答将会保密。

这里有张表，需要您允许我们对讨论进行录像、并用于研究。

### 受访人介绍

请说出您居住的城市

您会看什么电视节目和出版物？





您最喜欢哪些网站？网络有没有让您觉得困扰的东西？

(迈克介绍自己的情况，其他人轮流介绍)。

### 修理过程(10~20 分钟)

目的：介绍车祸后主题，并且让人们感到舒适。尝试确定人们认为的修理/索赔过程(即他们对汽车被撞和修理或者被毁之间所发生的事情的心智模型)。

(2 分钟)您可能根据在电话上问您的问题已经猜到了，我们今天要谈论的是车祸。但请不用担心，我们不会谈论事故本身——我肯定事故会让大家不愉快，不值得讨论。我想谈论的是事故后发生的事情。

首先，我想知道从拖车到达整个过程结束之间发生的事情。这样我们能看到哪有问题，看看有没有方法能把事情做得更好。

(10 分钟)整个过程的第一步是什么？[写下来]大家都同意吗？[如果有人不同意，再列一张平行清单。]请用几句话描述一下。

接下来发生了什么事情？[记下来]

我们来看一下清单。有没有人觉得漏掉了什么？

有没有不合适的东西，看不懂的东西？有没有人愿意把这些步骤换一下顺序？

### 一般印象(15 分钟)

目的：了解最高层的痛苦点，看看车祸很少的人和车祸很多的人是否一致。

今天邀请大家来，因为大家[“去年”，“几年前”]发生过事故。有没有人是第一次发生事故？

(3~5 分钟)[如果有，向回答的人提问]您觉得过程中最让您不高兴的是什么事情？[向第一次发生事故的人提问]您觉得是什么事情？

您觉得呢？[向不止一次发生过事故的人提问]

(5 分钟)[向特定的人提问]从 1 分~5 分——5 分代表您永远都不会、都不想再经历的事情，1 分代表，比如说小划痕——您如何让您恼火的事情打分？[向每个人提问]有人打 5 分吗[如果有，向特定的人问问为



什么]

(3 分钟)现在看看我们之前写的清单,我想给每个点快速打分。5 分制,请举手说明您觉得认为每部分有多让您讨厌;5 分代表超级讨厌,1 分代表不是真的讨厌或者稍微有点讨厌。我好计算平均值。[给每个点打分]

(3 分钟)好的体验呢?有没有人很高兴看到让人惊奇的东西?有没有让您觉得惊讶的东西?为什么它体验很好?还有没有其他东西?

如果时间允许:

还有没有其他正面体验,比如说,保险公司来的机械师或者什么人?

[描述]

(1 分钟)剩下的人都至少还发生过一次事故。有没有超过三次的人?超过四次的有没有?超过五次的呢?[继续提问]

(2 分钟)[向体验最多的人提问]在那些情况下,什么事情一直让您觉得讨厌?

### 车身修理店(15 分钟)

目标:确定修理体验中的痛苦点,以及人们如何选择修理店。

好了,我们换个话题,下面来谈谈如何修车。

(5 分钟)我们来给车身修理店的体验打分。1 分代表良好体验,5 分代表糟糕体验,3 分代表体验还行,我们不是在做娱乐节目。请举手来打分。[如果某人打了 1 分或者 2 分,对那个人进行提问]这个体验好在哪里?

(5 分钟)谁有过糟糕体验。[向打 4 分或者 5 分的人提问]最糟糕的事情是什么?其他人也有这种体验吗?[挑一个人问问]有没有最糟糕的事情,或者您认为最大的困难在哪?

和修理厂打交道有没有碰到其他困难?

(5 分钟)假设有人能给您一些关于车身修理店或者修理厂的建议。您觉得什么时候最希望能够得到别人的建议?[向特定的人提问]您关心什么建议?

如果有时间:

(5 分钟)[四处看看]您怎么选择去哪里修车?如果再选择一次,您会





选择不同地方吗？

(5分钟)我们来列一张清单。选择车身修理店时哪些事情最重要？

(3分钟)[挑几个人] 如果有朋友或者同事发生了事故，他们问您如何选择修车的地方，您会给他们什么建议？

### 保险(15-20分钟)

目标：确定理赔过程中的痛苦点，人们如何选择保险公司。

好，现在我们换个话题，谈谈体验中的保险部分。

(3分钟)首先，有没有在事故之后换过保险？

有没有考虑换保险？是因为事故吗？[如果是]为什么？

(7分钟)这次体验中，请想一想您认为车辆保险中最重要事情。想一想您要做的所有事情，还有过程中发生的所有事情。什么事情让您觉得最意外？最大的困难是什么？

[四处看看，继续讨论]

(5分钟)[向一个人提问]假设您要给我一些选择保险公司的建议。选择保险公司时，您觉得什么事情最重要？[向另一个人提问]还有没有其他我应该知道的事情？有没有人知道？

(1分钟)如果我想做一些比较，有没有地方我能了解到不同的保险公司和保险政策。

(3分钟)假设曾经有人能给您一些保险公司的建议，什么时候您最想找个人听听建议？[向人们特定提问]

如果有时间：

(2分钟)哪些人立即获得了赔付？哪些人等了两周？哪些人等了一个月？有没有等更长时间的？

有没有人通过用信用卡获得赔付？

(5分钟)谈到赔付，有没有谁觉得保险公司有问题？有什么问题？

(3分钟)我们用一句话总结一下，比如说对于保险公司，您现在是否知道了以前不知道的东西？[快速问一圈]



**AutoSafe.com 的想法和一般互联网情况(10 分钟)**

目标: 介绍 AutoSafe.com 的想法, 看看人们能否理解并欣赏; 通过简单头脑风暴获得一些特性; 发现人们排列现有特性想法的优先级别。

好了, 大家的问题我都记下来了。现在我想谈谈解决这些问题的方法。

(1 分钟) 假设有个网站能帮助您跟踪修理情况, 并告诉您所处的修理阶段情况。假设这个网站还能链接到其他保险公司和车身修理店的不带偏见的信息。

(2 分钟) 您会感兴趣吗?

(5 分钟) 我们来几分钟的头脑风暴。不用考虑现实性, 您希望此类服务能为您做什么? [把想法记下来, 如果人们不知道怎么开口, 可以说一些现有想法, 借此抛砖引玉]

(5 分钟) 好了, 有多少人想要不带偏见的信息来源, 能对保险公司进行评级, 就像消费者报告一样? 现在, 假设您要找这样一个网站, 您发现网站上有广告。有没有觉得在消费者报告类型的网页上看到广告会有问题? 假设是福特的广告, 您觉得怎么样? 如果是汽车保险公司的广告, 您觉得怎么样?

如果有时间:

(2 分钟) 如果这项服务对您称为“一站式事故店”, 根据您所知道的, 您觉得有意义吗?

现在我想问几个一般互联网问题。

(3 分钟) 有没有人注册过的在线服务需要输入名字和电子邮件地址? 有没有输入家庭地址? 电话号码呢?

(5 分钟) 把什么个人信息给网站大家会觉得不舒服? [写成清单] 如果用个人信息换取服务, 您觉得怎么样? 例如, 如果 AutoSafe.com 能为您代办理赔需要的所有文件, 然后邮寄到您家里, 您愿意输入家庭住址吗? 清单上有哪些信息您绝不会提供给网站的?

**总结(5 分钟)**

目标: 总结焦点小组, 并告诉人们联系信息, 为进一步的想法做准备。





我想大家现在都很兴奋，想谈上几个小时，但不好意思，我现在要问车辆和保险的最后一个问题。如果在回家的路上，您突然想到了一些非常重要的东西，请联系我们，让我们知道。您可以发留言给[写电子邮件地址]或者打电话给她[写电话号码]。

最后一个问题，与保险、交通事故和网页都无关：有没有什么建议让我们有何改进，无论是时间安排还是开展焦点小组？我们应该放点小吃、啤酒或者其他食品吗？

好了，感谢大家参加。听到大家的想法非常有趣，对我们非常有帮助。



## 第 10 章 可用性测试

一对一的可用性测试能快速提示人们如何使用原型的大量信息,无论原型是功能型、模拟或者只是纸质原型。在产品发布前,可用性测试也许是发现可用性问题最快、最简单的方式。

可用性测试是针对界面原型的具体特性进行的结构化访谈。访谈核心是由界面评估人(通常情况下指和产品理想受众相匹配的人)执行一系列任务。访谈磁带和笔记日后可用于分析评估人的成功、误解、错误和意见。经过多次测试,比较观察结果,收集最常见问题,形成功能和显示相关问题清单。

通过可用性测试,开发团队能立即明白人们能否像他们假设的一样理解他们的设计。不幸的是,这种方法已经被套上了项目完成之前最后检查的光环,而且可用性测试通常安排在开发周期的末尾——而此时特性设置已锁定,目标市场已确定,而且产品就要发布。尽管如此,可用性测试仍能更深入地了解下一版产品,但这种方法的潜力还有待开发。最好能更早采用可用性测试,给整个开发周期提供反馈,既可以检查具体特性的可用性,又可以探讨新想法并评估预期。

### 何时测试

可用性测试的优势在于发现人们如何执行具体任务,因此应该用它来检查每个独立特性的功能点向预期用户展示的方式。可用性测试更适合用于突出潜在误解或者特性实施方式中的固有错误,而不是用来评估整个用户体验。从开发周期的早期到中期,可用性测试对于定义特性、功能点和它们的开发起着关键的指导作用。但一旦锁定某个特性的功能,确定此特性和其他特性的交互,基本上已经不可能做出任何根本改变。此时的测试





### 忠告：

完全开放式的测试，或者称为“钓鱼”（fishing），很少有价值。如果在第一轮用户研究时进行的钓鱼——常常听到有人说“我们把全部都测试一下”——其结果既不是特别清楚，也没有见地。明确测试目的之后，再有的放矢地开始测试。

更多是对下一版本而不是对当前版本的投资。

此外，与本书中提到的其他一些方法不一样，可用性测试对于产品从来都不会只是一次性任务，也不能这么看待可用性测试。每一轮测试可以关注一小组特性（通常不超过 5 个），因此可以通过一系列测试来测试整个界面或者微调一组具体特性。

开发团队需要做的第一件事情是确定目标受众和要检查的特性集合。

也就是说，如果开发周期已在进行，启动可用性测试是好时机，但不能太晚，测试得太晚，如果测试发现了必须更改的特性，那么这个测试会阻止整个项目实施进程。可用性测试偶尔会发现需要很多工作来纠正的问题，因此如果需要，开发团队应该做好准备，重新思考和重新实施（理想情况下，重置）特性。在互联网世界中，一般需要几个星期的时间，这便是为什么往往要在两周的间隔期间实施迭代可用性测试的原因。

稳固的可用性测试计划包括对每个主要特性的迭代可用性测试，在整个开发过程中安排测试，强调并加深人们对人们行为的了解，并确保它会随开发设计而变得更加有效。

### 迭代测试过程实例：Webmonkey 2.0 全局导航

Webmonkey 是一家领先的互联网开发杂志，它使用了这种技术和方法。在重新设计周期，网站团队决定要给主界面创建一些全新的东西。1.0 界面的许多内容都进行了广泛测试，并带入了新设计中，因此团队希望把测试和开发精力的重点放到新特性上。

纳入考虑的新要素中，最有雄心、也最有问题的是一个 DHTML 全局导航面板，通过面板能访问整个网站（见图 10.1 和图 10.2），但面板并不会永久占用屏幕空间。相反，它会根据用户的需要在屏幕上出现或者消失。Webmonkey 以前的导航机制很好，但网站团队分析后认为，它使用不够频繁，不需要占用那么多空间。他们不想强调导航（毕竟导航不如网站内容重要），因此他们决定尽量减少面板占用的屏幕空间，而不是要增加其使用。他们最初的设计采用传统的垂直导航栏，和 1.0 网站中的左侧空白处相同，但有自己的面板。大部分时间该面板都处于隐藏状态，但如果点





击左侧条形栏的顶部箭头，面板就会出现。为目标受众做开发的人员希望他们能够注意到条纹栏和箭头，从而在好奇心的驱动下进行点击。

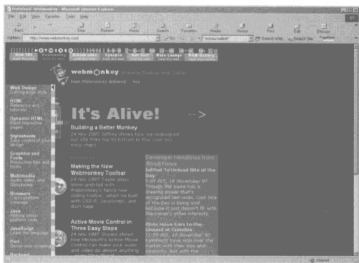


图 10.1 Webmonkey 2.0 导航面板设计(打开状态)

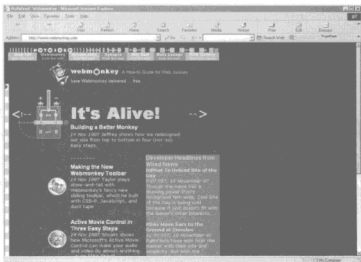


图 10.2 Webmonkey 2.0 导航面板设计(关闭状态)





Webmonkey 的开发基于迭代开发周期，因此互联网开发人员和高级用户受邀参加一系列测试，每个测试阶段后面都跟着设计阶段来整合测试结果。测试目的是要检查参加者的整体用户体验，但开发人员特别关注滑动面板。第一轮测试中，六位评估人没有一位打开面板。被问及是否看到条纹栏和箭头时，大多数人都说看见了，但他们认为条纹栏是个图形，而箭头是装饰物。

两周后，视觉设计并没有多大改变，但在网页首次加载时，设计师把面板从默认关闭状态改成了打开状态。测试时，评估人自然发现并理解了面板用途，但他们要看被面板遮住的内容时又遇到了关闭上的麻烦。有些人想像拖动窗口一样拖动它；其他人想点击来关闭。大多数人都看到了箭头，但不知道箭头和面板有什么关系，因此都没试着点击一下。进一步提问发现，他们没有发现面板是可以打开和关闭的滑动窗口。因此，有两个相互关联的问题：人们不知道如何操作面板，也不知道箭头是功能要素。

第三次设计想解决这个问题，他们提供例子来说明面板功能，作为页面的初次体验：页面载入后，会短暂停顿，面板会自行打开和关闭。设计人员希望动态显示面板能更说清楚面板的功能。接下来的一轮测试中，这种方法的确起到了作用，评估人正确描述了面板的内容和功能。但还是没人能再次打开面板。新的设计仍然没有解决箭头问题，多数人在条纹栏内点击并拖动，想打开面板。设计人员观察到这个问题，（经过一些辩论）并意识到在技术上无法实现面板拖曳机制，设计人员就把整条彩色栏设置成可点击，因此只要有人点击条纹栏上的任何地方，面板都能跳出来（在面板的打开的情况下，点击条纹栏的任何地方面板就会关闭）。

最后，人们仍然不知道箭头有什么用，但是他们点击条纹面板想打开面板时，面板就打开了，这就足以让此特性用起来，而且再没有观察到人们使用面板时会碰到打开或者关闭的麻烦。



## 如何测试

### 准备

尽管与第 2 章叙述的“朋友和家人”测试相类似，但完整的可用性测试计划、执行和分析需要的时间明显要长得多(见表 10.1)。至少要在希望拿到结果的前三周就开始准备可用性测试。

设定日程表

时间选择	行 动
$t-2$ 周	确定测试受众；立即开始招募。
$t-2$ 周	确定要测试的功能特性集合。
$t-1$ 周	编写第一版剧本；创建测试任务；和开发团队讨论；检查招募工作。
$t-3$ 天	编写第二版指南；检查任务；和开发团队讨论；完成招募。
$t-2$ 天	完成指南；安排练习测试时间；设置并检查所有设备。
$t-1$ 天	上午进行练习测试；适当调整指南和任务。
$t$	测试(通常需要 1~2 天，根据时间安排)。
$t+1$ 天	和观察员进行讨论；收集所有笔记抄本。
$t+2$ 天	放松一下，放一天假做点其他事情。
$t+3$ 天	观看所有录像带；记笔记。
$t+1$ 周	整合笔记；编写分析。
$t+1$ 周	提交给开发团队；讨论并写出进一步研究的方向。

设定时间表之前，要明确招募对象，以及请他们评估哪些特性。测试开始前的几周需要确定这两件事情。

### 招募

招募是要尽早开始的最重要的工作。招募时间要恰当、准确，尤其是外包招募的时候。需要找到合适的人，他们的日程安排要符合你的日程表。招募要花费时间和精力。投入招募的时间越多越好(但提前两周多就开始





招募一般还为时过早，因为人们往往不会那么早就知道自己的日程安排)。还需要仔细选择筛选标准。最初的冲动往往是想招募属于产品理想对象的人，但总是过于宽泛。需要关注能提供最有用反馈的目标受众代表。

假设要建立一个在线销售高端餐叉的网站。理想受众由打算购买餐叉的人构成。

为可用性测试进行招募，人的范围相当广泛。缩小重点有助于保持清晰，因为不同群体对相同的基本可用性问题会有不同行为表现。年龄、经验和动机对于相同的潜在问题会造成不同用户体验。从长远看，选择“最有代表性的”群体可以减少必要的研究数量，并关注结果。

最佳受邀人是需要你将来提供服务的人，或者是最近使用过竞争对手服务的人。这些人对主题兴趣最高、知道的最多，他们可以关注界面有多好，而不用管信息细节。对内容没兴趣的人仍然可以指出交互存在的缺陷，但他们并不擅长指出信息架构问题或者任何内容的具体特点，因为他们没有动力去关注，并让它发挥作用。

假设对餐叉市场的研究表明，餐叉市场有两个购买力强的子群体：要替换旧银制餐叉的人和购买作为结婚礼物的人。根据研究，第一组主要是 40 多岁的男性，而第二组男女基本对半，主要集中在 25 岁左右和 30 岁左右。

你确定购买套餐叉代替已有餐叉的人能代表用户群体核心。他们了解主题，并且最近做了一些研究。他们正积极使用这项服务，而且会定期使用。因此，你决定要招募最近会购买或者已经购买过餐叉的 40 多岁男性。此外，要排除互联网新手，想找有网上购物经验的人。有了所有这些条件，形成如下最终招募标准：

- 男性或者女性，最好是男性
- 25 岁或者更大，最好在 35~50 岁
- 在家中或者上班时可以上网
- 每周上网时间不少于 5 小时
- 互联网经验不少于一年
- 至少在網上买过三样东西



注意：

招募师会努力严格遵守你制定的标准。但如果告诉他们标准有灵活性(和灵活程度)以及可以不遵守哪些标准,他们操作起来会比较轻松。你最终操作起来也会更轻松。

此外,杰瑞德·斯普尔和威尔·施罗德指出([www.winwriters.com/download/chi01\\_spool.pdf](http://www.winwriters.com/download/chi01_spool.pdf)),如果希望评估人满足广泛目标,而不是执行具体任务,至少需要5个人。但在在我看来,广泛目标研究是比有重点的情境访谈(第8章)较为松散的可可用性测试,就应该如此进行。

忠告：

要尽量对每个主要用户市场实施不同测试,因为——从定义而言——每个用户市场会用不同方式使用产品。第7章定义了用户市场。

- 最近三个月在网上买过东西
- 有兴趣在线购买餐具

请注意,年龄和性别标准有一定的灵活性。这是为了让招募师的工作轻松一点。你可能会坚持要求所有参加者都必须是男性,必须40~50岁,但如果候选人符合其他所有标准,而又恰好是33岁的女性,你可能不希望立即取消她的资格。另一方面,采购经验需要准确需求,因为找到对电子商务概念不会感到迷惑或者惊奇的人是测试成功的关键。找一个没有任何网上购物经验的人来测试电子商务系统,测试他对电子商务概念的理解所做的工作会和测试一个具体产品一样多。很少会把测试脚本做到这么细;因此避免这种情况的产生,最好提前防患于未然。

如果要关注基于任务的可用性测试,每轮测试至少应该有5位参加者,应该招募6~10人来准备进入这五个位置。雅克布·尼尔森说(游击人机交互法:采用廉价的可用性工程缓和消费高消费的惊吓障碍,可以从[www.useit.com/papers/guerrilla\\_hci.html](http://www.useit.com/papers/guerrilla_hci.html)下载)可用性测试的成本效益截止点是每个目标受众大约有5位用户。如果超过5人,虽然仍然能产生有用的结果,但招募费用和实施测试以及分析所需的额外精力会导致回报迅速降低。8、9位用户测试完后,执行特定任务时发生的大多数问题会多次重现。但为了防止发生失约的情况,应该在基本的5位用户之外再安排几个机动名额。要绝对确保人数足够,可以在每个时间段都预约两位用户。虽然招募费用和激励费用会翻倍,但能保证测试的中断时间最少。

此外,要检查对主要受众的了解,可以从次要目标受众中招募一两个人——例如对于餐叉的例子,可以是较年轻的买家或者不精通互联网的人——看看那些小组中是否存在完全不同的观点。你会得到结论,但如果你找的人有理性,而且这些人一直与主要群体唱反调,这就说明可能要重新考虑招募标准。如果次要受众特别重要,那么无论如何也要受众有自己的测试。

确定招募对象,接下来就开始编写筛选文档,并发给招募师。本书第6章描述了筛选文档和招募。请务必和招募师讨论筛选文档,并至少应该和内部的两个人进行走查,来检查可行性。





### 忠告：

如果是第一次测试，安排的人可以少一点，并且在每人之间安排额外时间。如果是新手，可用性测试会很耗体力。

挑选几个测试日期，向符合标准的人发出邀请。访谈应该安排在对大家都方便的时间，而且每人之间至少间隔半小时。主持人有足够时间允许有人迟到，测试有足够时间运行，主持人能喝杯水，并且能和测试观察员进行讨论。60 分钟访谈，也就是说一天可以有四五次机会，有时能到 6 次。而 90 分钟访谈，一天可以有三四位评估者，如果抓紧时间不吃午饭的话能有 5 位。

### 选择特性

接下来，确定要测试的特性。这些特性最终确定要创建的任务及展现顺序。应该选择有足够准备时间的特性，以便可以微调测试步骤。60~90 分钟的访谈可以测试五个特性(或者功能特性组)。典型测试要花一两个小时。两个小时的测试可用于初次测试或者广泛的基础测试，而时间较短的测试最适用于深入研究具体特性或者想法(尽管 90 分钟内完全可以进行广泛的基础测试)。

单独功能应该放在特性组的情境中进行测试。如果只测试一组特性中的要素而不管整个组，几乎不能发挥作用。我的经验做法是，30 秒画一个界面示意图，如果能画在白板上东西，就可以进行测试。如果把要画的一团东西标记为“导航栏”，应该思考对导航栏进行测试，而不能仅仅测试到首页的新链接。

选择特性的最佳方法是和开发人员开会(至少要有产品经理、交互设计师和信息架构师)，并列岀要测试的最重要的五个特性。如果开始讨论要包含的特性，可以看看具有以下特点的特性：

- 常用特性
- 新特性
- 深受关注
- 根据先前版本反馈存在麻烦
- 如果使用错误，会有潜在危险或者会有不好的副作用
- 用户认为很重要的特性



## 特性优先级练习

这项练习通过结构化方法来形成特性优先级列表。如果可用性测试小组没有安排特性优先级的很多经验，或者有很多麻烦，这种练习能起到作用。

步骤1：请小组把界面上最重要的新东西或者最后一轮后发生显著变化的东西列出来。不要纯粹根据是否突出来定义重要性；重要性可能与企业底线或者管理层的优先事项有关。因此，如果每周餐叉新板块能否成功决定着下季度盈利能力，那么该板块就很重要，即使只是界面的一小部分。

步骤2：画一列，标记为“重要性”。看看每项特性，打上1~5分，5分表示对产品成功至关重要，1分表示不是非常重要。

接下来再画一列，标记为“疑问程度”。看看每项特性，根据团队对设计的满意程度打分，最满意的设计打1分，最不满意的打5分。小组可能会产生一些辩论，因此，可能需要像对待开发人员的焦点小组一样来处理。

步骤3：两列中的分数相乘，并写入第三列。数字最大的特性代表要测试的特性。把这些特性找出来，用一句话总结小组最想了解此特性功能的情况。

前五个餐叉类型需求特性的优先级

	重要性	疑问程度	总分
购买机制	5	5	25
对单个物品和整套物品都起作用吗？			
搜索引擎	5	5	25
人们能用搜索引擎找到特定餐叉吗？			
目录导航	5	4	20
如果人们不知道要买什么，他们会使用目录导航进行浏览吗？			
每周餐叉	4	4	16
人们看见了吗？			
愿望清单	3	5	15
人们知道它的用途吗？会使用吗？			





有了最需要测试的特性清单，接下来创建运用这些特性的任务。

此外，还可以包括竞争性可用性测试。比较两个界面比测试一个界面更花时间，但可以发现不同产品的长处和短处。例如，用已有界面和新原型执行同样任务可以发现新设计是否更有用(或者能发现新设计——每位设计师都害怕的事情——更没用)。同样，执行同样任务或者用两个互相竞争的产品进行界面比较能发现两者之间的相对优势。但这两种情况下，不要左右评估人的意思，认为一个界面比另一个界面更好。第14章将深入讨论竞争性研究。

### 创建任务

任务需要能代表典型用户的行为，并且足够关注产品的单个特性(或者特性集)。良好的任务应该具有以下特点。

- **合理。**任务是人们要做的典型事情。不会有人订购90种不同餐叉，并运往37个不同地址，而且每个模式都不一样，因此这不是典型任务。而订购一打餐叉并运往一个地址就是典型任务。
- **以最终目标进行描述。**每件产品、每个网站都是工具。但其本身不是目的。即使人们花费大量时间使用它，也是在使用它完成某件事情。因此，就像有了角色动机，演员可以演得更好一样，如果有了栩栩如生的情况，界面评估人会更现实地执行任务。描述任务时要能联系到评估人的生活。如果他们要找到一些信息，就告诉他们为什么要找这些信息(“贵公司考虑在莫斯科开设办事处，您想知道莫斯科的再保险业务环境。您认为最佳方式是看看今天关于莫斯科再保险公司的商业头条新闻”)。如果他们要买东西，就告诉他们为什么要买(“米莉阿姨的紧凑型轿车听起来像喷气式飞机。她需要新的消声器”)。如果他们要创造东西，就给他们一些情境(“这是弗雷德叔叔的照片。您想开个玩笑，用数字化的方式给他加上胡须，然后发给家人”)。
- **具体。**为了让评估人之间保持一致，并让任务关注要测试的产品部分，任务应该有具体的最终目标。因此，不要说“买些餐叉”，





而应该说“有天您在商店橱窗看到一套漂亮的路易十四餐叉；这是图片。请在目录中找到这个设计，并购买一打吃鱼用的餐叉。”但应该避免使用界面上有的术语，因为这会提示参加者如何执行该任务。

- **可行。**如果网站上只有餐叉，就不要让人们去找刀。有时很希望能看看他们如何使用信息结构来找一些不存在的东西，但是，这是带有欺骗性的，令人沮丧的，而最终不会对你的设计带来很多有价值的信息。
- **按照现实顺序。**任务流应像产品的实际环节进行流动。因此，购物网站可以先有浏览任务，然后进行与选择任务相关的搜索任务，然后再进行购买任务。这种方式能让环节更现实，并能指出任务之间的交互，而信息架构师能用这些信息来确定产品中高质量的任务流程。
- **保持领域中立。**理想任务指每个测试界面的人都知道一些情况，但没有人知道很多情况。如果有评估人了解任务的程度远远超过其他人，他们采用的方法很可能会不同于组里的其他人。他们知道更多技术词汇、更多方法来完成任务。相反，所建任务也不应该对有些评估人完全陌生，因为他们可能甚至会不知道该怎么开始。例如，要测试一般搜索引擎，我请人们搜索黄油鸡图片：每个人都对鸡有些了解，但除非是饲养班塔姆莱母鸡的农民，否则不太可能有人知道黄油鸡。对于真正重要的任务，并不存在明显的领域中立解决方案，招募过程中应该排除懂特定知识的人(例如，招募筛选时问“您知道什么是黄油鸡吗？”能排除对鸡了解太多的人)。
- **时长合适。**对于大多数特性，都不会复杂到要超过10分钟才能完成测试。任务时长取决于三点：访谈总时长、结构以及要测试特性的复杂程度。以任务为重点的90分钟访谈中，50~70分钟用于任务，因此，平均每项任务大约需要12分钟才能完成。60分钟访谈中，大约有40分钟的任务时间，因此每项任务不要超





过 7 分钟。对于较短访谈，任务花 5 分钟，而对于较长访谈，任务花 10 分钟。如果发现有些事情需要更多时间，可能就需要将其分成子需求，并再安排优先级别(但要注意例外情况：有些重要任务需要更长时间，而且不好分解，但它们仍需要进行测试)。

### 估算任务时间

卡罗琳·施耐德推荐了一种估算任务时间的方法。

- 问问开发团队专家——比如开发团队成员——需要多长时间来执行任务。
- 把上述时间乘上 3 ~ 10，得出从未使用过产品的人执行同样任务需要多长时间。较小数字用于简单任务，例如能在普通受众网站上找到的任务，而较大数字用于复杂任务，例如需要数据输入的专门软件或者任务。

对清单上的每项特性，至少都要有一项任务来运用。对于最重要的特性通常要有两三项任务，以防止出现有更多时间可以执行更多的情况、第一项任务难以完成或者无法提供有用信息的情况。

在合理范围内，人们还可以自行创建任务。开始可用性测试时，可以请参加者描述他们最近发现的使用你的产品可以解决的具体问题(即新用途)。到执行任务时，请他们使用产品，就好像他们在解决访谈开始时所描述的情况一样。让任务感觉真实的另一种方式是用现金。例如，有一家电子商务网站给每位可用性测试参加者 50 美元，并告诉他们无论买不买东西，他们都能得到这 50 美元(参加测试的现金奖励另外计算)。如果他们搞不清找什么东西，这种方式能更好激励他们找到他们真正想要的东西。

测试过程从本质上属于定性过程，但也可以给每项任务增加一些基本的量化指标(有时也称为性能指标)，以便能调查不同设计的相对效果或者比较竞争对手产品。基于互联网的一些常见量化指标如下：

- 人们完成任务的速度
- 人们犯错的次数



- 人们从错误恢复的频率
- 成功完成任务的人的数量

这类数据收集给出的结果在统计学上无法使用,也无法超越测试过程进行总结,这类数据只能帮你设定任务需要花多长时间完成的想法。因此,往往应该采用相对数量规模,而不是具体时间。

对于餐叉的例子,可以采用以下任务,与前面所列的需求相对应。

餐叉任务	
特 性	任 务
搜索引擎 人们能否用搜索引擎找到特定餐叉?	<ul style="list-style-type: none"> <li>● 路易 14 餐叉风靡一时,您决定要买一套。您能在目录中找到全部路易 14 餐叉设计清单吗?</li> </ul>
目录导航 如果人们无法准确知道要什么东西,他们是否会使用目录导航来浏览?	<ul style="list-style-type: none"> <li>● 有一天您在商店橱窗看到了这把好餐叉(拿出照片)。请在目录中找一个跟它非常相近的设计。</li> </ul>
购买机制 能否对单个餐叉或者成套餐叉都起作用?	<ul style="list-style-type: none"> <li>● 假设您真的很喜欢我们刚才看的一个设计(挑一个),您决定买一打这种款式的餐叉。您会如何购买?</li> <li>● 假设一个月后,您喜欢这些餐叉,但弄坏了一把。从网站首页开始,您将如何购买到替代品?</li> </ul>
每周餐叉 人们看见了吗?	<ul style="list-style-type: none"> <li>● 这个任务有点困难。要把看见变成任务不太容易,但可以建立能引起讨论的情况。几个月后,您又要买餐叉,这次是用作礼物。您首先会看哪些地方来找到有价值的有趣餐叉?</li> <li>● 请人们不要看界面来画出界面或者描述界面,能发现人们觉得哪些东西值得记住,这些东西会和人们看到的东西紧密关联。 [关掉显示器]请根据记忆,画出我们刚才看过的界面。</li> </ul>
愿望清单 人们是否知道它的作用?	<p>购物时,您希望能够有一份感兴趣设计的清单,也许以后会购买一个,但现在您想要记住有趣的东西。您会怎么做?</p> <p>[如果他们自己找不到,应该指给他们看,并问他们是否明白它的作用,他们会怎么用。]</p>





### 提示：

传统意义上，人们将可用性测试任务描述为可以计算时间的小型不连续行动（例如“保存文件”）。可以收集大量这类任务的时间，并且和预定理想时间进行比较。对于有固定长期用户的专门应用工具的低级别可用性任务，这种方法能发挥作用。而通过本书描述的较大颗粒任务能容易分析万维网上的任务，因为使用网站和经验较少的人使用专门软件常常不相同。此外，高性能的时间计算能转移对立即理解和满意度问题的注意力，而后两者对网站设计的作用比应用设计的作用更大。

列出清单后，接下来需要计算任务时间并检查任务。可以自己来，也可以请不接近项目的人来做。可以作为测试前演练的一部分，但如果有可能，总是应该自己来走查这些任务。

此外，随着测试进行，应该不断评估任务质量。根据用来创建任务的相同准则，看看任务是否能满足这些准则。在每个环节之间想想任务效果，并与主持人和观察员进行讨论。在测试中大幅度改变任务不是件好事，但可以在测试之间进行小幅调整来提高任务准确性，同时要准确记录每次的变化情况。

### 编写剧本

有了任务，接下来开始编写剧本。剧本有时称为“用户协议”，有时称为“讨论指南”，但其实就是主持人要遵循的剧本，以保持访谈一致性，并能完成所有事情。

剧本分为三部分：介绍和初步访谈、任务及总结。以下实例来自一家电子商务网站的典型 90 分钟可用性测试，参加测试的人从未使用过要检查的网站。大约三分之一的剧本都用来了解参加者的兴趣和习惯。这些主题通常属于情境调查过程或者焦点小组系列，但可用性测试也应该包含一些这类调查。另外三分之一关注任务表现，其中会使用最重要的特性。最后三分之一和管理有关。

#### 介绍（5~7 分钟）

介绍是一种打破沉默并将评估人引入一些情境的方法。这个过程能把评估人带入角色并使他们在整个测试过程感到自然。

[关掉显示器，关掉视频设备，重新设置计算机]

您好，欢迎，感谢您能来。您好吗？（您觉得这里不错吧？对 NDA 有没有问题？等等。）

我是\_\_\_\_\_。我是来帮助\_\_\_\_\_了解他们的一个产品能对受众发挥多大作用。这是\_\_\_\_\_，他会观察我们要做的事情。今天请您来是想看看对产品的看法：哪些地方对您有用，哪些没



用等情况。

评估会持续约一小时。

我们将对今天的评估情况进行录像，但录像只用于分析工作。主要原因是我不能只坐着记笔记，我可以专心和您说话。有一些开发团队成员，几个其他人，还有我会看录像。录像严格用于研究，不会用于公共传播、宣传、促销或者圣诞联欢会。

如果有视频设备，总会有些突兀，而且有些吓人。说出来能缓解很多紧张气氛。同样，如果有双向镜子，应该说出来——包括镜子后面有人的事实——也有助于减轻大多数人的焦虑感。说出来之后就不用再提及。它迅速融入背景，再次提及会造成注意力分散。

还应该注意用对话风格来编写剧本。没必要逐字阅读，但能提醒主持人保持随意的访谈口气。此外，每节都应该有一段时长，这样主持人能知道每节应该花多少重点。

就像我说的，我们希望您能帮助产品开发。产品是为您这样的人进行设计，所以我们真的想知道您对产品看法，哪些东西起作用，哪些不起作用。目前产品处于开发早期阶段，因此，您看到的不会都有用。

不管产品团队说产品处于什么阶段，如果要对产品进行可用性测试，都应该说产品处于早期阶段。告诉评估人产品还在开发早期阶段有助于让他们放松，他们能有更多授权能对产品整体发表意见。

今天要做的事情是这样的：我们开始会谈谈您对互联网的使用，您喜欢什么，遇到过什么问题等。然后我会给您看一个产品，这个产品一直在开发，并还会请您用它做几件事情。然后进行总结，我会问您几个问题，然后就结束了。

您有没有疑问？

要清楚说出整个过程，评估人能预测要发生的事情，同时他们能有更多角度来了解过程。





现在我要向您读一下保密声明。我对访谈的每个人都会读一下。您会知道参加此类研究应有的权利。

作为参加研究的人，具有以下权利。

- 可以随时终止。
- 可以随时提问。
- 可以随时离开。
- 不会受到欺诈。
- 回答会保密。

开始以前，您有疑问吗？

下面我们就开始。

保密协议告诉评估人他们的输入有价值，他们在一定程度上可以控制过程，而且不会有任何欺骗性的东西。

### 初步访谈(10~15分钟)

初步访谈用来为参加者之后发表意见提供情境。它从一般性问题开始，它可以逐步进入用户的使用体验，并且缩短与用户之间的对话，直接谈论产品设计这个话题。对于从未参加过可用性测试的人来说，初步访谈向他们提出“简单”问题，能增加了他们的好感，建立信任，并让他们了解过程。

这种情况下，初步访谈还能广泛调查人们的背景和习惯。提一半问题很常见，花5分钟时间建立初始情境设置访谈，不用花10~15分钟。

[打开视频进行录像]

您一周一般花多长时间上网？

其中有多少时间用于工作，多少时间是个人使用？

除了电子邮件外，您最经常在网上做什么事情？

您在网上购物吗？您购买过什么东西？您多长时间会在网上购物一次？

您有没有在网上做研究，而最终在实体店铺购买？是不是有些种类的商品发生这种情况的次数比其他种类要多？为什么？

有没有东西您不会在网上购买？为什么？



如果适用，应该问问人们的线下习惯，然后再重新讨论在线环境。把他们说的线下情况和你观察到的他们在线上做的事情进行比较，能深入了解人们对界面的认识。

稍微换一下东西，您会购买银制餐具吗，除了在网上外？多长时间购买一次？

您在网上购买过吗？为什么？

[如果购买过]您有没有喜欢的购买银制餐具的网站？

[如果购买过]您最喜欢[网站]的什么地方？有没有什么地方经常会搞得您烦？

评估指南 (3 分钟)

不要让评估人感觉到产品的轻视。任何产品背后的目的是让产品成为从属工具，但人们已经严重受限于设计拙劣工具和傲慢公司，从而认为责任全在于自己。尽管很难打消软件的不安全感，但评估价格能帮助评估人感到自然，无所顾忌地叙述自身体验、正面、负面评论和他所有的感受。

稍后我会请您打开显示器来看看产品，但请让我先告诉您如何使用。

使用产品时，请记住一件最重要的事情，我们是在测试界面，而不是测试您。您做的事情绝对不会是错的。请记住这点。如果发生损坏、错误、奇怪、尤其是迷惑的事情，肯定不是您的错。但我们想知道发生了什么。因此，不管什么时候发生问题，都请告诉我们。

同样，如果您喜欢什么地方，也请告诉我们。即使是特色，颜色或者布局方式，我们也想请您说说。

尽量坦率。如果您觉得有什么东西很可怕，请告诉我们。不要觉得不好意思，这不会伤害任何人的感情。因为产品是为您这样的人设计的，我们真的想知道您的准确想法，哪些地方有用，哪些没用。

此外，使用产品时，我会请您大声说出想法。这样我们能知道您对所做的事情有什么想法。只是详细叙述您正在做的事情就行，告诉我您正在做的事情，为什么要这么做。





请人们在思考的时候说出正在思考的东西，这是有效可用性测试的重要组成部分。要在一开始就说明这种方法，但在实际访谈过程中还要强调这种方法。

您能明白吗？是否有疑问？

请打开显示器[或者是打开笔记本电脑]。开机时，可以把键盘、显示器和鼠标放到使用方便的位置。

第一印象(5~10分钟)

产品第一印象对网站的重要性令人难以置信，总是应该明确测试第一印象，而且很快完成。问人们在看什么地方，在看什么内容，这样既能发现界面上突出的东西，又能深入了解页面载入和渲染如何影响焦点和注意力。

访谈从打开浏览器开始，但应该把浏览器设置为空白页。加载顺序会影响人们看到网页要素的顺序，而且往往会影响他们对这些要素的关注程度。了解页面加载过程中他们的注意力重点有助于解释为什么有些内容受到关注多，而有些少。

现在电脑已打开。请大家从“Favorites”菜单中选择“Forks”。

[快速]您第一眼看到了什么？第二眼看到的是什么？您看到页面时首先想到的是什么？

[1、2分钟后]这个网站有什么用途？

您对这个网站有兴趣吗？

如果是头一回访问，您接下来会做什么？会点击什么？想对什么进行调查？

此时剧本可以往两个方向发展。可以是一个基于任务的访谈——用户可以立即开始执行任务——可以是混合体，一半是基于任务的访谈，另一半是基于观察的访谈。

基于任务的访谈关注少量的具体任务或者特性。混合访谈可用于首次测试，以及开发周期早期的测试。混合访谈中，评估人浏览界面，看看界





面主要组成部分的每个要素，迅速发表评论，然后执行任务。

基于任务的访谈可以如下所示。

#### 任务(20~25分钟)

现在请您在界面上做几件事情。像平时一样操作，请在操作的同时说说您的看法。

请您完成这几件事情。[发任务列表]

第一个场景如下：

任务1 此处填写任务描述\_\_\_\_\_

[阅读第一项任务，发任务1描述纸]

第二件想请您做的事情是\_\_\_\_\_

任务2 此处填写任务描述\_\_\_\_\_

[阅读第二项任务，发任务2描述纸]等等。

如果有方法可以远程观察参加者，有时可以请他们自己完成几个任务，而主持人不用待在房间里。这样能获得人们在没有知识来源的情况下如何解决问题的有价值信息。此外，主持人可以利用这段时间和观察员讨论一下测试。主持人离开房间时，要再次强调评估人叙述所有想法的必要性。

列一张需要探索的具体问题清单，有助于确保所有重要问题都会得到回答。只要在访谈合适的时候，主持人要随时提出要研究的问题。

#### 探究问题(在合适的时候进行研究)

- 导航要素的名称是否有意义？
- 界面要素的功能是否和评估人预期的一样？
- 有没有没有意义的界面要素？
- 什么东西吸引了评估人的注意？
- 某项特性里面最重要的要素是什么？
- 有没有地方评估人需要额外信息？
- 评估人对特定要素/屏幕行为/内容期望是什么？

混合型访谈如下例所示。从快速的一般任务入手，看看人们与产品的体验，然后再仔细检查界面。





### 首批任务 (5 分钟)

现在请大家在界面上做些事情。

就像您平时操作一样，请叙述出您操作时的想法。

第一个场景如下：

任务 1 此处填写任务描述

[阅读第一项任务]

### 界面浏览(10 分钟)

现在来看看界面，一次看一个东西，请说说您希望每个东西能做的事情。

[查看

- 首页的大部分
- 目录页
- 购物车页]

[关注

- 网站导航要素
- 搜索要素
- 主要特性的标签及行为
- 模糊要素
- 期望]

细查每个要素[在合适的时候，对每项有意义的要素进行提问]：

- 请用几句话说说您对这个东西的想法？
- 这个[标签，名称]是什么意思？
- 您觉得从这会去哪里？
- 不要点击，您觉得会到达什么类型的页面？应该有什么内容？页面看起来是什么样子？

细查每屏[在合适的时候，针对每屏提问]：

- 屏幕上什么东西对您而言最重要？
- 是否缺少您需要的信息？
- 填写表后，下一步您会做什么？



- 从网站首页您会去哪里？您会点击什么？
- 您怎么去[其他一些主要板块名称]？

#### 任务(10 分钟)

第二件请您做的事情。

任务2 此处填写任务描述

[阅读第二项任务]

最后一件请您做的事情。

任务3 此处填写任务描述

[阅读第三项任务]

所有任务完成后，信息收集和访谈的核心部分就结束了。但观察员和分析师应该总结本次讨论的重点和关键性问题，这对于他们来说是有意义的，对产品进行发散讨论可以发现评估人的内心想法，并能产生一些好想法(或者也可以用这段时间请人们画出他们记得的界面样子，同时主持人离开房间问问观察员最后是否还有问题要问参加者)。

#### 总结及发散头脑风暴(10 分钟)

请关掉显示器，我们用几个问题来总结一下。

##### 总结

如果要用几句话向电脑水平和网络经验与您差不多的人介绍这个产品，您会怎么说？

这项服务有趣吗？您会使用它吗？

您会把它推荐给其他人吗？为什么推荐/为什么不推荐？

您能总结一下我们讨论的内容吗？请说说产品三大优点和三大不足。

##### 发散头脑风暴

好了，我们已经知道了这个产品能做的一些事情，现在我们随便聊一会。如果不考虑实际情况，有没有什么您希望这类系统能做但现在还没有的任务或特点？您是不是说过：“我希望有些工具能为我做 X 事情”？是什么事情？

您还有没有问题？有没有建议？





非常感谢，如果您回去时有什么想法，明天、下周都行，请随时给我发邮件\_\_\_\_\_。[递名片]

最后，应该获得一些对测试及时间安排的反馈。

原型相关问题到此结束，还有最后一个问题：

无论是时间安排，还是实施方式，您对如何更好实施测试有无建议？  
非常感谢。

[关掉视频]

在用户研究的每个阶段，产品利益相关者都应该参与编写测试剧本内容。完整剧本草稿应该经过利益相关者的核实，以确保事项的优先级别及技术报告的准确性。第一份草稿应该在测试开始前至少一周就提交给开发团队。第二个版本加上开发团队的意见至少应该提前几天给他们看。

### 实施访谈

进行用户访谈有两个目标：从评估人那里获得最自然的答复，并获得最完整的答复。用户访谈环境中的所有东西——从物理空间到提问方式——都是为了这两个目标。

#### 物理布局

尽量让物理布局看起来不像实验室，而像产品设计后的使用环境。如果是供工作时使用的产品，就把产品放在办公环境中测试，最好能有窗户。如果是供家庭使用，就把产品放在家庭办公室中进行测试。如果不必面面俱到，只需精心挑选几个道具，产生适当感觉就可以了。例如，对于家庭办公，软照明，办公桌的桌布，即刻就能让测试看起来更随意。

但通常情况下，可用性测试必须在预定的会议室或者租用的实验室中进行，而那里的很多条件都无法改变。此时，应该确保空间安静、整洁、并免打扰。

如果有可能，每次访谈都应该录像。理想情况下，应该采用视频扫描转换器(一种将计算机视频输出转换成标准视频的设备)和视频混合器来



建立“画中画”过程，一幅画面显示人，另一幅显示屏幕(图 10.3)。初次访谈，摄像机摆放的位置要能看见评估人的脸和手，这样能看到用于访谈任务部分的屏幕。主持人没有必要出现在屏幕中。

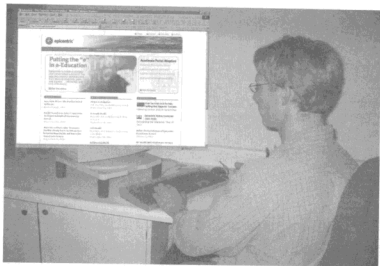


图 10.3 画中画视频记录

音频准确、清晰至关重要，摄像机要带内置良好的麦克风，过滤外部噪音，也可以购买佩戴式麦克风，评估人可以夹到衣服上，也可以采用可以贴在显示器上的小麦克风。能捕捉评估人声音，但无法一直捕捉到主持人的声音，这是佩戴式麦克风的缺点。理想情况下，可以采用两个无线佩戴式麦克风和一个合并两个声源的小型混音器，或者采用单个外置麦克风，要能敏感到足以捕捉到双方谈话，而且不会录入外部噪音，摄像机内置的麦克风无法做到这一点。但这样一来设备就会很多。

如果没有双向玻璃镜的房间，也可以采用闭路视频。很容易实现，在隔壁房间放一根长视频线和电视(但房间的隔音效果要好，观察员可以畅所欲言，不必担心被在测试房间的人听到)。一轮典型可用性测试的最终布局如图 10.4 所示。



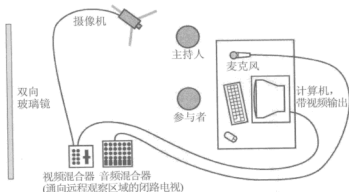


图 10.4 典型可用性测试设置

### 主持

主持人需要让用户感到舒适，并在合适的时候征求有用反馈，而不要突然中断用户的叙述，也不要改变其观点。第6章深入介绍了非引导性访谈的风格，所有用户访谈都应该采用这种风格。

除了第6章叙述的一般访谈风格外，所有访谈主持人还须注意以下几点。

- **细查期望。**参加者会点击链接，选择选项，或者在界面上执行任何操作，而在此之前，他们对将要发生的事情有一些想法。尽管对接下来要发生的事情不一定有完整想法，但他们总会有一定期望。执行动作之后，他们对动作效果预期就永远改变了。而在此之前要想知道他们的想法，只能在他们要开始执行任务之前请他们停下来，问问他们对结果的期望。例如，有个超链接，请评估人描述一下如果点击链接，他们觉得会发生什么事情，这样能了解很多他们对网站功能的心智模型情况。应该在一个动作之后立即提问：“这是您预期的吗？”这样能发现体验是否符合期望。
- **提出大量“为什么”。**在合适的时候提出简单、直接、没有偏见的问题也可以了解到态度、信仰和行为方面的情况。5岁的小孩一直都这么做：就是一遍又一遍地问“为什么”，逐步深入问题，却从来没想过他们觉得会有正确答案。例如，如果有人说“我不



想做这些事。”询问“为什么”能获得更好信息，而不会仅仅知道那人不做什么事情。

- **有时建议解决办法。**访谈时不要设计，但如果有特定想法(即使现有产品没有)能解决他们的问题，可以进行细查。这样能检查访谈人对问题的理解，也能健全检查可能的解决方案。例如，测试中有很多人说他们用 Outlook 和 Palm Pilot 来管理个人日程。他们对在线日程管理没兴趣，因为他们觉得会做很多重复劳动，即使认为日历基于互联网很方便，他们也没兴趣。主持人建议如果离线日程可以同步到网上，他们都感到兴奋，并表示如果具备同步特性，他们很可能会使用整个服务。
- **调查错误。**如果评估人发生错误，先等一等，看看他们是否已经发现犯了错误，然后马上调查他们的想法和期望。他们用哪种方式做事情？希望能做什么？期望怎么做？发生了什么事情让他们明白没有做成功？
- **探讨非语言线索。**有时，人们会对某种体验作出身体反应，但不会说出来。如果出现的东西令人惊讶、意外或者不愉快，有人可能会往后退一下，但不会说什么。同样，微笑或者身体前倾则可能代表着满意或者兴趣。观察这种行为，并且如果可以的话就采取后续行动。例如，“对话框出现的时候，您皱眉了。是不是有什么事情让您皱眉？”
- **保持访谈以任务为中心。**人们会自然偏离某些出现的话题。有人正在执行任务，他们可能会想探寻某个想法或者体验。要允许人们探寻体验，但同样要保持对产品 & 任务的关注。如果有人往后靠，手离开键盘，不看显示器，说话开始含糊不清，此时一般需要引入新话题或者回到手头任务上。
- **尊重评估人的想法。**如果人们跑题了，可以让他们跑一会儿题(也许一分钟左右)，看看他们是否可以自己结束。如果没有结束，应该把对话引导回手头任务或者主题。如果还是不起作用，就应该更明确地说：“很有意思，或者我们晚点再讨论，现在我们来

#### 提示：

本章中，我一直采用“评估人”和“参加者”来指评估界面的人，而没有采用“对象”、“测试人”、“豚鼠”等或者诸如此类的词。我是有意这么做的。招募来的人或者评估界面的人此时都是你的同事。要检查的不是他们，而是产品。很容易把测试搞成了心理学试验，但这实际并不是心理学试验。这是对产品进行引导性评估，而不是调查人性，在所有层面上都应该如此对待。





看看每周餐叉页面吧。”

- **专注个人体验。**人们有一种倾向，即理想化体验，并把自己的体验推论为其他人的需求或者他们未来的需求。但直接经验更多源于人们的实际态度、需求和行为，而且通常比期望更有用。如果彼得说：“我认为这对人有用”，应该问他是否对他有用。如果英戈说她理解了，但其他人可能不理解，应该告诉她了解她的看法非常重要，而不是了解怎么为其他人进行设计。如果汤姆说有个东西“可能将来有用”，应该问他现在对他有没有用。

### 管理观察员

要想关联测试发现并赢得开发团队的支持，最快的方法之一是邀请开发团队的成员观察测试。

#### 提示：

有些研究人员声称，在不损害观察质量的前提下，房间里可以有多位观察员。我还没有碰到过这种情况，大部分时间我也没有邀请室内观察员。或许可以在房间里安排一大堆观察员，而同时请参加者舒适而自然地执行任务，毕竟舞台演员一直都是这么做的，但是，我会通过避免提问尽量避免可能需要对人们的陈述做出解释的复杂情况。

如果有可能，应该邀请合适的工作人员实时观看可用性测试。对开发人员(或者甚至是产品开发副总裁)而言，用户错误使用界面、误解假设、在界面上什么事情也做不了，这些对他们更有启发。

让观察员参与的最佳方法是通过双向玻璃镜或者闭路电视视频传送。带够食物(比萨就行)。开发团队一边休息，在测试进行的同时一边进行讨论(同时不要忘记看看参加者的实际行为)。开发团队知道产品的内外两方面，他们能看到主持人和分析师都无法看到的行为和态度，这对于分析师和对于开发团队理解客户都是宝贵的原始资料。

如果既没有双向玻璃镜，也没有闭路电视传送，可以请开发小组成员直接观察测试。但每次测试绝不应该多于一位观察员。评估人在实验室中已经够慌的了，再加上有几个人坐在后面，有时画画东西，有时窃窃私语，即使最冷静的人也觉得恐怖。如果观察员在房间里，应该介绍一下他的名字，因为这能确认其存在，并在测试过程中给观察员一个角色，而不是“坐在角落里默默看着我的家伙。”

要告诉观察员可以接受哪些行为，并设定他们对测试过程的期望。





## 可用性测试观察员指南

(1) **仔细听。**立即对观察的东西进行讨论很诱人，但首先一定要听人们真正在说的内容。可以随时讨论你看到的东西，但别忘了仔细听。

(2) **可用性测试在统计学上没有任何代表意义。**如果4个人里有3个人说了什么，并不代表75%的人都会这么认为。它的确说明有人会这么认为，但没有数量上的意义。

(3) **不要字字当真。**只是几个人的观点。如果他们有好想法，那很好，但除非有重要证据能证明，否则请凭直觉来判断重要性。因此，如果有人说我“我讨厌绿色”并不是说要改变颜色(但如果每个人都说我“我讨厌绿色”就说明需要进一步研究)。

(4) **人是矛盾的。**听听人们如何思考主题，他们用什么标准得出结论，不一定听他们说具体愿望。人们可能没有意识到两个愿望不可能同时实现，或者人们根本就并不在乎。要准备接受偶尔的无聊或受混淆。人们的行为并不总是有意义或者有见地。

(5) **不要期望革命。**如果得到一两个好想法，可用性测试就算是成功的。

(6) **注意观察。**观察人们做的事情和注意到的事情，同样要注意观察没有做的事情和没有注意到的事情。

如果是内部观察员，应该增加以下指南。

(7) **如果主持人提供明确机会，应该随时提出问题。**提出的问题不要以任何方式对产品产生价值暗示。因此，不要问“这是同类产品中最好的吗？”而应该问“是否还有其他产品能这么做？您对这些产品有什么看法吗？”

(8) **不要说你和产品有直接关系。**如果人们认为同一间屋子中的人与产品并无太多关系，他们会更愿意对产品效果发表意见。

如果观察员是开发团队成员，应该鼓励他们一直要观察完所有参加者，再进行总结和设计解决方案。一看到问题，人们就会自然想要开始解





决问题,但在花精力解决问题之前应该要知道问题的情境、规模和普遍性。在确定问题的所有方面之前,一般不要推荐解决方案设计。

### 提示与技巧

- 应该总是提前一两天进行访谈演练。为真实测试准备好一切,安装好所有合适的硬件和原型。然后找一个你会进行再次访谈的人。利用这段时间来确保剧本、硬件以及所有工作任务都能作用于设计。应该检查整个访谈过程,然后给评估人买午餐。
- 每场测试之间应该重启计算机,并重新设置实验室。清除浏览器缓存,重设历史记录(所有链接都是新连接,同时删除 cookie),重新启动浏览器,保持空白页(大多数浏览器都可以设定成首页为空白页),确保每位用户的测试环境都相同。拿走前一位测试者的笔记和任何纸张,并关闭监视器。
- 如果可能的话,应该同时提供苹果机和 PC 来进行可用性测试,这样评估人可以选择用起来更舒服的电脑。甚至还可以在筛选文档中包括这个问题,能提前知道参加者通常使用哪种计算机。
- 不要在测试时做大量笔记。应该专注用户正在做的事情,并探究特别行为。参加者也不会把他们的行为和匆忙记笔记联系在一起,而且他们常常会觉得这说明他们做错了事情。
- 测试结束后立即做笔记,把有趣行为、错误、喜欢和不喜欢都写下来。然后立即和观察员讨论 10~20 分钟,并然后用笔记形工记下他们的观察结果。

## 如何分析

有些事情会很明显,但仍有必要进行正式分析来获得根本原因,并从访谈中获取最大价值。分析结果分为三个阶段:收集观察、组织观察以及从观察中提取趋势。



## 收集观察结果

提示:

主持人和分析师指单独的人,但实际上两个角色常常由同一个人兼任。

有三组观察需要进行收集:主持人的观察、观察员的观察和分析师的观察。

收集主持人和观察员的笔记非常简单。拿到他们的笔记(或者副本),并请他们走查一遍,请他们解释每条笔记的意思。此外,针对其他不属于他们的笔记访谈他们。最后一次测试后的几天或者几小时内的这些笔记通常都代表着大型观点。

分析师的笔记最重要,也是最耗时的数据收集过程。分析师至少要看4盘录像,并记下所有发生错误、混乱和评估人对产品或者其特性发表意见的情况。分析师应该记下评估人对哪些特性有问题,他们在什么情况下碰到这些问题,同时要详细说明问题。随着人们的行为和期望形成,可能会在这一阶段发现产品的大多数可用性问题。

定量信息无法推广到整个目标市场,但往往能用来总结和比较行为(但这样会充满潜在问题,因为阅读报告的人会把大量无意义的数字当成某种绝对真理)。要收集定量信息,首先要为每个问题确定度量范围,分析团队的每个人都要同意。不要用秒表,只要记住确切数字。可用性测试中小样本量表现出来的统计学错误会超过秒表的精度。最有用的指标就是最通用的指标。英国用户体验设计和评估咨询公司 Flow Interactive([www.flow-interactive.com](http://www.flow-interactive.com))采用以下范围来度量人们执行任务的时间:

- 0—失败
- 1—以缓慢迂回方式获得成功
- 2—成功有点慢
- 3—成功很快

大多数情况下,这种精度就足够了,因为数量级的度量足够做出关键比较。每个规模应该有3个或者5个级别(不要使用2个、4个或6个级别,因为很难找到中间值;不要超过5个级别,因为往往会造成混乱),另外还有1个单独值表示失败。





给每位参加者建立一个要收集的任务指标网格。一边看录像带，一边在每格中标记严重程度(在适当的时候，可以采用开发团队定义代码缺陷而使用的相同语言和尺度来定义严重程度)。对于餐叉任务，下表能反映一个人的表现情况。

马龙的任务表现			
用户：马龙	阅读时间	错误	完成时间
<ul style="list-style-type: none"> <li>找到路易十四餐叉</li> <li>购买替代品</li> <li>找到类似餐叉</li> </ul>	<ul style="list-style-type: none"> <li>1</li> <li>3</li> <li>1</li> </ul>	<ul style="list-style-type: none"> <li>3</li> <li>1</li> <li>2</li> </ul>	<ul style="list-style-type: none"> <li>1</li> <li>2</li> <li>0</li> </ul>
数字说明	<ul style="list-style-type: none"> <li>0—未读</li> <li>1—阅读很慢</li> <li>2—阅读有点慢</li> <li>3—阅读很快</li> </ul>	<ul style="list-style-type: none"> <li>0—因为发生错误而失败</li> <li>1—很多错误</li> <li>2—有些错误</li> <li>3—极少错误或者没错误</li> </ul>	<ul style="list-style-type: none"> <li>0—失败</li> <li>1—以缓慢迂回方式获得成功</li> <li>2—成功有点慢</li> <li>3—成功很快</li> </ul>

如果要分别做最终分析，应该对每项指标创建一张表来总结整体用户体验。完成时间指标表可以如下所示。

任务表现时间指标						
	马龙	伊娃	马克	巴布	乔恩	平均值
找到路易十四餐叉	1	2	1	0	2	1.2
购买替代品	2	3	2	1	1	1.8
找到类似餐叉	0	0	1	1	0	0.4

从绝对情境而言，平均值并无意义，但它是比较任务、比较设计的一种方法。

记下特性要求，还有评估人说的话，特别是那些包含特定行为的话(例如，“我不明白“Forkopolis”是什么意思，所以我会点击那里”)。特性要求说明评估人无法用其他方式来说明问题。但也可能是对相同问题的创



新解决方案，因此不管什么情况都应该收集起来。

### 2x 视频座很酷

为了加快视频检查进程，我建议采用视频座(或者数字视频播放器)，可以用1.5或者2倍的自然速度来播放视频和音频。此时仍然可以听得懂讲话内容(但听起来会有点愚蠢，因为人们的声音听起来像金花鼠的声音，除非声音是从高音降低，索尼的专业级视频硬件能做到这点)，通过这种方式，回顾磁带会快很多。

如果时间和预算允许，应该对整个环节进行文字录入，但它应该只能作为观察磁带的辅助方法，因为它忽略了能真正澄清一些情况的声音变化和行。例如，有5秒钟暂停，而此时评估人把鼠标指针划过屏幕上的每一个视觉元素，寻找点击的地方，不足以表达出他的发言“啊哈！在这里！”

## 组织观察结果

首先应该通读一次所有笔记，能对材料有大概感觉。寻找由于常见问题而造成重复的地方和东西。

接下来把所有观察结果放在一起(从字面上理解，或者只形成一个大文档)。用文字处理器打开一个单独文件，看一下每个观察结果，并和其他相似的观察结果合并成新文档。相似可以是表面相似(“不理解术语”)、特性集相似(“购物车问题”)，或者根本原因相似(“信息架构让人混淆”)。根据最广泛的根本原因对观察结果进行分组。找出引言，并根据它们最能解释的原因进行分组。

## 提取趋势

对所有观察结果分好组，接下来检查所有组，巩固分组，把无关联话题组分开。舍弃只有一两条观察结果的组。对每组尝试用一个短句来分类问题，再加上几句话来充分描述现象。尽可能解释其根本原因，把对现象的解释和对原因的假设分开。关注说明问题，它对用户体验产生的直接影

提示：

和分析情境调查信息或者焦点小组观察非常相似，可以和开发团队(以及其他利益相关者，酌情而定)一起完成组织可用性测试信息并提取趋势。小组可以利用收集到的知识来充实对问题的理解，并开始着手解决。





### 忠告：

用户对严重程度的衡量很容易与项目开发的优先级别混淆。通常都不恰当。对用户成功使用产品最重要的东西不一定对产品成功贡献最大。产品团队要从用户角度知道问题严重性，并用来确定项目的优先顺序，但两者不是一回事。

响，以及发生问题的地方。提出解决方案时应该非常小心。开发团队最终可以更了解技术和产品中的假设，他们还要负责独立出根本原因并寻找解决方案。你提出的建议应该作为能找到解决方案的指南，而不是规定必须做什么。

应该从用户角度说明问题的严重程度，但不要给观察加上严重程度数值等级。如果希望或者要求对观察特征进行速记，应该从它们对用户体验产品产生的效果进行分类，而不是任意给它们加上严重程度。这种尺度会“妨碍行动”，“造成混乱”，“不符合期望”，“觉得没必要”。

做完所有这些事情，将得到一份清单，内容包括观察结果，造成现象的假设，以及加强和总结观察结果的简介。准备向团队提交结果！第 17 章将讨论有效提交结果。

## 实例

以下内容是针对 Webmonkey 另一个原型测试而为网站开发团队准备的简短报告。测试建立在网站之前的测试基础上，并关注对网站首页的变化以及对网站各板块的重新命名。

### 执行摘要

5 位互联网开发人员看过了 Webmonkey 4.0 的功能原型。他们都表示喜欢，尤其是教程和配色方案，但有些组织方式让他们感到模糊(具体而言是 Category 和 Tutorials 板块之间的差别)。新的文件夹式导航采用的暗喻，每个人都能看得懂，他们希望每个页面上都能采用。每个人都看到了 Cool Tools 板块，但觉得是个广告，因此将其忽略，他们喜欢 Inspiration 板块，但从长远看，他们期望它除了动画外，还能有其他东西。

最终，有几个人说如果能不带偏见地把 Webmonkey 链接到好的、有用的外部内容会很酷，因为它有用，而且能强调 Webmonkey 的草根血统。

执行摘要对于传达结果非常有用。产品开发副总裁可能从未读过报告，但很可能会读几段给出 5 万条可用性测试结果的内容。如果在电子邮



件中附加报告，执行摘要也应该包含在电子邮件中，而报告的剩余部分——包括执行摘要——都应该作为附件。

### 程序

五位花费大量时间开发网站的人受到了邀请。首先问他们一些初步问题、一般互联网使用情况、他们去哪寻找开发人员信息(包括网上和一般情况)。然后给他们看看网站原型，并请他们仔细看看，并集中讨论具体细节，包括文件夹式导航和 Cool Tools 板块。请他们说出对首页的想法，然后请他们看一条头条新闻，请他们谈谈对界面的体验和对内容的想法。接下来请他们寻找一些具体内容，用来衡量他们对布局网站的理解。最后，请他们总结，并提出一些开放性问题，然后结束测试。

快速说明测试过程的程序能阐明过程，并为报告接收人理解结果提供重要情境。

### 评估人背景资料

#### 迈克尔

迈克尔每周花在互联网上的时间超过 40 小时，其中 20 小时花在制作网页上，包括设计、编程和制作。所有开发网站中，他喜欢 Webmonkey，因为它“内容广泛”。他还定期阅读“Flash 地带”，因为这块内容可以提供他没有出版的教程。而对 CGI 工作，他访问另一个网站，“CGI 资源”(是指“CGI 地带吗”？访谈没有说明具体网站)。

#### 约翰

约翰每周花在互联网上的时间有 30 个小时，一般都与工作有关。他每周至少花 10 小时做网站，包括设计、标记和代码。他利用参考书和 Webmonkey 获取网络相关的技术信息。他还访问“SGML 大学”网站，但从没去过 builder.com。他访问这些网站时大部分时间都带着具体问题。一般情况下，他希望开发者网站最好能根据主题来组织内容。

#### 大卫

大卫每周花在互联网上的时间有 20~30 个小时，75%的时间和工作有关，5%~10%的时间用于 Web 开发，大部分是设计。他的主要技术信





息来源是 Webmonkey 和学校笔记。他从未访问过 builder.com，访问 Webmonkey 了解技术更新并回答具体问题。

[余下背景资料略去]

评估人背景资料既能帮助报告阅读人理解人们做出陈述的情境，同时无法观察测试的人也能通过个性化的方式了解参加者。和第7章创建用户背景资料一样，评估人背景资料有助于个性化产品用户的抽象概念，并使结果更直接。

### 观察结果

#### 一般观察结果

(1) 和其他东西相比，人们更喜欢教程(tutorial)。所有评估人都被教程所吸引，有时会排斥其他内容。在被问到接下来会点击哪里时，评估人往往会首先提到教程板块。教程板块也是人们喜欢去寻找一般资料的地方，即使 Category 板块中有更广泛的内容。

(2) 几乎人人都说喜欢配色机制。在提及这个问题之前，就有很多评估人主动说他们真的喜欢首页的配色机制。

(3) 人们去开发网站时一般都带有具体问题，他们不是要去看“最新信息”。被问及去 Webmonkey 类似网站是否要看最新技术或者获取具体问题答案的时候，人们一般都说是获取具体问题答案。

同样，被问及喜欢怎样浏览 Webmonkey 类似网站时，人们说他们就喜欢搜索，而不是浏览，因为搜索能让他们更接近要找的具体信息。

#### 特性

(1) 人们在 Category 板块和 Tutorials 板块中找到的内容会有混淆之处(在较小程度上，Tutorials 和 Guides 板块也存在混淆)。部分是因为 Category 名称含糊不清，部分是因为两个板块中的类似——但不完全一样的——标签，因此人们对会在每个版块中找到的内容感到困惑不解。







只要有可能，务必尽量提供屏幕截图。

(2) Cool Tools 板块几乎每个人都能一眼见到，但被当成大广告，大多数人都把它忽略了。他们看起来都没有注意到，直到有人告诉他们 Cool Tools 板块中有非广告内容。多数人都指着图片和价格来说明为什么他们觉得是广告内容。

(3) 有几个人看到了 First Time Here，也喜欢这个想法。

(4) 人们不知道可以访问 Backend 找到 CGI 主题内容，也不敢肯定能找到什么内容。有个人提到他喜欢把这块内容称为“服务器相关”或者类似名称。

(5) 人们没有注意到首页上会首先跳出的参考框，而发现的时候，他们又不知道其内容和访问左边空白处文件夹内容的关系。但大多数人都认为它带的内容是个有用工具(除了“国际标准化组织实体”)。有几个人建议把它和左边空白处的文件夹放在一起。

### 导航

(1) 每个人都能理解首页上的文件夹暗喻。

(2) 左侧导航的内容和表现形式不一致造成了一点混乱。很多人提到他们对左侧导航从首页变成子板块和教程很惊讶。有几个人提到他们喜欢能延续首页的文件夹暗喻。

(3) 人们一般都能理解在页面顶部的类似路径名的面包屑导航，但不是人人都注意到了这一点。人们从首页直接跳到“教程”时(人们预期路径会像“home/tutorial/javascript”)产生了最大断层，而实际上是“home/categories/javascript/tutorial”，这不符合他们的期望。

### 命名

(1) Category 这个名称的意思不清楚。人们不敢肯定 Category 指什么，还有个人甚至都没有看出来它能点击。





(2) Guides 作为板块名称不清楚。大多数人都搞不清楚 Guides 板块和教程。

(3) 同样的，eBiz 也不清楚。没有问每个人这个问题，但有几个人不知道点击后会出现什么内容。

(4) Heard on the Street 的意思含糊不清。不让参加者看到，请他们定义该板块会是什么内容，和其他板块有什么区别，大多数人都说，该板块的内容应该是“最近内容”或者强调一些开发新闻。

### 结论

教程概念具有的吸引力，加上 Category 名称模糊不清，造成人们会频繁地完全忽视 Category 板块。

很多混乱来自单个词名称的意思含糊不清。Category 和 Guides 在人们看到板块后能充分表达出板块内容，但在人们看到之前几乎不能提供什么信息，因为从字面上看意思太宽泛。因此，板块命名(也许网站上的所有东西)都必须要根据用户脑中的情境来决定。回想起来，看起来完全有道理的东西在定义产生之前可能会令人困惑、模糊不清。

### 引言

#### 迈克尔

“你知道功能就在那儿，就是想知道如何把它放在一起。”

“我从这开始的时候，这是个起始网站，在这方面它做得很好，但是网站貌似就停留在那个层次，所以我离开了。”[回复: *Builder.com*]

“我看到了 Cool Tools，我觉得可以是 GoLive，可以是 Dreamweaver，可以是其他东西。”

“如果从两者中拿掉一个，会更容易[区分两者]。”[回复: *tutorials vs.categories*]

#### 约翰

“它很打眼，而且看起来也不像让人讨厌的《连线》(Wired)杂志。”

“假设我要在这里找到一些问题的具体答案，我会去 Tutorial 板块。”

“Categories 可以是任意东西，这些只是子主题。”



“我喜欢在首页看到最新教程[而不是通过启发]”。

[剩余引言略去]

评估人实际所说的寥寥数语往往能比长篇大论的解释更清楚地说明观点。读者能看到你用它们来说明的模式。在每个观点旁边放上模式，有助于加强每个要点。如果马上提交所有要点，它们能传达可用性测试的感觉。

可用性测试是用户体验研究的“驿马”。它可以快速实施，而且成本低廉，并能提供大量可以立即采取行动的信息。很多情况下，可用性测试都只被当成一种用户反馈形式，但如果使用得当，就是无价工具。



## 第 11 章 调 查

确定理想目标受众,是否就代表你知道真实用户是谁呢?你可能知道了网站哪些部分正在被使用,人们抱怨了什么问题,这些情况甚至可能会说明实际受众符合你预期这一事实,但你有多肯定?

定性方法,比如焦点小组、可用性测试和情境调查,能表明人们使用产品时为什么会做现在的事情,但不能准确地从一般用户人口中区分出用户的特点。只有定量技术能预测有多少用户是青少年,或者他们是否想要你考虑开发的特性。了解受众的构成表明应该对哪些人集中进行定性研究,而且更重要的是根据什么特质来定义受众。

谁在使用网站?他们是你预期的目标受众,还是完全不是你预期的人?他们看重服务的哪些地方?他们使用服务是如你的预期,还是另有原因?除非能得到大部分受众的意见,否则无法知道产品受欢迎的根本原因。

**调查**是发现用户是谁和他们有哪些意见的最佳工具。调查是一组问题,创建结构化的方法请广大群体描述自己、兴趣和偏好。有了调查结果,就可以采用统计工具来检查受众,发现其广泛特点,还可以提炼出感兴趣的子用户群体。如果正确完成,调查能产生更精确的用户背景资料,相较于其他定性研究方法或者非直接用户行为分析方法(如日志文件分析)。

调查可以回答如下类似问题。

- 用户年龄有多大?
- 他们通过什么方式使用互联网?
- 用户人口是同一类型,还是包含很多不同群体?
- 他们想要什么?产品有没有提供?
- 他们喜欢产品的哪些特性?不喜欢哪些?



提示:

可以采用简单调查来获得对受众的基本了解,但创建准确统计调查是一个复杂过程。本章仅涉及基本问题。如果你要发现的问题不仅仅是简单的背景资料,或者如果重要产品决策需要准确数值结果,我建议找一家专业的互联网调查公司和他们一起工作,也可以阅读下列书籍:厄尔·芭比的《调查研究方法》(Survey Research Methods),乔舒亚·格罗斯尼克尔和奥利佛·拉斯金的《在线市场调研手册》(The Handbook of Online Marketing Research)。

但调查很容易出错。如果不精心设计,调查会向错误的人提出错误的问题,导致产生的结果不准确、不确定,甚而具有欺骗性(最糟糕的情形)。由于无法直接接触到受访者本身,因此基于互联网的调查特别经不起推敲,调查取决于人们自身的认识和他们准确认识报告的能力和意愿。没有直接联系,他们就不能告诉你他们真正用了什么,也无法说明他们的自我描述是否准确,他们只能告诉你他们的想法。

## 何时开展调查

调查一组人,需要三点:1)一组问题;2)收集回答的方法;3)找到这组人。这是最重要的。这对调查时机的影响最大。如果无法找到受众,就无法开展调查,没有自己的稳定网站,就很难得到受众。竞争对手肯定不会让你来调查他们的用户,传统电话、电子邮件或者当面调查,费用都很高,电子邮件清单往往带有偏见,而且电子邮件调查很容易被当成垃圾邮件。因此,如果没有网站来开展调查,最好能做好准备,一直等到网站运行再开展调查,或者计划用昂贵方式开展调查。

如果已经有受众,调查时机就取决于想了解受众的什么情况,不同调查需要在不同时机进行。例如,背景资料调查可以在任何时候进行,能大致了解当前用户群体的构成。满意度调查可以在重大重新设计之前开展,以了解人们觉得产品哪些地方让他们觉得失望,以便重新设计能够解决这些问题。调查人们觉得哪些东西重要的价值观调查,可以在重新设计广告推广时候的作为大型调查,广告设计师能明白向受众营销时要关注什么。

调查有多种规模和结构,而时机最终取决于你想开展哪种调查,希望得到什么结果。

## 如何实施调查

开始设计调查之前,需要目标明确。先写一两句话,想想为什么要做





调查——即目标。例如，假设有一个成熟产品，运行了好几年的一个网站，收入还行，但发现网站的新用户增长速度在下降。你也许了解市场规模，但基于一个假设：市场并未改变，并且现在发现网站的人和当初网站发布时也一样。此外，网站有一些独特特性，但几个竞争对手也有类似特性。你希望网站的独特特性会吸引客户，日志文件说明已经有人在使用了，但很难判断这些特性是不是人们使用产品的原因，也无法知道这些特性是不是只是锦上添花。把“我们将开展一项调查”这句话写完整，原因可能如下：

目的是了解用户群体从去年以来是否已经发生变化。如果发生变化，又发生了什么变化。另一个目的是了解他们觉得哪些特性最有吸引力。

## 设定日程表

确定调查总体目标，接下来需要建立日程表。调查准备是调查成功的关键，因此典型的调查研究要有足够时间来编写和检查问题，必要时还需要进行测试并修订调查(表 11.1)。调查一旦发出，就无法改变以保证统计的有效性，因此准备是关键。

表 11.1

时间选择	行 为
$t-3$ 周	确定测试受众及目标。
$t-3$ 周	开始列出问题，准备收集软件。
$t-2$ 周	完成列出问题，找几个人一起检查。必要时重写编写报告草稿。
$t-1$ 周	完成收集软件准备，开始建立调查网站，进行认知测试，采用收集软件和汇总软件。根据认知测试结果编写报告。
$t-3$ 天	根据认知测试反馈重写问题；找几个人一起检查。
$t-2$ 天	通过软件并根据问题完成网站，在多种情况下测试网站功能。
$t$	进入调查(通常需要 1~14 天，根据是每天否有重要变化而定)。调查完成后，立即移除调查网站，并结束数据收集。
$t+1$ 天	开始分析。



续表

时间选择	行 为
t+3 天	完成分析, 开始编写报告。
t+1 周	完成报告, 提交给开发团队, 讨论并记下进一步研究的方向。

## 编写调查

应该根据主要原因, 尽量具体地列出调查目标。

- 创建受众的人口统计数据、技术和互联网使用背景资料。
- 得到调查受众对主要特性使用的优先级别。
- 得到他们共同使用的其他网站清单。

有了清单, 就应该有足够信息选择要开展哪些类型的调查。

### 对问题进行头脑风暴

脑海中有调查目标, 接下来要对问题进行头脑风暴(可以自己, 也可以和一组人进行)。不要停止, 写下能想到的需要调查回答的每个问题。不要把问题写成“调查形式”, 写下想知道的内容即可。

进行头脑风暴的同时, 要记住有两种调查目标: 描述性目标 and 解释性目标。

描述性目标旨在建立受众的背景资料。描述性目标采用受众的个人特征、计算机设置、他们想要什么以及他们声称如何行动来总结受众的构成。这类背景资料会非常广泛和复杂, 但并不是了解特征之间如何互相影响。

解释性目标通过答案之间的关系来解释人们的信念和行为。例如, 描述性调查设法了解人们使用哪些特性, 他们的平均收入是多少, 而解释性调查则试图解释收入规模如何影响人们喜欢的特性。这类目标的目的是找到特征之间的内在关系, 被独立出来的关系越多, 解释就更准确。

调查问题本身有多种风格。一般性问题类别可分为目标对象描述, 其物理和软件环境如何的特征问题, 概述人们行动步骤的行为问题, 探讨人们想要什么和相信什么的态度问题。每个大类可以有很多子类。





### 特征类别

- **人口统计数据。**这些问题与调查对象是谁有关。他们年龄多大？职业情况？教育程度如何？
- **技术问题。**这些问题与他们的计算机设置和经验有关。他们使用什么操作系统？显示器分辨率是多少？

### 行为类别

- **互联网使用。**这些问题询问人们如何互联网，已经使用了多久？每周会使用几次？用互联网做什么事情？有多少电脑使用经验？
- **产品使用。**他们(声称)使用产品的哪些特性？使用频率多高？他们为什么会访问你的网站？已经使用网站多长时间？这些都是行为问题。
- **竞争对手。**他们访问过其他哪些网站？频率多高？已经使用了多久？使用的是哪些特性？

### 态度类别

- **满意度。**他们喜欢你的网站吗？网站符合他们的期望吗？能做他们希望做的事情吗？
- **偏好。**他们认为最引人注意的东西是什么？他们对朋友说了哪些特性？认为哪些东西多余或者分散注意力？
- **希望。**他们想要什么？觉得缺少哪些特性？

脑海中有了解问题清单，接下来请开发团队的成员拿出一份他们自己的清单。你可能想把清单给他们看，但千万打住，应先请他们从零开始，这样他们不会被你的想法所左右。接下来收集大家的问题，进行分享，并看看是否会出现其他任何问题。

清单如下所示。

您年龄多大？

您收入多少？

您的教育水平是什么？





您使用什么操作系统？  
 您是否听说过 Linux？  
 您是否用过 Linux？  
 您多长时间使用一次自动报时？  
 您是否用过 eBay？雅虎拍卖呢？  
 您的网络速度是多少？  
 您花多少时间上网？  
 为了获得答案，您愿意等待多久？  
 您是否用过 Quicken？  
 现在的电脑您用了多长时间？  
 您上网多长时间了？  
 对于网站，您觉得哪些事情最重要？  
 您住在哪个时区？  
 等等。

### 列出问题

应该只问人们可能知道答案的问题，应该只问和他们相关的事情，应该说清楚要问的问题。风险是人们会告诉你答案——无论可信还是不可信。

——厄尔·芭比，《调查研究方法》，第 133 页

现在需要编写问题。第 6 章介绍了很多提出非引导性问题的规则，非引导性问题是问题不会让回答问题的人认为有“正确”答案。第 6 章提到的大多数建议都集中在主持人的直接行为，并有助于保持非判断状态，那些建议也同样适用于调查问题。但如果定性研究的问题需要保持灵活而又不会限制受访人回答，调查问题就需要更准确和限定，以便可以计量问题。大多数访谈都应该是开放式问题(换句话说，问题不应该受到访谈列出的清单限制)，而调查则不一样，大多数调查问题都是封闭式问题。一般情况下，对于开放式问题，回答问题的人和分析师都要花费更多精力。对于长时间、只有几个人参加的访谈中可以做到这点，但在调查环境中可能潜在有成千上万参加者，这种方法用起来就要难得多。开放式问题可用于调





查——如果你不知道怎么用封闭式方式编写问题，它们能提供答案——但使用时要仔细，要有节制。

最常见的封闭式调查问题是单一答案的多选项选择题。我们都见过这类问题：给受访者多个选项，但只能选择其中的一项。

您使用互联网电子邮件有多长时间？

- 不到一年
- 1~2 年
- 2~5 年
- 超过 5 年

另一种常见问题形式是核对清单。这类问题包含答案清单，可以多选。

上周您看了《连线》新闻杂志的哪些新闻？(可多选)

- 电信
- 新产品
- 黑客
- 时尚
- 旅游
- 硬件评测
- 软件评测
- 预测
- 体育
- 行业分析
- 政治评论
- 艺术
- 小说
- 投资
- 个人档案
- 其他
- 我上周没看《连线》杂志的新闻



为了让受访人能顺畅回答封闭式问题，提供给他们的答案要具体、详尽，而且相互排斥。具体的问题能减少分析解决问题时的不确定性。如果受众包括舞蹈音乐 DJ，要问的问题是他们会放什么样的音乐，可能不能只问“舞蹈音乐”，因为 DJ 会区分不同的音乐子类型。详尽的问题能减少受访人受到的困惑，还能减少错误，因为人们如果无法找到能代表他们真正意思的答案，就不太可能会选择一个随机答案。详尽的问题应该要包含所有可能的答案，如果不能列出详尽答案，就不应该采用封闭式问题。答案还要尽量相互排斥，减少选择之间的不确定性。如果问人们去哪里买东西，你给他们的选择是“在我的社区”或者“附近”，他们可能不知道你的想法中有什么不同。

很明显，封闭式问题会限制要提问的问题类型，因此要改变问题措辞，才能以封闭方式提出问题，并去除无法重新写出来的问题。

现在画一个四栏表格。对于每个问题，写出问题、为受访人提供说明、可能的答案和提问的原因。“原因”这一栏尤其重要，因为需要确保每个问题都有具体理由：问题为什么重要？信息用于哪些地方？需要哪些人知道？

调查问题表格样本			
问 题	说 明	答 案	原 因
您的年龄(岁)?	无	弹出： <ul style="list-style-type: none"> <li>• 小于 13</li> <li>• 13~18</li> <li>• 19~20</li> <li>• 21~24</li> <li>• 25~34</li> <li>• 35~44</li> <li>• 45~54</li> <li>• 54 以上</li> </ul>	<ul style="list-style-type: none"> <li>• 和去年的调查进行比较。</li> <li>• 和经验进行比较。</li> </ul>
您的性别?	无	弹出： <ul style="list-style-type: none"> <li>• 男</li> <li>• 女</li> </ul>	和去年的调查进行比较。





续表

问 题	说 明	答 案	原 因
上周您看了《连线》新闻杂志的哪些新闻?	可多选。	检查清单: <ul style="list-style-type: none"> <li>● 电信</li> <li>● 新产品</li> <li>● 黑客</li> <li>● 时尚</li> <li>● 旅游</li> <li>● 硬件评测</li> <li>● 软件评测</li> <li>● 预测</li> <li>● 体育</li> <li>● 行业分析</li> <li>● 政治评论</li> <li>● 艺术</li> <li>● 小说</li> <li>● 投资</li> <li>● 个人档案</li> </ul>	<ul style="list-style-type: none"> <li>● 衡量读者的信息需求。</li> <li>● 把已知的阅读习惯和基于日志分析的实际行为进行比较。</li> <li>● 总结广告销售。</li> </ul>
您最频繁采用什么操作系统访问互联网?	无	弹出: <ul style="list-style-type: none"> <li>● Windows XP</li> <li>● Windows 2000</li> <li>● Windows 95/98</li> <li>● Windows NT</li> <li>● Macintosh</li> <li>● Linux</li> <li>● 其他</li> <li>● 不知道</li> </ul>	<ul style="list-style-type: none"> <li>● 和去年的调查进行比较。</li> <li>● 和日志分析进行比较。</li> <li>● 总结设计。</li> <li>● 总结编辑。</li> </ul>

**提示:**

有些情况下,人们可能很少或者没有适当的既往来选择答案。例如,购买房屋的过程(或者开胸外科手术、或者大学报名)对大多数人而言都不常见,他们过去的经验可能和将来的行为没什么关系。此时采用假设性问题胜于询问过去行为的问题,因为过去行为不大可能带来任何有用信息。

可能还要创建第二份清单,列出能自动收集到的信息。网络服务器日志文件能收集填写调查的时间,填写调查的机器操作系统,人们使用的浏览器,机器所处的网域等。cookie 能跟踪哪些人以前访问过网站,哪些人以前从网站买过东西,他们有什么偏好等。

**编写问题时,不要让人们预测他们的行为。**人们过去的行为通常比其



陈述能更好预测他们将来的行为。除非要问的问题属于全新概念，否则问题应该关注人们过去的实际行动，而不要关注他们想象自己在未来的行为。如果想知道某人是否会使用在线日程系统，则不要问下面的问题：

您是否打算使用在线日程系统？

而应该问：

您曾经使用过或者想过使用在线日历系统吗？例如 Yahoo 日历或者 eCalendar。

**避免反面问题。**相对于正面版本问题而言，反面问题更难理解，更容易出错。

您对下面哪些特性没有兴趣？

很容易被理解为：

您对下面哪些特性有兴趣？

相比之下，下面的提问方式更容易阅读：

您对下面哪个特性有兴趣？

也更容易推断出受访人对于没有未选择的选项没兴趣。

**不要让问题超载。**每个问题都应该最多只包含一个要调查的概念。多个概念可以联系在一起，但仍然应该分为单独明确的独立问题。混合问题会让只同意一半问题的受访人感到沮丧，会让要从问题两方面推断受访人观点的分析师感到复杂。

因为带宽问题，您是否对互联网性能感到沮丧？

可以重写成：

您对互联网性能感到沮丧吗？

如果是，下列哪些互联网性能方面让您感到沮丧？

- 页面长度
- 下载时间





- 图片大小
- 等等。

**尽量具体。**避免使用有多个意思或者意思模糊的词(“有时”，“左右”，“大约”，“任何”)。除了货币以外的单位、百分比或者其他常见缩写，其他单位都要写出完整名称(“hours”而不是“hrs”，“thousands”，而不是“K”等)。要使用准确时间。

因此，如果要问人们阅读新闻的频率，如下形式的问题：

您最近在網上花了多少時間閱讀新聞？

- 一些時間
- 很多時間
- 每天
- 等等。

可以改写成：

您上周花了多少時間在網上閱讀新聞和信息？

- 沒有
- 0~5 小時
- 6~10 小時
- 11~20 小時
- 超過 20 小時

**永远不要设限。**提出的問題應該總是讓人們能找到適合自己的答案。例如，下面的問題對受訪人及其態度做了很多假設。

您最喜歡《連線》雜誌的哪些特點？

應該重寫成以下形式：

《連線》雜誌的哪個(些)特性對您很重要？(可多選)。

後者的措辭避免了前者的大部分問題——特別是如果提供了“沒有”選項——並且提供了大部分相同答案。



**保持一致。**每次都要用相同的方式提出问题。这不仅意味着对类似问题采用类似措辞，还应该尽量保持答案选项顺序的意义一致性以及问题展现方式的一致性。

**避免极端情况。**极端情况极少发生，而且大多数人都不会频繁看到自己的极端行为，因此要避免需求或者暗示需要极端行为的情况。

您每次上网都查看网上的新闻吗？

这个问题很可能会导致几乎罕见的反面回应，因为很可能只有几个人每次上网时会查看新闻。更好的提问方式如下：

您上网看新闻的频率是多少？

- 每天几次
- 每天一次
- 每周超过一次
- 一周一次
- 一月一次
- 每月少于一次
- 从不

还可以和人们说的上网频率的回答进行比较，看看比例。

**保持问题相关。**如果人们面对的一系列问题都和他们的体验或者生活无关，就不大可能完成调查。例如，和一组出租车司机相比，一组计算机服务技术人员对计算配置细节会产生不同反应。如果受访人不能回答其中的许多问题，或者对答案不感兴趣，他们就不太可能兴致勃勃地完成调查。

**使用李克特量表。**李克特量表是人们都熟悉的一种方法，用来展现多项选择答案。它们包括一项陈述或一系列陈述，后面跟着 3、5 或者 7 个选项(大多数调查采用 3 个或者 5 个)，这些选项确定了答案的可能范围，包括一个中立的中间选择。

根据你感兴趣的程度，给新闻和信息网站的以下方面打分。





	非常有趣	有些有趣	中立	有些无趣	非常无趣
一定主题不同新闻的数量					○
报道不同主题的数量					○
页面下载速度					○
新闻媒体的名声					○
每条新闻覆盖的广泛性					○
独特社论观点					○
网站搜索引擎质量					○
网站的视觉表现					○
新闻发生后报道的速度					○
浏览网站的容易程度					

**注意：**

附录提供了常见问题清单，但请不要犹豫尝试采用新问题和问题类型。许多网上研究仍能看到纸质调查世界的起源，网上问题和其印刷版本的祖先基本相同。采用这种方法很可能产生创新问题。有个简单例子：如果要问人们网络浏览器窗口（许多人可能都不知道）的大小，可以在屏幕上放尺子的图片，并问“在上图中，不用滚动屏幕，您能看到的最大数字是多少？”

**创建跟进问题。**如果技术上有可能，应该提出能进一步扩展某一答案的问题。理想情况下，跟进问题只能出现在给定问题的具体答案之后。有些网上调查系统能从技术上实现这点。如果无法做到，应该明确跟进问题与特定答案相关，否则不应该回答跟进问题。

如果要问：

请选择您定期访问的所有网站。

那么跟进问题可以包括定期访问的网站清单：

根据重要性，从“至关重要”到“不重要”给下列网站打分。

**包含退出选项。**总是应该包含一个选项，人们可以表明问题不适合他们，或者表明没有一个建议的答案是恰当的。通常可以表达成“以上都不是”，“不知道”或者“没有答案”，这取决于什么样的问题。

**为意见留出空白地方。**大多数人都不会用，但应该在调查末尾留出空白地方，以便人们对调查发表意见。

**编辑并安排问题顺序**

首先需要削减调查问题规模。保持调查简短是调查回复数量最大化的办法之一。大多数人都应该能在 20 分钟或者更少的时间内完成调查。超





过 20 分钟才能完成的调查会感觉像负担，而受访者必须安排时间来完成任务。阅读说明大概要花 5 分钟，阅读并回答每个问题大概要花 30 秒，因此问题数量限制在 30 个左右。调查问题保持在 20 个左右很安全。此外，过去的实践已经证明，只要仔细选择问题，20 个问题就足以了解一个人的几乎任何方面。

编写并挑选问题之后，接下来开始编辑和组织问题。问题顺序和措辞同样重要。调查是一种对话，每个问题都向接受调查的人揭示一些情况，同时向接收调查结果的人提供信息。问题顺序应该有规律、有重点并有选择性地透露信息。

就某些方面而言，调查就像一段小故事。调查开头要吸引读者注意，吸引他们进来。他们一边阅读调查，一边开始了解调查想要发现的信息。到中间，调查开始探索大的想法并根据基本情节引入“曲折”，后面再通过一定渠道进行询问。最后，零星问题组合在一起，调查结束。诚然，即使是最好的调查也无法有令人兴奋的故事，但编写调查的时候能加上一点叙事弧度能让参加者对调查更感兴趣，从而降低因为感觉枯燥而放弃调查的人数。

正如故事一样，为了获得人们对一般主题的反应可以逐步揭示调查目的，而不是直接解释太多。较早的主题可能会影响人们的期望和想法。如果调查是要了解人们如何在线购买玩具，你可能不要说出调查和玩具有关，直到已经对他们的购买行为略有了解再说出来。

例如，有一项调查想了解人们购买食物和购买玩具两者习惯之间的关系。以下两个问题从逻辑关系上可以放在一起：

- A. 您去年一共购买了多少钱的杂货？  
B. 您跟踪购买情况有多仔细？

如果先问 A 再问 B，人们可能就会认为 B 只是针对杂货，而他们对杂货会跟踪得非常仔细，但调查更感兴趣的是人们的一般购买行为。先问 B 再问 A，调查就能收集人们的一般印象，而不会限制他们对调查范围的看法。





典型调查包括以下四部分。

- **介绍。**说明调查目的、填写说明、时间以及有问题时的联系方式。
- **采用难题作为开头。**接受调查的人应该觉得这些问题很有意思，能吸引他们。不应该出现人口统计问题，人们会觉得枯燥，而且在该早期阶段，人们会觉得问题有攻击性。
- **中间部分。**此时可以交替采用受访人感兴趣的和不感兴趣的问题，以保持调查前进。可以根据主题划分问题，例如“一般新闻阅读行为”、“在线新闻阅读行为”、“新闻来源质量”和“未得到满足的需求和愿望”。
- **结尾部分。**包括所有余下的人口统计问题，还要提供开放空间用于一般反馈，并再次说明联系信息。

采用这种结构，调查应该具有灵活性。如果多项选择问题的答案清单没有逻辑进度(最好到最坏，最多到最少，第一到最后等)，只要有可能，清单应该随机化。这样能降低答案顺序对人们选择机率的影响。有些网上调查软件产品能自动做到这一点，但也可以创建多个版本调查，答案采用不同顺序，并在创建邀请时随机把问题分配给人们。

### 编写说明

调查中的说明有两种：一般说明和个别问题说明。

一般调查说明要简单明了。最长可以有一段，应该包括如下内容。

- **调查很重要。**“我们希望 vFork 能为您提供更好的服务。您参与此次调查对我们非常重要。”
- **调查目的。**“本次调查有助于我们了解人们使用 vFork 的需求和愿望。”
- **人们的回答为什么很安全。**“您的所有回答都将保密，并且只用于研究目的。不会有任何针对您参加调查的推销或者营销行为。”
- **有赏奖励。**“完成调查后，为了表达对您的谢意，您将有机会获得 Palm 个人数字助理，中奖机会是 1%。”
- **如果不是调查中的公司，谁对调查负责。**“本次调查是由 Adaptive



Path 公司为 vFork 公司所开展的。”

- **调查持续时间。**“本次调查从 2003 年 7 月 17 号开始，2003 年 7 月 24 号结束。”
- **有问题时联系谁。**“如果您对调查有任何问题或者意见，请填写调查底部的表格，或者发送邮件给迈克 <mikek@adaptivepath.com>。”这样还能进行个性化调查(调查不再是匿名表格)，会提高答复比例。

忠告：

许多地方，如美国大多数地方，对抽奖都有法律规定(也就是调查中说的“X 人中有一人”有机会赢得一些东西)。要在调查描述中清楚说明抽奖规则。复制另一个抽奖的规则能满足大部分要求，但务必要在向公众发布调查之前请律师看一看规则内容。

同样，许多地方对儿童访谈也有法规。应该明确说明只有 13 岁及以上的人才能参加调查。如果目标受众中有很多孩子，应该向律师咨询如何让这些孩子能参加调查(通常需要他们父母的书面许可)。

问题说明也要简单明了。大多数封闭式问题都不需要特别说明，但开放式问题应该准确说清楚你想要什么答案，而不是像下面这样。

列出您最喜欢的网站。

说明可以写成下面这样。

请列出您经常访问或者您真正很喜欢的网站 URL(网址)。最多 10 个。

如果有个问题只需要单一答案，但人们觉得可能有多个合适答案，应该清楚说明人们必须只能选择一个他们感觉最强烈的选择。

一般需要几句话来说明李克特量表。

在 vFork 和其他在线餐饮网站可以找到下列特性清单。如果您觉得在 vFork 使用过其中一项，请根据您在线购买餐饮时该项服务对您的重要程度进行打分。如果在 vFork 从未使用，请选择“从未使用过”。如果不知道是否使用过，请选择“不知道”。

还应该适当的时候做出更多说明，比如错误页或者问题类似于其他但功能不同的部分。例如，如果显示错误页面而需要重新载入，重新载入页面的原因以及任何新说明都应该在新页面的第一屏高亮显示。

### 设计报告

报告？调查都还没开始啊！是的，没错，就是设计报告。想知道应该问什么问题，最佳方法就是列出想要的答案。

收集到的数据应该取决于要执行什么分析，要回答哪些问题。不要问





“以防万一”的问题(但可以问一些你不知道人们怎么回答的探索性问题)。

报告应该从目标和方法着手。就好像调查已经完成一样来编写报告,描述一下为什么要开展调查、目标以及研究设计。要包括对进行抽样的总人口大小、方法、样本大小、完成率以及分析数据的方式。当然,在开始分析数据之前,无法知道实际结论是什么样子,因此这些都是假设,但能明白一些想要什么样的结论。因此要把这些写下来。

上个月购买过餐叉的人中,有 65%的人会一年购买三次餐叉,有些人次数更多。

收集的信息应该是最后一次购买餐叉的时间和人们一年购买餐叉的次数,上个月购买餐叉的人数需要足够大才能分析得到有统计学意义的结果。

写下预留位置结论,应该通过所有要用到的表格和图表来支持结论。如下例所示,表格应该能准确说明哪些变量相对,显示了哪些变量。

基于在互联网的存在时间,对特性的观点				
特 性	在互联网上的存在时间			
	少于 6 个月	6 个月~1 年	1~2 年	2 年
餐叉搜索器				
购物指南				
主要目录				

编写报告这个过程会影响调查内容,因此要提前确保留足够时间来写报告,以便对调查方法或者要收集的数据做出更改。

开展试验性调查后,可以使用报告模型,并采用真实数据的实际过程对所有试验性数据运行一遍,得到相同表格,并得出结论,就像处理真实情况一样。这样有助于从分析程序中找到问题。

## 互联网调查提示

准备在互联网建立调查后,我不会讨论建立网上调查技术方面的问题,因为如果你正在使用能完成很多展示和汇总工作的软件包,这个过程



现有系统就可解决。但所有互联网调查的基本结构都相同：一个含有表单的 HTML 页面、后端 CGI 程序或者基于电子邮件的脚本来收集结果。汇总工作可以通过自定义程序和统计程序来完成，或者对于较简单的调查，用电子表格来汇总。

#### 互联网免费调查工具

截至 2002 年，除了众多商业产品外，还有几家免费的在线调查展示和汇总服务，它们不是全功能商业产品，但也能满足简单调查研究的大部分需求。

- Sparklit([www.sparklit.com](http://www.sparklit.com))是一家基于广告的投票服务网站，可以提出基本的多项选择题，并将结果汇总成简单表格。
- Zoomerang([www.zoomerang.com](http://www.zoomerang.com))提供更多问题类型，预制调查选择，还可付费使用更高级的分析工具。
- Survey Monkey([www.surveymonkey.com](http://www.surveymonkey.com))提供类似更多问题类型以及更多调查选项。

此外，还有调查创建和汇总开源包 phpESP([phpesp.sourceforge.net](http://phpesp.sourceforge.net))，如果网络服务器安装了(同样是开源的)PHP 服务器端脚本语言，就可以运行 phpESP。

创建任何基于互联网的调查时，要记住以下几件事情。

**错误检查。**后端系统可以检查人们如何回答。如果填写有遗漏或者不正确(例如，如果选中了复选框的具体选项和“以上都不是”选项)，网页应该能清楚标出填写不正确的地方。

**功能性。**要像检查网站上任何其他网页一样检查调查网页的 HTML 功能，所做的检查甚至远远超过往常。调查网页下载速度要快，要兼容尽可能多的浏览器、操作系统、屏幕尺寸和调制解调器速度。如果觉得用户群中有一部分会使用 AOL 或者处在防火墙后面，一定要检查这些系统下的调查页面性能。如果网页采用 JavaScript，务必确保网页在所有浏览器都能正常显示，而且如果人们禁用 JavaScript(或者根本无法解释 JavaScript)，应该提供合适的错误消息。





**可用性。**对调查进行可用性测试，像对待新特性那样对待调查。这样有助于确保一切正常工作，以及人们对调查的体验符合预期。

**时机。**既然能跟踪到答复返回的时间，就应该进行跟踪。这样有助于跟踪答复在调查期间达到的时间，还能对用户人口的行为提供另一种信息。如果有大量受访人说在美国东部标准时区，而他们的答复是在美国东部时间从上午9点~下午5点，就可以假设他们是在上班时间访问的网站。

**退出率。**应该跟踪中途退出调查的人。记下他们在哪里退出了调查，并试着根据他们完成的答案来推测原因。

**答复率。**和退出率相类似，答复率是受邀参加调查的人数和实际答复的人数的比例。把分析结果推广到更大人口时，答复率是关键信息。

### 表 单 技 巧

很难区分网络调查中的跳跃答复和预设答复，除非能设置表单来预见答复。可以在默认选项中加入空白字段以示区分。这样并不会阻碍人们选择特定选项，同时又能清楚表明所选选项(空白字段能确保文件中有答复，能简化数据的处理和交换)。空白下拉菜单的HTML代码如下：

```
<select name="news_quality">
  <option value="1"> Strongly agree
  <option value="2"> Agree
  <option value="3"> Neither agree nor disagree
  <option value="4"> Disagree
  <option value="5"> Strongly disagree
  <option selected value=" ">
</select>
```

对于单选框，不要定义默认项。

```
<P>What is your gender?</P>
<blockquote>
<P>
  <input type="radio" name="gender" value="M">
  Male
  <br>
  <input type="radio" name="gender" value="F">
  Female
</P>
```



而非以下定义:

```
<input type="radio" name="gender" value="F" checked>
```

### 测试调查

不管现有测试条件如何,都应该抓住每个机会预先测试研究设计的每个方面。

——厄尔·芭比,《调查研究方法》,第235页

调查常常只能开展一次,再次进行整个调查几乎和第一次进行一样昂贵,不能匆忙改变调查。预先测试调查(也称为 pilot testing)是调查开发的关键部分,能消除大量代价高昂的错误。

#### 提示:

如果有资源,或者特别关注的调查,可以考虑运行预先测试前的调查,以开放方式编写调查问题,并给5~10个人看看。如果所有答复都符合已经准备好的调查中的答案,就能确定调查已覆盖合适的人。

预先测试就像真实情况一样运行,和最终调查采用的软件、招募方法、数据收集方法和分析方法都相同,甚至应该用报告模板来编写预先测试报告,就像会根据最终数据一样做出结论。两者的区别在于样本大小。实验样本包含5~10个答复,在理想情况下,这些答复能代表会真实情况中答复调查的人(如果找不到符合背景资料的人,也可以找只符合一些关键标准的人)。不要告诉他们这是预先测试,采用为常规测试准备的相同邀请材料,看看他们提供的答案是否符合预期。

然后或者同时可以对调查进行两个或者三个用户当面测试,观察填写调查的过程,看他们要花多长时间才能填满,调查他们有什么困难和问题,还可以采取后续行动,发电子邮件给参加者,请他们谈谈对调查的看法,他们碰到了什么困难,等等。

### 奖励

除非调查极短,否则都应向参与调查的人员提供奖励。提供的奖励和调查复杂程度及受众的兴趣相关。了解受众对于挑选合适的奖励非常重要。网站为青少年提供的奖励和为中年百万富翁提供的奖励会不同,但两者目标相同:这是对参加者放弃一部分时间和提供个人信息的回报。有一家在线游戏信息网站提供能得到微软 Xbox 游戏机的机会,来换取网站访问者参加调查。95%的访问者参加了调查,这一比例高于很多普查数据。





奖励的方式很重要。如果对完成调查的人提供的东西价值 2 美元，而需要一个半小时才能完成调查，许多人会觉得不值得花时间来填写调查。但如果有 1% 的机会能获得价值 200 美元的东西，你可能会得到更多答复。对你而言，最终花费都一样，但这说明人们会为了价值更高的东西花时间来赌博，而不愿意直接拿价值较低的东西。为什么？没人知道。

### 实施调查

实施调查指邀请人们参加调查的过程。听起来简单，但实际上不然。调查不是普查，普查是从全体人口获得答复，调查也不是偶然一群人要回答的问题。调查是通过可控方式联系随机选取、均匀分布的人口子集合。

因此，在思考如何实施调查之前，要讨论一下从人口中挑选样本的意义。

#### 抽样和抽样框

填写调查的一组人称为样本，但它是什么的样本？它是根据抽样方法可以联系到的、随机选择的人口子集。这个组称为抽样框(见图 11.1)。换句话说，在所有用户中，根据所采用的方法，能联系到一个子集，还有一个子集不能联系到。假定能联系到的子集就是抽样框，那么从抽样框中随机选择出来的人就是样本。

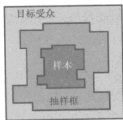


图 11.1 样本和抽样框

理想情况下，抽样框和所有你感兴趣的人相一致。如果不一致，结果就可能不准确，并会产生误导，因为调查结果无法代表整体受众的观点。例如，如果只有抱怨产品的人参加调查，可能就不能很好了解整体受众的

#### 提示：

因为大量调查分析可用来了解收集方法的不确定性，调查设计的一个主要目标就是减少需考虑和控制的变量。从哪些人会看到调查，到内容页面下载的速度有多快，每件事情都会影响人们如何答复调查。因此，在创建调查、召集受访者以及收集答复的机制时，要不断想办法减少未知变量的数量。





观点。同样，如果只和订阅了邮件列表或者已经从网站购买过东西的人进行讨论，就只是和对产品有积极兴趣的人在进行讨论，而和构成所有对你提供的东西有兴趣的人相比，这些人可能只是其中很小的一部分。

抽样框不一致还会发生各种情况。如果一家网站只在周一和周二进行一项滑雪情况的调查，而大多数用户是在周四和周五访问网站，调查结果就不具有代表性，因为在周末滑雪的受众很可能不同于在一周中段滑雪的受众。同样，只在周四和周五进行调查可能会错过铁杆滑雪爱好者关键市场。了解抽样框构成情况是了解调查结果有效性的关键。不幸的是，有很多很多方法会造成糟糕抽样，并得到错误抽样框。

首先，定义最感兴趣的用户群体。使用频率是最理想的着眼点。他们是定期用户？（如何定义“定期”？每天？每周？）是首次用户？是不常来的用户？

添加其他一些重要而且能影响调查人员招募方式的特征。他们曾经在网站上购买过产品吗？是否来自特定地理区域？是学生吗？是正在考虑从竞争对手产品转移过来的人吗？是高级用户吗？

调查最终会包含若干子群体（高级用户和定期用户、欧洲采购商等）。

确定用户之后，接下来便是广泛邀请构成目标受众的人，不会漏掉任何群体，也不会过多代表任何群体。每项特征都能创建一种不同的邀请方法，能对数据收集产生不同的潜在问题。

例如，假设你感兴趣了解频繁用户的可用性问题，因此想建立一直在使用网站的人的背景资料。如果知道经常访问的人和访问并购买产品的人之间有明显的区别，这虽然很有趣，但要关注的是目标受众的选择。因此，确定目标受众是“每周至少访问网站一次，并买过东西的用户。”

如何联系这些人？可以利用对 cookie 的跟踪，对 cookie 短于一周的人进行提问。但这取决于人们没有更换机器，没有删除 cookie，也没有更换浏览器。更重要的是，是假设采用了跟踪 cookie。随机邀请用户，同时如果有人说没有一周至少访问一次，把这些人的答复也过滤掉，这样会容易一些。此外，需要至少一周再次执行调查，或者按照周级别的时间长度再次执行调查，这样不会过度代表在任何特定一天出现的访问者。这样还





能保证联系到符合目标受众背景资料的所有用户吗？无法保证。其中有些人可能在度假；还有人可能会一周不访问网站而错过调查，但人们是随机做这些事情，因此不太可能会影响调查结果，可能也还有其他用户群体——如不经常购物却买很多东西的人——对你也很重要，但你不关心这些群体的存在。漏掉这些人的确会影响调查结果。怎么才能了解并联系上所有这些人？你不可能、也没能力了解并联系到这些人，但要尽量联系到答复能对产品产生积极改变的人，要足够多。

说了这么多，意思就是越了解用户人口及其习惯，就能更好地选择抽样框。但如果对受众知之甚少，也不用发慌，实施第一次调查，可以用来揭示受众的大量背景资料，而定性研究(情境调查和焦点小组)可以揭示他们的行为情况，这样能更准确地定位随后的调查。

### 样本规模

要邀请多少人？调查总是包含一定的不确定性。样本总是存在误差，因为你不会用相同方式对用户人口中的每位成员提出同样问题。可通过数学方法估计不确定性的规模，但可以接受的不确定性数量要由你自己决定。

最终，调查取决于用户人口中存在的变量。如果所有用户都相同，只需要问问其中的一个人，就能了解他们会如何答复。但所有用户都不相同，因此必须要根据用户人口的变化程度来估计要问多少人。这是先有鸡还是先有蛋的经典问题：不做调查，就不可能了解受众的变化程度；不知道他们的变化程度，就无法进行调查。幸运的是，可以通过假设人口变化非常“标准”来估计需要邀请多少人，如果估计得太少，再对将来的调查进行调整(除了需要额外工作外，向太多的人提问很少会成为问题)。

假设整个用户人口有 1 万人。也就是说，如果把每个人都算成网站现有用户或潜在用户(在合理范围内，例如今后一年内)，你就会得到 1 万人。现在，你想知道需要调查多少人，以便更好地了解所有人(或者大多数人)回答问题的情况。你永远得不到“真正的”数值，除非调查所有 1 万人，但如果设在 300 人以上，则有 95% 的可信度，就可以认为从这 300 人得



到的答复和向每个人提问得到答复的误差比例在 10% 以内(“ $\pm 5\%$ ”，就像他们在电视上说的一样)。“95% 的可信度”是什么意思？本章稍后也会提到，但现在可以理解为“非常可信”。

300 人没问题，但 300 不是个大数字，如果受众有一百万怎么办？30 万人对于调查是一个大数字，人口中的人数和需要交谈的人数并不是以这种方式成比例。根据数学方法，你只需要 2000 个样本，所以调查明显容易得多。因此，不需要通过咨询统计员，就可以用下表来估计需要多少个样本。

样本估计大小(假设标准误差 5%，可信度 95%，正常变化)	
人 口(人)	样本数量(人)
1000	150
10 000	300
100 000	800

这些数字是非常粗略的样本数量，可以用来把一项调查统计推广到大用户人口。统计员会因为我去掉了统计数字而惩罚我(相信我，没必要那么多)，但这一数量足以做出一些猜测。

情况还要稍微复杂一点，这些数字并不一定适用于整个样本规模，但适用于要研究的任何单一群体。如果把受众细分成子群体(按照性别、经验或者地区)，并对子群体进行计算，每个子群体都需要那么大。例如，如果产品有 10 000 位独立用户，你打算把定期用户从非定期用户中分开进行分析，这两组数据都必须至少有 300 条记录。这意味着总样本至少需要有 600 条答复。

### 偏差

抽样偏差是调查研究的恶魔。抽样偏差很狡猾，很隐蔽，而且无处不在，也无法消除。偏差产生于你认为会答复的人的身上——抽样框成员——而不是你要进行抽样的人口成员身上(如图 11.2 所示)。这种情况非常危险。



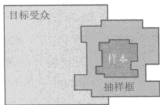


图 11.2 抽样偏见

如果希望样本提供整体用户人口的有用信息,就要求样本和用户人口及其观点相类似。如果不类似,某些子群体或者观点就会过度代表整体,而同时其他群体及观点会被扭曲。如果汇总有偏差的调查所产生的数据,调查结果将无法用来描述全体。调查的目的是为了描述用户,因此错误展现受众的信息会破坏整个工作。

在一定程度上,某些人始终应该排除在外。通过互联网实施调查,就排除了非网民。用英语编写调查,就排除了不懂英语的人。关键在于清楚自己应该排除谁(或者错误的代表),并确保他们不属于目标受众,不对他们进行分析。

最常见、最重要的偏差是不答复偏差。总有些人会忽略邀请,这就是偏差。如果调查让特定群体感到厌恶,那么他们肯定不会出现在研究中。例如,如果要针对使用时间进行调查,调查要 45 分钟才能完成,可能就会漏掉市场的最重要细分部分,因为人们没有时间来填写调查。还有其他许多不同类型的答复偏差。

- **时间选择偏差。**邀请人们的时间会影响他们的回答方式吗?在圣诞节前一天询问人们有多享受购买礼物和在人们不疯狂抢购玩具的时候提出相同的问题,得到的答案肯定不同。
- **持续时间偏差。**有些行为呈周期性:在一天、一周或者一个月的特定时间处于高峰。如果目标受众由商品交易员构成,调查时间是从上午 9 点到下午 5 点,而此时大多数交易员都太忙以至于没空填写调查,你就可能漏掉大部分受众。
- **邀请偏差。**如何邀请人们参加调查?地点、时间、奖励和措辞都有影响。如果要邀请滑冰一族,而邀请内容采用“亲爱的先生或



者女士,欢迎您……”作为开头,可能就会失去很多受众(除非邀请内容明显是反语)。如果提供 T 恤衫作为对高级经理的奖励,就不太可能获得代表性群体。

- **自主选择。**网络调查中有一种特殊邀请偏差和共同错误,即让人们选择是否希望参加调查,但并没有明确邀请他们。“点击这里参加调查!”是种常见表现形式。谁会参加调查?为什么会参加?在可以选择的情况下,这种方式显然无法了解不会点击链接的人。他们和点击链接的人或许一样,但值得怀疑。可选择参加与否的调查会吸引到的人有极端观点或者有特别兴趣,最好能让人们必须参加调查,而不是选择参加,这样就能知道随机分布的人能看到邀请,多少人受到邀请,有多少人选择不接受邀请。
- **展现偏差。**调查的展现方式。调查的外观和感觉也会决定谁有兴趣回答调查。从技术看,如果使用 JavaScript 来执行某个关键功能,而一部分关键用户人口禁用 JavaScript,他们就无法参加调查。从美感而言,如果用一群可爱的卡通人物展现调查,商人会拒绝参加(相反,如果屏幕上满是纯黑白文字,展现给青少年看,他们会觉得太无聊,从而视而不见)。调查的视觉修饰要符合调查所链接的网站。如果有个很好的站点和其所链接的调查在风格上非常不同,就可能会造成混乱(如果转变过于剧烈,甚至可能影响人们对网站品牌的认知)。
- **期望偏差。**人们对调查内容和参加调查的动机会期望。如果这些期望在参加调查期间没有得到满足,他们就可能退出调查。在一项调查预测试中,有位用户说:“噢,我以为调查内容完全与我有关,但问的所有问题都离不开我怎么换油。我不关心这个。”

调查设计的艺术在于通过跟踪答复率尽量减少和理解所有上述因素,同时让更多的人参与调查。相较而言,汇总结果及得出结论非常简单。典型的在线调查答复率各有不同,从 20%~40%(即受邀参加调查的人中有 20%~40%参加并完成了调查)。如果远低于 20%,调查结果就非常值得怀疑。正如《在线市场调研手册》(The Handbook of Online Marketing Research)一书的





合作者乔舒亚·格罗斯尼克尔所说：“一项只有 1% 答复率的调查是毫无价值的。”测试调查，并根据调查未被包括的群体来了解问卷设计的含义是尽量减少偏差的关键方式。

### 邀请

减少偏差后，接下来的任务是考虑如何随机选抽样本。找到真正的随机样本很困难，而且要取决于邀请方式。邀请人们参加调查有很多共同方法，每种方法各有优点，也各有问题。

### 邀请链接

邀请链接是最简单但最不准确的调查邀请方法。关键页面(通常是首页)上有一个链接，邀请人们参加调查。其好处是廉价，并且不引人注目，但受自主选择偏差的严重影响。无法保证随机选择的网站用户会看到链接，而且也没有试图在访问者中分发邀请，因此参加调查的人可能会是想对网站传达一些情况的人。这些群体可能是带有强烈观点的熟练用户。他们的观点有用，但几乎不足以代表典型用户的观点。

如果使用首页邀请链接，要跟踪所有独立的首页访问者，并且和做出答复的人的数量进行比较。这样能知道有多少人看到了邀请、多少人填写了调查。如果答复率很低(比如只有 5% 访问网站的人参加了调查)，很可能说明做出答复的人并不代表受众。这种调查仍能产生有用信息，但结论不大可能推广到广泛用户人口。

### 电子邮件

如果有已知用户清单，就可以随机选择一个子群体，并邀请他们参加调查。这类邀请清单不会发给数据库中没有的用户(因此没有使用网站服务的人也不会进入数据库)，也不会发给拒收电子邮件的人，但你能联系到几乎可以肯定使用网站的人。更重要的是，你可以发出明确邀请，这样能知道所有潜在用户都收到了同样邀请(从而减少自主选择问题)，你还可以准确跟踪到底有多少邀请产生了答复。

有两种基本类型的电子邮件调查邀请：一种是邀请访问网络调查，另

#### 忠告：

如果采用跟踪识别符，要确保调查软件具有这种能力，这样预期匿名调查的受访者能知道不是匿名调查。还要明确即使不是匿名调查，答复仍然会保密。



一种是电子邮件调查。电子邮件调查能接触到更广泛的人,但数据收集需要更多工作,因为必须要解析答复(无论是手工方式还是通过软件)。网络调查通常是首选。

理想情况下,每份电子邮件邀请都带有独特识别符,可以用来跟踪哪个人、有多少人做出了答复。

### 中断邀请

中断邀请是在用户访问网站的时候邀请随机用户参加调查。每位用户的受邀机会均等(无论是他们每次访问网站的时候,还是每位用户受邀一次),邀请会中断他们对网站的体验,因此他们必须主动接受或者拒绝邀请,然后才能继续使用网站。这样能准确知道有多少人受到邀请,确保他们都会同样注意到调查,并保证他们是从所有用户中随机抽取的样本。

随机中断邀请可以通过概率或者系统实现。通过概率进行抽样涉及根据概率选择访问者,符合此概率的总访问者数量大致等于预计的答复数量。在实践中,通常涉及为每位访问者挑选一个随机数;如果此随机数符合预先确定的“魔幻数字”,然后就邀请此人参加调查。否则,他们就继续访问网站,也不会知道其他人也会参加调查。因此,如果已经确定为期一周的调查需要 1000 个答复,而网站每周约有 10 万独立访问者,每位访问者在调查周受邀参加调查的概率就应该是 1%。

#### 随机选择的 JavaScript 代码

建立一致概率的一个有用方法是在每个页面上都包括一段代码。此段代码会执行必要的计算,以确定当前访问者是否需要收到调查邀请,并在访问者的浏览器中留下 cookie,这样每次调查访问者只会收到一次邀请(当然,除非调查是要度量每次访问的行为,而不是度量每位访问者的行为,但此类调查相当罕见)。

以下是《在线市场调研手册》(*The Handbook of Online Marketing Research*)的作者乔舒亚·格罗斯尼克和奥利弗·拉斯金编写的一段 JavaScript。这段代码在大多数情况下都起作用(当然,受访人浏览器禁用 JavaScript 的除外)。





```
<script language="javascript">
//<!--
//the variables sLocation and sRate
//are set to the location of your survey and
//the desired sampling rate

var
sLocation="http://www.adaptivepath.com/survey.html";
var sRate = 15;
var cookieName = "SurveyCookie";

var cookieValue = getCookie(cookieName);
if (cookieValue == null) {
    setCookie(cookieName, "Sampled");
    var sampled = Math.random();
    if (sampled < sRate) {
        window.open(sLocation, 'sWindow', 'scrollbars,resizable');
    }
}

function getCookie(Name) {
    var search = Name + "=";
    if (document.cookie.length > 0) {
        offset = document.cookie.indexOf(search);
        if (offset != -1) {
            offset += search.length;
            end = document.cookie.indexOf(";", offset);
            if (end == -1);
            end = document.cookie.length;
            return
            unescape(document.cookie.substring(offset, end));
        }
    }
}
function setCookie(name,value,expire,domainname) {
    document.cookie = name + "=" + escape(value) +
    ((expire == null) ? "" : ("; expires=" +
    expire.toGMTString()) + "; path=/")
}

//-->
</script>
```





为了准确度量所有访问者，而不是仅仅度量访问主页的人，这段代码需要插入人们可以通过输入网址到达的每个网页。动态生成的网站，或者有统一导航头的网站(此时代码可以作为头文件(header)的一部分插入)都很容易做到，但即使是用纯 HTML 生成的网站，也应该可以把代码插入到所有可能进入网站的页面。

系统中断邀请有时可以更简单。每第  $n$  位访问者，系统样本就邀请一位。本例中， $n$  是 1~100 之间的一个随机数字，然后选择每隔 100 位的访问者，直到邀请了 1000 位访问者。这属于更传统的纸质调查方法，其优点是更容易实施，但不够灵活。

### 电话、当面调查和标准信函调查

传统调查方法，如电话、当面调查和纸质信函调查，不在本书讨论范围内，但不要忘记这些方法，因为这些方法是得到真正用户人口随机样本的最有效方法。这些方法都是政治和市场调研采用的典型调查方法。有一个典型工具是随机打电话：有一台机器根据一定区号由电话交换设备随机拨打电话，直到有人应答为止，然后有人接管电话，并开始阅读电话调查剧本，但很难通过随机拨打电话的方式找到特定软件产品的用户或者特定网站的访问者，而通过同样方式就比较容易找到属于某一政党或者购买某种产品的人，因此，这种方法不能有效找到用户。尽管当面调查收集数据看起来违反了互联网的全球性质，但如果很容易找到受众——比如有个网站服务于当地学区的教师——当面调查的灵活性和收集信息的潜力将无可匹敌。这些方法通常比网络调查更花时间和费用，但不要仅仅因为这些方法不是数字化就忘记了。

## 如何分析调查答复

调查分析和解释既是艺术又是科学。要处理的是数字、比例和关系，但同样也要度量陈述，而陈述和行动之间的关系很难理解。从招募偏差的含糊不清，到对问题措辞的误解，到人们的夸大倾向，整个过程都涉及估





## 忠告：

世界上有很多糟糕的调查研究，这些调查研究就是因为所包含的数字而受到批评。不要被数字欺骗了，也不要想利用这些数字进行欺骗。因为数字并不代表这些数字是采用严格程序来获得的。

计和判断，归根结底，每个人都不相同，而最终分析会漏掉任何个人观点、行为或者体验中的细微之处，但这并没有关系。大多数情况下，结果并不需要完全肯定才能有价值、有用，而且往往为了做出明智决定只要知道可能性就可以了。

因此，调查数据分析要努力追寻准确性和即时效用。高级方法有时能从数据中提取重要的微妙之处，但大多数情况下较简单的分析更受欢迎。更简单的方法能减少错误的可能性，并足以回答典型产品开发情况中出现的大多数问题。

有两种通用分析方法可以简单地总结为统计和比较。

## 统计

处理结果最简单、而且也常常唯一能做的事情就是统计结果(调查语言中称为汇总)。统计基本答复后，就可以显示出简单趋势，并能发现数据输入错误。它由统计每个问题的所有答复的统计构成。

要从查看原始数据着手。通过数字总结结果之前，原始资料能告诉你数据展现出来的趋势。结果如何分布？是否有些答复存在明显的集中趋势，是否有任何明确的虚假答复或者不典型的答复(比如个人收入 15 万美元的青少年或者 80 岁的滑板者)？花些时间看看原始数据能给你一些直觉，以后都能用得上。

看过原始数据，简单统计某个问题的所有答案就有用了。例如，对“您的家庭收入属于下列哪个类型？”这个问题计算答案会产生下表。

低于 20,000 美元	0
20,001~29,999 美元	2
30,000~39,999 美元	3
40,000~49,999 美元	10
50,000~59,999 美元	12
60,000~69,999 美元	20
70,000~79,999 美元	25
80,000~99,999 美元	28



续表

100,000~119,999 美元	22
120,000~149,999 美元	17
150,000 美元或者更高	5
无答案	10

采用简单的柱形统计图可以看出受众收入分布的一些有趣信息。

如图 11.3 所示,受众的收入顶点在 8,000~100,000 美元之间,收入在 60,000~150,000 美元的人构成了用户主体。如果网站面向的是中等偏下收入的参加者,很明显,网站正按照你希望的方式吸引着这些人。

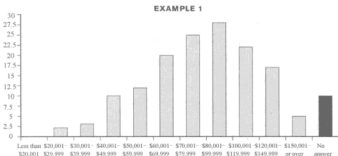


图 11.3 正态分布

把数据从表中拿出来,可以计算出数据的中间值(mean)和众数(mode)。中间值指传统数学意义上的平均值。把某一个问题的所有答复加在一起,再除以答复数量就得到平均值。对于数字范围的情况,如上例,可以使用每段范围的中间点作为起点。本例中,平均值约 86,000 美元,计算如下:

$$\begin{aligned}
 &= \frac{(25,000 \times 2) + (35,000 \times 3) + \cdots + (150,000 \times 5)}{144} \\
 &= \frac{12,425,000}{144} \\
 &= 86,285
 \end{aligned}$$

之所以是“约”86,000 美元,因为范围广,而且最高范围“150,000 美元或者更高”无法控制(所以采用下限)。从实用角度出发,这样通常够





用了。

但中间值很容易受到少数极端结果的扭曲。这属于“亿万富翁”问题：如果要年薪实际值进行抽样，而你恰好对比尔·盖茨进行调查，平均值就很可能要明显高于样本中大多数人的收入。这时就需要查看原始数据，因为对结果会有直觉判断。直觉判断仍有可能是错的，但如果看着收入为40,000美元和50,000美元的人的答复，而中间值是120,000美元，很可能有东西推高了平均值。要开始寻找极端值(outlier)，或者在数据普遍变化值之外的答复，因为少数极限值影响到了中间值。

众数，即最常见值，可以和中间值进行比较，看看中间值有没有受到少量极限值的扭曲(本例中，“80,000~99,000美元”有28个答复)。如果答复属于正态分布，其中的数据有最大单值，然后对称回落(形成所谓的钟形曲线)，中间值和众数相同(如本例所示)。样本越大，就越可能得到正态分布。但有时候也得不到正态分布。如果由于某种原因促使网站要努力吸引两组不同的人，中间值和众数就可能会不同。例如，有资质的医生和医学院学生广泛使用一个网站，收入分布看起来如下所示。

根据下表，收入中间值约70,000美元，但众数是90,000美元。差别足够大，说明答复分布呈不均衡钟形曲线(事实上，这是所谓的双峰分布)，因此需要进行额外分析。柱形统计图(图11.4)能清楚说明问题。

低于 20,000 美元	20
20,001~29,999 美元	17
30,000~39,999 美元	14
40,000~49,999 美元	6
50,000~59,999 美元	10
60,000~69,999 美元	12
70,000~79,999 美元	18
80,000~99,999 美元	22
100,000~119,999 美元	20
120,000~149,999 美元	15
150,000 美元或更高	9
无答案	3







## 比较

单一变量汇总会带来丰富信息，也很有用，但调查研究的真正强大之处在于互相比较若干变量的内容。例如，你可能关心人们使用网站的频率如何影响他们使用的特性，一直使用网站的人使用的特性是否不同于偶尔使用网站的人使用的特性？知道了这个情况，就能更好强调特性，并创建介绍性帮助。如果只看数据，很难发现这种关系，因此需要采用比较方法。最常见的比较方法是交叉汇总。交叉汇总通过比较两个值来发现两个变量之间的关系。

创建交叉汇总有很多方法，但典型方法如下所示。

(1) 首先确定独立变量。即你觉得会产生“影响”的因子，也就是问题的对象。在“访问频率如何影响用户使用的特性？”问题中，自变量是访问频率，因为访问频率很可能会影响人们使用特性(而不会影响其他方式，使用某些特性有可能造成人们更多访问网站，但根据你所了解的人们使用网站的情况，这种可能性比较小)。

(2) 根据独立变量的值，对问题答复进行分组。例如，如果问题是“您多长时间使用一次[网站]？”，多项选择的答案是“少于每月一次，每月一次，每月数次”等，根据答案对答复进行分组就是最佳起点。

(3) 在每个独立变量组内，分别汇总其他变量和独立变量答案。因此，如果另一个调查问题是“上次访问[网站]时，您使用过下列哪些特性？”，人们给出的答复会形成独立变量。如果问题的答案是“购物车、新闻页和比较助手”，就可以对每一组中有多少人选择了其中一项特性进行汇总。

(4) 利用汇总形成的数值创建表格。例如，下表比较了不同群体的人报告他们上次访问网站使用的特性。

	少于每月一次	每月一次	每月数次	其 他
购物车	5%	8%	20%	
新闻页	20%	25%	15%	
比较助手	2%	10%	54%	



此时,如果两个变量之间有关系,应该就能看到这种简单关系。例如,每月数次使用网站的人比不频繁使用网站的人使用比较助手的次数明显要多,这可能意味着使用网站的次数越多,比较助手特性对人们的价值就越大。(为什么?调查不能轻易回答这个问题。)同样,和其他人相比,对于定期用户而言,新闻页看起来不太重要,这种情况并不奇怪,因为定期用户的访问次数足够多,因此每次访问对他们来说新闻页就不算新了。比较其他变量还可以发现更多关系(例如,比较使用时间长度和使用频率:是不是人们使用网站越频繁,使用的时间就会越长?)。

下表比较了“您多长时间访问一次网站?”和“您为什么今天访问网站?”两个问题的答案,以了解访问频率是否会影响人们访问网站的原因。只汇总了对“您为什么今天访问网站?”超过数量 500 或者更多的答复(因为这个数字是确定有统计学意义的最低人数)。

	第一次访问	少于每月一次	每月一次	每周一次	大于每周一次	总计
查找具体广播节目信息	260	229	167	129	115	900
其他	220	159	104	56	78	617
想听广播节目	344	245	251	298	630	1768
想看新闻或者信息	140	120	106	96	109	571
总计	964	753	628	579	932	3856

用百分比来显示,信息性更强。

	第一次访问	少于每月一次	每月一次	每周一次	大于每周一次	总计
查找具体广播节目信息	29%	25%	19%	14%	13%	100%
其他	36%	26%	17%	9%	13%	100%
想听广播节目	19%	14%	14%	17%	36%	100%
想看新闻或者信息	25%	21%	19%	17%	19%	100%
答复中间值	25%	20%	16%	15%	24%	100%





如果采用比例图(如图 11.5),能完整说明问题。

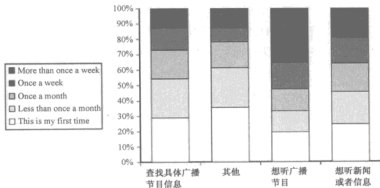


图 11.5 交叉汇总图样例

看看图就能发现几个问题。

- 定期访问者比随意访问者收听的广播节目更多。这在暗示驱动“黏性”的因素之一可能是因为网站能提供更多音频流。
- 相较于定期用户,非定期用户更倾向于寻找更多节目信息。也许是因为他们不知道网站有更多节目。如果是这样,可能导致网站的两种用户也会寻找“其他”东西,更进一步表明网站没有充分表达它所包含的内容,因为对于访问网站超过一次的人来说这个值只下降一点点。

当然,还能得到其他结论,还有更多高级方法能操作和显示变量之间的关系,但这些主题不在本书讨论范围内(可参阅爱德华·塔夫特所著的信息可视化优秀书籍)。

构建表格时,要始终明确每个独立变量群体的总答复数量。习惯上会增加一个“总计”栏(但讨论百分比时,这种做法没有太大用处,此时总计通常都是 100%;此时采用  $n=$ ,其中  $n$  是答复数量)。

此时,计算答复最低数量就有用武之地了。如果调查开始时,任何独立变量群体中的答复总数量低于最低样本量规模,就不应该对那个变量做出任何结论,结果没有意义,而且要表明出来,或者从报告中完全拿掉。

提示:

如果使用 Excel 来汇总,数据透视表(pivot table)能完成交叉汇总的大部分繁重工作,利用数据透视表,可以得到原始调查数据清单,每行一条完整答复,并自动交叉汇总两个变量。这样能节省大量时间,但要注意:很容易迷失在比较中,很容易去比较——在逻辑上——不应该进行比较的数据。





但可以把群体合成能满足样本规模的更大群体。这样做的时候，要清楚标出“超级群体”。因此，如果“18~24岁”群体没有足够结果能得出结论，可以把这个群体从报告中完全拿掉，或者可以和“25~30岁”群体进行合并，并建立“18~30岁”群体。当然，只能有意义的群体合并才有用。

## 判断误差

每个调查都对全体用户人口进行抽样，因此每次度量只是估计值。如果不进行普查，就不可能知道实际值，数字所具有的精确还会产生很多混乱。不幸的是，精确并不代表准确，仅仅因为可以计算到小数点后六位并不意味着实际有那么准确。幸运的是，有方法能估计出观察到的数据和实际数据有多接近，以及和重要的精确有多接近。这虽然无助于计算和度量，但能说明精确性。

标准误差(standard error)是一种测量的不确定性。它是计算所得值的模糊定义，也可以度量计算的精度。标准误差越小，测量越精确——标准误差越大，就越不知道确切值。标准误差是从调查样本规模计算得到，也是测量值占整体值的比例。计算公式如下：

$$\sigma = \sqrt{\frac{PQ}{n}}$$

其中，P 是相对于整体的百分比值，也可以用小数表示，Q 等于(1-P)，n 是样本数量。

因此，如果对 1000 个人进行抽样，其中 400 人说他们喜欢晚上从网上购物，而不喜欢其他形式的购物，标准误差可以计算为 $(0.4 \times 0.6)/1000$ 的平方根，即 0.016。也就是说，实际价值很可能在测量值的 1.6% 范围内（“测量值 $\pm 1.6\%$ ”）。大多数情况下，这个值已经足够准确。

标准误差也有助于搞清楚精确度能产生多大影响。因此，如果计算能精确到小数点后 6 位(即 0.000001)，但标准误差为 1%(即 0.01)，精确则无意义，因为数据内在的含糊不清会造成从小数点后第三位开始的数字都不会有意义。





减小计算标准误差最简单的方法就是对更多的人进行抽样。和上例相反，不对 1000 个人抽样，对 2000 个人抽样提供的标准误差为 1.1%，对 5000 人抽样，标准误差能减小到 0.7%。但请注意，标准误差不会是零，除非能问到所有人，否则总是存在不确定性。

标准差(standard deviation)可以度量可信度。它能告诉你在标准误差定义的扩散范围内能发现真正答案(永远无法肯定知道)的概率。在正态“钟形曲线”分布内，这些数字都标准：实际值有 68% 的机会在一个标准偏差(一个标准差扩散)内，有 95% 的机会在两个标准差内，有 99% 的机会出现在三个标准差内。

标准差根据测量值确定出来的范围称为置信区间(confidence interval)(见图 11.6)。它表明标准误差和标准差有什么关联。标准误差定义了值的所属范围，而标准差能说明值出现在那里的机率。

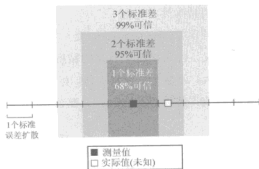


图 11.6 置信区间

例如，假设有个调查度量出 50% 的人口，标准误差是 3%。这说明，可以有 68% 的可信(1 个标准差，如图 11.7)认为实际男性人口比例是在 47%~53% 之间，有 95% 的可信认为实际比例是在 44%~56% 之间(两个标准差)，99% 的可信认为比例在 41%~59% 之间(3 个标准差)。无法知道在这个范围内的精确值(也许是 54.967353%)，但如果需要根据这个数量做出决定，至少能知道猜测离现实有多接近。



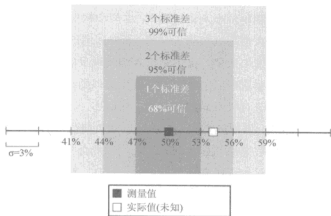


图 11.7 置信区间样例

### 度量误差

通过计算样本的标准误差和置信水平，可以了解到所度量数据和(常常不可知的)客观事实的接近程度。但并不是说，数据能真实代表应该代表的东西。系统收集数据，采用合乎逻辑、统计学上准确的方法认真收集数据，但仍有可能会完全错误。

有一个例子说明用统计数字无法解决或者预测问题，这个例子就是美国国家航空航天局(NASA)的火星气候轨道器坠毁。该轨道器是世界上一一些最聪明的人建造出来的最先进设备。1998 年 12 月，轨道器从地球发射，原计划 1999 年秋天到达火星轨道，这类飞行的误差幅度非常小，因为飞行器几乎不携带用于修正航线的燃料。在航行的大部分时间内，一切都很顺利。飞行器飞行完美，系统定期向地球报告位置和速度。在开始执行科学任务之前，最后的关键阶段是进入火星轨道，而这一过程由系统自动执行。随着飞行器接近火星，它开始自动减速。起初，它按照预期离火星越来越近，但随后开始越来越低，先从指定轨道坠落，又从稳定轨道开始坠落，又从可生存高度坠落，最终跌落进火星大气层，并永远消失。在随后进行的调查中，人们发现，有一个开发团队测量推力采用的是英制单位 lb/s 磅每秒，但另一个开发团队采用的是公制单位，1N/s(牛顿每秒)，这





比英制单位小 4 倍。因此，这两套软件都根据设计进行工作，但飞行器认为它在使用这套测量系统，而实际用的是另一套，造成价值几亿美元的飞行器撞上火星，而不是进入围绕火星的轨道。

调查研究中会很容易发生类似的事情。金融网站可能想知道人们常常如何进行投资新股。“您多久购买一次新股<sup>①</sup>？”这个问题对这种情况有意义，但如果要向一群小商店业主问这个问题，而首先又没有让他们做好要回答金融问题的准备，人们可能会把问题理解成是指他们的库存<sup>②</sup>。因此，尽管分析师认为能度量一件事，但参加者认为度量的东西完全不同。

这就是所谓系统误差，因为它会影响所有数据。火星气候观察者轨道器遭受了极端的系统误差。无论多么精确，都没有在测量工程师认为应该测量的东西。但也还存在随机误差，即答复的自然变化。从某种意义上说，度量标准误差是补偿随机误差的一种方式，它们能告诉你根据收集的样本数量，估计大概有多少随机误差。随机误差可以出现在任何方向，因此会相互抵消，这就是为什么随着样本数量的增加标准误差会缩小的原因。

### 得出结论

从结果得到的结论要侧重回答研究开始提出的问题，而那些问题对产品的未来最重要。但很难通过数据获得意料之外的知识。

在开始下结论之前，需要重新整理开始时放在一起的表格，重新了解作为分析的一部分而填写的表格。它们显示了什么变量？这些变量能度量什么情况？为什么这些度量重要？如果通过分析，事项优先级别发生了变化，要根据需要重新定义表格。

比较数据时，你可能想通过数字测试来确定两组答复之间存在的差异是否有明显区别。在上例中，每周阅读一次新闻的人和每周阅读新闻超过一次的人之间的差别是 2%。这个差别显著吗？有两种测试可用来确定这一点，即卡方检验(Chi-square)和 Z 检验(Z-test)，但其幕后的数学解释已超出本书范围。

① 英文中 stock 一词既有股票的意思，也有库存的意思。



根据数据做出结论时，要避免下面几个常见问题。

- **令人混淆的相关性和因果关系。**因为两件事情的发生时间非常靠近并不意味着一件事情触发了另一件事情。公鸡在黎明打鸣并不会导致太阳出现，虽然太阳通常是在公鸡打鸣之后出现。这是最常见的错误(找找这类错误，你会发现在所有媒体和糟糕研究中都有这类错误)，也可能就是调查和统计声名狼藉的“元凶”。问题很简单，但很微妙，很容易混淆，就因为有一群人都喜欢一个产品，并用了很多次，并不意味着喜欢产品会让人更多使用产品，也不意味着频繁使用会让人们更喜欢它。这两种现象可能毫无关联。
- **不细分子人口群体。**有时候看起来像单一趋势的情况实际上是由于多种不同人口多种趋势综合的结果。想知道是否属于这种情况，要看看答复的分布情况，而不仅仅是综合数字。答复分布情况常常和总结数据说明的情况不一样。例如，如果正在进行满意度调查，有一半的人说他们“极其满意”，而另一半的人说他们“极其不满意”，那么只看中间值显然无法充分了解受众的看法。
- **用事实来混淆观点。**调查问题度量的是观点，而不是事实。如果一项调查问：“您看过这个横幅广告吗？”得到的答复和人们相信的东西有关，但他们相信的东西可能和现实没有任何关系。这就是为什么针对未来行为的问题很少能代表人们会实际采取的行为：在填写调查时，人们认为他们会以某种方式采取行动，而实际很少会这样。

即使能找到答复之间的明显区别，并能通过适当的描述来展现这些区别，但人们的答复仍然会存在若干性质问题，解释调查数据的时候必须考虑这些问题。

- **人们总是什么都想要。**如果用户人口足够大，就会有一群人想要任何可能的特性组合，如果从实际产品中抽象出可能的特性清单，每个人都非常想要一切东西。为什么不要呢？即使知道不可能，希望东西能便宜、好、速度快又有什么错呢？因此，调查不





能用来确定人们不想用哪些特性——不存在人们不想要的东西——但调查能告诉你人们如何确定特性的优先级别以及他们最看重哪些特性。

- **人们会夸大其词。**介绍自己的时候——即使是匿名——我们几乎总是会按照自己希望变成的样子来介绍自己，而不是按照我们实际的样子来介绍自己。因此，我们会夸大积极特点，而弱化我们的失误。从表面来看，人们所描述的观点和行为几乎总是好过实际想法和行动。
- **即使对答案感觉不强烈，人们都会选择一个答案。**有一种强大社会压力要求人们形成观点。如果请人们从选项列表中进行选择，即使人们觉得自己的感觉、想法或者体验不在可用选项中，他们都会选择一个答案。这是多项选择调查的失败之处之一，而且问题选项需要仔细研究和撰写，这就是为什么提供“无”、“不适用”或者“不关心”选项如此重要的原因。
- **人们会猜测调查意图。**回答任何问题的時候，人们常常会想了解提出问题的人为什么要提出这个问题，这个人想听到什么内容。人们会猜测调查者“真正”想要的答案。这就是为什么要避免提出引导性问题的原因，但解释人们答复的时候也要牢记这点。最好能对受访人进行预先测试和访谈，以避免提出带有歧义的问题。
- **人们会撒谎。**当然，并不是所有人都会在所有时间对所有事情都撒谎，但如果说出事实不会获得奖励或者感觉不舒服，人们就会夸大和伪造信息。例如，如果要求人们提供地址，以便寄送奖品，人们不大可能会对收信地址撒谎，但询问他们的家庭收入，如果他们觉得如实回答不会带来好处，就不太可能会说出真实收入。

最终，分析调查最好的方式是聘请拥有合适调查研究经验的专业统计员，和专业统计员一起工作能回答关于产品的问题。理想情况下，甚至可以在编写调查之前就可以和统计员一起开始工作。和统计员一起工作的时间越多，就越能确定问题、发现答案。



但如果无法找到统计员，也不要回避调查。没有专业统计员，仍然可以展开调查，并产生有效、有用、有意义的结果，但要采用直接问题和简单分析。采用问题较少的有限调查，对一群已充分了解的客户进行调查，可以揭示关于用户群体的很多情况，足以形成进行其他研究的基础。

## 跟进和持续研究

执行完一个调查后，不要认为调查就完成了，不能认为调查过程也结束了。随着网站发展和变化，受众以及你对受众的了解也在发展和变化。跟进定性研究，并跟踪受众的变化有助于指导其他研究，并能预测受众需求，而不仅仅是对他们的需求做出反应。

### 跟进定性研究

调查研究指出人们如何认识和思考自己、行为和你的产品，但这种方法太受限制，无法清楚解释人们为什么会有这种想法。因此，需要跟进定性研究。

要了解人们的价值观和原因，最好的工具之一是焦点小组(详见第 9 章)。例如，如果在进行满意度调查，受众说他们对某项或者某些特性不满意，几乎不太可能知道他们为什么不满意。是因为特性的想法？是因为具体实施？是因为和其他特性的交互方式？如果不直接询问人们，就很难理解为什么，但如果没有首先进行调查，焦点小组系列就可能去关注不同的、不太重要的功能特性组，而不关注对受众真正有用的特性。

要了解人们的实际行为，而不是了解人们在调查中报告的行为，直接观察很重要。情境调查(第 8 章)可以揭示人们做出决定的一定情况，而日志分析(第 13 章)能揭示人们行动的纯粹模式。如果人们在调查中说每小时阅读在线新闻两到三次，在一天中观察一组人几小时就可以知道这个数字的准确性。如果只有几个人遵循“每小时两到三次”的模式，解释结果的时候就要对这种说法就要保持怀疑了。





可用性测试(第 10 章)和其他大声思考方式可以揭示人们的决策,以及哪些功能会引起他们对产品的认知。如果他们不喜欢,也许是因为他们不能使用,或者也许他们喜欢,是因为速度快,或者速度并不重要,他们不喜欢,是因为按钮上的黑底红字,他们无法找到要找的内容。很难通过调查知道观点的形成过程,但一旦你知道这些观点,肯定有助于关注之后要研究的问题。

### 跟踪调查

通过定期以相同方式展开相同的调查,可以跟踪网站受众的变化。例如,随着某种服务越来越流行,就很可能吸引越来越多的主流用户,但能有多少人呢?如何定义“主流”?对以相同方式受邀的相似数量的人重复进行相同的调查能发现用户背景资料是否发生了变化,如果发生变化,变化过程又如何。

### 精细调查

如果确定了定义受众的一组“核心”特征,就可以实施更多调查,提出更多问题,以加深了解。因此,如果确定了定义受众最重要的因素是他们的计算机水平、使用电脑的频率和使用的软件,就可以实施调查——除了这些问题之外——提供更多问题来了解偏好、满意度、使用产品的共同方式等等。从问题长度看,在一个调查中提出所有这些问题也许无法实现,但把“非核心”问题分布到类似规模的组中,同时采用类似的编写方法,能比用其他方法了解更多情况。

### 前/后调查

有时候需要知道受众对特定变化的反应如何变化。这种变化可能是重大界面变化,或者可能是广告活动。在网站或者营销的变化前后,进行相同调查可以揭示产品变化会造成用户观点或者用户人口构成产生如何变化。

顾名思义,前/后调查是在某个事件前后进行的调查。对结果进行比





较，如果有用的话，可以发现这些变化对用户人口产生了什么作用。进行广告活动后，产品是否吸引了新的用户群体？重新设计后，用户是否更满意？

前/后调查之前，要确定观察哪些变量。对拟实施的变化有什么期望？不想改变什么？撰写调查的时候，脑海中牢记这些问题，确保调查中包含的问题能解决这些问题。

还要了解这些调查的时机选择所产生的影响，以便在这些变化的效果影响受众之前进行“前”调查，而“后”调查要在效果最大的时候进行。你期望最显著的变化会发生在什么时候？变化会立即影响用户人口，还是需要一段时间才会产生影响？你期望拟实施的变化引发人们讨论吗？在变化之前仔细考虑这些事情能减少对两组进行观察的“噪声”。

一般情况下，多个调查不仅可以监测受众发生了什么变化，还能监测到如何变化。理想情况下，要在变化前进行两次调查，并加以比较，以便知道人们答复方式的自然变化方式(人们答案中的自然偏见)。变化后的若干调查有助于跟踪变化如何发展。例如，变化后一周进行了一项调查，变化若干月后又进行调查，可能会告诉你哪些变化属于短期，哪些属于长期。如果产品在你进行测试的方面没有发生显著变化，即使在预先调查一年后进行后调查都是可能的。

实施多个调查时，最关键的事情是尽量保持调查简单，不要改变措辞和展现方式，也不要改变邀请人们参加调查的方式。以同样方式进行分析。然后比较分析，并注意你认为两次调查中发生了变化的要素。假设变化是为了获得不同市场的信息——你的市场实际上有什么不同吗？它和预期的方式不同吗？

再强调一次，分析多个调查时，最重要的事情是确保已经提前设定问题，并且整个调查都以回答这些问题为重点。否则，就会迷失在调查产生的数据风暴中。

本章仅介绍了调查可以做什么的层面。调查方法的可能组合形式没有限制，但如果仔细使用调查方法，加上支持性的研究，就可以深入了解受众真正是谁以及他们在想什么。





## 实例：新闻网站调查

以下调查用来建立电台网络网站的用户背景资料，以发现驱动人们访问网站的一般信息类型。其目的是揭示访问者的期望，以优化内容展现形式，为后面的定性研究提供指导。调查的次要目标是对网站功能划分优先级，并进行竞争格局的基本分析。

问 题	答 案	原 因
您多长时间收听一次新闻广播？ [弹出形式]	<ul style="list-style-type: none"> <li>○ 超过一周一次</li> <li>○ 一周一次</li> <li>○ 一月一次</li> <li>○ 少于一月一次</li> <li>○ 从不收听</li> </ul>	和前一次调查保持一致 验证新闻电台的听众关系
您多长时间访问一次网站 [网站名称？] [弹出形式]	<ul style="list-style-type: none"> <li>○ 第一次访问</li> <li>○ 少于一月一次</li> <li>○ 一月一次</li> <li>○ 一周一次</li> <li>○ 超过一周一次</li> </ul>	和前一次调查进行比较 汇总对比功能 汇总对比访问原因
您为什么今天要访问网站？ (仅选一项) [单选框]	<ul style="list-style-type: none"> <li>○ 想阅读新闻或者信息</li> <li>○ 想听广播节目</li> <li>○ 进行研究</li> <li>○ 想购买唱片或者稿件</li> <li>○ 想购买唱片或者稿件以外的东西</li> <li>○ 看看网站上有没有新内容</li> <li>○ 和其他听众聊天</li> <li>○ 和工作人员和在线人物联系</li> <li>○ 其他（请说明）：_____</li> </ul>	发现访问网站的一般原因
如果您不是首次访问网站，以上原因是您访问网站的典型原因吗？ [弹出形式]	<ul style="list-style-type: none"> <li>○ 是</li> <li>○ 否</li> <li>○ 不适用(我是首次访问)</li> <li>○ 不适用</li> </ul>	交叉汇总原因



续表

问 题	答 案	原 因
如果您想阅读新闻或者信息,今天您访问网站发现了什么内容? (仅选一项) [单选框]	<input type="radio"/> 不是来看新闻 <input type="radio"/> 当前头条新闻 <input type="radio"/> 特定当前新闻事件的信息 <input type="radio"/> 在广播节目上听到的当前新闻报道信息 <input type="radio"/> 最近新闻事件的深度分析 <input type="radio"/> 评论或者观点 <input type="radio"/> 新闻制造者介绍 <input type="radio"/> 对特定话题的深度研究 <input type="radio"/> 文化或者艺术新闻报道 <input type="radio"/> 娱乐 <input type="radio"/> 广播节目时间表 <input type="radio"/> 特定广播节目的信息 <input type="radio"/> 其他(请说明): _____	如果一般原因和新闻或者信息相关,发现对访问原因的更多具体信息
如果您访问的目的是为了收听特定广播节目,请从下列列表中选择您收听的节目。 [弹出形式]	<input type="radio"/> [节目名称列表] <input type="radio"/> 不适用(不是来听节目) <input type="radio"/> 其他(请说明): _____	了解人们明确收听哪些节目
请选择您定期寻找信息的主题。 (请选择所有可能的主题。) [复选框]	<input type="checkbox"/> 政治 <input type="checkbox"/> 娱乐 <input type="checkbox"/> 体育 <input type="checkbox"/> 当前事件 <input type="checkbox"/> 商业 <input type="checkbox"/> 科学技术 <input type="checkbox"/> 有趣的人 <input type="checkbox"/> 本地文化活动 <input type="checkbox"/> 本地新闻 <input type="checkbox"/> 地区深度报道 <input type="checkbox"/> 旅游 <input type="checkbox"/> 时尚 <input type="checkbox"/> 其他(请说明): _____	发现感兴趣的一般主题





续表

问 题	答 案	原 因
请选择您至少每周访问一次获得新闻或者消息的网站。(请选择所有可能的网站)[复选框]	<input type="radio"/> www.npr.org <input type="radio"/> www.cnn.com <input type="radio"/> www.nytimes.com <input type="radio"/> www.news.com <input type="radio"/> www.bloomberg.com <input type="radio"/> news.yahoo.com <input type="radio"/> www.msnbc.com <input type="radio"/> www.ft.com <input type="radio"/> www.wsj.com <input type="radio"/> www.usatoday.com <input type="radio"/> www.espn.com <input type="radio"/> www.salon.com <input type="radio"/> www.slate.com <input type="radio"/> 其他(请说明): _____	竞争性分析
对于在线阅读新闻,或者阅读报纸新闻,您认为下列内容的价值有多大?[单选框网格,“没有价值”,“有些价值”,和“绝对有价值”。]	<input type="radio"/> 显示新闻报道提到的具体位置的地图 <input type="radio"/> 总结和解释新闻报道中信息的图表、表格、图解 <input type="radio"/> 显示新闻报道或者特别报道所描述的时间的照片 <input type="radio"/> 从视觉上陪伴你阅读报道的照片集 <input type="radio"/> 报道中的人物特写照片 <input type="radio"/> [等等]	了解对不同内容的渴望程度



续表

问 题	答 案	原 因
根据访问网站时的使用频率, 请对下列网站功能进行评级。 [单选框网格, “从不”, “有时候”和“经常”按钮]	<ul style="list-style-type: none"> <li>○ 今天、本周或者今年[网站名称]用户阅读或者收听的前十位新闻报道列表</li> <li>○ 与某一报道或者主题相关的书籍列表</li> <li>○ [网站名称]读者的投票或者调查</li> <li>○ 和记者、主持人或者新闻制造者在线聊天</li> <li>○ 对主题在线讨论</li> <li>○ 与某一报道相关的到其他网站的链接列表</li> <li>○ 把报道用电子邮件发送给朋友的能力</li> </ul>	了解对网站不同特性的渴望程度
请对下列您在网站上阅读的新闻特点的重要性进行打分。 [单选框网格, “不重要”, “有些重要”和“非常重要”单选框]	<ul style="list-style-type: none"> <li>○ 它们有最新突发新闻信息</li> <li>○ 它们提供了足够背景信息, 能帮助我理解新闻的真实意图</li> <li>○ 新闻属于原创, 对共同新闻的报道角度没有料到</li> </ul>	人们看重新闻的哪些特性? <ul style="list-style-type: none"> <li>● 时间性</li> <li>● 背景</li> <li>● 原创观点</li> </ul>
您上网使用的显示器分辨率是多少? [弹出形式]	<ul style="list-style-type: none"> <li>○ 1600×1200</li> <li>○ 1280×1024</li> <li>○ 1024×768</li> <li>○ 800×600</li> <li>○ 640×480</li> <li>○ 其他不知道的分辨率</li> </ul>	
您是否拥有或者定期使用 PDA, 比如 Palm Pilot, iPaq, 或者 PalmPC? [弹出形式]	<ul style="list-style-type: none"> <li>○ 是</li> <li>○ 否</li> </ul>	





续表

问 题	答 案	原 因
您连接互联网的速度属于哪种下列哪种类型? [弹出形式]	<ul style="list-style-type: none"> <li>○ 28.8 Kbps 调制解调器</li> <li>○ 56 Kbps 调制解调器</li> <li>○ ISDN (128K)</li> <li>○ DSL (128K+)</li> <li>○ 电缆调制解调器</li> <li>○ T1 或者更高</li> <li>○ 其他</li> <li>○ 不知道</li> </ul>	
您是男性还是女性? [弹出形式]	<ul style="list-style-type: none"> <li>○ 男性</li> <li>○ 女性</li> </ul>	所有人口统计问题会和之前的线上线下调查研究所得的背景资料进行比较
您属于哪个年龄段? [弹出形式]	<ul style="list-style-type: none"> <li>○ 18 岁以下</li> <li>○ 18 ~ 24</li> <li>○ 25 ~ 34</li> <li>○ 35 ~ 49</li> <li>○ 50 ~ 64</li> <li>○ 65 或 65 岁以上</li> </ul>	
您完成的最高学历是什么? [弹出形式]	<ul style="list-style-type: none"> <li>○ 文法学校</li> <li>○ 高中</li> <li>○ 高中毕业或同等学力</li> <li>○ 大学</li> <li>○ 大学毕业</li> <li>○ 学士/硕士学位</li> </ul>	



## 第12章 持续关系

大多数网站使用的次数都不会只有一次。几个月或者几年内，这些网站都对客户有用。人们使用产品以及与产品的关系会随着时间而改变。人们会逐渐习惯如何使用产品，学会关注哪些东西，忽略哪些东西；而随着对产品的了解不断加深，他们形成了产品的使用习惯。如果一切顺利，客户的忠诚度和好感会提高，而怨恨和沮丧会减少。

几乎所有情况下，你都希望人们使用产品时感到舒服，并能了解产品更多微妙的方面。但为了让产品能支持用户的长期使用，开发团队要明白人们和产品什么时候有关系、有什么关系，并要了解这种关系的变化。本书提到的大部分方法不是为了支持这种观点。这些方法能了解人们的直接体验，即曲线上的一个点。了解曲线上的点是有价值的，但并不能定义曲线。了解曲线形状有助于预测人们想要什么，有助于设计出合适的体验方式。

因此，需要一些有助于了解行为和态度如何随时间变化的方法，以及这些变化中的模式。

问题的解决方案起初看起来很简单：一次招募一些用户，定期邀请他们，并观察他们的使用和看法如何变化。不幸的是，实际没那么简单。如果通过常见方法——可用性测试，焦点小组，调查——对相同的人采用相同研究，这些人会开始失去客观性，研究过程也会失去有效性。

此外，用户研究通常是在强烈的人工环境下进行的。在这种环境中，研究会改变人们在日常生活中的看待产品的方式。他们对产品和他们自己产生了解，而这些往往无法通过其他途径了解。熟悉程度的增加改变了他们的反应能力，而如果没有经过测试，这些能力并不会改变。为了弥补这种效果，需要通过不同于定期反馈工作的方法来处理长期反馈。

本章将回顾几种方法——日记、咨询委员会、Beta 版测试计划和研究





嵌套(Research Telescoping)——这些方法能了解产品对人们体验的长期效果，体验如何随时间变化。

### 背景

使用产品一段时间之后，人们对产品和产品在自己环境所处位置的了解也会发生变化。根据人、产品和能用产品做什么，对这些变化的准确性的看法也会变化。家庭办公室用户浏览办公用品网站想更换碳盒，不断购买产品的办公室经理，两者会用不同角度看待网站内容。后者的起始角度很像前者，但后者对网站的长期使用会完全不同。

从新手到专家的进步过程中，会发生几件事情。

- **犯错误。**很多人(可能大多数人)通过到处提问来了解软件产品。他们猜测功能和特性。有时候能猜对。但经常都猜不对。这些错误信息丰富而且频繁发生。在了解过程中，人们会犯各种错误。
- **建立心智模型。**产品属于黑盒子。在学习使用产品时，人们会创建黑盒子如何工作的模型，以便在意想不到的情况下预测黑盒子的用途，以及怎么让黑盒子做自己想做的事情。
- **设定期望。**随着熟悉程度提高，人们学会了预测产品提供的体验。他们期望屏幕特定区域会出现特定内容。他们期望这些情况能遵循特定模式。
- **形成习惯。**人们习惯以某种方式来做事情。他们会找到能起作用的命令模式，即使这些模式效率不高，也不是在理想情况下产品应该发挥的功能。
- **产生观点。**使用过产品的人很少会完全接受或者拒绝产品。他们喜欢一些地方，不喜欢另一些地方。有时候，随着人们经验的增长，开始看起来混淆或者不必要的事情会变得优雅而且有用。还有些时候，轻微刺激会造成人们彻底厌恶某个特性或者子系统。
- **发展情境。**产品和用户目标、工作以及他们和所用其他工具的关系会随着时间变化。有些产品不可替代，而有些则被边缘化。

#### 提示：

所有这些情况都发生在人们的整体体验情境下。随着人们使用不同类似产品，这些知识针对整个类别产品都在发展，而不只是针对你的产品。很难跟踪全系统的作用，弥补则更困难，但其重要性并不亚于人们对单一产品的体验。





所有这些变化都会影响产品的用户体验，而且除非追踪进程，否则很难以系统方式捕捉并理解这些变化。

## 日记

随着时间推进，观察人们使用产品既困难又费时，而且可能会影响他们如何使用。顾名思义，日记研究是建立在使用产品时一群人不断写日记的基础上。日记能跟踪人们犯的错误，学会了什么，经常如何使用产品(或者体验研究员感兴趣的其他任何东西)。然后对日记进行编码并分析，以确定使用模式和要检查的常见问题。即使是最小量的分析，日记研究也能提供通过其他方式难以发现的模式的反馈来源。日记最适合用于全功能产品。

出于实际原因，间接观测也能提供良好知识，而且花费要低得多。日记研究是最不昂贵的方法，又能随着时间跟踪特定用户使用网站的变化情况(两种其他方法虚拟可用性测试和点击流日志分析，分别详见第 16 章和第 13 章，但两者会涉及更多的数据处理、劳动和费用)。让人们可以跟踪自己的进度，你就能了解他们的体验，而不用留意盯着他们看——无论是从字面理解还是比喻理解——而且一次就要看几个月。

日记也是一种进行地理分布定性研究方法。可以请全国各地的人(或者全世界各地的人，如果重要的话)填写日记，而不必离开办公室。这样就可以进行国际化研究，并研究文化和地理差异会如何影响人们对产品的体验。在开始日记研究之前，需要根据人们使用产品的频率确定几件事情。

- 研究跨度。可能要有至少 6 本日记用于观察趋势(但当然要取决于很多因素)。如果产品每周使用一次，大约需要 2 个月的研究才能度量变化。但如果预计产品每天都使用，一周时间就可以足以发现人们的使用变化。
- 研究抽样比率决定了能观察到的趋势详细程度。人们填写日记越频繁，就越能发现体验的微妙变化。但变化会以不同级别的细节发生，而且一天使用多次产品会存在更重要的月度和年度趋势。

人不是日记填写机器，因此，如果抽样和时间跨度挑选不会占用他们





### 提示：

招募能长时间坚持写日记的人可能会使你寻找的人带有一定偏见。因为伴随愿意写日记这种勤奋而来的其他特质并不是一般用户人口都有的。虽然日记研究所带来的价值能抵消这种偏见，但在分析结果的同时应该明白这一点。

时间、也不会让他们感到厌倦，可能会得到质量更好的信息。例如，研究搜索引擎中级用户如何学习新的搜索引擎，可以从招募大约每天都会使用搜索引擎的人着手(要和所研究的搜索引擎不同)。挑选每日搜索者，定义出最高抽样率，因为要求他们每天填写一条以上的日记不会产生任何额外信息。

日记研究有两种。一种是参加者驱动的非结构化日记研究。非结构化日记的研究结构松散，“写日记的人”和日常体验相关，跟踪学习情况和遇到的问题。另一种是结构化日记研究，类似扩展调查或者自我管理的可用性测试。在主持人的远程指导下，填写日记的人执行特定任务，检查产品的特定方面，并以预先确定的日记格式报告体验情况。

5~10 位参加者足够小型研究团队(或者甚至单一研究员)来管理。参加者应该在整个测试期间都有时间，他们愿意承诺定期参与，他们有能力表达自己的体验，并且——在大多数情况下——所有参加者的体验水平要类似。守时固然好，但很难筛选。

## 非结构化日记

完全开放式日记研究很少见。即使所谓无结构的日记研究也有某种结构关注对于开发者而言的重要事项。

经过招募、筛选并受邀参加，写日记的人接受了期望他们写什么内容的指导。指导要具体、简洁并且完整。在鼓励积极参与的同时，指导是告诉写日记的人要记录哪些行为的一套原则。

### 日记电子邮件指导样本

感谢您参加评估 Wired 的搜索引擎 HotBot。我们正在为重新设计进行评估，我们希望您能参与，以便我们能根据人们对 HotBot 的日常体验来确定设计改进。研究期间，我们希望您尽可能把 HotBot 用于互联网信息研究需求。您的想法和经验对我们非常宝贵。

在接下来的一个月里，您每周会两次(周一和周三)收到本指导页和日记表格的电子邮件。



应该告诉人们他们的义务并设定期望。在每封邮件中都附上指导说明能提醒参加者思考如何填写日记。

我们想请您在收到下一封日记表格之前填写本次日记表格，并用电子邮件发送到 [diary@adaptivepath.com](mailto:diary@adaptivepath.com) (在收到下一封日记表格的当天发送也可以)。我们估计完全填写需要花 20~30 分钟。

如果您无法在收到下一封日记表格之前发送本次日记，请尽快联系我们。联系信息请看本页底部。

联系信息要很容易看到，应该在合适的时候提醒人们联系信息的位置。

本日记表格用来帮助您描述使用 Hotbot 寻找信息的体验。您不必在任何特定时间填写，想写多少都可以。但我们想您尽可能多写一些。

填写日记时，请记住下列事情。

- 日记内容尽可能与您的体验相关，正面负面，大小都可以。我们对所有体验都感兴趣，无论当时看起来多不起眼。我们特别想知道让人惊奇或者出乎意料的体验。
- 我们的目标是让 Hotbot 能更好为您工作。如果您无法使用某些特性，不是您的错。请详细描述任何这种情况。

在人们报告的所有体验情况中，要让人们知道不是在测试他们，要测试的是产品，但他们有责任详细报告这种失败的情况。

- 无论什么时候您使用了以前从未使用过的 Hotbot 特性，都请告诉我们，无论您是否能让其发挥作用。请详细描述使用情况。

人们可能不知自己在用新东西，但鼓励他们发现并记录新情况有益无害。

- 如果您使用 Hotbot 有问题，也无法解决，请详细描述问题和您的解决方案。
- 请说明您在任何时候采用的搜索字词。

在可能的情况下鼓励人们填写细节。搜索关键字词和产品比较容易记录下来，因此人们可以记录下来，但要避免较抽象或者劳动密集型任务。如





果让人们写下他们点击的每个东西、每个选择的复选框，他们很可能要花费所有时间来填写体验的细小要素，而不是更重要的问题。

- 如果您不确定是否要在日记中填写某些内容，请把这些内容放到日记中。
- 如果您在日记表格到期日时之前没有进行任何搜索，请照实记录下来，并发回给我们。

如果有任何问题，请随时发电子邮件给迈克·库涅夫斯基 <mikek@adaptivepath.com>，或者拨打他的电话(415)235-3468。

指导发出后，要和写日记的本人或者通过其他实时方法(电话、即时通信等)一起检查指导，以便能说明清楚细小的地方。

Sample Unstructured Diary Email Form

**HotBot DIARY**

Please return this diary entry on or before Thursday, June 22, 2003

Today's date: \_\_\_\_\_ The current time: \_\_\_\_\_

1. Approximate number of searches since your last diary entry: \_\_\_\_\_
2. Of those, the approximate number of searches using HotBot: \_\_\_\_\_
3. Please describe your experiences while searching with HotBot. Your description may include
  - what you were searching for.
  - your search procedure.
  - whether you found it.
  - any difficulties (if you were able to find a solution, please describe how you solved it) or unexpected incidents (describe what you had expected and what you got).
4. How well is HotBot working for you? \_\_\_\_\_  
(Please rate your experience from 1 to 5, where 1 means that it's not working at all and 5 means that it's working very well.)
5. What is your opinion of HotBot as a service?
6. Has your opinion changed since the last diary entry? If so, how has it changed and was there a specific experience that precipitated the change?
7. Other comments. Are there any other issues you'd like to tell us about or questions you'd like us to answer?

When you've completed this form, please email it to [diary@adaptivepath.com](mailto:diary@adaptivepath.com). Thank you very much for helping us make HotBot a better product.

If you have any questions or comments about this form, please contact Mike Kuriavsky at [mikek@adaptivepath.com](mailto:mikek@adaptivepath.com) or (415) 235-3468.



根据指导,要采用标准化表格来报告体验。日记研究的表格相当灵活,一般来说,应该随时进行实验。表单的目标是向写日记的人提供相对开放性的论坛,同时鼓励他们能够分享目标、问题和见解。表格可以用纯文本电子邮件、HTML 电子邮件(只有在你知道所有参加者都可以接受 HTML 邮件时才使用),或者是指向到单独的 HTML 表格网址。

日记表格的具体内容取决于产品和研究目标。首先要测试表格的一致性,填写时间不要花费写日记人的太多时间(典型的日记记录要花 20 分钟或者更少的时间,最长 30 分钟)。

和所有体验研究方法一样,在研究开始时确定具体目标是研究成功的关键。随着答复不断发送回来,要仔细阅读答复。如果收到的答复类型看起来并不符合研究目标,就要调整研究格式。幸运的是,调整非常简单,因为日记表格可以很容易改变,指导也很容易进行调整(但所有变化都要在说明中突出,避免产生混乱)。即使没有给人们具体指导的日记研究,也还是很可能会提供一些有趣结果(虽然很少有研究不给予任何指导)。

## 结构化日记

给日记增加结构是确保检查特定方面体验的一种方式。非结构化日记会产生大量你并不关心的产品相关信息,而不涉及你最想研究的信息。对写日记的人进行明确的指导能使他们更好地对重要部分做出评论。但同样的结构会导致答复带有偏见。参加者可能会看到之前从未看过的产品部分,以他们通常不用的方式进行描述,或以新方式使用产品。

有三种不同的结构化日记:结构类似于调查的日记;看起来像可用性测试的日记;类似于问题报告的日记。

### 调查结构的日记

提出一系列问题,写日记的人通过回答这些问题来完成每次日记填写,这种结构性日记看起来更像扩大化的调查。问题可以重复,以度量写日记的人的体验如何直接变化,或者每次日记关注产品的不同方面。

例如,Hotbot 日记表格的调查结构化版本看起来会如下所示。





Sample Survey-Structured Diary Email Form

**HotBot DIARY**

Please return this diary entry on or before **Thursday, June 22, 2003**

Today's date: \_\_\_\_\_ The current time: \_\_\_\_\_

1. Approximate number of searches since your last diary entry: \_\_\_\_\_
2. Of those, the approximate number of searches using HotBot: \_\_\_\_\_  
Please describe what you searched for most recently, providing the search topic and the exact search terms you used, along with any + or - modifiers.
3. How successful was this search? \_\_\_\_\_  
(Please rate the search from 1 to 5, where 1 means that it was unsuccessful, 3 means that the information you found was adequate, and 5 means that you found exactly what you were looking for.)
4. How well is HotBot working for you? \_\_\_\_\_  
(Please rate your experience from 1 to 5, where 1 means that it's not working at all and 5 means that it's working very well.)
5. In your recent searches, did you use any of the search options in the left-hand margin of the main search page (the first page you see if you go to [www.hotbot.com](http://www.hotbot.com))? If so, which ones?
6. If you used any of the tools in the left margin, how well did they work? \_\_\_\_\_  
(Please rate their effectiveness from 1 to 5, where 1 means that they did not help your search at all and 5 means that they were critical to its success.)
7. Please describe your personal strategy for narrowing your search, if at first it is unsuccessful. Has this changed in the recent past?
8. Have any of your views about HotBot changed since the last diary entry? If so, how have they changed, and was there a specific experience that caused the change?
9. Other comments. Are there any other issues you'd like to tell us about or questions you'd like us to answer?
10. When you've completed this form, please email it to [diary@adaptivepath.com](mailto:diary@adaptivepath.com). Thank you very much for helping us make HotBot a better product.

If you have any questions or comments about this form, please contact Mike Kuniavsky at [mikek@adaptivepath.com](mailto:mikek@adaptivepath.com) or (415) 235-3468.

如前所述，只要不给人们带来负担，就可以请人们记录具体行为。因此，不要让参加者记录下上次日记以后所有的搜索内容，每张表格只要求记录最近的搜索情况。

开放式问题给参加者提供了深入解释体验的机会。虽然容易受到选择性记忆的偏见，但请人们提交一段体验变化描述能直接深入了解人们的观点和想法如何变化。

上表针对首页上的特性簇群增加几个子问题。这样，研究员就可以得



到针对产品特定部分的反馈。

### 可用性测试日记

另一种日记研究更像远程可用性测试，请写日记的人根据特定特性执行特定任务，或者让他们看到特定内容。请他们评估网站的这些区域，并描述使用情况。随着时间推移，就可以跟踪到参加者实际使用了向他们介绍的哪些特性。如果要研究特定特性使用随时间的变化情况，这种研究最有用。

#### Sample Usability Test Email Diary Form

#### HotBot DIARY

Please return this diary entry on or before Thursday, June 22, 2003

Today's date: \_\_\_\_\_ The current time: \_\_\_\_\_

1. Approximate number of searches since your last diary entry: \_\_\_\_\_
2. Of those, the approximate number of searches using HotBot: \_\_\_\_\_
3. Please describe what you searched for most recently, providing as much detail as possible about your search procedure, including approximately how long it took.
4. How well is HotBot working for you? \_\_\_\_\_  
(Please rate your experiences from 1 to 5, where 1 means that it's not working at all and 5 means that it's working very well.)

Please examine the Advanced Search page for 2-3 minutes. This can be found by clicking on the "Advanced Search" button in the left margin of the main search page (the first page you see if you go to [www.hotbot.com](http://www.hotbot.com)). Read through the options on this page then go back to the main page.

5. Find a photograph of a Silkie chicken. This is a kind of fluffy white Bantam hen with bluish feet and a blue beak. We are aware of several photos on pages belonging to agriculture schools.
  - You do not have to use any of the advanced options, we just wanted you to be aware of them.
  - Don't spend more than 5 minutes looking for the chicken.
  - Whether or not you can find a picture of the chicken, please describe your search process in detail.
6. Please describe your personal strategy for narrowing your search, if at first it is unsuccessful.
7. Have any of your views about HotBot changed since the last diary entry? If so, how have they changed, and was there a specific experience that precipitated the change?
8. Other comments. Are there any other issues you'd like to tell us about or questions you'd like us to answer?

When you've completed this form, please email it to [diary@adaptivethought.com](mailto:diary@adaptivethought.com). Thank you very much for helping us make HotBot a better product.

If you have any questions or comments about this form, please contact Mike Kuniavsky at [mikek@adaptivethought.com](mailto:mikek@adaptivethought.com) or (415) 235-3458.





再强调一下，为了便于让人们提供你需要的信息，并且如果他们无法提供也不会感觉糟糕，要针对希望他们提供的信息来设定他们的期望。(基本根据约翰·赖曼的著作《日记研究指导实验室工作的面向工作场所研究工具》(The Diary Study: A Workplace-Oriented Research Tool to Guide Laboratory Efforts), INTERCHI 会议, 1993 年阿姆斯特丹, ACM, 纽约。)

上表向参加者介绍了一组特定特性，并创建了利用这些特性可以快速完成的任务。通过检查写日记的人对搜索过程的描述，就可以发现人们是否能理解高级搜索界面，以及人们在初次介绍后是否使用了高级搜索。问题 3 还进行了修改，来要求获得最近搜索的时间，以便看看根据额外特性人们的搜索持续时间是否发生了变化(当然，这项措施太粗犷了，无法发现什么情况，只能发现最极端的行为变化)。

### 问题报告日记

还有另一种日记表格，类似于问题报告或者重大事故报告。无论什么时候有了问题或者见解，写日记的人就进行填写。这种日记并不拥有像远程可用性测试日记所具有的重点或者深度，但能最大限度地减少研究对用户行为的偏见数量，还能突出界面上最令人痛苦的部分。这类问题报告研究可以在整个产品生命周期不断进行，以文档形式记录人们碰到的各种问题。虽然听起来没有“灵魂”，但这种日记很容易大规模部署，从而提高总反馈率。这种日记不能很好地用于系统跟踪使用情况，因为要取决于人们有问题或者想法时能否记得要填写日记。

根据勤奋程度和对问题或者见解的看法，写日记的人填写此表会变成大麻烦。定义开发团队在找的具体参数，向写日记的人传达，并建立强有力的奖励(无论是物质还是提醒)都能降低一些无答复比率。总是会存在一些模糊不清的地方，但要尽可能合理地说明参加者应该写下的范围。

需要不断评估日记形式表格的作用程度来完成研究目标。如果日记表格不能发挥作用(如果写日记的人不是在讨论你所希望的内容，如果他们做出的答复没有完全符合你想要的方式等)，要想办法找出问题，并适当改变表格。





*Sample Problem Report Diary*

**HotBot DIARY**

Diary start date: \_\_\_\_\_

Please make an entry in this diary whenever

- *HotBot fails to perform as you had expected or intended.*
- *You have an insight into how to better use HotBot to find content.*

Please describe the situation in detail.

You do not have to describe the same problem more than once, but please enter it into the diary every time it happens.

a) Date of entry: \_\_\_\_\_  
 Search strategy description: \_\_\_\_\_  
 Problem feature name/description: \_\_\_\_\_  
 Problem: \_\_\_\_\_  
 Severity (1-5): \_\_\_\_\_  
 (A severity of 1 means that it's an annoyance or an observation, and 5 means that it's a catastrophic problem that prevents you from accomplishing your goal.)  
 Insight/Solution: \_\_\_\_\_

b) Date of entry: \_\_\_\_\_  
 Search strategy description: \_\_\_\_\_  
 Problem feature name/description: \_\_\_\_\_  
 Problem: \_\_\_\_\_  
 Severity (1-5): \_\_\_\_\_  
 Insight/Solution: \_\_\_\_\_  
 Etc. \_\_\_\_\_

Please email this form to [diary@adaptivepath.com](mailto:diary@adaptivepath.com) on Thursday, June 22, 2003, even if it's blank.  
 Thank you very much for helping us make HotBot a better product.

If you have any questions or comments about this form, please contact Mike Kuniavsky at [mikek@adaptivepath.com](mailto:mikek@adaptivepath.com) or (415) 235-3468.

## 管理答复

日记研究成功的关键之一在于管理答复和答复者。由于大部分内容都是自我发起，人们很容易把写日记当成可选可不做的事情。为了最大限度提高答复数量，必须鼓励人们，并提醒他们履行义务。

如果参加者没有得到充分鼓励，退出率会很高，相当于高自我选择率





并会歪曲答复。完成日记研究却没能充分鼓励的人可能无法代表目标受众。

奖励要和参加者完成工作所花费的时间成正比。完成6次日记可能总共需要3小时，还需要另外1小时进行管理。奖励支付给日记参加者和支付给实验室的参加者，前者要是后者的60%~100%之间。假设标准费率为每小时60美元，也就是说，为期两个月的研究，每周要求写两篇日记，研究结束时，每位参加者应该得到约200美元酬金。可以有办法便宜一些（我参加过一次日记研究，我得到了50美元和一件T恤），但退出率可能会很高——8个人中只有1人完成了研究。你可能想在整个过程中通过更多奖励（T恤和餐厅礼券等）来鼓励完成。如果和一家公司一起工作时发现奖励措施不合适，可以考虑请管理参加者的人分配时间来填写日记。

提醒可能是弥合非面对面研究裂缝的重要工具。提醒参加者他们的承诺以及承诺对产品的重要性，可以激励他们提供更好的反馈——并把日记放到他们的任务清单上。对没有填写的日记马上打电话跟进，告诉参加者你仍然在等他们答复，他们的反馈意见对你非常重要。定期发送电子邮件向他们表示感谢，并回答他们的问题，也能加强这一点，但不要一天给他们发送12封“请继续好好工作”的垃圾邮件。

### 如何分析日记

日记分析类似于焦点小组分析。第9章讨论的方法大部分都能适用，而且不会重复。从研究目标着手，然后根据共同主题系统检查人们的答复内容。这些主题之后可以组合成趋势。

分析的第一部分就是阅读日记。从一收到答复，就要尽快开始阅读。阅读填写内容，并比较参加者之间的答复。推测人们的行为，并解决初步目标。注意出现的其他趋势。随着收到的日记越来越多，要根据新的信息来测试和修改这些推测。

日记分析可以只是仔细阅读，但也可以采用更严格的方法。正式分析虽然要花更多时间，但可以发现常规分析不能发现的不明显的使用模式，



**提示:**

如果日记条目含糊不清,要发电子邮件或者打电话跟进。当然,答复会经过答复者的记忆过滤,但具体问题(或者请他们看看网页或者正在讨论的产品部分)能直接切入问题根源,或者能提供有价值的见解,省得你费尽心思去猜。

并使数据中的一般推论变得严谨。正式分析后,可以说“在研究中,6位参加者中有4位5周定期使用高级选项”,而不是说“大约一个月后大多数人都了解了高级选项”。

**严格趋势提取**

如果要让趋势更清晰,可以从日记条目中提取问题,以便可以用相同名称称呼相同现象,而不用管什么人在什么时间填写的条目。

如第9章和第13章所述,日记编码是采用统一标签对观察结果进行分类的有组织方法。建立了一套可行的编码,就能根据编码针对条目中出现的评论类型,对每条日记条目进行检查并分类。你可能想对支持或者反对假设的条目进行编码。此时也是找到代表性语录的好时机。

对所有条目进行了编码,接下来开始组织条目。组织答复的主题取决于你想回答的问题,并且有无限多种方式。单个人的体验中可能会存在有趣的群体意见,比如当答复人学会了更好地使用工具,并能记下甚至更高级评论的时候,或者如果每个人的学习速度都一样,在若干人中会出现簇群的时候。在人们所说的内容、人们评论之间的关系、他们所说的时间以及评论所指的特性中(有些特性会有很多评论,而有些则很好),都能发现行为相似之处。

时间表常常能马上告诉你一些其他方法很难观察到的有趣模式。例如,在 Hotbot 研究中,为了看看人们提到高级选项的次数是否多于他们使用的次数,可以把“提到高级搜索选项”的次数(其中一个编码类别)和时间对应起来。

编码和组织的过程要能让你知道数据中有哪些明显趋势类型。但如果有了信息簇群清单,就必须要把它们和人们的行为关联起来。

有些关系会显而易见。图12.1中,李并没有经常讨论高级搜索选项。但随着研究进展,他讨论到的次数越来越多。同样,即使刚开始没有用户讨论高级搜索选项,但到第4周,5个人中有4个人都至少谈论了一次。



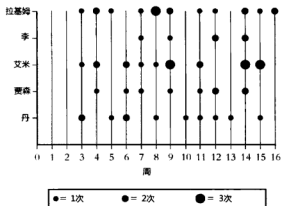


图 12.1 HotBot 日记中提到高级搜索选项的次数

能发现多少趋势，就列出多少趋势。有些趋势可能显而易见，你对它们代表实际行为模式的信心就会高。有些趋势可能很难发现(人们在第12周使用 HotBot 高级搜索选项的次数比第6周多吗？也许多，但很肯否定)。应该用标记说明各种趋势(“第3周效应”)，以区分并在以后讨论这些趋势。

### 提出假设

有了趋势清单，接下来就要来了解原因。由于不是纯科学(而是高度实用)，这种假设的真实性不必特别完整。为了做出产品开发决定，如果假设很可能正确，差不多就够了。

因此，现在就需要做出假设。李谈论高级搜索选项的次数看起来为什么会随着时间增加？这可能就表明他使用高级搜索选项次数越多，就对高级搜索选项就越有好感。也可能表明他寻找搜索的次数在增加。快速检查他的日记能说明是哪种情况。如果是因为他对高级搜索选项越来越有好感，通过查看其他用户的日记就可以指向能进行研究的学习曲线。

如果把日记研究和焦点小组或者访谈结合起来，就可以检查假设和人们实际体验的对应情况。结合在一起，你和记日记的人可以独立出产品体验的最重要部分。



## 自行报告

如果请人们报告自己的体验,就是要请人们走出对自己的正常看法,并对他们自己的行为发表评论。有些人这么做没问题,能提供准确诚实的答案。而另一些人就会比较困难。即使他们想帮助你,也觉得把所有事情都告诉你,他们可能对承认失败或者承认自己不知道情况而感到不舒服。如果阅读任何自行报告的信息,比如支持评论、调查问题或者访谈答复时,必须要考虑上述情况,但在日记中特别重要。日记条目常常是能和写日记的人只有联系上的唯一方法,而且也没有办法知道写日记的人的真实情况。

就像在新闻报道中一样,要尽可能独立核实自行报告的行动。请参加者对重要陈述解释清楚。通过额外访谈或者研究进行跟进。如果感觉的确模糊不清,可以比较不同日期的日记,看看所报告的行为是否吻合。要一直保持怀疑态度。

## 咨询委员会

创建用户咨询委员会(有时也称为用户“评判小组”,“团队”或者“委员会”),这种方法能把用户输入转化到软件开发过程,而且受到广泛欢迎。咨询委员会由一组用户构成,在任何时候,如果觉得需要最终用户对产品变化的输入,开发团队都可以向这些用户寻求帮助。

从理论上说,这种方法为开发团队提供了能从产品直接受益(或者因为产品想法糟糕而受到伤害)的人们那里获取想法的简单方法。咨询委员会成员能很快熟悉产品特性、产品和公司正在努力解决的问题,熟悉开发团队成员。这种熟悉可以缩短参加者对产品 & 公司的熟悉过程:大家都知道产品目前的问题,重点只会放在新想法上。一方面,可以节省大量时间。但另一方面,让这种方法看起来有吸引力的各个成员对产品的相同熟悉也是其最大问题。

用户和开发团队讨论产品越多,他们就越会像开发团队一样考虑产





### 忠告：

如果咨询委员会主要由买家构成——尤其是大买家——而不是由用户构成，会造成开发人员和销售及营销人员之间关系紧张。销售和营销人员自然会向委员会提出新产品理念来推销这些想法，而委员会成员可能希望在销售环节中得到优先处理。需要技巧来平衡，既不能导致委员会产生偏见，也不能失去他们的工作。这种情况下，要向两组人解释委员会的目的，并设定合适的期望。提前说明对委员会成员不会有销售行为，而同时委员会成员“都带着委员会成员的帽子”——向委员会和销售人员都说明。另一种良好方法是不在现场见面，谈论数小时，吃点东西。这样能减少参加者预期成为销售对象的可能性，并减少他们成为销售对象的机会。

品。如果他们的观点和开发团队的关键相融合，而这很快会发生，他们就不会再是局外人。他们也不会再提供公正用户的角度。实际上，他们已作为专家顾问加入了开发团队。

因此，这种关系绝对有价值，但只能把咨询委员会的观点当成领域一小组专家内部人士的观点。

除了上面的忠告外，对任何有稳定长期用户基础的产品而言，都要保持某种形式的历史咨询会。委员会可以在产品发布后就很快形成(如果预发布用户群体相当大，甚至可以在产品发布之前就形成)。小组在开发中的产品形成后，就可以引入更多模糊的产品想法。如果过早在过程中引入模糊的想法会导致产生大量希望清单、几乎没有跟进和失望的咨询委员会。

## 委员会成员

如何选择咨询委员会成员取决于和用户的当前关系。软件公司 SAP 让其每组(自发组织的)用户群体选择一位代表加入公司的咨询委员会。其中大部分人都是经过投票选出，整个过程有点像参议员选举。但大多数公司没有这种有组织用户群体，因此无法进行选择。他们必须用其他方式找到成员。

一种常见方法是邀请大客户代表。这种方法很有效，但会遇到如下事实，即大多数大客户的联系人都是指定或者购买产品的人，而不是实际的最终用户。这些人(例如管理信息系统的管理人员)是目标市场，但不是用户市场。此外，邀请大客户往往会忽略规模较小客户的需求，而较小客户的需求和大客户的需求常常非常不同。

邀请现有雇员的“朋友和家人”也是种常见做法。招募会更容易，但也会碰到委员会成员无法对产品开发提供公正角度的问题。如果我朋友露西知道我花了4个月时间开发一项不能正常工作的新特性，她可能不会对其缺陷发表看法，就好像她不认识我一样。

相对公正的委员会招募方法是把委员会当成焦点小组。根据第6章和第9章所述的邀请方法筛选参加者，但其承诺是持续变化，而不是为了单一群体。这样能保证最大化委员会的公正性，但如果不小心进行招募，就



会导致委员会软弱无力或者效率低下,因为招募到的人可能对产品在未来不会有太多投入或者兴趣。

招募过程和焦点小组的招募过程有点相似,同时委员会成员需要有一些特定特质,而大多数焦点小组不需要这些特质。

- 他们必须知道任务。他们不必知道具体软件或者公司,但应该是产品所要解决的领域的高级用户。
- 他们需要是最终用户,至少在某些时候是。
- 他们应该表达清楚。好的主持人在焦点小组中能对付一两个表达不清楚的人,但委员会的生命周期要长得多,因此,参加者能表达自己的想法、体验和感受就特别重要。
- 他们要有空。委员会成员保持一致有助于保持势头,并保持和开发团队之间的联系。参加者要能承诺至少一年的时间(或者能参加三四个既定会议)。

委员会规模要和焦点小组规模差不多大。总人数不要超过10人。

为了保持角度的新鲜程度,可以定期组成新委员会。通常每年会邀请所有新人参加一次委员会,有些委员会虽然会持续多年,但有些会根据日程表轮换委员会成员。

提示:

要为委员会设定合适期望。作为用户,他们处于特权位置,但他们的输入是——或者应该是——最终用户研究项目成员之一。

## 和委员会共事

咨询委员会会议可以类似于董事会会议或者焦点小组。有些是每月举行会议,有些只能一年开几次。委员会的目的确定了会议频率和主题。

如果已完成的产品有广泛研究事项表,可以每年适当召开数次会议来讨论产品面对的“大画面”战略事项。

对于采用轻量级研究策略快速发展的产品,可以采取安排定期会议日程表,并详细讨论产品。委员会可以讨论当前问题、希望以及产品特性。

所有情况下,都要从整体上提前确定每次会议以及委员会的目的和结构。这就是委员会的“章程”或者“权限”。章程不必广泛,但要具体到委员会的角色和目标。例如,电脑顾问公司(Computer Consultants Corporation)用户咨询委员会的章程如下。





## 电脑顾问公司用户咨询委员会章程

### 任务陈述

用户咨询委员会(UAB)是计算机顾问公司(CCCORP)客户服务承诺的重要组织构成部分。通过帮助计划双方成功的共同未来, UAB 能加强 CCCORP 在信用联盟信息系统服务市场的定位。

### 价值观

UAB 立志为公司未来或者现有或将来客户提供相关的诚实且完整的输入。

UAB 成员资源贡献时间给所有用户支持小组成员, 同时理解 CCCORP 希望并计划为客户提供最佳价值。

UAB 成员会获得个人利益, 而其信用联盟也会从其参加中受益, 但成员立志代表用户支持小组整体的最佳利益, 在讨论和决定时把个人或者特定信用联盟的利益放置一边。委员会内容的讨论、辩论及异议都将视为隐私, 如果没有取得委员会的预先同意, 不会和他人讨论进行内容。

### 政策及程序

1. CCCORP 用户咨询委员成员必须来自信用联盟, 必须是 CCCORP 用户支持小组的当前成员。
2. CCCORP 执行小组提名 UAB 成员。CCCORP 应当决定增加成员及向 UAB 提名的需求。UAB 应该和被提名人见面, 并向 CCCORP 提交建议。CCCORP 应当根据 UAB 的推荐选择候选人。UAB 应当在正式会议上确认并任命候选人。
3. 成员可以一直服务到辞职或者如程序 5(如下)所述的退职时候。当前成员及其约定日期如下: [成员名单略去]。
4. 理事会应当在冬季会议时选择决定官员。官员应当至少有主席和秘书。委员会设定官员职责, 最小职责为他们要确保每年至少有一次年会, 必须记录和保存时间。
5. 可以有原因召回成员, 或者委员会可以根据多数投票处于任何原因退职成员。
6. 所有委员会的行为都要求有法定人数, 在职成员多数定义为法定人数。
7. 最低要求, 使命陈述、价值观、政策及程序至少每年冬季会议检查一次, 所有 UAB 当前成员的三分之二多数通过, 可以修正。

美国农业部的用户咨询委员会也有章程。他们的“用户”是广泛服务的用户, 而不是特定软件产品, 但章程仍能产生启发性对比。它对职责和会员身份的规定更为具体, 但并未具体规定程序。





## 美国农业部用户咨询委员会章程之责任与结构

美国国家农业研究和推广用户咨询委员会(UAB)是依据美国国家农业研究、推广和教学政策法案建立,并经1977年、1981年、1985年和1990年农场法案再次授权的法定委员会。

委员会负有准备对食品、农业和自然资源科学发表观点的一般责任。委员会成员应当做到以下几点。

- 回顾政策、计划、研究目标、教学及美国农业内部推广、其他联邦机构、州立机构和学院及大学。
- 评估在私人领域,以及私人领域和联邦支持的农业科学和教育之间关系的农业研究、教学以及推广活动范围。
- 评估资源配置及研究、教学和推广基金的充分性,并提出相关建议。
- 确认新兴研究、教学和推广事项。
- 推荐计划和技术——向公众和私人农业科学及教育机构转让解决方案。
- 在两份年度报告中做出评估和建议。
- 作为总统、农业部长和参众议院农业和拨款委员会的顾问。
- 向农业部长进行口头简报,报告资源配置及研究、教学和推广基金的充分性情况。

委员会成员资格有意多样化,以便能对预期通过计划受益的人形成综合观点。21位UAB成员来自以下农业领域的13个部门。

- 生产部门(8人):4位成员代表农产品、林产品和水产品;1位代表农场合作;2位代表一般农场组织;还有1位代表农场供应商。
- 消费者利益(2人):其中一位成员必须代表非营利消费者运动小组。
- 食品及纤维处理商(1人)。
- 食品营销利益(1人)。
- 非营利环境保护组织(1人)。
- 城市开发工作(1人)。
- 动物健康利益(1人)。





- 人类营养工作(1人)。
- 食品和农产品运输(1人)。
- 与农业相关的劳工组织(1人)。
- 私人的非营利组织/基金(1人)。
- 发展中国家的私人领域计划(1人)。
- 无研究能力的美国农业部机构(1人)。

## Beta 版测试

**Beta** 版测试是一种用时间证明质量保证(QA)的方法。如果软件已经进行调试,也达到了所有关键指标,就可以邀请有限数量的最终用户进行使用。他们使用的时候——假定是在实际环境中以真实方式使用——会发生其他问题。使用结构化反馈系统,这些用户可以向软件开发者报告这些问题发生的情况。

最终,不会再有关键功能缺陷,软件就可以发布了。

**Beta** 版测试是了解软件问题的好方法,但并不是那么好的用户研究方法。诚然,质量保证和用户研究之间的界线很模糊,但两者之间的确存在界线:一边是用户需求、愿望和能力;而另一边是软件需求、愿望和能力。有两个主要区别:研究发生的时间和参加研究的人。

**Beta** 版测试一般是在开发周期的结尾进行,也就是在网站要上线或者压制最终光盘之前进行。此时,所有功能已经获得一致同意,已经做出界面决定,形象也已确定。更多了解用户也不可能对产品做出任何根本改变。此时,已经假定正在做的东西就是人们想要的,而且最重要的东西也不存在缺陷。如果对产品或者其用户存在根本疑问,就很可能就无法进行 **Beta** 版测试。

此外,受邀测试软件的人不是典型用户。他们充其量不过是早期采用者和高级用户以及受新功能吸引甚至不惜牺牲稳定的人。例如,在视频游戏开发领域中有传奇的专业测试者以及攻破他人软件的专家。他们对制造



提示:

建立所有用户的“支持”邮件列表,鼓励他们分享体验并互相帮助,这是收集有关用户心智模型和期望信息的好方法。这些互动会揭示人们预期如何使用产品、他们如何使用和错误使用。稍加修改意想不到的用户快捷方式或者方法,就可以从整体成为产品的合理特性。

高质量软件非常重要,但无法代表用户。他们的体验和看法只能代表一小部分实际用户。

也就是说,用户研究员不能忽视 Beta 版测试反馈,并要和质量保证小组一起开发 Beta 版测试材料。就像客户支持评论(见第 13 章)一样,Beta 版测试能有效提供更多“免费”的用户人口信息。可能不是典型用户人口采用典型方式使用该产品的信息,但也是部分用户人口,不管怎么说,都能从这一过程中获得有用见解。

从用户研究角度看,可以从 Beta 版测试反馈知道以下情况。

- 用户人口中高级成员的构成。谁是为了新功能可以牺牲稳定性的高级用户? Beta 版测试计划很少调查用户人口,也很少和“普通”人口进行比较。采用标准调查,可以了解 Beta 版测试用户人口真正是谁。
- 他们使用的特性。通过跟踪 Beta 版测试评论,可以形成特性集和报告的各种问题,可以揭示高级用户的兴趣和习惯,而这些人常常是用户人口重要的细分部分。
- 对广大用户人口需求的疑问。Beta 版测试的反馈可以触发其他用户研究主题。例如,如果大量 HotBot 的 Beta 版测试反馈都指向“相关网页”链接,而“相关网页”链接没有被认为是高级用户的重要特性,也许就需要来研究其他类型用户如何期望和理解此特性。

## 嵌套

由于各种原因,有时会无法和用户保持长期关系。可能开发周期很短,可能用户流动性很大,也可能因为其他原因无法与他们保持关系。无论如何,都会需要知道不同经验水平、不同使用体验时长的人随时间对产品的体验情况。

收集这些资料的一个方法是嵌套研究。光学望远镜能让角度扁平化,让遥远的东西看起来就在旁边。嵌套研究能同时把不同体验水平的用户放





在一起来压缩时间。

一般来说，为研究进行招募时，可能会得到不同体验的参加者。有些可能是新手，有些是老手。通常希望就是这样，因为能知道在使用网站过程中影响人们的问题类型。通过邀请人们，可以看到哪些问题影响到了委员会的每个人，以及用户看重哪些东西，不论他们已经使用产品了多长时间。有时候，可能要为了这组人或者那组人有意打乱招募名单。

嵌套可以把这种想法再往前发展一步。如果跟踪受邀经验的人的体验水平，就可以招募到特定专业的人，或者是已经使用了产品很长时间的人。通过跟踪人们根据其经验水平提出的问题和观点，就可以推导出人们将来长期使用产品的变化类型。

例如，通过使用一个月服务的人来进行一个焦点小组系列，而另一个系列由使用了一年的人来进行，能告诉你人们对服务的愿望随着时间推移会如何改变。它不能告诉你何时会发生变化，也不能告诉你在什么情境下会发生，但你能发现人们的看法有什么不同。

## 看得更远

用户体验研究领域还非常新，没有很多现成的方法能让你研究人们的体验随时间的变化情况。大多数研究都侧重初步体验或者典型体验，并不关注体验随时间的变化情况。这很不幸，因为这些研究扁平化了用户能力和观点的复杂方面。应该了解随着时间推移产品如何变化，人们和产品的关系如何变化，这样一来，产品设计才可以随着对用户了解及需求的增长而成长。



## 第 13 章 日志文件和客户支持

很多网站的用户体验研究都是了解人们如何体验以及期望如体验网站，但很难了解人们现在的体验情况。满意度调查可以揭示人们的满意程度，但根据人们的预测和喜好并不能很好地预测他们的行为。情境调查可以发现人们使用网站时存在的问题，但一次只能有一个人。可用性测试能提示很多可能的问题，但无法根据人们在实验室环境中的行为来推测实际生活中的行为。上述方法中，没有一种能实际说明人们现在如何使用网站，他们在真实环境中使用网站到底有什么问题。了解用户的实际行为可以结束起始于研究阶段的循环，还可以结束网站开始开发前建立背景资料的工作。不了解用户的实际行为，就永远无法知道设计是否真的奏效。

幸运的是，已经有两个信息来源能表明用户的当前体验：客户支持意见和网络服务器的日志文件。把人们的行为及意见和你对用户及公司的其他了解合并在一起，能揭示行为的深层模式，并有助于让人们更高效地使用你的产品，让每个参与人都受益。不幸的是，这两种来源都以大型数据为特色，其功能经分析后才能凸显。虽然是能发现用户行为和兴趣的金矿信息，但金子通常不会就在掘金队周围。它需要花工夫来提取，而且要花很多工夫。

收集用户体验的首要来源是用户自己产生的意见。客户支持意见和用户的思想、兴趣、身份和问题直接相关。这个研究领域不需要研究员提问就会有答案。

如果客户反馈的意见客观且有代表性，就没有必要进行任何其他类型的用户研究。人们会告诉你他们想要什么，哪些有问题，哪些没问题。但客户反馈并不客观。人们联系客户支持的原因很多，但都太负面，因此只看用户意见不可能得到对产品所提供体验的平衡观点，但不管怎么说，客





户反馈都是一种基本不需要刻意培养、不应该被忽视的资产。通过客户反馈，能深入了解用户体验的若干关键方面。

- 有助于了解人们如何理解产品，并了解他们使用的语言。有助于了解他们对产品和任务产生的心智模型。
- 能大量发现人们的期望，期望在哪里得到满足，在哪里没有被满足。
- 强调“痛苦点”，有助于建立问题的优先级别，确定要关注哪些特性。
- 能指出研究中要提出的特定问题，指导未来研究方向。

换句话说，这些发现有助于指导研究朝着用户认为的最重要的问题方向进行，如果要理解客户支持意见，应该从了解支持过程开始。

## 客户支持过程

无论正式与否，每种服务都有支持过程。如果人们有疑问，或者他们要提问题，他们就会联系某人，即客户支持，不管他们是否有此称谓。有时候，支持过程的组织形式较正式，但总存在一个过程，能回答人们的问题(或者至少能收集)。

理想情况下，一个有组织的系统能收集问题，进行回答，并产生相关统计数据(典型问题被问的频率，回答的肯定性，更多资料链接等)。情况最坏的时候，这种系统只是记住提问者的电子邮件。不管什么系统，都应该有能收集并分类支持意见的方法。

遵循人们留下的反馈的过程，以确定这种过程如何影响余下的反馈，这是有指导意义的。邀请语气如何？是否对每条意见都有相应的跟进？是否有自动答复机制？如何创建答复？是否需要收集或者鼓励更多信息？每条反馈的收集过程都会导致收集的意见有偏见。如果有人留下反馈的确有困难，收集到的信息只能是来自最有驱动性、最持续、最聪明(在他们思考如何留下反馈的时候)和最不安的用户，这些都将影响他们的反馈方式。



## 收集意见

在自动化系统中，应该已经收集好意见，所以收集过程就容易一些。如果有自动化系统，问题和答案的数据库就是着手的好地方。这类数据库的数据可能经过了很长时间，并且经过了许多版本，对于用户支持，这种归档数据非常自由，保留它们是为了以防万一，但要想知道现在发生的情况，就不会那么有用了。过去几周或者几个月收集到的问题，一般都足以提供需要关注的主题和特性清单。

如果没有自动化系统，收集数据可能会比较困难。可以从访谈支持工作人员开始，了解他们如何回答问题，并收集对常见问题的“库存”答案。这样能从更高的层面上了解问题。

此外，应该看看人们的实际意见。看看所有意见，而不仅仅是被认为是必须回答的重要意见，能更全面地了解用户的体验。没有必要收集大量反馈数据。从上个月中随机抽样 100~200 条意见应该就能很好了解人们的问题和答复，但如果内容非常不同，或者意见量特别大，可能必须收集更多意见。如果有可能，要保存现有的任何其他信息，比如有意见的板块，或者是否在反馈期间伴随着点击流。

## 阅读意见

最快速的分析方法就是阅读用户的意见。我打印过大量批评意见(再说一遍，100~200 条就足够了)，然后去一家安静的咖啡馆待了几小时，找出好的用户意见或清楚的例子。阅读时，我记着以下几件事情。

- **从用户的角度阅读。**对于产品，用户并不知道我知道的所有情况。他们使用的词汇和我不一样，他们不会带有相同感情。要用书面方式来阅读这些意见。
- **关注事实。**什么有问题？怎么有问题？提到了产品哪部分？有些





意见是在愤怒或者兴奋的时候产生的。如果是坏话，不要觉得有夸张之处，如果是好话，也不要就更重视。但让人们高兴或者愤怒的东西能告诉你他们的体验中非常强烈的部分。

- **不急于下结论。**几条信息不足以定义出趋势，但即使单一信息也会包含有用想法。仅仅因为有几个人说了他们喜欢或者不喜欢某个东西并不说明他们代表了相当大的用户人口。
- **不要把常见问题清单搞成必改清单。**投诉指向问题，但并不不是问题描述。有人没有找到“餐叉”链接，他们的投诉并不代表必须立即把“餐叉”链接改得更醒目。这可能意味着划分银制餐具的方式不符合目标受众的期望，或者说明用户不理解用来传达这种结构的方式。要抵制了解情况时想解决问题的迫切性。
- **不要完全相信意见。**大部分人都保持沉默是有原因的，很多人的感觉或者行为方式和发送意见的人并不相同，不要轻视反馈，但也不要把它当成福音。有时抱怨是常见问题的放大版，有时只是没道理的抱怨。

提示：

很容易忽略怒火（“你臭死了！”），而对赞美特别注意（“猫猫，我爱你！”）。不要这样，要停留在反馈的个人性质上，并通过它来了解受众的观点，首先看事实，然后从观点产生的情境看，最后再从情绪反应看。

出于分析考虑，我跟踪了受众体验的多个方面，并想发现其中的模式。

- **用户是谁？**人们在解释或者请求的时候会给出大量自身信息。不同群体的人也许会有不同意见。
- **他们想做什么事情？**有没有他们想完成的目标？中间目标是什么？（根据用户所描述的，并不是交互模型的一部分）
- **他们如何处理问题？**尽管具体问题不同，可能还是有共同策略能帮助人们了解如何使用系统。过程中在类似地方发现的类似描述和问题能更深入了解人们的心智模型。
- **他们碰到了什么问题？**人们碰到的问题有相似性吗？有没有明显的瓶颈？

## 组织和分析意见

如果系统保存了统计数字，应该从那里开始探寻。这样可以设定初始





期望，并能快速知道用户觉得什么东西值得发表意见。不要把意见的数量和比例当成意见所描述问题的严重程度，但这些意见能告诉你受众的一些情况和他们关心的东西。

正式意见分析从组织开始。组织是通过意见主题对意见进行分组的过程。分好的族群可以让你减少人们说话中的偏激意见，从而转变成更容易管理的東西。

组织意见有一个快速方法，即通过意见主题和严重程度。主题指意见所针对的服务部分，而严重程度代表了分析师判断意见主题对用户体验产生的影响。组织意见更正式的方式是对意见进行编码。

## 编码意见

对任何类型的内容进行编码是为相对非结构化的信息创建组织机制的过程。在为美国劳工统计局工作时，卡罗尔·A. 赫特和加里·马基奥尼尼提出了一种编码([ils.unc.edu/~march/blsreport/mainbls.html](http://ils.unc.edu/~march/blsreport/mainbls.html))。本书所描述的过程基于此编码。编码需要两个人参与。

(1) 一个人开始查看消息，并创建类别，直到不会出现明显的新类别。不需要对所有意见进行分类。意见一般可基于“句子层”，分析师可以看看意见中的每个句子。分类可以是任何东西——意见主题、意见性质(问题和特性要求等)、基调(“生气”和“赞扬”等)，或者是任何对分析有意义的东西。单个句子可能会激发多个类别。给每个类别指定名称，名称就是代码。

(2) 使用代码清单，两人分别对相同的子集消息进行归类。子集要足够大(几百条信息常见于客户服务分析)。

(3) 两人比较分类。他们采用了相同方式来解释代码名称吗？有没有过于笼统的代码？有没有过于具体的？有没有代码没意义？两个代码都指同样的东西吗？通过比较子集编码方式并讨论代码，两个人定义出代码的意思及其范围。确定了以后，接下来两人创建代码使用指南。

(4) 建立编码机制，并形成文档，剩下的意见——虽然有很多——可以让更多人参与分离和分类。





### 编码软件

有几款特别设计的软件有助于编码访谈数据。使用这些软件，可以用代码标记出文稿中的特定段落，然后把已编码文档作为数据库，根据特定代码和代码组合提取内容。如果有大量文本，或者需要细化分析，这种方法可能就有用。

这些产品的说明可参见 [www.textanalysis.info/qualitative.htm](http://www.textanalysis.info/qualitative.htm)。

美国疾病预防控制中心提供了一个免费分析软件 EZ-TEXT([www.cdc.gov/hiv/software/ez-text.htm](http://www.cdc.gov/hiv/software/ez-text.htm))。

### 汇总意见

意见经过分类和组织后，接下来进行汇总。

数一数落入每个代码的意见数量，能揭示出有趣或者有摩擦的地方。此外，一旦确定编码机制，一段时间内收集到的意见或者一群人发表的意见都可以应用此编码机制，并且可以比较不同群体如何看待你的服务，或者比较看法如何随时间变化。例如，如果产品变化前后一个月产生的投诉数量翻倍，很可能说明相当多的受众体验发生了定性变化。翻倍并不代表整个用户群体中感到不愉快的人也翻倍了，但它表明有变化，而且是糟糕变化。

### 分析意见

分析支持意见，创建组织机制，接下来便是和在第一线收集用户反馈的人分享。如果客户服务对问题已做了分类，这些基于内容的分类会有助于组织将来的意见，同时深入了解用户的背景资料 and 心智模型会有助于制订答案。完成一次分析后，定期重新访问支持意见更容易关注受众的变化，以及发现受众对网站改进的反应。

但是，人们对所面临的问题的看法可能不会反映出他们体验中的真正核心问题。知道他们的实际行为，并和他们的期望进行比较，能暴露出实际发生问题的地方。要想这么做，可以借助于日志文件。



## 日志文件

只要有人达到网站的一个页面，服务器就会记录下来。在传统商店或者办公室中，管理员只能猜测人们要去哪，要看什么，但网络服务器能确切知道谁在什么时候看了什么东西(当然，单凭这一点还不能说明用户理解或者有兴趣)。网站访问者必须要求询问并看看他们感兴趣的每件东西，就像在珠宝店一样。这和在超市不同，在超市他们可以整天捏捏每个西红柿，没人会知道。这就是为什么珠宝店职员很可能比超市收银员更了解客户行为。

网络服务器收集到的信息多得让人吃惊。对于每个请求(包括每页、每张图片 and 每次搜索)，服务器都会保存请求来源、内容、时间、请求发出的浏览器、浏览器运行的操作系统和其他一些连接情况。这会产生大量数据，对于每天都会产生十亿字节大小日志文件的繁忙网站来说，很正常。

如何挖掘这些数据？认真挖掘。事先的计划可以减少数据挖掘时的混乱。回答几个初始问题，就能发现应该检查、收集和处理其他什么数据。

## 日志文件中有的和没有的

日志文件是程序员认为值得记录的所有东西的原始记录。有点像程序的意识流。网络服务器能够记录各种信息，并和各种软件进行交互，因此可能有很多不同日志文件。有些可能是采购清单，有些可能是授权登录记录，还有一些可能是软件内部错误的冗长记录。能提供最多用户体验信息的日志类型一般是访问日志。此文件包含记录了对服务器请求的每个项目，同时还在互联网服务器能收集的范围内尽可能地记录了用户浏览器信息。

很多日志文件都有以下一些要素。





### 典型访问日志要素

- 发出请求的计算机 IP 地址(大多数服务器都能把数字地址——如 192.0.34.72——转换成人们可读的 DNS 名称,比如 [www.example.com](http://www.example.com))。
- 请求发出的日期和时间。
- 发出的请求类型(获得 HTML 网页,提交表单答复,等等)。
- 被请求的文件名称。被请求的任何东西都会以跟在域名后 URL 的形式出现,经常包括内容管理系统生成的动态网页请求串。
- 服务器状态码。状态码为数字形式,代表服务器是否能传递文件,或者服务器是否被重定向,或者服务器产生了错误。臭名昭著的“404,网页不存在”就属于这类错误。
- 用户代理,即识别客户端使用的浏览器和操作系统。
- 客户端发出的 cookie。cookie 指客户端程序(如浏览器)发给服务器的身份文件。
- 推介链接(referrer)。做出请求之前,客户端浏览的网页或者网站。

例如,以下是访问名为 example.html 网页的典型访问日志。

```
192.0.34.72 -- [05/Apr/2002:04:50:22 -0800] "GET /example.html HTTP/1.1" 200 18793
```

以下是一条推介链接日志记录,说明了上次访问的最后一页(很可能此页链接到当前页)是 Google 搜索结果页,关键词是 adaptive,网页被提取为/adapt.html。

<http://www.google.com/search?hlen&qadaptive->/adapt.html>

Webmonkey 网站上有几篇理解和解释日志文件的好文章,网址如下:

<http://hotwired.lycos.com/webmonkey/e-business/tracking/index.html>

从理论上讲,有了这些信息,就可以知道什么人在什么时候做了什么事情。但日志文件也有问题。

首先,如果浏览器请求的网页带有图片,请求就不止一个。浏览器会分别请求网页上的每个元素。每个图形和 HTML 部分都会被分别请求并记录。因此,看起来是单一网页的网页会产生出 12 个或者更多请求和日



志条目。同样，如果网页包含框架，浏览器通常会分别请求框架文档(定义框架如何显示的文件)和 HTML 文件(以及所有内容)。这样可能就会产生大量日志条目，以及混乱交错的框架文档、导航和内容网页，所有这些反应都是由于一次点击引起。

服务器日志并没有记录在文件完全交付前浏览器是否停止了文件传送。查看标准日志文件不可能知道人们放弃页面是不是因为内容太多，也不知道他们想脱离困境是不是因为服务器太慢。如果没有包嗅探器(监测所有进出服务器信息的设备)，你只能知道被请求的网页，但无法知道网页是否已经到达。

动态生成的网站(通常是数据库驱动或者内容管理系统驱动的网站)也会产生大量令人混淆的日志文件，因为这些日志文件可能只包含一个事实，即某个软件被激活，但没有该软件返回的内容。除非合理配置该软件，访问网站的所有动态内容看起来就像访问了一页一样。

最严重的问题来自网页缓存。缓存是把文件存储在快速存取的中间位置的过程，不必再次从服务器下载文件。这样能大大改善用户浏览体验，点击后退按钮不会导致前一页再次下载，因为前一页是从本地硬盘载入。从用户角度看这样很好，但这样一来就会给分析网站日志带来一定复杂性。如果有人点击了返回按钮回到前一个页面，或者利用浏览器的历史记录机制，浏览器就可能会从缓存中读取页面，而不会联系服务器。从服务器角度看，看起来就像是有个人花了很长时间看一个网页，而事实上这个人是在网站上完全不同的板块。

有两种缓存：个人(Personal)缓存和机构(Institutional)缓存(见图 13.1)。所有常见网络浏览器都内置某种个人缓存机制：把过去数天或者数周内值得浏览的内容保存到本地磁盘缓存。现代浏览器的设置可以检查网页的更新版本，默认情况下这种设置都处于打开状态，因此浏览器通常都进行快速检查服务器，以确保要载入的 HTML 网页自保存到缓存后没有被更新过。这样就解决了最严重的问题，但仍然存在浏览器没有向服务器说明用户正在查看不同网页的情况。



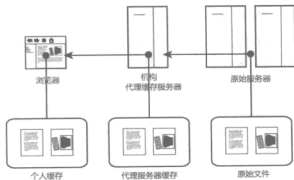


图 13.1 网页缓存机制

机构缓存的问题更多。这些缓存由位于用户机器和网络服务器之间的计算机进行维护，并在到达服务器之前拦截浏览器请求。它们对浏览器请求的网站数据进行本地复制。当另一个请求和已存储的数据相符时，就直接提供，而不是向信息原始来源的服务器传递该请求。

从用户角度看，这样可以节省带宽，并能显著提高网站速度，但它回避了网站服务器跟踪使用情况的方式，并对所有日志分析增加了不确定因素。例如，美国在线(AOL)为所有用户提供了庞大的机构缓存。如果一位美国在线用户请求的网页其他人最近也请求过，该缓存会把网页直接提供给该用户，而绕过“原始”服务器。网站的原始服务器可能永远都不会知道有人正在查看其网站内容，因为人们永远不会联系它。数以百计的美国在线用户可以查看你的网站，但你的网络服务器只记录了一次访问，即把内容读到美国在线缓存的那次原始访问。

并不是要诋毁缓存。缓存在用户体验方面提供了许多优点，并创建了更有效的互联网，但它也确实让统计信息无法完美，因为有些流量总是被缓存所掩盖了。缓存的这种性质导致日志分析更偏向采用定性分析方法，而不是采用定量分析方法。

如果销售代表需要知道有多少美国在线用户点击了广告横幅，准确数字会很重 要，但幸运的是，谈到理解用户体验时，准确并不是绝对必要。把日志文件当成对用户行为的随机样本一般就足够了：你需要知道人们如



何表现、他们的行为如何受到网站变化的影响，而随机样本能告诉你这些情况。知道人们某种行为方式的准确数字对于做出一般用户体验结论很少有用。

## 日志和 cookie

提示：

Apache 和微软的 IIS Server 都内建了基于 cookie 的会话追踪，只需打开出现在日志文件中的 cookie 信息就可开始追踪。请查看服务器软件文档了解如何实施。

Apache 用户追踪模块的相关信息可参见 [http://apache.org/docs/mod/mod\\_usertrack.html](http://apache.org/docs/mod/mod_usertrack.html)。  
描述 IIS 4.0 用户追踪的文章可访问 <http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnw2/mag01/html/EventLogging.asp>。

如果只收集匿名使用日志，就不可能了解到人们行为中特别不易察觉的地方。因为日志只是收集被请求的文件，只可能汇总出所有页面受欢迎程度的庞大统计数字。为了深入了解这些信息，如人们访问网页的顺序，他们在每页花费的时间，访问网站的频率，有必要从日志文件中独立出个人，并跟踪这些人的访问。

最初，通过记录独立 IP 来跟踪个人，但使用此方法，很难跟踪到不同个人。即使能识别出来，也几乎无法跟踪不同会话中的同一个人。因此，人们发明了网站 cookie。cookie 是一种令牌环——一般是一个小文件——通过浏览器和网络服务器进行交换，每次浏览器连接到服务器，就会发出该文件。对于大多数 cookie 而言，cookie 内嵌一段独特识别码，以确定它属于某一特定访问者的浏览器。cookie 身份常用于个性化(例如，亚马逊利用 cookie 来识别你，以便提出建议)和订单跟踪，但也能用于跟踪用户行为。亚马逊给你个性化体验的相同的 cookie 也可以用来知道你多长时间访问一次网站，你都看了什么，购买了什么。

就像 cookie 的名字看起来是吃的东西一样，cookie 有过期日。日期可以从短期(“从现在起 10 秒”)到数年之久。虽然主要作为安全特性，但过期日可以用来最大限度地发挥 cookie 作为了解用户网站使用体验工具的作用。有一个技巧可用于轻松分析用户行为分析，即使用两种不同过期日的不同 cookie。

- 会话 cookie，到期时间从几分钟到几个小时，能识别单独会话。例如，如果一个搜索引擎网站 10 分钟内没有页面被请求，用户很可能已经没在使用。同样，报纸网站可能要给人们 30 分钟的会话 cookie，而在网络游戏可能会给人们几小时的 cookie。





- 身份 cookie 过期时间更长，数月甚至数年。这些 cookie 能识别出独立用户，并能用来跟踪同一个人在多个会话中的行为。

采用这种方法，每位用户在任何特定时间都从网站获得两个 cookie：识别当前会话的短期会话 cookie 和识别计算机的长期身份 cookie。通过会话 cookie，再加上访问日志中的网页清单，就可以提取出会话过程中的网页顺序，从而形成点击流。点击流能说明会话过程中的网页访问顺序，访问了哪些具体网页，每页停留了多长时间。点击流是用户体验的一个模块，许多争论声称互联网具有为用户创造前所未有体验的潜力(也能为理解点击流的公司创造前所未有的利润)，点击流正是这些争论的基石。后面将进一步讨论点击流。

### 日志分析道德

一谈到跟踪，人们就会变得神经紧张。这很自然。没人喜欢被跟踪，更不喜欢被销售人员或者老板跟踪。如果人们觉得自己被跟踪了，就会对你失去信任，对你失去了信任，你就已经失去了他们的生意。

因此，要创建并遵守的严格用户隐私条款，并要告诉用户你正在这样做。以下是一些指南。

- 保持保密。不要以不会直接对人们产生好处的方式把人们的身份和行为联系起来(要从研究中确定好处，而不要只是假设人们对销售宣传感兴趣)。
- 不要出售或者共享可能透露个人身份的任何行为信息。
- 保护行为数据。个人信息对公司来说和知识产权一样属于高度机密，要同样严格地保护个人信息。
- 保持所有结果匿名。不要编写可能把用户身份和行为关联起来的报告。
- 保持分析师的道德准则。公布这些准则。让每个人都同意遵守这些准则，再允许人们访问这些数据。
- 告诉用户收集了他们的哪些信息，并告诉他们可以拒绝提供这些信息。
- 加入用户隐私组织，并遵守这些组织的指南，如 TRUSTe 和电





子疆界基金会(Electronic Frontier Foundation)。

还可以查看互动营销研究组织的道德准则(网址为 [www.imro.org/code.htm](http://www.imro.org/code.htm))和 TRUSTe 网站 [www.truste.org](http://www.truste.org)。

## 一些有用指标

在制定分析日志文件计划前,要确定网站具体有哪些不足,或者确定开发团队是否认为可以改进网站。问题和假设原因清单有助于限制对数据提出的问题(见第 5 章针对收集研究目标和问题的建议)。这样的分析过程侧重于亟待解决的最紧迫问题,并避免伴随日志挖掘所产生的数据分析困境。

以下是一些典型的开放式问题。

- 界面最近所做的改变是否有利于用户的购买?
- 在人们找到要找的东西之前,我们失去了多少人?问题出在哪里?
- 人们在哪里花的时间最多?
- 多少访问者最终变成了购物者?
- 不同群体用户的行为有什么不同?

如果要更系统地制定问题,可以看看要完成的分析种类。在此有四种不同分析:聚合度量、基于会话的统计、基于用户的统计和路径分析。

### 聚合度量

最容易回答的问题是占据最大窗口信息的问题。看看网站流量的最原始数据,合并页面浏览记录就能从“50,000 英尺”高度创建人们使用网站的具体情况。这些记录包括以下内容。

- 一段时间内的被浏览网页总量。“2000 年 12 月与 2000 年 6 月相比,页面浏览量是多少?”注意,不要比较重新设计前后的页面浏览量,因为这就像比较两个不同的网站,没有任何意义。





- 特定时间内页面浏览的分布情况。“人们在上班时间的页面浏览量是否大于非上班时间？”
- 整个网站的页面浏览分布情况。这是对网站某个板块的总浏览量进行分解。可以是单个页面或者根据内容类型的群组。

这些都是原始度量，不能只看数字。例如，如果基于时间来观察页面浏览分布情况，周三和周四的页面浏览量可能一直低于其他工作日。虽然这很有趣，但并没有真正说明浏览量下降的原因。可能是因为用户行为，也可能那两天互联网服务提供商(ISP)很忙，导致有些人无法访问你的网站。

其他聚合统计数字有以下内容。

- 操作系统和浏览器比例。这些数据能反映操作系统和浏览器在网站用户群体中的受欢迎程度。从用户体验角度看，这对于确定网站如何符合浏览器特点是最有用的。
- 客户端域名。对大多数网站而言，点击量要么来自.net，要么来自.com 网站。而这些信息极少能说明情况，因为地球上任何地点的互联网服务提供商在这些域名下都可以有网址。一旦从等式中去掉这些网址，就可以利用余下国家的比例来知道网站国际受众的规模，或者能知道来自单个互联网服务提供商的受众(比如 Earthlink 公司或者美国在线)。
- 推介链接网站。这些网站在你的网站之外，而且是用户在请求特定网页之前直接访问的网站。大多数情况下，这些网站是链接你的网站，知道人们如何来到你的网站有助于了解人们使用网站的情境。
- 新用户/重复用户。寻找绝对数字时会产生有问题的统计数字(因为 cookie 的不可靠性)，但仍然可以使用新用户/重复用户数字来理解长期的用户行为。如果重复用户和新用户的比值一直很低，用户在返回的网站上可能没有找到足够有价值的内容。
- 搜索引擎推介和关键词。推介日志包含访问的最后一页 URL。许多搜索引擎在 URL 中包含生成结果页的关键词，因此可以看



到哪些搜索引擎给网站带来了流量，还能看到导致网站出现在结果页的关键词。这样可以帮助营销部确定搜索引擎广告。

## 基于会话的统计

创建大量分析结果时，最有用的指标是可以使用会话信息的指标。这些指标能发现出更丰富的用户行为，而不是简单的大量统计数字。

一些最有用的指标如下。

- 每次会话的平均访问页面数量。典型会话中的页面数量可以度量人们访问网站的广度。例如，如果大多数会话只包括一两个页面浏览量(搜索引擎往往就属于这种情况)，所有导航和整个信息架构都需要关注这些“短跳”的导航要素。但如果人们平均访问 5 页，导航就需要支持不同使用情况。这项指标的略微变化就能测量出每次会话的不同页面数量：如果每次会话平均有 20 页，但所有访问平均只有 3 页，也就是说如果 20 页都不同的话，人们将会有不同体验。
- 平均会话的持续时间。日志分析中的页面持续时间没有得到充分利用，主要是由于缓存和后退按钮造成的混乱。即使有这些限制，仍然能互相比较基于时间的指标。分离出每页的停留时间也能用于观察哪些网页被当成了内容页，哪些被当成了过渡页。
- 第一页和最后一页。有时也被称为“进入”和“退出”页，能告诉你人们有没有按照你预期的方式访问网站。人们看到的第一页决定了他们与网站关系的很多情况。如果有人以意想不到的方式进入网站，可能就需要考虑如何展现、在哪里展现信息。

## 基于用户的统计

身份 cookie 能进一步告诉你单个用户聚合行为的信息。

一些指标如下所示。

- **访问次数。**用户访问网站的次数是衡量忠诚度和信任的重要指标。





- **访问频率。**人们的访问频率能确定内容需要更新的频率。如果内容每天更新 5 次，而大多数访问者每周只访问一次，可能就没有必要那么频繁更新了。
- **网站总停留时间。**人们在一周、一月或者一年内在网站上花费的时间是说明忠诚度和使用度的另一个指标。如果网站的平均停留时间是每月 5 分钟，而竞争对手的用户平均停留时间是 40 分钟，则说明你的网站要么效率更高，要么趣味性更低。
- **保留率。**首次访问后返回的人数。这也可以通过周、月或者年来衡量，取决于网站目的和典型用法范围。
- **转化率。**最终转换为购物者的访问者比例(或者定期用户，或者其他对于网站成功的用户)。转化率可以计算在某一时间点最终变为购物者的用户占有所有新用户的比例，并将这一数字和没有“转换”的用户进行比较来得到。转化率是衡量网站成功的最直接指标。

除了度量整体用户人口的行为之外，还可以把身份 Cookie 和点击流分析结合在一起，根据用户行为对用户进行细分。例如，与其他板块相比，有个人更经常访问餐叉板块，这个人就可以称为“餐叉迷”。餐叉迷的行为可以和汤匙爱好者及杯子爱好者的行为进行比较。这些比较反过来又可以和开发阶段开始时定义的用户背景资料进行比较。这种方法有利于了解行为和个性化，但这种方法目前仍处于初级阶段。

## 点击流分析

除了一般指标之外，还有其他度量方法能了解用户体验。当然，和所有间接用户研究一样，这些方法都不能告诉你人们“为什么”以某种方式发生的行为(要了解这个，需要通过情境调查、可用性测试和焦点小组直接接触用户)，但这些方法肯定可以缩小用户如何使用的可能性。

这些综合方法中，最有用的方法是点击流分析。点击流分析是指分析人们在网上如何访问，从而发现人们访问的共同方式。点击流分析可以产生如下有趣结果。



- **一般路径。**一般路径是对网站“典型”路径的分析。当然，和所有指标一样，被认为典型的东西要依赖于数据和确定“典型性”算法的质量，但它仍然可以提供有用见解。
- **“下一步”页面。**这些统计数字能说明对特定网页直接继承页的比例。例如，可以从搜索引擎结果页看看“下一步”链接，观察与回到搜索页的比较，有多少人去了具体信息网页。这样可以知道搜索结果能多有效帮助人们找到要找的东西。

针对需要专门关注的产品可以进行更专业的分析。例如，电子商务网站可以定义出若干专业类型的点击流分析。

- **购买路径。**购买路径是人们买东西时的一般路径。通常可以从进入页面到“感谢惠顾”页计算出一般路径。
- **购物车放弃率。**如果并不是在“感谢惠顾”页结束，这些数据的其常见形式能知道在购买路径中购买过程经常终止的地方。可以研究这个脱离点来解决可用性(人们只是改变了想法，还是有什么东西阻碍他们完成购物?)。此外，如果把放弃路径和购物车数据结合在一起，就可以看到有多少人从来没有购买购物车中的东西，从而确定销售损失价值。

另一个好方法是内容分类。网站的网页可以根据主题(“银制餐具”、“碟子”和“新产品”等)或者功能(“导航”、“帮助”和“购物”等)抽象分成块，然后把这些分类用到网页中，得到内容分类结果，而不是单个网页结果。例如，看看厨房用具网站上所有银制餐具的流量，比看汤匙页的统计数字更有用，然后就可以通过比较链接增加前后的银制餐具和碟子流量，调查“每周餐叉”链接能如何有效提升银制餐具的销售额。

## 从数据中提取知识

如果很容易度量，很可能就没用。如果有用，就不容易度量。

——特雷莎·埃尔姆斯，《互联网统计入门》，发表于 builder.com 网站  
有日志和要回答的问题之后，接下来便是根据日志内容回答问题。令





人遗憾的是，回答过程很困难。

日志分析一般都十分耗时，手工无法完成。因此，有两个选择，要么自己编写分析软件，要么购买现成解决方案(当然也可以请别人来分析)。高端分析软件包会花费几万(甚至几十万)美元，但也有能满足很多中小网站需要的高性价比解决方案。

### DIY

不推荐完全从头开始分析。有很多廉价或者免费的产品可以提供基本聚合统计数字(本书写作期间，市场上的路径分析或者行为细分软件很复杂。我不知道有哪个廉价产品可以进行任何复杂程度的路径或者行为分析)。

如果采用低端产品工作，要预先过滤数据来减少结果中的干扰。可以去掉四类点击量来改善最终统计结果(阅读软件包文档，因为有些软件包能进行过滤)。

- **来自搜索引擎的点击量。**搜索引擎定期发出蜘蛛来收集网站内容。这些蜘蛛和用户的行为不一样。它们根据链接在网页上出现的顺序访问链接，而不管这些链接是否有用或者有趣。因此，蜘蛛填满日志的数据和人们如何使用网站无关。幸运的是，大多数蜘蛛都使用独特的用户代理字段，这样就可以迅速从日志文件中去除蜘蛛。可以访问 [info.webcrawler.com/mak/projects/robots/robots.html](http://info.webcrawler.com/mak/projects/robots/robots.html)。
- **离线浏览代理。**有些程序能把网站所有文件都下载下来，以便离线浏览。这些程序像蜘蛛一样，也存在相同问题。可以访问 [www.jafsoft.com/misc/opinion/webbots.html#offline\\_browsers\\_and\\_other\\_agents](http://www.jafsoft.com/misc/opinion/webbots.html#offline_browsers_and_other_agents) 获得清单。
- **极端值。**极端值指出现在大部分所收集到的值之外的值。极端值会破坏复合统计数字，比如破坏了平均时间和路径长度的中间值(第 11 章将其描述为“亿万富翁”问题)。要测试统计数字是否受到极端值的影响，可以把数据的中间值和中位值进行比较。如



果两个数字非常不同,那就说明有大量极端值,要去掉这些极端值,然后重新计算结果(再次声明,第 11 章详细叙述了这种方法的思路)。

- **你的网站。**网站的很多流量来自你的公司,因此这些数据很容易破坏推介统计数据。

想知道免费或者廉价的日志分析产品清单,可以访问 [www.uu.se/Software/Analyzers/Access-analyzers.html](http://www.uu.se/Software/Analyzers/Access-analyzers.html) 和雅虎的 Computers and Internet > Software > Internet > World Wide Web > Servers > Log Analysis Tools 目录。大多数这类产品都能产生聚合统计数据。这些产品一般都不能处理会话,或者进行用户追踪,但其中有几个产品能分析点击流。

检查了基本统计数据,接下来可以观察个人点击流来获得更多使用信息。即使没有自动路径分析,也可能从直觉知道人们如何使用网站的情况。看看随机点击流中较大的数字(50~100),并记下来,有可能会知道人们经常如何访问这个或者那个链接,路径有多深,以及进入页和退出页。

但手工日志分析不仅耗时而且不准确。商业工具能获得更深层次、更准确、更一致的知识。

### 大型工具

柯克对计算机:“关联所有已知数据。”

计算机:“工作中……”

就本质而言,工业级的分析产品和低端工具的工作方式相类似,但工业级的分析产品把分析过程完全变成另一种领域。通过关联多个数据库信息,这些工具能够比不太高级的普通工具产生更复杂的用户画面及其“消费者生命周期”。Aberdeen Group 把这种产品描述为能够“处理日志、客户购买历史、人口统计信息和其他数据,以了解客户访问和倾向的情况。可以把客户行为和商业类型建立映射,分析重叠的地方,最终可以把客户细分成小组,并解释创建细分的规则”。实际上,这样就让公司能为符合细分的人创造高度个性化的体验(和营销计划)。

一般通过搜索互联网“数据挖掘”或者 CRM 信息都能找到进行这种





数据整合和分析工作的产品。我写本书的时候,有些公司能开发这类产品,包括 TeaLeaf, NetGenesis, Accrue, Web-Trends, Coremetrics, Limelight 和 Personify。这些公司的产品有很多重叠之处,但每个都有一些独特的地方,有些还针对体验个性化,体验、广告和战略三者中,它们往往更注重体验。

大多数软件包的主要目的是为了协助战略营销和商业决策,但开发人员和设计人员也可以利用其中的工具来理解用户行为,并判断用户体验变化的效果。例如,如果知道看餐叉的人中有很大比例的人也会看餐巾圈,这样会有助于调整网站信息架构。把这种分析扩大到所有类别之间的关系,可以比定性研究方法更能精细地调整整个网站。

此外,如果分析工具能把底线财务信息和用户体验数据整合在一起,可以有助于缓解可用性和业务需求之间的紧张关系。例如,有一家搜索引擎在搜索结果页移动了横幅广告的位置,横幅广告的位置会更显眼。这是故意吸引用户注意,并产生预计的更高点击率和更高收入。但应该知道,额外收入是否属于短期效果,是否有破坏性,最终导致用户讨厌新广告位置而一去不复返。分析广告位置变化前后的用户保留率,发现没有明显变化——与位置改变之前相比,人们离开的数量并没有增加——而收入显著上升,该搜索引擎决定保留位置变化。后来,还是该搜索引擎做出改变,让广告更加突出,但保留率开始下降,而收入没有显著增加,于是又回到之前的改变。

这些大型工具很昂贵(通常情况下,要花几万或者几十万美元才能购买到一个能和你现有数据库集成并能分析结果的工具),因此要精挑细选,而且脑中要有具体研究目标。和来自提供这些服务的公司的顾问评估这些工具时,要问问软件包可以回答研究中的哪些问题,能回答到何种程度。

不管分析的先进水平如何,都不要忽略人们的现有体验。客户支持和日志分析作为强有力的工具,可用于产品生命周期中的几乎任何时间,以发现有价值、可立即采取行动的信息。它们应该用在完整研究计划的情境中,但不能忘记它们。





### 客户关系管理提示

本章只提到客户关系管理(Customer Relationship Management, CRM), 客户关系管理是跟踪客户与公司所有互动的快速增长领域。客户关系管理先从客户意见和日志文件入手, 但会包括用户/客户行为深层定量分析, 并试图把用户行为和公司产品及收入流联系在一起。例如, 玛特尔公司的挠痒娃娃(Tickle Me Elmo)上市时非常受欢迎, 也很贵。它玩起来还不错, 但如果坏了, 几乎没有修复支持(至少和我交谈的家长是这么说的)。许多买家发现玩具无法修好时会非常难过。玛特尔公司显然没有满足特定群体的期望, 但有多少人? 这种情况会如何影响这些人对玛特尔公司和芝麻街的看法和行为? 短期内如何影响玛特尔公司的底线? 长期影响呢? 客户关系管理的关键就是理解客户和组织每个方面之间的关系。

客户关系管理的许多想法和定量方法与用户体验研究有重叠, 将来这两个领域一定会有很多跨专业人士。



## 第 14 章 竞争性研究

产品不会存在于真空中。任何特定的服务或消费市场，都存在产品互相竞争的完整生态环境。每个产品都采用不同方法来满足用户需求。但哪些差异是关键？功能、展现或者形象存在的微小差异有时候会严重影响人们对产品的看法，而有时这些差异又根本不重要。

了解竞争对手的哪些战略有效、哪些无效，对于了解如何开发产品和在哪里投入开发精力至关重要。

传统竞争性分析是由业务开发部门或者独立审计员来完成，其目标是检查公司提供的产品或者服务，并从财务或者营销角度进行分析。通过审查公开文件和实地研究，他们能够了解一些情况，比如竞争对手的产品成本是多少，谁购买他们的产品，顾客住在哪里，广告中要强调哪些方面。这些信息对于帮助管理人员针对公司及产品方向做出战略决策非常宝贵。不过，这些分析无法揭示人们实际使用产品的哪些部分、喜欢产品什么地方或者是什么促使人们成为回头客。这就是竞争性用户体验研究的领域。

本书讨论的大多数方法既能应用于竞争对手的工作，也能应用于你的工作。这些方法能揭示竞争对手的优势和劣势，并有助于你看到机会和威胁。从某种意义上说，竞争性研究是所有研究中最纯粹的用户研究，因为它忽略了其他产品建立的假设和限制，纯粹地关注用户看法。这些都让竞争性分析成为所有研究中最有用的研究之一。

### 竞争性分析何时有效

[竞争性分析]不可或缺。它能为你确定你的世界，并让你的生活轻松很多。

——约翰·希普尔，雅虎用户体验架构师



竞争性研究可以在产品周期的任何时候进行。无论是在开始规划产品，还是把最终产品和竞争对手的产品进行比较，竞争性研究都能提供有用信息。为产品做定期的和全面的研究要比挑选研究最佳时刻重要。为了获得最大收益，要在产品生命周期中不断持续检查竞争对手的情况。不管什么时候，如果没有测试自己的工作，就应该研究竞争对手的工作。即使有产品完整测试计划，偶尔把自己的产品放到一边，调查研究对手的情况也有意义。

也就是说，竞争性分析有一些关键点能真正影响开发过程。

- **收集需求的时候。**研究竞争对手的客户觉得哪里有用、他们认为哪里有吸引力、产品哪些地方让他们觉得失望，能对你自己产品的特性选择和优先级别起到指导作用。
- **在再设计之前。**随着产品发展和进化，竞争性研究能回答设计方向问题，因为可以把竞争对手的产品当成你自己一些想法的功能原型。
- **竞争对手做出重大改动的时候。**研究人们看法和行为的变化，能深入了解提供的效果，进一步了解如何应对这些变化。

## 竞争性研究方法

竞争性用户体验研究主要关注人们如何使用和感知产品，并不关注产品的受欢迎程度、经营模式或者收入来源。传统竞争性分析或者品牌分析试图通过观察收入及其他“基本要素”来了解公司成功的根源。除了偶尔有特性表格或者品牌认知研究外，这种分析很少关注消费者或者用户的观点。当然，用户体验研究工作可以和传统竞争性分析一起使用，但其关注点是自下而上，而不是自上而下。

除了这个根本区别外，竞争性用户体验研究的过程和传统竞争性分析非常相似。在两种情况下，分析师都必须确定谁是竞争对手，要比较哪些方面，如何进行比较。其一般步骤和金融分析师了解竞争格局时遵循的步骤基本相同。





- 确定竞争对手。
- 列出导致竞争对手具有竞争力的关键属性。
- 竞争对手之间的互相比拟(和你的产品进行比较)。
- 利用比较结果创建“可采取行动的情报”。

### 确定竞争对手

道理似乎显而易见,但在开始竞争性分析之前,需要知道竞争对手是谁。不幸的是,这比起初看起来更难。很明显,有公司或者产品在你所处的行业进行竞争,但这个行业的确切定义常常很模糊。

可以从查看在线目录入手寻找潜在竞争对手。如果你的产品列在雅虎目录或者开放目录(Open Directory)中,可以看看同一目录中的其他网站。他们很可能就是竞争对手。转到上一级别目录。选择一个可能的竞争对手,并在目录中进行搜索。该竞争对手还出现在哪些类别中?这些目录的成员可能也是竞争对手。看看你的产品所属目录的同一级目录。这些目录可能也包含了竞争对手的产品。

此外,第8章所讨论的情境调查,明显竞争可能只占竞争格局的一部分。在其他媒体还可能还有其他竞争对手。自行车旅游网站的竞争对手(至少对于某人的休闲时间而言)不仅仅是其他骑自行车信息网站,还有关于骑自行车本身的网站,因为网站用户的最终目标不仅是寻找有关自行车的信息,还要有骑车出去旅行的内容。健康咨询产品会和《医师药师药品使用手册》、家庭医生、甚至奶奶产生竞争。同样,虽然巴恩斯·诺布尔书店明显是亚马逊竞争的对手,但街角书店和亚马逊也是竞争对手。街角书店有什么东西而亚马逊书店没有?和亚马逊书店相比,书店的体验在哪里,差在哪里?要考虑人们不使用你的产品时执行的其他任务。除非是付钱请他们使用,如果产品不是更有用、更有趣,信息也不比竞争对手的任务更多,否则他们不会使用你的产品。

把竞争对手列成清单。从其他目录列出最有可能的竞争对手,进行头脑风暴,从其他媒体和其他场景中找出竞争对手。如果需要,可以写到索引卡片或者即时贴上,这样组织起来会更容易。



接下来进行排序。当然可以采用任何你喜欢的方式，但以下的典型组织方式，我觉得很有用。

- 第 1 级竞争对手是和你的产品最直接竞争的产品。巴恩斯·诺布尔书店和亚马逊的图书销售业务直接产生竞争，雅虎拍卖和 eBay 直接产生竞争。这些产品尝试用相同方式捕捉相同受众，提供相同(或者非常相似)的服务，而这些都是分析的主要目标。清单中的数量不要超过 5 个。可能实际不止 5 个(例如很多地方都卖图书)，但 5 个足够让你很好了解竞争格局。竞争对手无处不在，即使你认为你的产品是全新，而且创新，仍然免不了和受众的时间进行竞争，而受众的时间就是竞争对手的产品。汽车和马匹从外观和工作上都不同，但在 20 世纪早期，它们是彼此的第 1 级竞争对手。
- 第 2 级竞争对手的产品和第 1 级竞争对手来自同一类别，但它们如果没第 1 级产品那么直接竞争，或者很相似，研究第 1 级产品就足够了。应该看看这些产品，以检查对第 1 级产品的理解，但了解第 2 级产品不是关键。清单中的数量不要超过 5 个或者 10 个。例如，Half.com 和 Alibris.com 都在线销售二手图书，它们可能互相是竞争对手，但它们都没有真正和巴恩斯·诺布尔直接竞争。
- 专门(Niche)竞争对手会与你的产品一部分直接竞争，但不和整个产品竞争。19 世纪 80 年代，马和火车在长途旅行领域展开竞争，但在本地运输领域没有竞争。截止到 2002 年夏天，巴恩斯·诺布尔没有与亚马逊的玩具和游戏业务进行竞争。但它卖书，因此它的图书与音乐业务和亚马逊竞争，而在玩具业务方面，它需要担心的是儿童图书业务。

## 勾画竞争对手

如果要关注竞争性研究，就需要了解研究对象。完成竞争性分析首先从所有竞争对手的完整背景资料开始。内容包括两个要素：产品描述与受





众背景资料。

### 产品描述

产品描述不能只是特性清单。产品描述应该是从用户角度陈述产品给用户带来的价值。从用户角度看，亚马逊的主要价值并不在于大幅折扣、众多选择、快速访问和高度相关的建议、这些特性很少能总结成单个特性或者准确数量。

产品描述不宜超过一段。在鸡尾酒会上，如果要向从未听说某个产品的人描述该产品，用户会怎么说？例如：

ZDNET 属于每日科技杂志。它提供硬件和软件评论，高科技业界的最新闻，深度分析，还有行业专家评论员专栏。它的文章提供了对技术世界变化的深入了解，写作风格直接、生动。它还提供技术世界流行的电子邮件列表。

### 受众背景资料

第7章列出了用户背景资料的属性。使用此清单，看看竞争对手的产品，并尝试着多给属性加上值，如果你觉得产品会服务多个受众，可以创建单独的背景资料。就像和你自己的产品一样，可能需要创建三份或者更多份不同背景资料，以代表不同类型的客户，突出这些人和受众之间的差异。如果他们直接竞争对手，你的产品任务非常相似，但他们至少都有一项关键区别：他们不使用你的产品。为什么不用？是什么让竞争对手的用户市场和你的用户市场不相同？

### 识别特性和属性

给竞争对手创建背景资料，最重要的一步是要创建可以进行比较的清单。清单可以是具体特性，也可以是抽象描述。特性识别要来自于用户角度。你的产品可能是在带有空调设备的数据中心中最先进的硬件环境下运行，而竞争对手的产品是在10年的Amigas公司的设备柜中运行，但用户并不关心。他们关心的是能多快完成工作，是否可以找到要找的东西，等等。



收集属性清单有两种方法：询问用户和自行寻找，两种方法组合进行。

对具体特性，用户的看法可能和你的看法不同，因此，你可能需要进行轻量级的情境调查或者小型焦点小组系列入手。从用户的角度识别服务中重要的特质及重要的功能。可能会像“我喜欢特性的集成方式，这样我的数字能从一个板块到另一个板块”一样复杂，也会像“搜索按钮很容易找到”一样简单。

此外，要审查特性，坐下来，列出产品所有的突出特性清单。从用户角度看，你的产品能做什么？有哪些突出特性？然后看看每个竞争对手的产品，并检查它们有哪些相同或者相似特性（再次强调，相似是从用户角度来衡量，而不是从你知道网站如何工作来衡量）。发现竞争对手提供的新特性后，把它们添加到清单中。也可以查看产品评论，看看评论者突出了哪些特性，他们是怎么说的。

搜索引擎的清单会如下所示。

- 反应快速。
- 能提供目录和具体网站匹配。
- 纠正简单的拼写错误。
- 根据相关性排列结果。
- 兼容 AltaVista 搜索语言（链接等）。
- 其他。

接下来，把两张清单组合在一起，按照优先级别把特性列入第 1 级和第 2 级清单中。关注研究时优先级别非常重要，要根据对产品功能、公司成功、或者用户满意最重要的程度来排列优先级别。可能要和管理层讨论竞争性研究计划，因为管理层对研究过程能回答的问题有自己的想法。

收集特性的时候，也要同时收集竞争对手产品的重要属性。属性不一定和功能有关，但属性是让产品难忘、可爱或者有趣的东西。属性是人们用来向朋友或者同事描述产品的对象。对于产品可用性而言，属性虽然没有实际特性那么重要，但属性能描述用户体验的感觉和环境。产品引起的共鸣能改变人们体验和使用产品的方式，不能忽视产品的共鸣。例如，人们认为某个产品“枯燥”“无味”，此产品的用户当然会受到人们观点的





影响,即使这些观点和产品工作的好坏没有直接关系,形式可以遵从功能,但不必受到功能的限制。

### 竞争性分析方法

能深入了解自己产品的方法——可用性测试,焦点小组,调查等工作——也适用于了解竞争对手的工作。两者的主要不同在于对创建平衡观点的侧重点不同。研究自己的产品时,通常会把重点放在产品存在的问题上——毕竟“哪些问题”比“哪些变化不需要”更容易回答——但在竞争性研究中,了解竞争对手做了什么正确的事情和了解他们在哪些地方举棋不定一样重要,除了知道对手产品的不足,还要知道这些产品的哪些部分用户是认可的?用户喜欢哪些方面?

#### 招募

就像为了了解人们对产品的看法而招募真实用户(或者潜在用户)一样,进行竞争性研究时也需要招募真实用户。

招募过程和第6章所描述的一样,但需要考虑更多目标受众对竞争对手产品的体验情况。这种体验是否重要取决于要进行的研究类型。如果对高级特性进行可用性测试,很可能就必须找到体验过更多常见特性的人,因此,没有必要让参加者熟悉基本功能。同样,当前用户的焦点小组可以告诉你什么东西吸引用户使用产品的情况,但研究新手用户——没有使用过你的产品或者没有使用过竞争对手的产品的人——常常能提供一些最宝贵的竞争性情况,因为对产品的初次体验决定着大量用户与产品的未来关系以及对产品的看法。

招募符合竞争性背景资料的人的过程和第6章叙述的方法没有差别,但如何才能找到竞争对手的用户?

有种方法是向大众发出邀请,并在产品筛选文件中包括你的产品用户和竞争对手的产品用户,但这种方法无法发现你(或者招募师)能代表其中的一个。清单中包括你的产品、竞争对手的产品和其他若干产品。可以发邀请给使用竞争对手产品但不用你的产品的人。

#### 忠告:

为竞争性研究招募用户时,要保持公司身份匿名。如果人们知道研究是比较的目的,几乎所有人都会产生不同行为。有些人会为他们选择进行辩护,还有些人会利用测试机会发泄不满。这两种情况都会夸大他们的实际反应,应该避免。





### 竞争性情境调查

针对竞争性产品进行一对一用户访谈能知道让产品的优势和不足。观察人们使用竞争对手的产品能揭示出你的产品会激发或者要避免的使用模式。

例如，观察人们在线阅读新闻，用于科技新闻竞争性分析，研究员观察到定期阅读新闻的人首先会去 CNN 或者 CNBC 网站阅读几条新闻，然后会跳转到较小的新闻网站。被问及原因时，人们都说他们通过主要新闻来源阅读一般头条新闻，而通过专门新闻网站来找到自己感兴趣主题的全面报道。专门新闻网站为他们最喜欢的主题提供了更好的观点，而一般新闻网站能更好报道“飓风和重大事件”新闻。该科技新闻网站原本计划从新闻通讯社获得一般新闻内容，但看到了这个情况后，他们确定新闻通讯社无法与 CNN 就“飓风和重大事件”头版展开竞争，因此他们决定突出突发性科技新闻。

### 焦点小组

邀请竞争对手产品的用户参加焦点小组可以揭示他们为什么使用竞争对手的产品、哪些方面吸引他们、哪些地方让他们感到失望。这些群体还可以揭示人们对竞争对手产品的品牌、形象以及他们赋予这些品牌的特质。可以分析焦点小组期间收集到的逸闻趣事，看看与使用产品 A(你的产品)相比，使用产品 B(竞争对手的产品)的最初动力是什么。

例如，有一家家具销售网站的几个焦点小组受邀访谈他们在线上线下如何购买家具。目标是要了解他们发现哪些信息最有价值、哪些展现形式最令人信服。结果很快发现、而且很一致，对某些特定类别的家具，大多数参加者只会购买值得信赖的品牌。参加者对大多数制造商的名字都不熟悉，又无法亲手摸到家具，因此他们需要信誉好的卖家。他们抱怨导航和网站上的图片质量，但他们对品牌的信任超过了担心。

竞争性焦点小组中可以讨论的产品不止一个。介绍几种产品，并积极进行比较，有助于进一步揭示人们看重的地方和觉得有趣的东西。这些群体可以由使用过其中一个产品的人构成，也可以由没有使用其中任何产品





### 提示:

根据记录,这两家家具网站——焦点小组讨论的那一家和提出研究的那一家——都失败了。焦点小组还发现,人们不会在线购买大量家具,而这两家网站显然都认为这种情况会发生变化,事实并没有改变。

的人构成(但不要将两个群体混在一起)。如果有多个产品,要变换讨论的顺序,以尽量减少由于每次首先看到的都是相同产品或者因为相同两个产品每次都以相同顺序出现所造成的偏见。

### 可用性测试

竞争性可用性测试是更有用的竞争性研究方法之一。观察人们使用竞争对手的产品能告诉你竞争对手的功能和你不一样的地方(检查界面的时候这些不同不会总是很明显),它成功和失败的地方。

你所关注的特性会驱动脚本和任务的开发。要关注竞争对手的产品和你的产品最与众不同的特性,并且要把最佳描述所有类别功能的特性分离出来。例如,调查竞争对手的导航选项卡(Tab)的发挥程度可能比调查人们是否理解购物车中的图标更有价值。尽管图标可能是华丽的区分,但选项卡可能会对人们对网站的体验产生更多影响。

创建的任务要和测试自己的产品的任务没有差别。实际上,可以采用相同任务。除了减少工作量外,还可以比较不同界面组织会如何创建不同使用模式。

偶尔会发生无法获得访问竞争对手产品的问题。在公共网站上,很容易打开浏览器,支付一些注册费用就能获得一两个账户。但如果无法获得产品,或者注册费超过预算,测试竞争对手的产品就具有挑战性了。此时,可以用创建的原型来回答很多问题,就软件而言,还可以通过演示版本回答问题(演示版本经常可以免费下载,但要仔细阅读演示版本的用户协议,以确保竞争性分析对产品属于合理使用)。

还可以同时测试多种产品。可以通过针对对不同界面给评估人相同任务来比较的方法(但只有在所有界面都可以执行相同任务的情况下)。再次强调,可以通过改变展现竞争对手产品的顺序以及不要招募没有使用过其中任何产品的人的方法来减少偏见。

例如,有一家通用目的互联网目录想了解其在竞争对手产品界面下的工作情况。许多基于分类的目录网站都基于雅虎,这家网站也研究过雅虎,因此他们决定选择两个没有太多人知道的目录网站。他们选中了



忠告:

偏见很容易溜进测试过程。对喜欢的解决方案下意识地点头,或者主持人顺口说出的话都可能影响参加者的看法。任何研究都可能发生这种情况,但如果主持人有强烈的竞争性忠诚,就更可能发生这种情况。关注行为和当前使用情况,而不是关注观点,能降低有观点倾向的主持人扭曲结果的机会。因此,偏见也是进行第三方分析时第三方顾问特别有用的原因。

LookSmart 和 snap.com, 因为两者(在当时)都对类似受众提供类似服务。他们请两家网站都没有使用过的人通过这两个网站来搜索类似东西。定位功能(或者通过指定邮政编码, 或者通过指定地区)被认为是第1级重要, 也是可用性测试的主要重点。大多数参加者都说这些特性很有意思, 而且他们都理解两个网站采用不同方式来传达这一功能(Snap 网站上是“Enter Zip Code”字段, 而在 LookSmart 网站上是“My Town”导航选项卡)。但使用两个网站特性实施方法的能力并不相同。LookSmart 采用的“My Town”导航选项卡试图预见人们的本地化需要, 但几乎没有预测准确, 而 Snap 采用的选项引导人们达到不同目录, 人们觉得这是项单独服务, 并能像理解 Snap 网站一般界面一样很好理解。该互联网目录网站决定采用类似 Snap 的模式。

### 调查

进行特殊竞争性调查之前, 要看看已有的客户和用户研究。通常可以从其他调查中提取有用信息, 可以对答复进行过滤, 抽出竞争对手的产品用户, 并且汇总这些用户的答复, 把这些用户当成整个调查人口的子集。结果很可能与一般用户人口并无明显相关, 但分析他们可以知道你的用户和竞争对手的用户有什么区别。

为了解竞争对手的情况而特别编写调查的过程遵循常规调查的同样步骤。要建立调查能完成的合理目标, 关注第1级的大多数问题, 并以中立方式展现特性。从答复者的角度看, 要采用独立的第三方形式来写出问题。

调查中的竞争性调查主题包括一些流行主题。

- **用户使用的其他产品。**通过这个问题, 你可以估计竞争对手的受欢迎程度, 通过已公开的第三方受众数字, 可以知道市场总规模的大概数字。
- **他们使用竞争对手产品的程度。**如果使用动态调查生成软件, 可以通过其功能创建个性化的跟进问题, 针对竞争对手的产品提出深入评级问题。人们多久使用一次竞争对手的产品? 使用了多长时间? 什么时候开始使用的? 他们有没有在竞争对手的产品





之间切换过？多久切换一次？什么要切换？

- **他们对竞争对手产品的忠诚度。**竞争对手产品对用户有哪些吸引力？他们的满意度如何？可以创建李克特量表，受访者可以针对产品的具体特性和参数迅速填写意见(详见第 11 章，深入了解李克特量表)。
- **他们使用的特性。**这些问题可以来自已知竞争对手产品的特性数据库，也可以是自由文字，让人们自由讨论产品对他们有什么用。还可以包括特性满意度组成部分。

想调查竞争对手的用户很困难。甚至都不一定知道他们是谁。和对来到家门口的用户进行调查不一样，你必须积极找到竞争对手的用户。可以对自己的用户进行调查，过滤出那些使用竞争对手产品的用户，也可以对广泛用户进行调查(例如，在雅虎或者某个你觉得会吸引大量竞争对手用户的网站投放横幅广告)，并希望有足够的受访者是竞争对手产品的用户来抵消招募用户花费。当然，获得有代表性调查的最好方式是进行随机电话调查，但大多数情况下这些方法花费很高。

像其他类型的竞争性研究一样，如果人们发现调查和竞争性有关，就很容易产生偏见。如果发现他们是为了另一个产品而在讨论一个产品，可以改变人们回答问题的方式，特别是满意度和偏好问题。在人们提交答复后，在调查结束时(或者在任何程序结束时)向他们表明自己的身份也没问题，但一般情况下以第三方来完成整个事情会更好。

## 分析竞争性研究

传统竞争性分析的目标是创建并列的特性比较网格，建立点对匹配功能的矩阵，这种假设对于竞争性分析非常重要。但从用户角度看，并非总是如此。两个产品被视为互相竞争通常需要有一组相似核心功能，但除此之外，实际特性集合可以非常不同，两个产品仍然会竞争。例如，许多网上企业直接和传统企业产生竞争，尽管实际用户体验完全不同。

竞争性用户体验研究要能发现竞争的根本优势和弱势，而不是创建记



分牌。

比较可以从数据收集阶段形成的研究报告入手。拿着报告,列举出竞争对手产品为其用户提供的优点以及给用户造成的阻碍。同样,这些情况要从用户角度出发。让人们停留在你的网站时间长一点对你而言是最大利益,但人们可能更喜欢快速访问并离开。同样,每日更新内容可能会被定位为产品优势,但可能对用户而言差别不大,而且事实上,如果要求人们重新学习导航的话,频繁变化的网页可能会被当成障碍。

进一步仔细观察结果,这些结果说明了公司及其发展过程的哪些方面?在哪些地方提出了一致的强烈体验?这些体验有哪些特点?定期如何变化?商业分析师把公司优势集合称为核心竞争力文件夹。同样,维护竞争性分析所创建体验过程中所列出的一致弱点文件夹也有用。

例如,CNET 是品牌形象和设计聚合方面的大师。去掉所有特色徽标之后,看过 CNET 网站的人还是能从一堆网站中立即认出 CNET 网站。用户认为这是个优点,因为他们信赖这家公司,使用其产品很顺畅。但在一致性方面 CNET 的导航不如其品牌。对其进行竞争性研究后(我为 Wired Digital 工作时进行了此研究)发现,用户觉得他们知道如何使用 CNET 产品,因为产品与 CNET 其他产品相似,而实际上他们浏览 CNET 某些网站比浏览其竞争对手的网站麻烦得多。研究很多 CNET 网站后发现,CNET 很多产品中都存在品牌强烈和糟糕导航的现象。

当然,传统竞争性分析也能发现问题。虽然它不是唯一的分析方法,但可以提供一個有用捷径。例如,如果用户明确看重一些特性,就可以针对这些特性建立比较矩阵。但这种并排比较需要在实际可比较项目之间。因此,位于网站某个地方的网站地图(Site Map)就可能无法和首页中心位置的地图(Map)进行比较。

#### 基 准

从建立用于评估竞争情况的捷径来考虑,有些竞争性分析方法会引入数字或者字母竞争性评级。例如,可以给特性矩阵中的每项特性打分,比如 0~5 分,其中 5 代表产品做某些事情不错,0 代表产品无法完成某些





事情。把所有分数加起来，就能知道网站的竞争性。这样就可以从理论上知道哪些东西最有竞争力，因为相对其他东西而言，这些东西最擅长做的东西。

如果不断有新的竞争对手进入市场，而你需要快速评估这些竞争对手，这种方法就会有用，但它掩盖了哪些方面影响着最终得分的重要信息，而这实际上可能比整体价值更重要。但这种方法能搞清楚哪些强项是严重威胁，哪些弱点是重要机会。

## 根据竞争性研究采取行动

最终，收集到所有数据，分析答复，并进行所有比较之后，接下来就需要采取行动。为了让研究可用——创建可采取行动的情报——需要重新关注工作。拿出优势和障碍清单，并对应到产品。对于清单上的每一条，解释和产品的关系。竞争对手产品提供的每个优势都可以看成对你的产品的危险，而每个问题都可以当成机会。竞争对手产品比你的产品每个有效的地方都暴露出你的产品弱点，反过来也一样。

接下来创建一套理论，解释竞争对手做出设计选择背后的动机，这些设计选择有好有坏。为什么产品采取了这种方式？有什么限制？理由是什么？你认同哪些选择？不认同哪些选择？什么地方令人费解？关注竞争对手决策过程中的原因，而不一定关注做了什么。

竞争性研究的最终结果对构建良好用户体验能有更深入了解尊重竞争对手解决问题的能力，对自己的产品感到自豪，建立利用知识发挥产品优点的计划，这些都是值得吸取的经验，也还有其他一些方式能学习到这些经验。

## 实例：ZDNet 网站快速评估

以下内容摘自对 ZDNet 一家网上新闻网站所进行的快速竞争性可用



忠告：

不要在广告中使用研究结果。在广告中使用内部竞争性研究结果很有诱惑力，但通常不推荐这么做。人们不会很相信广告，如果竞争对手觉得你的研究结果不正确，可能还会起诉你。例如，如果内部研究表明五分之四的人喜欢你的产品胜过竞争对手的产品，你确信这个数字能准确反映全部人口的比例，那就雇一家独立研究机构来复制研究。独立机构很可能会采用更严格的标准进行研究，其研究结果对消费者(和法官!)而言份量更重。

性测试。为节省篇幅，略去了研究过程描述以及这些结果融入更大核心竞争力清单的方式。

- 人们一般认为，ZDNet 网站的主要内容是软件和硬件评测，而不是文化信息。被问及网站是关于什么的时候，评估者一般都关注网站的技术评测方面。他们并不认为网站会提供技术如何影响世界的内容，而看起来觉得网站内容是关于最新的主流硬件和软件方面的信息。
- 因为订阅了电子邮件列表，许多人从未去过首页(如图 14.1)。至少有一位评估者说他最近去过首页，但给他看 ZDNet 首页时——已重新设计好几个月了——他对设计感到很惊讶，因此，从首页直接来的流量与来自其他网站的链接或者与来自邮件列表的流量相比，深层网页的导航和信息与首页上的导航和信息一样重要，或者更重要。

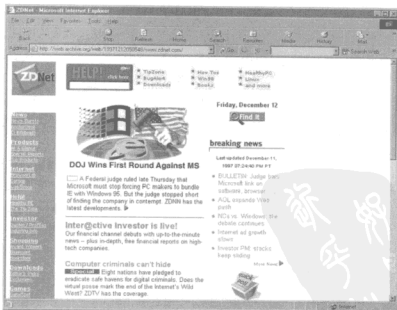


图 14.1 ZDNet 界面(约 1998 年)





- 板块命名混乱，而且缺乏邀请探索之义。“CI”指计算机情报(Computer Intelligence)，属于 ZDNet 的财产，但没有一个参加者知道或者认识这个缩写。Infobeads 和 Garage 被挑出来，人们觉得特别困惑。Help! 可能是最夸张的例子(甚至有点欺骗性)，因为人们觉得是网站需要帮助，而实际是对使用其他软件的提示。事实上，即使浏览 Help! 板块一段时间以后，人们期望看到的内容还是让他们感到迷惑，因为他们仍想要把实际内容和期望内容联系起来。
- 评价者没有看到导航栏，因为导航栏放在广告和头条新闻之间。有一位评估者一开始觉得是广告。由于位置太近和形状相似造成了广告和内容之间产生混淆，这种情况以前就发生过，也可能就是评估者没看到 Excite 搜索结果页上方的建议词的原因之一。
- 到具体新闻的导航让人迷惑。有位评估者说：“[这个网站就有]我想要的东西，如果我能再找到的话。”
- 股票报价导航几乎无法看出来，这也是设计出来的界面总是被误解的绝佳例子。想找到 Netscape(NSCP)的股票价格时，每个人(包括我自己)都对股票报价界面犯了同样错误。
  - a. 在提供的输入框进行输入时，每个人都输入了实际股票代码。但因为界面无法理解股票行情代码，搜索结果只出现了一个空白页面。
  - b. 接着，每个人都看见了输入框上方的一行数字和字母，并点击了 N。出现了以字母 N 开头的公司列表，一次出现 20 家。
  - c. 然后，每个人都点击了 Next 按钮，出现了数据库的 NA 部分。
  - d. 意识到数据很大，要找到需要的信息会点击很多次，每个人都在输入框输入 Netscape 来生成想要的信息。

所有这些事情(大约 30 秒的工作)都可以通过在输入框下方增加一行“请输入公司名称，不要输入股票代码”就可以解决。
- 搜索是网站导航的主要方法。当请评估者定位特定主题的一篇文章





章时——虽然不是特定文章——人们都普遍使用了网站搜索引擎。幸运的是，Find 弹出按钮令人难忘，人们能很快使用。但有位评估者觉得这个按钮和页面左侧列出的导航列表中的搜索有点混乱(用来搜寻整个网站)。

- 在下载板块，用户赞赏了评级系统，但不知道系统在给什么评级。因为共享软件非常多，而且会重叠，人们欣赏对共享软件质量的“专家”观点。但这种质量指标对教人们知道四星程序和三星程序之间的差异及差异根源可能很重要。



## 第 15 章 其他艰苦工作：公开信息和专业顾问

并非所有研究都要从零开始。可以不用进行研究就能发现受众的重要信息。想认识受众，了解他们如何体验产品，通过公开信息和专业顾问能节省时间、精力和(偶尔情况下)费用，甚至还可以得到自己无法收集到的信息。

走出公司去了解受众有很多原因。很明显，如果你感兴趣的信息已经有了，通过寻求外部资源常常能节省费用和时间，但这不是采用外包研究最理直气壮的经济原因。走出公司的开发环境，能很快获得更高层的观点，而在公司内部无法得到这种看法。很可能没有资源针对所有可能的目标受众进行广泛调查，以发现他们是否有基本需求，是否满足使用产品的基本要求，但市场研究公司就有这种资源。另外，不同组织可能会从不同角度看待问题。从不同目标和假设看待相同受众产生的结果和你收集到的结果不相同，这对于深层分析和理解你自己的结果非常有价值。

但这个过程并不简单。要找到值得信任和合适的信息很困难，解释这些信息就更困难了。专业顾问也不是能帮你解决所有问题的人，要获取他们的所有价值，必须像管理自己的研究一样管理他们的工作。外部研究是一种重要资源，但和所有强大工具一样，必须谨慎采用。

### 公开信息

有很多种公开分析可供选择。即使这些信息不能准确用于你的问题，也仍然足够成为决策基础。也许这些分析能指出从哪里入手寻找问题，或者能指出谁不是受众。这些分析或许能让你更好地了解如何进行自己的研究。



请记住，公开信息和在公司内部进行的研究不一样。公开信息比你研究的范围广得多，但更肤浅，也不会有你能在研究中投入的质量控制。这些信息可能与你的研究目的相比太过格式化，或者太没有重点。但购买别人的研究常常比自己研究快得多，从长远看也便宜得多。

### 独立分析

大多数研究公司的商业模式都非常简单：独立深入分析特定主题或者特定行业，然后向属于被分析行业公司出售研究报告。这些研究报告都并不便宜，但常常都能代表一组行业专家(或者至少是研究专家)的很多思考。

作为市场调研行业的产物，这些研究公司倾向于关注特定市场、行业或者公司的财务健康情况，但作为研究的一部分，这些公司常常会研究那些市场的目标受众的需求和愿望，把产品卖给目标用户所需的大部分知识向用户提供良好体验所需的知识基本相同，创建用户背景资料或者为情境调查设定期望时，这种营销关注的信息可立即产生价值。

以下几家公司专业进行这类研究。

- 朱庇特通信(Jupiter Communications)，网址为 [www.jup.com](http://www.jup.com)。
- 弗瑞斯特研究公司(Forrester Research)，网址为 [www.forrester.com](http://www.forrester.com)。
- Zona Research(Intelliquest 公司分支)，网址为 [www.zonaresearch.com](http://www.zonaresearch.com)。
- IDC，网址为 [www.idc.com](http://www.idc.com)。
- Gartner 集团(The Gartner Group)，网址为 [www.gartner.com](http://www.gartner.com)。

另外，还有些其他公司的研究关注公司或者行业的用户体验评估，而不是仅仅关注他们的商业指标。这些公司如下。

- 尼尔森、诺曼集团(Nielsen Norman Group)，网址为 [www.nngroup.com](http://www.nngroup.com)。
- 用户界面工程(Nielsen Norman Group)，网址为 [www.uie.com](http://www.uie.com)。

这些报告提供了现成财富，但要仔细阅读这些报告。编写这些报告的研究员往往不如业内人士了解行业，他们可能会错误判断用户行为和公司动机。要仔细阅读他们的研究方法。





## 流量/人口统计

知道哪些人在使用竞争对手的产品、他们的使用程度、用来做什么，就可以避免竞争对手犯的错误，并利用竞争对手的力量。不幸的是，能访问到竞争对手的日志文件和调查数据固然是好事，但通常情况下都属于非法行为。幸运的是，有些服务可以独立收集其中的一些数据，报告聚合信息，并出售对特定数据的访问权。利用这些服务提供的数据和工具，就可能深入了解竞争对手的用户构成和行为。

以下公司提供这类研究。

- ComScore，网址为 [www.comscore.com](http://www.comscore.com)。
- Nielsen/Netratings(AC Nielsen 分支)，网址为 [www.netratings.com](http://www.netratings.com)。
- NUA，网址为 [www.nua.com](http://www.nua.com)。
- WebSideStory，网址为 [www.websidestory.com](http://www.websidestory.com)。

解释这些结果以及如何将其应用于产品比阅读分析师的报告难得多，而且这些服务提供的纯粹数据量很大。这些报告通常包含两种数据：公司跟踪的参加者行为以及向公司报告的参加者的背景资料。把这些数据联系起来能产生强烈的比较。例如，(理论上，不是所有网站都允许直接这样做)可以获得某个市场最受欢迎的网站信息，然后能得到使用这些网站的人的背景资料。

在所有可能提取到的数据中，竞争对手(或者平行行业的公司)产品的用户人口构成、技术能力和互联网使用描述往往是最直接有用的信息。可以马上(从很多其他变量中)看到受众年龄、经验和富裕程度。当然这些信息是你在创建自己受众的背景资料时可能就会考虑的方面，但一组独立数据能确认假设或者怀疑假设。

和所有研究一样，需要仔细检查数据收集的过程，因为检查至少是你能够控制的。有时候，数据收集方法会引入微小的偏差，这些偏差也需要纳入考虑范围。例如，Comscore 公司的 Media Metrix 要求参加者在计算机上安装一个软件：该软件会跟踪他们访问的网站和访问时间。通过这种方式能知道最有代表性的用户样本，但这种方法会漏掉一个关键群体：由于



会违反公司软件安装规定而无法在计算机上安装该软件的人。这就意味着这项服务收集到的数据偏向家庭使用计算机和内容安全标准松散的公司。在很多情况下，这种偏见不会影响最终数据的适用性；例如，这些数据可能不足以代表财富 100 强公司的 B2B 网站。应该跟踪 B2B 用户上班时间的使用情况，但由于 MIS 安全规则，在大多数目标受众公司中都无法跟踪。

## 市场调研

从上述描述中可以明显看出，市场调研的工具也可以用于用户体验研究。公司的营销部门关心什么东西让人们想去你的网站、使用你的产品。而这些都与用户体验直接相关。

营销部门完成的研究常常都能马上用来了解用户人口。例如，以下内容是服务目标对象是 20 多岁~30 多岁女性受众的一项市场研究的背景资料。

部分目标受众的背景资料

### 人口统计

女性，24~34 岁

教育：大学教育

收入：30,000~70,000 美元

78% 的人上网

### 动机

和同住的人及家人有社交活动

拜访朋友

了解当前事件

### 优先级别

在工作和职业之前建立人际关系

### 家庭

25~30 岁以下：通常会和另一个人共享住房

30~34 岁：大多数人都有孩子





### 互联网行为

25~34 岁的女性构成了互联网新用户的最大部分  
每周上网 3~10 小时  
不浏览网站,但依靠他人给出网站指导  
生活中主要动机和那些无法上网的女性的动机一样  
约 50%的人通过美国在网上网  
电子邮件是主要在线工具

### 电子邮件使用

使用邮件和生活中最重要的人保持联系  
91%的人上网时主要使用电子邮件  
78%的人打开电子邮件时感到兴奋  
像打电话一样,用电子邮件和家人及朋友保持联系

### 最喜欢的互联网活动

寻找健康信息  
玩游戏  
寻找精神或者宗教信息  
寻找工作信息  
还喜欢有趣的网络活动,如聊天室和兴趣网站

### 最常去的网站或数字媒体网络(针对 18 岁女性)

AOL 网络  
微软网站  
雅虎网站  
Google  
About.com 网站  
亚马逊  
NBC 互联网

当然并不是所有研究都一样奏效。收入情况并不能真正告诉我们人们会怎么使用,但能提示我们人们可能会使用哪种计算机系统。同样,最喜欢的网站可能是用户习惯的信息架构和交互模型,但不是关键信息。“像打电话一样使用电子邮件”到底代表什么意思?对营销者而言,可能足以



定义出网站的市场定位(如广告)，但无法知道很多网站应该提供的体验方面的情况，意思是希望交互和打电话一样，还是要像电话一样随时能用？可能还应该移动手机界面选项？营销文档没有回答这些问题，但从体验角度看是要调查的重要问题。

但其他信息能立即有用。因为女性是发展最快的互联网受众，高比例可能说明对于互联网属于新手，即暗示了对体验、专业和其他的某些情况。相类似的，她们最喜欢的网站活动会影响信息架构，强调了和我们已知群体偏好相吻合的网站。

### 出版物

熟悉影响你所处领域的出版物，可能是进行任何研究的良好开端。泡图书馆(或者搜索引擎)的时间不会白费，因为常常能发现需要花很多工作才能得到的信息来源，图书、商业类杂志显然是候选者，但还有一些没有预料到的潜在有价值信息来源：你可能会考虑采用新闻剪报服务(也称为“媒体监视”)的旧式方法，来找到分析竞争对手产品或者受众的文章，然后寄给你。

“白皮书”本质就是分析报告，但也会有偏见。编写白皮书常常用来评判特定公司的观点，解释技术，但并不代表无用。在维护观点的同时，白皮书常常包含有价值的信息，但应该注意检查有无偏见。

很多用户体验学者会写一些网络日志或博客，这些日志或博客虽然也包含很多错误，但也包含对用户体验设计和研究的(但未经整理的)深度看法，之后他们会发表在正式的刊物上。

#### 一些免费的互联网资源

##### 一般资源

Usable Web(基思·因斯通的可用性目录)，网址为 [www.usableweb.com](http://www.usableweb.com)。

Usability.gov(美国国家癌症研究所关注网站可用性的网站)，网址为 [usability.gov](http://usability.gov)。

信息架构资源(杰西·詹姆斯·加勒特的信息架构目录)，网址为





[www.jjg.net/ia/](http://www.jjg.net/ia/)。

Boxes and Arrows 期刊, 网址为 [www.boxesandarrows.com](http://www.boxesandarrows.com)。

尼尔森诺曼集团(包括雅克布·尼尔森, 唐·诺曼, 布鲁斯·托格纳兹尼和其他人定期更新的文章), 网址为 [www.nngroup.com](http://www.nngroup.com)。

IBM 研究开发者作品可用性研究论文及文章, 网址为 [www.106.ibm.com/developerworks/usability/](http://www.106.ibm.com/developerworks/usability/)。

微软可用性研究出版物, 网址为 [www.microsoft.com/usability/](http://www.microsoft.com/usability/)。

CyberAtlas (internet.com 运营的互联网使用统计数字), 网址为 [cyberatlas.internet.com/](http://cyberatlas.internet.com/)。

Adaptive Path 出版物, 网址为 [www.adaptivepath.com](http://www.adaptivepath.com)。

### 可用性博客

WebWord, 网址为 [www.webword.com](http://www.webword.com)。

卢·罗森菲尔德, 网址为 [louisrosenfeld.com](http://louisrosenfeld.com)。

查理斯·霍奇, 网址为 [www.challishodge.com](http://www.challishodge.com)。

Gleanings(克里斯蒂娜·沃德克的博客), 网址为 [www.eleganthack.com/blog/iaslash](http://www.eleganthack.com/blog/iaslash), [www.iaslash.org](http://www.iaslash.org)。

Black Belt Jones(马特·琼斯的博客), 网址为 [www.blackbeltjones.com/work/](http://www.blackbeltjones.com/work/)。

本书出版时, 有些链接可能已经无效。开放目录项目(Open Directory Project)的互联网可用性目录会有最新链接。

[dmoz.org/Computers/Internet/Web\\_Design\\_and\\_Development/Web\\_Usability/](http://dmoz.org/Computers/Internet/Web_Design_and_Development/Web_Usability/)。

### 邮件列表

<sigia-l@asis.org>, 网址为信息架构邮件列表。订阅信息和存档网址 [mail.asis.org/mailman/listinfo/sigia-l](http://mail.asis.org/mailman/listinfo/sigia-l)

<CHI-WEB@acm.org>, 网址为 ACM SIG CHI 社团维护的互联网特别邮件列表, 但对所有人开放。订阅信息和存档网址为 [sigchi.org/web/](http://sigchi.org/web/)。

UTEST, 撒伦·霍华德博士维护的可用性测试邮件列表。订阅信息请和沙伦·霍华德博士联系: <tharon@hubcap.clemson.edu>

#### 忠告:

完全解密: 我是用户体验咨询公司 Adaptive Path ([www.adaptivepath.com](http://www.adaptivepath.com))的合伙人。





## 聘请专家

有时候由于资源有限，无法建立必要的专业内容来进行研究。本书提到的几乎每个例子，从招募参加者到竞争性研究到设置摄像机，都可以雇用专业人士来做。根据一定价格，专家能立即需要的几乎所有知识和经验。

但和专家工作并不是写张支票其他事情一概放手不管那么简单。想有效利用专家，需要在合适的时间聘请他们，设定合适的期望，并谨慎管理。

### 选择适当的时机

用好专家的关键是在合适的时间召集他们。顾问经常都是在游戏后期才被召集来提出完美解决方案，而此时所有内部方法都失败了，最后期限也在不断逼近。这时候多半都是希望能出现奇迹。不幸的是，有些广告是这么吹的，但顾问不是圣人。

专业顾问的工作和内部职员所做的工作并没有太大不同，也需要像内部工作一样安排日程。实际上，它需要的时间比你给内部项目的时间还要多一点，因为专家需要了解你的产品及所涉及的任务。

我曾多次受邀“在发布前几周进行一些用户测试”。我很遗憾地告诉招募师火箭已经发射，现在不是去寻找月亮在哪里的时候。测试已完成的产品能发现的结果可能有助于修正小错误，但火箭指向了错误方向，再多测试和修正都是徒劳的。这时需要的是另一种专业。

另外，在特别早的时候就需要把和其他类型专家相对的顾客召集进来。技术专家不会教你做事情，而是为你做事情。在进入之前，他们不了解你的业务，而且离开之后也不会了解你的业务。另一方面，优秀顾客会充分了解你的业务，推荐解决方案并努力把他们的专业技能传授给你。优秀顾客可以教会你一些东西，以后公司再出现问题的时候就不用去找他们。而优秀技术专家会快速准确完成工作，但如果要再做同样工作，你很可能必须再次邀请他们。

#### 提示：

我对专家的定义包含专业顾问、承包商和咨询代理。大量咨询公司(Accenture/KPMG/IBM 全球服务模式)都从单一承包商那里开展工作，但他们的产品和你的产品及公司的关系相类似。他们都是来解决特定问题的，充当团队的助手，在需要的时候进行互动，但保持负责雇用他们开发的特定方面。团队的工作可能会根据需要遍及整个产品，但专家很少脱离其专长来解决其他问题。





幸运的是，很难太早召集用户研究员，但还是要在合适的时间进行合适的研究。如本书第一部分所讨论的，优秀的迭代开发过程几乎会在每次迭代中引入用户输入。挑选需要什么研究既是项目开发者的责任，也是研究员的责任。例如，如果在产品特性定义前对产品交互进行可用性测试，很多信息都是浪费，因为人们对产品的使用会随着现有选项而发生变化。同样，如果没有首先研究用户需求，对建立出来的产品进行测试会导致大量多余工作；如果产品受众对产品不感兴趣，他们就没有动力了解产品如何工作，也不会以现实方式使用产品。

### 找到专家

如果任务需求相对直接，比如设定特性优先级别的一轮情境调查或者一些焦点小组，找到专家的过程和为住房找木匠相类似。

(1) 写出研究需求和目标。要进行哪种研究？会如何利用结果？这和第5章描述的如何准备自己的研究相类似。

(2) 列出能联系的专家名单。向同事询问推荐，或者联系某一个可用性行业组织，这些组织都有专业顾问和承包商名单。有下面这些知名机构。

- 可用性专业协会 (Usability Professional's Association), [www.upassoc.org](http://www.upassoc.org)。
- 美国信息科学与技术协会，特别是其信息架构特殊兴趣小组，[www.asis.org](http://www.asis.org)。
- 美国计算机协会的人机交互特殊兴趣小组 (ACM SIGCHI), [www.acm.org/sigchi/](http://www.acm.org/sigchi/)。
- 旧金山湾区交互设计特殊兴趣小组 (BayCHI)，维护着许多地区的咨询师名单，[www.baychi.org/general/consultants.html](http://www.baychi.org/general/consultants.html)。

(3) 检查资质。在雇用之前，要调查研究公司的具体经验。你不会想要专业修房子的木匠来做家具，也不会请做家具的人来盖房子。即使方法非常类似，用户体验专家可能也没有任何市场调研的经验(反过来也一样)。



(4) 获得报价，并了解对理念和方法的解释。如果有可能，要从所有纳入考虑的咨询师那里获得结果文档样本。阅读结果，了解方法解释，并寻找针对产品和客户需求的敏感性内容。

(5) 要求几位最近客户的证明，并跟进这些证明。调查工作质量和时间，还要调查服务质量。咨询师有没有履行承诺？他们倾听了吗？他们反应快吗？

### 正式方法：需求方案说明书(RFP)

如果任务较复杂(大型焦点小组，调查，采用多种方法的多次迭代)，物色专家的过程更像从零开始盖房子。由于任务规模和任务之间的内在关系，找到合适专家的过程会非常复杂。

(1) 编写需求方案说明书(RFP)。需求方案说明书包含问题描述和解决方案的结构化说明。不仅能为评估专业顾问投标设定参数，也是管理和组织项目的第一步。它用具体条件来说明问题以及你期望通过外部解决方案中获得什么结果。

(2) 分发需求方案说明书。可以把说明书发送给开始联系的专业顾问，或者也可以发送给更大的群体。不要滥发，但有一些邮件列表和公告栏允许张贴这种说明书(向管理员询问需求方案说明书张贴政策)。

(3) 评估答复。专业顾问要能快速做出答复，并提供具体解决方案，而不是销售宣传。注意，有一些公司会把关键工作分包给其他公司，如果有这种情况，要像评估主承包商一样来评估分包商。

以下是一项大型长期多部分项目的需求方案说明书样本。较小项目的需求方案说明书不需要下面的全部内容，也不需要这么详细。

需求方案说明书样本
需求方案说明书
B2B 工业积压产品网站用户体验研究
2003 年 1 月 12 号
答复截至时间：2003 年 2 月 12 号





### 第一部分：项目概况

我们代表一家世界领先的积压原材料网站，该网站正在进行系统重新设计。我们致力于采用以用户为中心的设计的最佳实践作为本次网站重新设计的一部分，并作为之后的持续设计基础。通过本需求方案说明书，我们希望找到能满足项目要求的用户研究所需要的提供商。

本需求方案说明书将说明我们对方案的远景、挑选标准和对答复的期望。

### 背景

我们公司运营着世界上最大的在线工业原材料贸易服务之一。2001 财政年度交易额超过 10 亿美元，30,000 活跃用户，我们是世界上最主要的积压原材料交易服务提供商之一。我们的用户依靠我们开展业务，信任我们进行交易。我们为了用户利益和业务利润而持续改善服务。

### 项目描述

为了改善服务用户体验，我们已经开始了主要的重新设计计划，关注用户需求和输入，自上而下地创建服务。

重新设计将会分成一系列阶段完成。每个阶段都会涉及对当前产品的再检查，为下一次设计修正产品理念。用户体验研究员是每个阶段的重要组成部分，每个阶段都会包含适用于该阶段的主要研究项目。

本项目将会于 2003 年第二季度到第三季度进行。

### 第二部分：需求方案说明书要素

我们希望您的说明书采用特定格式，包括下面提到的所有章节。如果您认为这些部分不足以说明您的核心竞争力，可以增加章节。

### 问题

我们希望能了解您对本项目的想法。请使用本说明书中的信息回答以下问题。

1. 本项目成功的关键之一是实施研究所确定的用户需求，但这些需求并不一定和服务的商业需求一致。您会采用什么步骤来解决此问题？
2. 项目成功的另一个关键因素是知识从用户研究员到产品人员的传递。在您看来，在公司内部传递用户知识的主要障碍是什么？您如何解决这些问题？
3. 您认为项目成功的最大障碍是什么？我们应该做什么才能确保获得可能的最佳结果？



续表

**案例研究**

请提交至多三个案例，以说明贵公司管理类类似本说明书中提到的项目能力。请提交最终交付成果和合适的附带结果。

**核心竞争力**

我们寻找能在下列领域显示成果的一位或者多位合作伙伴。如果有可能，请提供您在下列领域的具体经验案例。

- 复杂信息搜索及购买任务分析
- 面向用户体验的焦点小组
- 网站用户测试
- 研究商业网站用户的需求，最好是大型工业和制造业用户
- 理解新用户、或者偶尔用户及定期长期用户的需求

**过程交付成果**

请描述一下贵公司完成项目的过程。要包括研究描述，详细开发说明书，产品，集成，质量保证(Q/A)和每一阶段达成共识和结束的方法。

**日程表**

整个项目时间为 200 天。请提供本项目的具体开发时间表。请用交付成果分开每段过程，并具体说明每部分的时间。

**客户角色**

如果有的话，请描述一下您的期望或者您需要从我们获得哪些交付成果。请描述一下每个阶段您期望的资源种类，您希望和公司哪些人见面。

**团队**

请描述分配到本项目的每个人的具体角色。您会分配多少人参加本项目？请尽可能详细提供参加本项目的每个人的背景资料和联系信息，并具体说明每个人的职责。请确认每位成员在此项目中负责的关键项目。另外请说明他们在本项目上能够投入的时间。

**预算**

请提供详细预算说明书。您可以采用熟悉的任何格式提交预算说明书，但要包括预算里程碑，每位成员的每小时花费和预计费用。

同时请说明在项目期间花费的说明方式，超额费用如何进行管理并获得批准。

**证明材料**

请提供最近项目的三家公司的名称和联系信息，他们要能对贵公司的服务发表评论。





### 过程

我们会邀请几家组织对本项目提交说明书。根据这些说明书的质量和性质，我们将邀请三家公司进行答辩。

您将有一个半小时的时间说明方案，并回答管理团队的提问。请对预算和过程的详细情况做好准备。请安排把参加本项目的具体团队成员带上，包括项目管理、定量研究、定性研究和分析。

[基于贾尼丝·弗雷泽创建的模板]

### 随意方法：电子邮件和电话

但对于大多数用户体验研究和设计工作而言，还没有复杂到需要用需求方案说明书的地步。

就在我们 Adaptive Path 的大多数工作而言，我们喜欢采用简短电子邮件来说明需求和问题，然后开个 1 小时或者 2 小时的电话会议来讨论细节。这种程序能节省双方时间——不用编写需求方案说明书及其答复——并能快速关注到最相关的要素，而不用提前预测这些要素。我们认为，咨询师的角色是从打第一个电话开始，我们努力帮助客户立即了解并格式化他们的需求，而不是只针对他们觉得他们的需求应该是什么而做出答复。但如果项目很大，而且有很多潜在公司都在争取获得项目，采用需求方案说明书是最好的选择。

### 设定期望

和专家、尤其是咨询师或者咨询公司建立关系需要双方都设定期望。

作为客户，你比他们更了解业务。如果他们知道了你知道的东西，他们很可能就是竞争对手。即使你把知道的所有东西都告诉他们，你还是比他们更了解业务。你的经验更多。并且应该对你期望专家知道的东西形成框架。他们扮演的角色不仅仅是替补，而是信息资源和工具。

另外，从整体看，用户体验研究员更擅长发现问题，而不是针对这些问题创建解决方案。很容易会依靠单一来源，请他们告诉你产品存在的问题，并请他们针对这些问题给出解决方案。咨询师会很乐意告诉你他们觉



得问题的最佳解决方案,但这种解决方案仍然是来自对业务了解有限的人的观点。听取建议、根据你的情况来理解这些建议并正确应用仍然是你的责任。

**专家根据经验提出看法。**由于有经验,因此他们知道有些通常解决方案比其他的更好,他们可以根据你的问题来定制这些方案。他们告诉你的不一定和你对世界的了解相吻合,但要听取他们的说法。开发过程中,你会根据某些假设做出结论。无论你和你的同事有多诚实直接,最终都会根据相同信息从差不多的角度发现问题。而局外人能带来完全不同的假设和信息,他们做出的结论会和你的结论非常不同。并不是说他们对了,你错了,也不是说你对了,他们错了,但他们带来的看法能丰富你的看法。要允许他们提出根本问题,并重新引入你已经做出决定的主题。

如果不雇用咨询师,就把他们当成是知道你两三个月之后要做什么事情的人,这是看待咨询师的一种方式。的确就是这样。把他们看成比你超前两三步的人,如果雇用他们,开发日程表上就能减少几个月的时间。几个月的开发时间真的很昂贵,因此招募专家一般都很值得。但你买的是时间,不是魔术。

这些期望可以提炼成一系列专家管理指南。

### 专家管理指南

- **清楚自己的需求。**如果你很熟悉业界的基本情况和方法,你就属于咨询服务行业更成熟的客户。如果把用户测试称为焦点小组,不仅在和专业顾问见面时会显得失礼;还会造成对研究目标和项目整体目标的混乱。知道了有什么东西,就能知道哪些问题你能获得答案。根据公司商业需求和提议的产品,提前确定研究目标能让结果更有意义,更有用。
- **仔细安排日程。**研究交付时,需要能解决公司的需求。如果交付太早,在需要研究的时候,项目可能已经发生了变化。如果交付太晚,可能会错过解决已识别问题的最佳时机。
- **提供先导时间(lead time)。**对任何主题而言,专家准备的时间越





长,结果就会越好。招募合适的受众就需要经过多次迭代才能得到合适的筛选文件。要提出合适的问题,研究员需要了解研究目标和你的产品。分析总是非常耗时,因此为分析留的时间越多,效果就会越好。专业顾问总是会受益于一定长度的前导时间受益。

- **对建议保持开放心态。**专家可能不如你了解业务,但如果他们建议的东西会挑战你对产品或产品市场的假设,还是要评估和考虑他们的建议。例如,假定青少年玩滑板的时候会在 PDA 上使用某个产品。用户测试发现虽然青少年会用该产品,但他们玩滑板时不会带 PDA。专业顾问建议把信息做成简单印刷格式,因为纸张更容易携带,更便宜,但客户说:“我们没法投放纸质广告,纸张不是收入来源。”
- **观察过程。**报告和演示说明都是有价值的总结,但它们能包含的只是通过观察咨询师所产生信息量的一小部分。只要有可能,都要请开发团队成员现场观察研究。如果无法进行直接观察(比如在大多数情境调查研究中),可以请求录像带的拷贝并进行研究。因为作为客户,你更熟悉产品以及要解决的问题,你就可能会注意到咨询师忽略的问题。
- **进行汇报。**专业顾问的书面报告容易得到,而忽略了演示报告。别忘了,想法在发展。而对报告的演示就是专业顾问向我们讲解报告,我们自己也可以阅读报告。好的演示说明远远不止这些。在好的演示中,专业顾问可以对事项分出优先级别和强调重点,详尽阐述要点,并回答问题。
- **初始研究结束后,把专业顾问当成资源。**专业顾问完成了研究,但不能忽略专业顾问的很多专业知识。如果对他们完成的工作有疑问,不要犹豫,马上去问他们(但要付钱给他们!)。有时候甚至都应该在大型项目之间每个月留他们几小时,从他们那里获得想法,搞清楚之前的观察。时间一长,他们可能会丧失一些客观性,但他们能更深入了解产品及其受众。





## 第 15 章 其他艰苦工作：公开信息和专业顾问

专家和客户之间的关系能变得有价值，如果这种关系能发挥作用，双方都能互相受益、互相学习，开发过程效率同时会更高。如果无法发挥作用，就会很容易破裂，双方都不会比以前更好。从长期来看，重用知识和经验最有价值的方式是提高整个行业的效率，每个人都有理由来思考真正关键的事情。如果不是必须，就不要重新发明轮子。



## 第16章 新兴方法

本书描述的每种方法都从不同角度了解人们的看法、希望和能力。可用性测试告诉你人们的能力和期望的精细信息。调查能广泛描绘出人们的愿望和希望。情境调查有助于了解体验发生时的全部环境。

正如本书所述,这些方法都在特定时间采用常见技巧来发挥作用。但本书内容并不是绝对真理。很多情况下最终需要采用的方法和我所描述的方法会非常不同,请不要受限于本书所述方法。要发明、试验各个方法并探索这些方法的可能性,根据具体情况进行调整。

另外,每种方法都可以单独使用,但每种方法都无法满足全部用途。就像锯木厂生产木板时需要注意锤子和钉子的力量 and 需求一样,用户研究工具存在于共生环境下。每个工具都来自于其他工具的集合,使用时要记住这些工具,把从不同方法获得的信息进行整合,博采众长,从而具体地、有目标地综合理解。

### 方法变体

本章介绍的每种方法描述了一些替代方法和融入其中的灵活性工作。你可以对3个人或者30个人进行可用性测试,你可以调查网站随机用户或者订阅邮件列表的人。焦点小组能讨论大型抽象问题,或者能紧密关注特定主题,有很多方法能让这些方法适合研究需求。另外,有很多流行的方法变体能引向明显不同的方向,同时还能收集到有价值的数据。

#### 虚拟焦点小组

焦点小组不一定要让所有参加者都坐在同一间房间里,可以通过任意



一种现有的“网真”(Telepresence)技术把一组人召集在一起,最简单的方法就是电话会议。电话焦点小组的优点在于不必把每个人都召集到房间,而且只需占用参加者参与焦点小组的时间。如果要考虑谁能来、什么时候能来,面对面的焦点小组就必须考虑旅行时间、交通、有空的时间和等待时间等因素,这些因素会把参加者限定于住在或工作在会议地点及附近的而且有大量时间的人。还会让招募更加困难,同时也会让小组的用户群有偏差。通过电话引导小组访谈能消除这些地理局限,可以通过电话进行焦点小组,而每位成员都在国家的不同地方,在他们最方便的地点进行发言。

但电话焦点小组比面对面的小组更具挑战性:如果陌生人互相看不见,就更难联系在一起并不受约束地进行讨论。如果他们看不见肢体语言,就更难明白对方的意图。焦点小组通常把视觉说明作为引发或者维持讨论的触发器,此时却很难引入。另外,在面对面的焦点小组中,能知道参加者的注意力:能看到他们,你给他们看到的東西,他们有没有反应。而电话焦点小组,你无法控制他们的注意力。虽然你能鼓励人们注意,但无法阻止他们检查电子邮件,也无法阻止他们购买杂物,而此时他们的注意力应该集中于焦点小组。

另外,如果无法找到特定技能的人参加(如繁忙的执行官)或者希望参加者能覆盖更广泛的地理区域,此时电话焦点小组会非常有价值。

第二种常见的虚拟焦点小组是在线焦点小组。这种焦点小组通过聊天室(常常需要专门软件)来进行小组访谈。参加者登录到焦点小组,并通过打字而不是互相说话以及向主持人说话来进行交流。从参加者的地理和时间概念上看,这种方法和电话焦点小组有相同优点。另外,在线焦点小组能提供其他方法所不具有的距离和匿名感,这是一个问题的两个方面:一方面,在线焦点小组允许人们说出一般无法说的东西,或者不会用一般的方式来说的东西;而另一方面,在线焦点小组更难让人们觉得他们是小组的一部分,所以更不可能说出自己的想法。因此,有些参加者会越来越开放,而另一些则越来越安静。

主持这种小组——特别在没有表明个人态度的语调或者身体线索的情况下——明显比主持标准焦点小组更难。另外,所有参加者都必须打字





才能实时表达想法，这种局限性确实会限制潜在受众。

在受众合适的情况下(如青少年或者长期上网用户)，这种方法非常有效。网站和软件能在最自然的环境下展示出来——人们自己的显示器——可以进行话题讨论。

有了更宽的带宽、运算能力更强的桌面系统，在不远的将来可能还会有新方法能用于远程焦点小组。更宽的带宽能让互联网电话应用系统比当前电话会议更好地集成到演示软件中。这样能进行更丰富的电话焦点小组，可能很快就会有直接基于桌面的视频流焦点小组，能消除聊天室和电话访谈的很多问题。

### 提名小组

焦点小组一般获得参加者的潜在价值观和愿望，通过请参加者深入讨论他们自己的问题，就可以了解问题，但很难理解问题的相对重要性。讨论最多或者争议最多的问题常常看起来最重要，而和人人都同意的问题实际相比，前者只是相对重要。为了尽量减少这种情况，德尔贝克和范德·维恩和古斯塔夫森在 20 世纪 60 年代晚期及 70 年代早期发明了提名小组(Nominal Group)的方法。这种方法是混合焦点小组和调查方法高度结构化的小组访谈。

提名小组基于以下主要想法：如果在小组讨论开始之前请人们写下自己的想法，人们会对自己的感觉最诚实，最忠于自己的看法。这样能减少群体思维的可能性，还能减少有魅力的参加者左右其他参加者看法的可能性。

方法如下。

- 首先引入主题。人们在短时间内就能对这个主题做出简洁答复。例如，“您觉得重要的银制餐具的特质”。
- 参加者写下对主题的答复。
- 每个人读出并解释自己的答复，一次一人，同时在中心位置(比如白板)写出所有答复。不用讨论答复，只要解释说明即可。给答复加上特定的识别符(数字或者字母)。



- 参加者对答复评出等级，写下他们认为重要的答复的识别符，采用李克特量表给每一项评级(见第11章)。
- 分享并讨论每位参与者的选择及评级，以评级为中心传统焦点小组中的关键的讨论在名义小组中甚至有时候可以略过或者缩短)。

这种方法以牺牲深度来尽量减少参加者的偏见。因为在特定时间内，人们一般写得都比说的少很多，因此针对人们的价值观和需求而言，这种方法比传统焦点小组产生的信息量要少。其优点是简单，事后几乎不需要分析，并且能确定问题的优先级别，可以通过设定的议程进行更广的讨论(个人之间的互动有限，因此可以在实际焦点小组之前通过电子邮件来完成设定日程表)，但涉及深层理解时，这种方法就显得有些苍白无力。

### 摩擦小组

招募好了焦点小组，需要仔细关注所有参加者在关键方面的相似度，使他们感到轻松自在，能从类似体验得出结论，保持用户和谐相处，这样能帮助研究员了解影响人们看法和体验广度时减少大量变量。

挑选相同世界观的人有一些缺点：和相似的人讨论问题会漏掉根本问题，完全是因为参加讨论的每个人都会从一套常见假设进行发言。另外，如果人们觉得他们是小组成员之一，就不太会脱离小组，说出异议。这就会导致群体思维，而在有重点的访谈情况下，群体思维非常危险。

为了减少上述影响，有些营销研究公司已经开始试验这种焦点小组形式，有种方法称为摩擦小组，(有点像)传统焦点小组的孪生坏兄弟。没有根据相同值来筛选所有参加者，而是在招募参加者时有意包括多种角度。然后鼓励参加者讨论并维护自己的观点。

招募的时候，参加者应该在单个细分因素上都相同。这样能减少研究时因为其他因素造成的人们观点的差异。细分因素可以是观点(“房地产是比股票更好的投资”)或者选择(“戴尔比宏基”)。

要选择相同数量的人来代表任何特定观点，这样就没有人会觉得其他人会结成一伙。因此，如果要在8个人的焦点小组中涵盖两个不同观点，每个观点要有4个人来代表。如果决定需要3个观点，可以安排9位参加

#### 忠告：

摩擦小组相对而言是个新概念。摩擦小组还没有很多经验，经常会真的发生错误，因为人们的不同观点会造成人们省略、而不是维护自己的观点。因此实施摩擦小组时要尤其小心。





者，每3位代表1个观点。

招募过程和常规焦点小组一样，但就差异和类似性而言，需要额外谨慎地来验证所有参加者都符合所有关键变量。

主持摩擦小组肯定会特别困难，摩擦小组的目标是通过观察人们如何定义价值观来了解他们的价值观。而同时不能让他们发怒或者生气，要在一组陌生人之中创造这种情况需要小心，而且要完全理解参加者动机和参加者的心智模型。

### 虚拟可用性测试

没必要在人们身后观察他们能否完成任务或者他们对产品有什么问题，可用性测试的基本概念是估计人们能成功使用产品，同时允许他们回退并评论产品的使用体验。在标准可用性测试中，可以通过主持人给用户任务，并在人们使用网站时提出问题来完成。这种方法可以针对任何类型的产品，不管是软件还是烤箱。但网站容易创建这种“识别距离”，而同时又不需要让人坐在旁边。因为可以从中央服务器下载下来，就可以通过框架或者另一窗口创建一种网站的“心理”包装，并提供主持人在实验室中会提出的任务和问题类型。另外，还可以将过程自动化，这样测试可以不是6或10人、而是60或100人来完成(或者6000或10000人)。这些数据可以产生复合可用性指标，而这些指标来源于比传统的可用性测试更广泛的基数。

“虚拟可用性测试”(Vividence 和 NetRaker 等提供这类服务的公司有时将其称为“远程可用性测试”)不仅仅是“虚拟”执行的可用性测试，它们是调查、可用性测试和日志分析的交集。在这些系统中，独立的框架(或者对话框)包括问题和任务。随着人们执行任务，系统会自动跟踪进程(就像日志文件)，并在合适的时候针对人们的体验提出问题(比如在人们觉得完成任务的时候，或者如果人们放弃了任务的时候)。另外，它能度量人们完成任务的速度，以及完成这些任务的成功率。

最后，自动分析数据，将收集到的数据根据不同指标进行计算。例如，Vividence 能计算参加任务的人数和点击第一个事项的人数的比例。这样



能更深入了解网站导航的有效性,还能提供点击流程图,并能突出显示成功及未成功的路径。

虽然这种系统能度量成功率,并提出基本问题——可用性测试的基础模块——但它缺乏弹性和真实可用性测试所具有的深层事件数据。如果要分析可用性测试录像带,大多数分析都不会只度量任务成功率,还会记下人们的评论。优秀主持人会在问题发生的时候观察到问题,并探究其原因,把可用性测试的注意力放到对参加者和产品影响最大的事情上。分析师能利用这些探究来了解参加者的心智模型和假设,这样能更基本、更广泛地了解用户体验。

对于直接性问题,这种方法还是不错的, (“人们想登录时怎么失败了?”),但无法解决更复杂的问题 (“人们为什么会在餐叉板块中寻找刀子?”)。回答基本问题很重要,能突出很多问题,但只是了解了网站交互问题背后原因的一部分,另外,很难了解用户的理由,或者很难根据他们对多项选择问题的回答以及任务成功率来建立心智模型。这些系统虽然也提供开放式问题,但无法调整来符合特定用户体验,因此人们的答复(和开放式调查问题差不多)容易变成夸夸其谈,没有太多用途。

还可以采用允许虚拟焦点小组的网真技术来进行可用性测试。利用视频会议设施配合远程计算机控制技术,如通过 Netopia 公司的 Timbuktu Pro([www.netopia.com](http://www.netopia.com))或者 WebEx 公司的桌面会议产品([www.webex.com](http://www.webex.com)),主持人和用户可以位于世界的不同地方,而又能保持面对面可用性测试的自发性和深度。

## 眼动跟踪

虽然眼动跟踪看起来是一种非常低层次的观察体验方法,但通过了解人们看过哪些东西能发现大量高层次现象。只有视网膜中央才能看见精细细节,这是人眼的属性。也就是说,如果需要检查某样东西,我们需要直接看着那样东西。例如,试试往左看的同时看看右边的东西:虽然能看见要看的東西,但看不清楚,因为没有足够的分辨力。另外,人们对给定物体的注目时间和对这个物体的思考时间成正比。利用跟踪设备(如 Eyetools

编者注:

眼动跟踪的一些有价值的研究数据,详见清华大学出版社 2010 年出版的《Web 表单设计:点石成金的艺术》一书,作者 Luke Wroblewski。





公司的 [www.eyetools.com](http://www.eyetools.com)), 就可以针对人们看过哪里和看了多长时间创建地图, 而且能细化到一定程度。如果和人们看过的东西对应起来(例如他们看过的网页), 就可以发现什么东西主导了人们的思考过程。

用于可用性测试时, 根据用户的视觉重点和交互要素组织形式, 眼动跟踪能发现某些行为和错误是什么样子。

可以从视觉上完成眼动跟踪信息的基本分析, 可以检查使用模式来看看界面的哪部分带来了最大流量, 这样能立刻知道人们的注意力如何分布到界面要素中。参加者有没有注意到左侧的横幅广告, 有没有注意到底部导航? 但很难知道注意力不同的原因, 人们花了很长时间看导航列表, 是因为真的关心导航还是因为觉得困惑想要搞明白? 他们没有看到左边的链接, 是因为没看见还是因为已经了解此板块不需要再看? 一般情况下, 眼动跟踪不能回答这些问题, 需要和能回答这些问题的方法结合在一起。

眼动跟踪最大的问题在于实施眼动跟踪方法的设备昂贵而且麻烦, 除了基本事实外(哪位用户看这个点最多), 任何解释还需要有受过培训的操作员和分析员, 从而进一步增加了费用。换句话说, 随着成本不断降低, 眼动跟踪能变成一种强大工具。

### 平行研究

任何研究过程都会受到某种偏见的影响, 这些偏见源于招募参加者的方式, 问题提出方式或者结果分析方式。理想情况下, 研究人员能知道哪里会发生偏见, 而且能对偏见进行补偿(或者至少能对偏见发出警告), 但无法完全避免偏见。

研究人员常常会给过程引入个人风格, 而且很难知道哪里会产生这些偏见, 尤其是对于负责研究的人来说, 有种方法能显示出研究被扭曲的地方: 即对相同主题比较多种研究结果。不同群体的人可能会采用略有不同的方式来进行研究。这些不同可能会从不同方向扭曲结果, 因为这些群体不太可能包含所有相同偏见。对结果进行比较, 就可能发现结果产生分歧的地方。

在有些情况下, 可以把结果与其他研究进行比较(如第 15 章所述)。





如果要收集根据地理组织的人口统计信息,可以把这些信息和美国人口普查局的数据进行比较。这种方式检查数据很快、很简单,但产生的偏见很少与产品或者要回答的问题密切相关。

进行平行研究是找到自己的研究和产品产生扭曲之处的最佳方式。但这样会导致研究成本加倍(需要分析的数据数量也会加倍),但能提供有价值的观点。

简单的平行研究方法是在内部同时进行两个研究。这种研究由两组(或者更多组)来进行,这些小组采用“净室”(Clean Room)研究方法。小组完全互相独立地进行研究,不讨论方法或者结果,但根据相同问题开始工作。每组会采用不同方法,并以不同方式进行分析。还可以让每组分析另一组的数据以及自己的数据,然后比较结果。

如果没有资源在内部进行两套研究,可以在内部进行一套研究,而雇用咨询师或者专业研究公司进行另一套研究,或者雇用两家公司,并比较结果。

这种方法要求能有完全独立研究小组的资源,会非常昂贵。而且,虽然能提高结果的准确性、使事件的证据和简洁的故事的数量翻倍,但结果数量不会翻倍。

通过不同方式应用平行研究方法,可以发现所有(或大部分)用户的现象,同时也可以发现特定用户在特定任务下的现象。然后将典型研究和这种有意不同用户在不同任务的研究进行比较,就能发现哪些问题是大多数用户共有的,哪些问题只出现在特定用户组和特定任务下。

## 参与性设计

作为一套方法学,参与性设计作为一种民主化产品设计方法之一,在 20 世纪 70 年代的北欧起步,并在 20 世纪 80 年代获得进一步发展,形成了一套完整设计方法(见 1993 年 6 月 ACM 通信特刊,阅读此主题的深层研究)。

参与性设计的核心,是利用焦点小组总结需求进行迭代设计的过程,通过任务分析把需求提炼成规格,然后通过用户建议委员会保持用户在整





个开发过程中的输入。头脑风暴、优先级别练习以及定期的最终用户评审参与性设计的大部分内容。

IBM 的联合应用设计(JAD)和爱德华·戴明的整体质量管理(TQM)是广泛利用参与性设计思想的方法。此类方法大多数都有一个常见现象,即一整天的研讨会。在研讨会期间,产品开发中不同利益相关者(直接受到产品影响的人或者影响产品的人)的代表,包括用户,要对产品理念及其需求达成一致看法。通常通过焦点小组风格的会议来完成,会议以大量需求分析和优先级别练习为特色,总结出问题后,采用任务分析把这些事项分解成具体问题,然后把这些具体问题作为集体编写的解决方案规格的基础,再把解决方案交给开发人员来实施。

在之后的常规时间中,通过原型来重复这一过程,而且焦点转化到其他产品和问题。

这种方法能很好地满足用户的功能性需求,其弱点类似于建议委员会的弱点:经过一段时间后,参加的用户会像开发团队成员一样思考,无法再跳出开发过程的局限。而且,其过程很容易把自己限制到少数用户代表所表达的观点,而用户代表常常都不是整体用户人口的代表。对于产品形象和财务职责,这种方法基本无法提供指导。视野狭窄的参与性设计很容易遗漏产品受欢迎或者盈利的关键要素。

## 组合使用

就方法本身而言,本书中描述的方法都很有用,而且信息量丰富,但如果结合在一起使用,使一种方法的产出成为另一种方法的输入,这些方法会更强大。采用多种方法来了解问题答案能让你从不同角度看待问题或者深入讨论问题。

### 焦点小组和日记

使用日记的最有用方式之一是混合使用有关联的焦点小组或者访谈。日记作为讨论触发器,它们维持着不同焦点小组之间的前后环境。



例如,招募一组用户进行4次系列焦点小组来研究某个搜索引擎,为期3个月。焦点小组的目标是了解有经验的搜索引擎用户在学习该搜索引擎时所面对的问题。此时,产品用户中的很多人都是从使用了很长时间的其它搜索引擎“叛逃”而来。该搜索引擎公司关心的是吸引更多这种用户,但想知道人们体验到了什么问题,以便能最大化体验价值,并提高停留时间(大多数用户都认为搜索引擎是一种商品,如果有一个搜索引擎让他们失望了一次,就会立即弃用,转而投向另一个引擎)。

第一次会议,向小组介绍研究概念。焦点小组的目的是了解用户使用该搜索引擎选择时碰到的问题,并发现他们一般看重什么搜索服务(因为要在研究中跟踪这些情况)。然后指导他们使用日记系统(每周寄给他们两次相对无结构的格式日记),并请他们在会议间隔期间填写日记。

过了一个月,召开第二次会议。在此过程中,研究小组跟踪在用户格式日记中的调查以及出现的问题。对这些情况的讨论构成了第二次焦点小组的主体。参加者澄清他们的想法,并分享上个月他们体验的更多细节。

再过一个月,第三次焦点小组用来关注已经发生的持续学习情况。阅读第二个月的日记,研究人员发现了人们如何使用该搜索引擎,并明白参加者学到了什么。本次会议关注澄清这些问题,了解人们对该搜索服务工作机制以及他们对搜索引擎的期望所建立的心智模型。探索用户所提特性的渴望性,并比较其他搜索服务。

第三次和第四次会议中间,日记频率下降到每月一次,而且日记格式变得更像可用性测试类型,以便能把参加者暴露到特定特性下面。第四次会议关注总结参加者对服务不同特性的体验情况,同时引入若干功能原型,而这些原型是根据前期焦点小组中讨论过的问题进行的针对性设计。

在焦点小组和日记之间进行反复研究,开发人员能更好地了解过渡用户在体验其他系统产品面临的问题。

同样可以把日记和访谈结合在一起。利用对人们体验的认识来结构化访谈能比标准脚本更好地针对访谈主题。





## 观察性访谈和可用性测试

传统的可用性测试几乎完全基于任务，希望人们执行任务并利用人们对这些任务的表现作为了解交互和产品架构问题的基础，传统性、可用性测试过程受此驱动，这种方法能很好发挥作用，但低估了访谈产品实际(或者潜在)用户的潜力。这些人能告诉你的不仅仅是保存设置是否有麻烦的体验情况，而且还包括什么吸引了他们？他们对产品用途有多深的了解？哪个特性对他们有用？为什么？

把可用性测试和深层态度问题结合在一起，比如在情境访谈或者观察性访谈中会发现的问题，这样能创建出更丰富的分析数据。第10章描述的概况访谈就属于此类访谈。在此过程中，在参加者使用网站之前，询问他们对网站的整体印象。通过这种主观的印象评价，就可能发现不仅仅是人们不能用什么和不能理解什么的情况，就可能了解什么吸引了他们，他们的价值观是什么，动机在用户体验中是强大力量，愿望和亲密度与动机关联。有一种方法能允许研究员同时能了解什么是广泛、功能性、愿望性和有吸引力的东西，能发现产品为什么会对受众起作用，为什么不能。

例如，有一家在线家具网站关心潜在客户能否接受他们的原型。时间是关键，几乎没有完成其他用户研究。因为原型已投入开发，因此太晚了无法研究目标受众的详细需求和愿望。有一组潜在客户受邀参加原型的可用性测试。这就有机会在向用户展示产品创意或实施的产品之前来研究人们的愿望、体验和期望。招募参加者时，没有向他们透露产品或者研究标准，研究是在匿名租赁的办公室空间中进行。

每90分钟环节的前半小时用于一系列一般访谈问题。所有参加者都受邀来讨论购买家具的体验，他们遇到的困难，以及他们在选择家具零售商和家具时的标准。访谈虽然简洁，但可以提供对客户的以及他们之前未见过的产品的看法，同时对产品定位产生深刻影响。

不幸的是，请人们检查并说出什么吸引他们的过程会对使用产品造成偏见。他们可能会更关注界面，注意他们本来会忽略的产品方面。这会导致结果扭曲：最初令人混淆的东西会变得可以理解，或者一开始看起来有



吸引力的东西会因为扩展暴露而变得过分。因此，综合可用性测试和访谈不能取代严格的可用性测试，如果时间和资源有限，而又需要快速和深层（如果不需要精确的话）研究的时候，这种方法能增加有价值的要素。

## 调查和焦点小组

相互关联的一系列调查和焦点小组是经典的营销混合研究方法。调查回答受众“什么”的问题，而焦点小组回答“为什么”的问题。将两者相互关联在一起，就可以用一种方法来回答另一种方法提出的问题。

调查揭示出人们行为中的模式。造成这些行为的原因可以用焦点小组进行研究，反过来可以用调查来验证其他趋势。交替发现有趣行为，并试图解释这些行为，这是一个循环模式。

- 首先，实施一般人口统计/技术统计/互联网统计调查来提供基本受众背景资料。
- 焦点小组根据背景资料招募人选。焦点小组研究人们的价值观以及使用体验，而这些体验也正是产品试图要解决的。
- 焦点小组针对功能和竞争对手提出问题，可以通过另一轮调查来实施。
- 等等。

这是一个经过实践证明的可行方法，而且已被使用了多年，用来创建你熟悉的产品和广告。像 Clorox 一样的公司采用永不结束的调查和焦点小组循环过程，每个循环都融入下一个，这些结果之后都应用于产品。这些方法的缺点在于几乎无法发现人们的能力或者实际需求。营销和消费产品开发常常更多是关于认知的需求，而不是实际需求，而后者正是这些方法最擅长的地方。但人们自己的观点不一定准确，还需要用更客观的角度来看待用户的实际行为状态。所以，这种研究的产出是了解人们动机、愿望的有价值的因素，但要和能力及需求评估方法结合使用，如可用性测试和情境调查。





## 日志文件和可用性测试

使用数据分析是解决交互问题的最佳方式之一，可以通过可用性测试来检查和分析交互问题。日志文件提供了解人们行为的直接方式，可以利用日志文件来发现哪里有未预期到的特定行为。例如，有一项任务是提交信息给脚本，如果所提交的信息不符合要求的格式，脚本就会产生报错页面。如果有一大部分用户都看到了报错页面，也就在暗示没有清晰定义任务，没有很好展现选项，要么就是用户误解了过程。日志文件本身并不能发现用户与产品交互中这三种情况的主导程度，或者需要复杂分析才能提取出关系。相反，只关注单个特性的可用性测试或许能迅速发现广泛/交互问题。

同样，也可以根据可用性测试中收集到的信息来检查日志文件。对人们的行为做出假设可以观察他们使用界面的情况，还可以通过分析日志文件来看看他们在实际生活中有没有遵循那种行为方式来进行测试。

两种方法都有助于了解人们和界面的低层交互。可用性测试的客观性较弱，更偏重于探索性方法，而日志分析具有客观、描述特征，但对能力、理解或者动机提供的信息很少。两种方法可以互补。

## 任务分析和可用性测试

任务分析是分解任务如何完成的过程。可用性测试揭示人们实际如何执行任务的情况。因此，任务分析能作为通过可用性测试来验证的理想模型。同样，可用性测试的方式和目标可以基于任务分析所定义的模型。

例如，任务分析发现在汽车修理店铺选择中，可能的店铺取决于动机（“我们一直都在那里取车”）、位置（“它在办公室附近”）、权威（“我总是去福特经销商那里”）以及推荐（“保险公司推荐的”）。选择几乎很少：人们也许会有一个或者两个候选店铺，他们选择的一般都是首先考虑的那个。而且，人们一旦选择了一家店，就会固定在那家店，即使那家店有问题。索赔管理网站中，这种实施方法会把这四个选项以修理厂选择模块显示出来，但即使这些是人们选择修理厂的标准，人们也不会采用这些术语



来考虑。可用性测试发现把这些选项展现出来会混淆参加者。他们把系统理解成比他们更权威的东西，因此他们更喜欢总是让系统来推荐一家店。

把任务分解成各部分时，很难预测到这些效果，但它们对于产品功能是关键。这些效果可以通过大声思考法的原型可用性测试中进行调查，而且现有产品能迅速瞄准影响人们完成任务能力的最主要事项，并能验证一般意义上任务分析的准确性。

最终，每种方法都只是起点。它们都是观察并解释许多相同现象的不同方式。通过使用这些方法并进行调整，就能更好了解研究目标 and 需求，以及这些方法的优缺点。如果知道要找什么，知道这些方法如何发挥作用，就能调整研究方法，就能准确得到需要的信息。

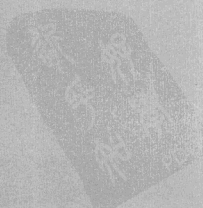








## 第 III 部分 交流研究成果





## 第 17 章 报告和演示

和任何产品一样，要想成功，研究结果必须对受众有用。优秀的研究报告，不应该只是一份产品问题清单；报告要能让开发团队对产品做出变化，并避免将来发生类似错误。结构化结果的展现形式——“实际上就是把输出物程式化”——这是让研究派上用场的最重要步骤，因为如果研究结果被误解或者忽略，最好的研究也毫无用处。

受众的需求决定着展现特定研究成果的最佳方式，它可以是一封很简单的邮件，也可以是 100 页的论文，可以是一个电话，也可以是长达一天的视频演示，可以是缺陷(bug)数据库。幸运的是，无论最终产品有多复杂，过程都一样。无论是歌舞片，还是简短对话，编写报告的步骤都非常相似，不同的只是每步需要花费的时间。

书面报告和现场演示是报告研究结果最常用的两种方法，因此本章主要关注这两种方法。另外，本章的大部分建议并不只是适用于用户体验，它几乎适用于任何面向商业的演示。

### 准备工作

准备报告不同于分析数据，刚开始报告时，应该确定报告内容(和第 11 章所描述的调查一样)。报告是信息的展现形式，知识的架构化体验，而且取决于已完成的内容。因此，在写报告和演示之前，首先要了解受众和数据分析过程。

#### 了解受众

受众的背景资料有助于定义采用何种措辞，内容应该详细到何种程





度，应该采用何种顺序展示结论。对于受众的背景资料，虽然不必采用建立背景资料的正式过程，但针对受众建立若干资料有助于报告的编写。以下是可以提出的若干问题。

- **谁会拿到报告？** 报告的受众构成会改变报告的重点。和质量保证(QA)团队相比，首席执行官当然更需要较短、更高层的报告。大量非技术人员需要的信息不同于程序员群体。某些情况下，甚至要考虑受众中特定个人的需求。如果产品副总裁，想要产品快速解决清单，她会期望看到报告。但如果可用性测试表明产品存在系统问题，而不是一系列小修小补，报告就需要满足两者的需求，例如，每个快速解决方案都能作为更大系统性问题的表象，这样，副总裁的偏好和更大问题都能得到体现。
- **他们的目标是什么？** 他们如何使用信息？有时候只需要对现有问题的解释及一组开发目标。其他情况下，受众需要判断来决定某些变化。还有一些情况下，他们需要从外部角度来评判他们对用户的了解程度。
- **他们知道什么？** 受众的技术水平在什么程度？这些情况不仅适用于考虑中的技术，当然这也应该纳入考虑，也同样适用于当前正在采用的技术。如果他们从未参与过用户体验研究，他们就需要对方法背后的原因有更多解释。如果演示供市场研究使用，就不太需要说明焦点小组和访谈之间的区别，但如果是为工程师或者设计师准备的，就需要解释。
- **他们需要知道什么？** 他们为什么会拿到报告？报告怎么简化他们的工作？研究项目有一系列目标，他们都是基于需求清单的，满足这些目标是报告的主要目的，但还需要传达其他信息。在研究过程中，项目的重点可能已发生转移，或者可能揭示了次要信息。这些情况都需要传达。例如，在为一家儿童教育网站进行情境调查时，父母的技术能力水平并不比儿童的技术水平更强，这一情况变得很明显。这种看法在原始研究目标之外(原始目标只是要研究使用模式和儿童的需求)，但知道的话会非常有价值。



- **他们期待什么？** 预测受众对报告中信息的反应有助于构建有效的演示。报告应该根据对受众期望内容的了解而写成。如果报告结论和开发成员相信的东西完全相反，报告就需要包含更彻底的解释，并对结论进行评判。同样，如果很多发现都符合开发成员对问题的理解，报告就可以更加强调不同于常规理解的地方。

如果对受众的限制非常严格——例如，必须向非技术执行官和核心开发程序员传递信息——可能需要对受众进行分类。以不同形式展示相同信息的不同方面，会受到每个参与者的喜欢；受众不会觉得被遗漏或者庇护，演讲人也无须回答太多问题。

## 了解过程

再仔细的研究也会受到限制。考虑如何展现研究结果时，要记住这些限制，认识到过程中存在的限制有助于受众理解结果、相信结果，并根据建议采取行动。对过程提出的疑问和对任何研究提出的疑问一样。

- **有哪些数据收集问题？** 要认识到数据收集方法中存在的限制和潜在失真。要在报告中提到实际数据收集问题（“我们只能和两个受众中的一个进行交谈”），但在准备工作期间要考虑方法中存在的潜在问题（“我们选择关注一个焦点小组，因为我们认为在有限资源的前提下，他们能给我们提供最佳反馈”）。如果有大量受众指出潜在缺陷，则需要预计并解决这个缺陷。
- **有哪些分析的限制？** 分析数据可以有多种方法，但分析过程中只采用一小部分。可以观看并记笔记来分析用户测试录像带，也可以一帧一帧播放来分析每个说话、敲击键盘和鼠标点击过程，我们要了解并认识到所选分析方法的优缺点。例如，汇总调查结果很快、很容易，或许并不能揭示多变量交叉汇总的结果，但这个程度需要更多技巧。同样，焦点小组能表明人们想要什么，但无法明确反映人们是否需要某个产品，也不能明确反映人们是否能使用这个产品。





了解研究过程中各个方面存在的偏见能在解释报告时发挥作用。招募过程、参加者的不同观点、研究条件以及分析人员的经验——每种情况都会影响信息收集过程。做好准备迎接这些偏见，了解这些偏见，有助于尽量减少偏见，揭示真实情况。

## 报告

书面报告是大多数用户体验研究的主要输出物。报告的结构是演讲得以顺畅进行的基础，而其他材料，如视频，也是必要的。

### 挑选格式并精心组织

写报告之前，要和受众讨论报告格式。如果时间紧张，往往可以用电子邮件发送报告。其他情况下，比如要向执行董事会演示，往往要求有插图和漂亮封面的更正式的书面报告。还有些情况下，最好能采用 HTML 格式的 report，观察能直接链接到问题区域。给预期的报告接收者看一份报告样例，看看报告格式是否符合其需求。

确定报告一般格式后，接下来要用最有效的方式向受众展现发现成果。次要发现结果可以和主要结果混合在一起，来支持“大”想法。要根据三种优先级别来确定结果的优先级别。“知道了很好”的信息只包含在报告的完整版本中，“应该知道”的信息要包括在一般报告中，而如果时间是关键，“必须知道”的信息必须要包含在电子邮件中。确定了优先级别，手稿中的引用信息就可以支持详尽阐述发现的成果。

#### 报纸风格

无论采用什么格式，所有报告都要采用类似报纸报道的结构。要假设有些人只有空阅读开头几段，有些人会阅读一两页，有些人会浏览所有内容，而有些人会仔细阅读每个字。报告内容要满足这些受众的所有需要。

经典报纸风格中，第一句话说出最重要事实(例如，“今日消费者信



心信号软弱，而股票市场大涨)。第一段说出基本事实，接下来的几段对第一段提到的若干要素进行详细阐述，报道的剩下部分提供背景，并采用报道的历史总结作为结尾。

因此，写报告时，永远不要“把最好的留到最后”。要把最不重要的放到最后。

## 报告样本：在线贺卡公司的可用性测试

### 执行摘要

体验过 Six Webcard 的用户受邀查看 Webcard 现有界面和 Gift Bucks 界面原型。人们一般都觉得 Webcard 导航容易使用，而且包含很多想要的特性。Gift Bucks 界面能被很快看懂，大多数参加者都觉得其特性让人兴奋。网站存在的大多数问题都和功能的侧重点及信息的组织有关系，而和导航或者功能没有关系。

本例中，最重要的观察结果是网站交互基本良好，人们能使用网站，并觉得网站有趣。后面就是一段简短而清晰的问题总结。如果有人只阅读这一段，就能明白最重要的发现。因为项目的目标是发现交互存在的缺陷，而不是要发现人们喜欢的理由，所以详细阐述了产品问题。

- 人们并不关心是否能找到特定卡片。他们关心的是分类，并且愿意改变想法接受已有的东西。即使没找到特定卡片，他们仍然有良好的浏览体验，这是好事。参加者都由于分类组织形式没有找到特定卡片，但他们在分类之间浏览几乎没有碰到麻烦。特定卡片位置进一步受到了搜索界面的阻碍，搜索界面很难理解，也不是像参加者预期的全文搜索。
- 人们忽视了首页上的若干要素，因为他们的视觉没有被吸引。
- 大多数人不明白 My Webcard 是什么，也不明白对他们有什么好处。单独告诉他们有何好处之后，他们也只表示有兴趣。
- 人们很快看懂了 Gift Bucks 界面，而且觉得很简单，但有些人想





进一步知道此过程。他们往往看不到指向更多信息的链接。

另外一些人想要的特性包括能通过 Microsoft Outlook 同步 Webcard 通信簿,能输入模糊 Gift Buck 数量,以及组织卡片的更多方式。

执行摘要能大致说明总体情况,下一节设置期望,并提供必要的背景来理解相关观察。如果受众精通相关方法,可能简单用几句话描述一下即可。本报告的受众并不了解方法,因此需要更详细的解释。

## 流 程

我们邀请了六位有电子贺卡体验的人来评估 Webcard 界面,并对 Gift Bucks 原型发表评论。我们根据他们最近使用网站的情况(必须上个月使用过网站)、在线购物行动(必须前两个月购买过礼物),以及在 1 月 20 号和 1 月 21 号来 Webcard 办公室的情况,从 Webcard 用户名单中挑选出这六位用户。

每次访谈时间为 90 分钟,从询问评估者的网络使用、网上购物体验以及网上祝福服务情况开始。主持人给他们看了当前的 Webcard 网站,并在看的过程中询问他们的直接印象。几分钟之后,主持人请他们给朋友搜索一张复活节贺卡。搜索几分钟之后,主持人请他们返回到主页面,完整看看界面,讨论首页上的每个要素以及大部分分类和个性化页面要素。他们的下一个任务是找到一张有旧金山照片的卡片。几分钟之后,主持人给他们看了一个带 Gift Bucks 的卡片个性化界面原型,并请他们讨论一下新界面中的要素。

主持人帮助评估者总结对产品的看法,对额外的特性进行头脑风暴,并结束访谈。

在整个过程中,主持人请他们大声说出想法,并偶尔请他们详细解释特定动作或者评论。另外,在适当的时候,请他们讨论特性以及额外特性。

所有访谈都录制成了视频。录像带来仔细分析参加者的行为、信念和陈述,部分手稿包括有趣的或有揭示意义的用户语录。然后组织和分组观察结果。这样能提供更一般的材料。

因为具体情况会随着项目发生变化,所以这种解释可以在任何情况下





都有用，但如果受众需要了解如何得到结果，这种解释是最有用的。

接下来，说明过程中的缺陷。

### 评估者背景资料

描述参加者不是关键，但对未观察的人可以提高研究的真实效果，并提供解释这些用户陈述的上下文。每个人都知道访谈的是真实的人，但访谈细节能真正说明参与者的看法、差异、以及(最重要的)他们存在的真实性。至少要一张表来总结参加者是否符合招募标准，或者要有更广泛的描述。

#### 提示：

招募时间短，而且要花一天上班时间去旧金山市区的办公室，因此，评估用户群倾向于在中心商业区工作和电子商务相关工作职位。

#### 忠告：

为了保护参与者的隐私和秘密，在报告中揭示任何个人信息时(特别是姓名)，要绝对慎重。如果从一般大众招募到的人不会再参加产品研究，通常可以在报告中使用他们的名字。但报告中涉及公司内部的人时(或者公司内部的人的朋友和家人)，去除任何可识别的信息通常最安全。可以把参加者重命名为“U1”，“U2”，等等。

#### 利 娅

在一家电子商务艺术网站的营销部门工作。每周上网超过 10 小时，其中一半时间用于个人冲浪。上网大部分时间花在投资、购物或者新闻事件上。每月会有一次上网购物，节日期间购买较频繁。买过酒、书、CD 和“健康及美容用品”，其中一半都属于礼品(特别是酒)。因为即时便捷而喜欢电子贺卡，电子贺卡很适用于偶然活动，可以同时发给很多人。曾经觉得发送电子贺卡会有障碍，但最近收到一张，从而改变了想法。会根据接收人、而不是感情来发送电子贺卡：有些人不上网或者(她觉得)对电子贺卡不会有正面反应。只使用过 Webcard，每月至少发送一张。

等等。

同样，如果创建的录像带包括观众可能认识的人，也要当心。如果有人会被认出来，一般就不要显示出来，而应该引用他们的评论。建立情境后，接下来就要尽量直接地充分展示结果，避免使用术语。

报告前两条记录叙述了深层次影响产品、跨越大量特性的用户行为。如果有数据支持，可以归纳内在原因，但要避免过分归纳或者缺乏支持的结论。





### 提示:

以下观察均基于访谈录像带回顾。根据主题或者重要性来组织观察结果。不能根据条目长度来判断严重性,有些情况本来就比其他情况需要更多句子来解释。

### 提示:

常见做法是给观察结果加上严重程度级别(“1代表精彩表演,2代表显著影响用户体验,等等”),但我更喜欢让受众自己来决定观察结果的重要程度。用户体验中出现的“光点”也许会对产品及公司产生深远影响。人们划过页面顶部的横幅广告在产品体验中是个小事件,但如果广告代表了产品整个收入流,而且点击量提高0.5%代表了收入提高20%,就不是小事件了。可能只有熟悉公司业务的人会明白观察用户这种行为的重要性,并推测外部咨询师也会同样知道这点。

## 观察

### 一般情况

(1) 人们看起来不会寻找特定卡片,也不会对特定卡片有兴趣。他们通过分类来浏览卡片,要么找到能符合他们要求的卡片,要么改变策略去找其他卡片。他们的标准很松散,而且常常不会受到所找卡片的限制。如果初始想法失败了,评估者就会开始在特定目录中浏览所有卡片,看看有什么能吸引他们的。如果没有卡片能满足他们的基本想法,他们就会调整期望,直到找到卡片。如果发现一些不太合适的卡片,人们在一定程度上也可以接受,只要这些卡片足够“接近”就行。

“如果想发送祝福,我觉得‘如果是一张旧金山贺卡不是很酷吗?’,但如果没找到这种卡片,我就会有其他想法。”

“我通常都不看第一页。我不想花很多时间找来找去。找到我喜欢的,我就点击,发送,就可以了。”

“如果没找到想要的,我就会改变想法,发送其他卡片给她。”

(2) 参加者寻找卡片的时间不超过10分钟,平均每次搜索5分钟。在此期间,大多数人都不断改变策略,根据现有卡片来调整自己的期望。

可以适当描述问题的严重程度,但不要把严重程度当成了优先级别。有些问题虽然非常严重,但如果出现不频繁,在用户体验中只会起到很小作用,描述问题时要考虑这些情况。

(3) 没有人在一分钟之内找到搜索页,到一分钟的时候,大多数人都放弃了寻找。我们问他们有没有看到搜索功能,有几个人说没有。再次追问他们,每个人都扫视整个界面来找,检查每个要素,直到找到为止。另外,参加者都有广泛使用过Webcard,但没人看起来有经验。但正如第1条中描述的,没有参加者关心找到特定卡片,而且所有人都都很高兴能浏览目录。

“如果有搜索页,我没看到它。”

“如果有搜索框,我会用搜索功能来找到卡片。”



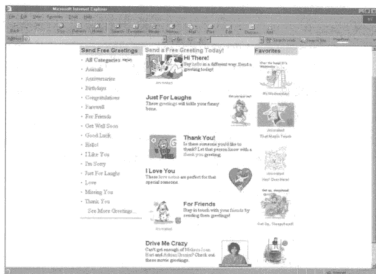


图 17.1 Webcard 首页

对于以下观察，可以推荐一些明显的解决方案。但决定是否要对这些问题提出建议解决方案取决于受众构成和分析人员的专业性。本报告的受众是产品交互设计师和产品人员，他们比我更了解产品领域和设计限制。他们负责创建合适的解决方案，因此我选择完整描述问题，并在适当的时候提供建议，但让设计人员自己创建合适的解决方案。

(4) 有三位评估者说有些小字体的字太小了。具体来说，目录页中扩展子目录中的问题以及 Gift Bucks 页面的链接信息太小了(而且在 Gift Bucks 页面上的链接常常都被忽略了)。

(5) 如果请人们“去网站首页”，有几个人找不到页面顶部的 Home 链接，他们更喜欢使用浏览器的后退按钮。

(6) 如果请他们“去网站首页”，但不能使用后退按钮，人们更常见的做法则是点击 Home 链接，而不是点击标识去首页。

有时候要注意期望发生的事件并没有发生，也要注意发生了非预期的行为。





(7) 首页左侧的大部分内容都被忽略了。看了左上角的特性之后,大多数评估者的注意力都马上转移到了主菜单,造成左侧的特性没有看到,即使初次扫描页面时看到了也一样。人们发现了页面下半部分的特性,也发现了靠近底部的东西——“Make your own Valentine”, “Sound Cards”和 “What's new”。

“这可能是最有趣的东西,但我没看到。” [指着左侧底部的链接]

“这个想法很好,但我没看到!” [指着 Make your own Valentine]

确定了所有观察结果,要通过结论和有用的背景信息来完善报告。结论可以讨论研究产生的更大问题,并建议广泛解决方案。要能帮助受众脱离细节,并从整体上了解产品和研究过程。

### 结 论

受众认为 Webcard 服务有用,而且用得都不错。人们不会看中任何特定贺卡,虽然改善了无法找到特定卡片的情况,但并没有消除搜索功能和信息结构存在的问题。从人们在执行寻找以旧金山为主题的贺卡任务中的表现来看,由于信息结构不符合人们的期望,网站所有板块都无法访问。我们建议通过卡片分类和观察方法来研究用户期望和自然组织机制,并采用全文搜索界面以及仔细选择的关键词来改善搜索过程。

Bonus Bucks 容易使用,但实际功能得到的反应却很冷淡。我们推测是因为提供的面额太大了。对用户而言,互联网贺卡是一种传递情感的快速免费方式。25 或者 50 美元的礼物没有被看成是免费的东西,大多数用户都不会使用。实际上,更大的面额会导致人们完全离开。我们建议把 Bonus Bucks 面额减少到 10 美元。

等等。

要包含额外信息来支持主要论点。为了对读者有用,很容易把收集到的种种信息都写入报告,但避免用信息淹没受众。



## 有趣语录

(可索取全部文稿)

**利娅**

“我曾经说不会[用 Webcard]发送祝福信息，但最近有人给我发了一张，祝福非常漂亮，所以我愿意尝试一下。”

“不是基于事件，而是基于接收人。有些人对电子交流保持开放心态，而有些人不是。”

“在网上买 CD 我已经有地方了。不需要另外再找地方了。如果我要在 Webcard 上买礼品，我希望能买到在网上其他地方买不到的东西。”

等等。

写好报告，接下来要进行测试。和任何产品一样，需要检查报告看看是否能满足受众的需求。也许开发人员需要明确标记出问题区域的屏幕截图。也许他们想用开发部门或者用户市场来划分问题。报告完成前，试探受众可以验证报告对受众是否有用。

报告受众中的有几个人要能看到 Beta 版报告。他们花一两两天时间看完报告后，对他们进行访谈能解释报告在多大程度上解决了受众需求。报告涵盖他们期望的主题吗？报告有没有对结果合理安排优先级别？他们预计会有其他同事对报告有问题吗？

## 演示

结果展现的方式几乎和总结结果的报告一样重要，甚至更重要。仅根据描述常常难以理解产品用户体验的复杂性，而且功能细节和精心准备的演示通常比文字描述解释得更清楚。

但演示不只是阅读报告。创建并传递有效的演示是一门艺术，而本书几乎无法进行评判。从高中的演讲课到激发积极性的执行官演说，世界上有很多演示建议，有几点特别适用于用户研究结果的演示。

- **请受众做好准备。** 讨论具体情况之前，要请受众进入一种心态。





这种心态有助于理解要看到的信息。除了给出流程的快速介绍，我通常还会提供下列发言：“本次用户研究的结果不会让产品没有问题，也不会让产品讨好所有人。报告内容是我们和人们交谈、观看人们使用产品时得出的观察结果。通过解决这些问题，你可以做出变化，我们认为这种变化会让产品更符合受众的需求和能力。但这种变化不会改变产品的根本性质，也不会改变产品的商业模式。如果产品不是人们想要的东西，无论产品多么讨人喜欢或多么好用，人们都不需要。”

- **仔细挑选要点。**仔细选择要重点说明的观察结果。如果演示要一小时，有一两个主题，5~10个要点，还有5个左右次要点，那对大多数受众而言意味着很多信息，基本没有时间跑题或者提问。也就是说，涵盖“必须知道”的主题一般就可以了。
- **采用实例。**能实际看到问题比描述问题更有效。受众越能接近看到准确的行为或者态度越好。几组人讨论同一件事情或者不断犯同样错误的录像带比描述更有效。如果没有办法，快速演示使用实际产品的问题也非常有效。如果还是不可能，也可以采用实际参加者的名字和语录。例如，“吉姆花了三分钟看首页，然后就放弃了，他说肯定没有搜索界面。把搜索指给他看后，虽然主持人说这不是他的错，但他又花了一分钟为自己的‘愚蠢’表示道歉。”但任何描述或者示范都不要超过一两分钟。
- **强调用户的角度。**无论什么时候开始准备解决方案，或者开始理解问题的性质，人们就会自然从自己的角度出发。但和用户的角度几乎总是不相同，会导致开发团队产生大量被歪曲的结论。展示结果时，要想重点强调用户角度，一定要有主题线贯穿演示，暗示用户想法和开发团队想法之间的差异。
- **谨慎使用术语，并且只能在给出定义之后才使用。**如果人们听到很多不能理解的词汇，他们会觉得失去上下文联系，或者觉得受到轻视(或者两种感觉都有)。但有时候，受众应该要知道技术概念术语，因为对于理解报告很重要。用一句简短的话先定义信



息架构，“以对特定受众有意义的方式分类和组织信息的原则”，然后用起来会更顺畅。另外，要了解受众的术语，并在报告中突出受众的定义和报告中采用的定义之间的差异。例如，有一组电台编辑认为单词“headline”仅仅指头条新闻，而设计师则用来指代文字的展示风格。在发现术语差异之前，对体验研究的隐藏含义还有很多混淆的地方。

- **小心使用数字。**即使已经说明了这些数字并不客观和绝对，但人类还是容易把数字看成客观绝对的东西。数字、柱形图和饼图在演示时很好看，但除非有确凿数据支持，否则应避免采用。否则，受众就会把这些数字当成精确度量，而不是用来演示想法。
- **留出三分之一的时间用于提问。**如果演示结束得早，也没有提问，不会有人抱怨，但如果受众想提问，又没有时间来讨论，他们就会变得沮丧。预测受众要提出的问题，制造“罐装”(canned)回答会非常有用。
- **一定要反复练习。**至少要从头到尾自己演示一次，还要至少在其他人的面前演示一次。

## 展示给特定群体

展示信息时，针对不同受众采用不同方法。每种受众的准确需求都会不同，但有些特定群体有一些相同的需求。

### 工程师

软件工程师的传统角色是作为问题解决者。“瀑布式”开发过程就是完全朝着这个方向，否则开发无从谈起。如果工程师花时间调查问题的原因，会被认为比花时间改善问题更浪费时间。大多数情况下，问题都被假定在特定阶段就已经理解，到工程师那里时不需要进一步研究。这种错误看法所产生的工程师文化几乎完全关注解决方案，而不关注问题。

向工程师演示时，务必预期他们关注解决方案的角度。指导该群体理解 and 解释结果。在刚开始演示时，我通常会说：“这不是一份要解决和被





遗忘的清单，而是指内在的症状。有时候可以通过治疗这些症状来解决问题，但常常需要理解根本，才能有效解决问题。”这样能让工程师考虑产品问题的较大情境，而不是在报告页边的空白处写代码。

但“较大情境”会被严重误解，让人觉得认为做出任何变化都没有意义，除非从头改造产品(而在大多数情况下这都不可能)。解释广泛情境问题和立即改善用户体验的短期变化之间务必平衡。

工程是科学原则，工程师常常比社会科学家更迷恋自然科学的程序和哲学。因此，他们常常怀疑很多用户体验中的非统计学方法。解释这些程序及其有效性，能有效预测并清除工程师的很多异议。同样，要关注事实，不要引入对观点的假设，因为具体事实会给受众带来更多影响。

要慎重展现已经经过讨论并且被工程师抛弃的建议，除非这些建议是根据完全不同的方式检查问题而产生的。例如，界面开发团队放弃了采用 Flash 创建部分界面的方案，因为他们觉得虽然能提供好用的开发环境，体验解决方案更漂亮，但用户的计算机和网络的连接速度会很慢。对网站受众的调查表明，绝大部分的主要用户使用高速网络连接和新款计算机。他们会用 Flash，但工程师团队已经抛弃了 Flash 解决方案，如果再次引入，则需要额外解释为什么要用 Flash。

工程师还容易针对所有可能的情况来解决问题，而这包括所有可能的用户和使用情况。“如果用户使用只支持文字的浏览器，怎么办？”这种问题是这种角度下最典型的例子，当然他们是对的，而且有意去除一部分用户也不好，但想满足所有可能受众所引发的制约通常比满足那些受众带来的奖励需要更多资源。告诉工程师用户背景资料的隐含意义以及用户研究提出的只是需求优先级别，并不是完整清单，能帮助他们将其接受为“所有可能用户”的一部分。

### 视觉设计师

一给他们看任务分析，设计师就会睡着，所以我会不惜一切代价让他们保持清醒。

——维多利·亚贝洛蒂，施乐帕洛阿尔托研究中心资深科学家(个人电子邮件)





提示:

如果把 e. e. 肯明恩和威廉·S·巴罗斯当成反例证,就要指出两人艺术的目标不同于大多数网站的设计目的。在这个文学类比中,大多数网站更像 Sharper Image 目录,而不是小说《裸体午餐》(Naked Lunch)。

即使视觉设计师看起来和工程师脾气相反,但两者的看法也有重合的时候。传统意义上,视觉设计师的角色是充作问题解决者。与工程师相比,他们要针对问题的不同级别创造解决方案,但仍然都是基于详细说明,他们会在报告页的空白处画草图,但他们会画出形象和界面要素,而不是对象层级。

除了面向解决方案外,许多视觉设计师都是从形象设计(常常来自印刷媒体)转变到视觉设计的。形象设计师面对的问题通常都涉及创造强烈的第一印象,并传递单个信息,而不是软件设计常常要求的系列相关信息。因为后者无法表达前者的优秀设计,因此常常会产生摩擦。为设计师受众做好准备,首先需要解释并强调产品形象和功能之间的差异。我有时候采用写作类比方法:写作风格和语法准确性并不抵触。语法是文档的功能,而风格是文档的形象。对文档进行拼写检查不会影响意义,而用户研究就像拼写检查一样。

确定了差异,设计师常常需要确认其在交互领域中的问题解决者的角色。对设计师(或者任何其他面向解决方案的职业)而言,最痛苦的事情之一就是通过委员会工作。如果由一组人来决定,解决方案常常是最不常见的共同特性,最不漂亮的选择。用户体验研究不会取代——也不应该暗示会取代——设计师解决问题的任何创造性或者视觉权威。研究并不是创造解决方案的方法,而是确定问题优先级别的方法。我有时候会说:“问题你可能都见过。这个部分要做的事情是有助于确定问题优先级别,然后再开始解决问题。即使 10 个用户中有 9 个说他们讨厌绿色的背景颜色,他们喜欢褐色色调,也并不意味你需要把界面改成棕色。关键是用户会注意颜色,而不是他们对颜色的选择。”

一般而言,设计师不会像关心表面有效性(感觉正确)和一致性一样关心研究的科学有效性。因此,观察结果和建议要能立即发挥作用,并适合网站设计所基于的一般模型。只要有可能,要采用视觉效果和交互来支持你的观察结果。理想情况下,设计师要能在实际生活中实际观察到一些研究,视频和现场演示和屏幕截图一样有效。用户体验研究师和设计师杰夫·维恩喜欢在演示而不是在讲座中创建工作环境。正如杰夫所说:“坐





在会议室写代码很难，但可以坐在会议室推动事情前进。”可以检查问题，并快速画出大致功能。可以根据研究结果快速迭代，给设计师提供具体情况进行处理，同时展现研究结果。

### 营销

营销实务关注了解愿望。营销涉及通过让受众获得或者使用产品或服务来满足受众愿望。

营销研究人员使用的工具常常是定性的，可比较的。调查用来创建竞争对手和竞争参数清单。用户体验研究常常关注选择背后的原因，而不是选择本身，因此，调查可以帮助营销研究人员理解搜集到的数据。营销沟通职能能找到有用信息，用于精心制作他们想要展示的产品画面。

因此，向一组营销人员展示用户研究结果时，要关注人们的选择动机。为什么人们受到产品吸引？什么东西导致人们离开？什么东西让他们回来？什么东西让他们永远离开？用户体验研究所展示的连贯一致的心智模型有助于解释用户的现有行为和希望，而同时又有助于预测未来行为，并发现阻碍人们购买产品(加入服务，或者任何对于产品短期成功重要的东西)的障碍。心智模型摘要在演示过程中很可能是最重要的关键点。

### 上层管理层

作为受众，管理层的观点会非常不同。有些经理只关心会影响公司和产品的长期重大问题。他们关心的问题具有战略性，主题只有影响他们的观点(5万英尺高度的看法)才重要。其他经理更关心产品如何工作、公司如何运营的细节。如果他们知道客户如何认知产品的细节，他们会觉得最放心。

有效演示要从参与的执行官的利益和角度出发。精心准备演示时，要考虑他们的日程表(就你所能知道的)。作为影响产品的专家源，要留意到对他们真正重要的东西。

发现对他们最重要的东西的最简单方法就是去问他们。安排时间打几个简短电话给要参加会议的关键利益相关者。向他们介绍讨论对象，并询问他们有没有特别关心的主题，或者他们有没有想知道答案的问题。经常



会有对数字指标的要求(见第 18 章), 这些数字指标用来度量成功和调查研究价值。找到哪些指标对管理层最有价值(是新商机的数量, 新注册服务的人数, 还是花费在支持工作上的资源?)能为报告和演示奠定基础。

结构化实际演示时, 不要掩盖或者简单化重要发现, 而要通过能帮助经理做出决定的方式来展现这些信息。用例子和建议来说明一般趋势的讨论常常比列出建议变化清单或者完全模糊的讨论更有效。

给提问留出大量时间。即使是针对受众兴趣的背景研究, 常常也很难预测到人们需要什么信息, 而且公司执行官的职责兴趣广泛, 因此特别预测他们关心的东西。

### 常见问题

演示中频繁发生的大量问题会导致演示者发生各种结结巴巴和逃避回答问题, 让人觉得不舒服。预计到这些情况并做好准备能减少大量现场的惊奇情况。

- **“这没有统计学意义!”** 展示焦点小组、用户测试或者情境调查结果时常常会听到这种说法。的确是这样, 这些研究方法带来的结果通常都没有统计学意义, 但这并不是研究的目的。定性研究的目标是发现可能的问题, 了解原因, 并创建可信的解释, 而不是确定广大用户人口中存在问题的准确比例。另一个办法是通过不同研究方法——比如可用性测试和日志文件分析——来显示相同行为。如果不同研究方法显示出人们的相同行为方式, 就更能确定他们的实际行为。
- **矛盾的内部事项。** 有时候公司内部互相矛盾的事项, 会通过用户疑惑揭示出来。例如, 在线销售团队和经销商支持团队之间会有摩擦。前者努力提高公司的直接销售额, 而后者则努力安抚经销商对直销感到的威胁。从网站上看, 就是设计师通过创建分支界面来缓解这种紧张, 但这样会给用户带来混乱。这种情况下, 介绍用户的混乱情况会引发内部紧张的关系, 而研究结果就会成为催化剂。如果你不是为这种公司工作, 最好不要参加这种讨论,





而应该作为两方的独立资源。但同样要确定两方都能理解研究的意义和局限性。维护你的发现，并说明研究的事实，但要让两方自己来确定这些事实隐含的意思。但如果你属于这种公司，维护研究的独立性就特别重要。

- **“用户很愚蠢”**。这种说法有很多表达方式：“用户不聪明”，“用户经验不够”，等等。如果是根据项目团队所接受的目标受众的背景资料为研究进行招募，用户就是目标受众成员，他们的经验要么需要被当成目标受众的经验，要么需要进一步明确目标受众。
- **“用户 X 不是我们的市场”**。类似地，如果在研究开始时明确定义了目标受众，并且招募过程也符合定义，那么用户就应该属于目标受众成员。如果受众有其他感觉，则应讨论目标市场的定义，用户如何不符合目标市场。讨论所观察到的问题是否是由于特定个人造成，或者他们在市场更大细分中是否也有用。
- **“用户 X 做了事情 Y；所以，每个人都必须做事情 Y”（常常会接着说“早就告诉你你这么做了！”）**。这种情况属于过度概括。从若干用户行为中观察到的行为暗含了需要研究的问题和需要留意的现象，但单一数据点还不能称为趋势。因此，一个用户的行为（即使是几个用户的行为）很有趣，并能指向潜在问题，但不能代替广大用户人口的体验。如果内部对某一现象存在争论，这样或者那样的单个例子都不能解决这种争论（但可用会是将来项目的好研究内容）。正如斯奈德咨询公司的负责人卡罗琳·施耐德所说：“从单个数据点能向任何方向进行扩展。”
- **“他们都讨厌绿色，所以我们需要把它全变白，像雅虎一样”**。人，特别是以解决方案为导向的人，常常会与人工观察相切，假设解决它就能解决内在问题。这种问题属于“治标不治本”，应该要避免。讨论时要远离具体解决方案，同时要强调内在问题。
- **小心隐秘问题**。这些问题通常都属于严重问题，会影响用户体验，但参加者并不会抱怨，或者这些问题对开发人员或者研究参



加者从未讨论的产品想法都非常根本。如果不怀疑产品的基本形式，会很难发现并讨论这些问题。例如，有一家网站合并了联系并向非盈利组织捐款的过程，但这家网站自身是一家为了盈利的公司。人们普遍会拒绝其盈利的性质，拒绝的阴影会笼罩每位用户检查网站特性的体验过程，即使这种形式很少公开提到。与讨论如何展现网站根本性质相比，比如受众不会拒绝它，任何针对可读性或者导航的讨论都显得苍白。

理解演示受众的希望和期望和理解软件产品用户的需求没有非常不同的地方。媒介不同，但问题一样。一般而言，克服人们的疑问和回答他们问题的最佳方式就是理解他们，并预计他们的需求。有时候，可以是期望设置的形式，但大多数情况下都是准备工作。知道受众的时间安排和问题不仅仅是良好的表演技巧，关键是把研究项目从有趣的信息转变成不可分割的工具。



## 第 18 章 打造以用户为中心的企业文化

要让产品更受欢迎、赚更多钱、更引人注目，用户研究的重要性不言而喻。但公司生产的是产品，不是用户研究小组。公司需要的不只是用户数据，还需要根据数据采取行动。只有将以用户为中心的设计及研究方法融入公司流程、工具和决策，才能真正了解用户数据。

以用户为中心的开发过程意味着要把角度从产品的制造过程转移到产品的使用过程。开发过程关注的是开发产品，而不是满足需求，这是公司和客户看待产品的主要差异，而用户研究往往能发现这种差异。产品功能和用户未被满足的需求，两者之间的差异能反映出开发过程中的缺陷，也能反映出产品自身的缺陷。如果解决方案完全基于用户研究，就会发生这种问题。如果只知道用户需要什么，往往也不足以持续开发出用户想用同时又能满足公司需求的产品。要想研究对用户体验产生持续作用，整个组织都需要了解用户，重视用户，并知道如何采取行动。

极力推广。人们通常把以用户为中心的开发过程的好处视为可望而不可及的长期目标，而且如果需要挤出一季度利润来做这个事情，就会把以用户为中心的开发过程当成“锦上添花”的模糊东西而去掉。但产品以用户为中心的用户体验并不是可以省钱的长期战略，而是产生直接价值的短期过程。良好的用户体验需要同时考虑用户和公司的需求，来建立一组需求和限制，这样才能生产出更成功、更赚钱的产品。

不幸的是，这种机制尚未建立。打造有效的以用户为中心的企业文化非常困难。让每个人都从原则上认同开发产品时“牢记用户需求”，这很简单，但要创建深刻了解用户同时又能平衡公司需求的开发过程的确相当困难。创建“开发者看法”次于“用户看法”的开发过程要求摒弃自己的

### 提示：

卡伦·多诺霍在其著作《电子商务大赢家》(Built for Use)中叙述了商业目标、特性和用户体验之间的关系，她的很多想法都与本章内容直接相关。



偏见，这很困难，而且需要认真进行，同时还要逐步改变并及时评判。

### 集成

业界普遍认为用户体验只是附属品。

——唐·诺曼，《隐形电脑》(The Invisible Computer)

刚开始的时候，把东西分离出来并重头开始貌似很容易，因为传统开发过程看起来和以用户为中心的开发过程是如此的不同。

别这么干。看起来很容易，做起来难。公司中已有开发过程通常非常强大，完全重构需要强大的破坏性努力。对公司生命周期基本层面提出质疑是一场豪赌，而豪赌往往会输得很惨。因此，一开始引入以用户为中心的思想时，要选择正确的起点。

提示：

本节的大部分内容都归功于用户体验咨询师贾尼丝·弗雷泽和莱恩·贝克尔的著作及想法。

### 了解现有过程

产品的产生会按照一定过程。有时，这个过程很正式，而且已经建立，并且被严格遵循。有时，这个过程又很特别。要了解开发流程、实际工作机制，操作，从而实际如何工作、如何操作，以便因此更好地开展工作，之后再以用户为中心的过程引入公司文化。

可以从“内部发现”着手引入新概念，即在公司内部了解产品如何开发、为什么开发，以及开发受到限制的过程。简单而言，这个过程就是要回答三个主要问题：

- 为什么要做这个产品？
- 这个产品能做什么？
- 谁关心这个产品？

对任何单个项目而言，从广义上回答这些问题有助于了解如何从整体上把变化引入开发过程中。像新闻界中存在的“紧跟财富”(Follow the money)一样，每个问题都是用来引导出其他问题的，而不是马上得到回答。

忠告：

本章中，我尝试着为公司中已有的社会知识和互动制定规范。了解公司流程，根据流程并围绕流程进行工作并不需要你变成激进的用户体验革命者。了解以用户为中心的设计价值的最佳传达方式之后，你可以自然而然地顺利改变，而且不会像在电梯门上刷口号那么乱。





### 为什么要做这个产品？

产品的根本商业情况是什么？创建或者更新一种东西的理由通常都不会很简单，这些理由包含了假设和期望。公司想做这个东西必有原因。是什么原因？如果这个东西能产生收入，到底如何才能带来收入？如果能改善公司形象，如何才能改善(现在的公司形象有什么问题)？如果它能满足人们的个人日常需要，人们希望能从它那获得什么，他们为什么需要它？

### 这个产品能做什么？

公司要提供的功能揭示了对用户需求的理解和优先级别考虑的大量情况。为什么这些特性比其他特性重要？什么东西让这些特性显得至关重要？为什么不增加更多特性？为什么不减少一些？

### 谁关心这个产品？

引入以用户为中心的概念意味着首先要确定做出产品决策的人，以及他们的职责、需求和日程。如果某个人是产品的相关利益者，产品成功就意味着他们获得成功，改变产品的开发过程就会影响到他们对产品的好感程度和忠诚程度。最想开发产品的人、推动产品开发的最佳人选是产品的资助者(Sponsor)。谁是项目资助者？其他利益相关者是谁，关心项目或者能影响项目的人是谁？

描绘利益相关者是了解如何向他们呈报的第一步，提前确认他们有助于了解应该在什么时候和谁进行讨论，以便创建策略在公司中引入以用户为中心的想法。在很多情况下，就是简单到要让首席运营官相信开发可用产品能有底线盈利(当然，说起来容易，做起来难)。还有一些情况下，需要进行草根活动，向工程师说明如果能提前而且正确知道用户的心智模型就基本不需要对产品进行修改。很多情况下，两种方式都需要，获得高层支持，同时又和产品团队及主管结成联盟。

### 内部发现

以上叙述都表明一点：如果了解产品的推动力，则需要了解公司的





当前定位、自身形象和政治结构。如果这些未曾公开，就存在于每家公司。发现这些细节就是内部发现的过程，即以系统方式并且一致地提出以下问题。

可以从早前定义的一般类别所建立的主要问题清单开始。常见问题如下：

- 整个利益相关者有哪些人？
- 如何做出决策并划出范围？
- 当前商业方向是什么？
- 如何定义关键术语和概念？

可以适当扩展这些问题。例如，以下问题可能有助于确定关键的利益相关者：

- 哪些人很关心项目？
- 哪些人应该关心项目？
- 项目要依靠哪些人？
- 哪些人会阻碍项目？哪些人可能会阻碍项目？为什么？
- 哪些人可能会帮助项目？

通过提出这些问题，能快速知道进行内部发现的时候要 and 谁进行交谈。他们可能是你的团队成员，也可能是兄弟部门的人，可能是高层执行官，可能是顽固的工程师。

知道对项目很重要的人、公司的商业方向以及项目划定范围的理由，就有了创建策略的建筑模块。从某种意义上说，企业文化是一种游戏。通过发现规则和玩家，再加上有点讽刺的超脱态度，就能玩得更好。

### 从小处着手

在有些公司，最理想的是从高层开始改革过程，有一位执行官进行宣讲，并不断推进，一直到整个公司发生变化。在其他一些公司里，比较合适在高层之外创建“臭鼬工厂”。先散播一些模糊想法，让这些想法渗透几个月或者几年，然后再引入更多方法，在有些公司这样做更合适。

很多情况下，最适合从小处开始工作。如果失败了，人们不会把小项





目当成严重挫折。如果获得成功,这些新项目就可以被推广到更大的项目。此外,这些项目能从方法、目标、用户体验研究及以用户为中心的设计的期望各方面教育每个人。

多小算小?初始目标要短,可操作,并且简单。这些目标还不能是项目成功所依赖的东西。第一个项目可以是根据已知问题对小板块进行重新设计:例如注册页。**BigStep.com**的前用户体验主管约翰·希普尔建议把用户体验研究强加到网站首页的重新设计中。首页定期进行重新设计,首页是进入公司的方式和公告牌。首页是用户体验的重要部分,但首页设计常常更多影响的是展现,而不是功能,因此不成功的设计很容易被重新设计。这样就能得到重新设计的范围和重要性,而而不包含对核心功能的重新设计。

例如,有一家新闻网站决定要重新设计首页,来改善停留时间和回访情况。设计师决定把一些用户研究集成到开发过程中,他们根据自己的经验提出假设,并进行工作。此时还有一个月就要发布新的首页,因此没有时间来研究受众的根本需求,也没时间创建受众的详细背景资料,但通过测试,设计师可以了解到提高内容密度并强调子板块导航的方案对潜在用户是否能起作用。他们招募了8位符合受众背景资料的人,编写了用户测试来创建用户背景资料,并测试了几个原型。

他们对参加者进行了访谈,发现了几个没有预料到的问题:在首页上提高数据密度说明网站有很多信息,但很多用户感到困扰,有几个人认为首页“裂开”了。此外,大多数参加者都没有使用屏幕顶部的导航条,而是选择在搜索框中输入关键词。对网站成功而言,设计团队意识到网站搜索引擎比他们之前认为的要重要得多。

同样,虽然没有研究人们对网站品牌的态度,但他们观察到大量用户非常忠诚,但这些人却不是日常用户。而在此之前,他们假设每个人每天都会访问网站。实际上,观察结果对产品设计以及公司如何定义客户停留和成功提出了根本问题。



### 准备迎接失败

开始在内部引入以用户为中心的方法后，就要做好准备，接受犯错。可能会在错误的时间对错误的产品部分进行错误研究，可能会以错误方式提出错误问题。可能有人会让你做些不合适的事情，但你还不得不做。研究结果可能对开发团队没有用处，开发团队也可能会完全忽略研究结果，管理层可能会要求提供无法得到的确凿数字。

这些都是要从小处着手的原因，也是需要哲学思想的例子。每个人都会发生糟糕研究，但每次用户研究，不管有多少缺陷，都能提供深层看法，即使只是对下次研究提供更好的结构。意识到会失败能显著减少将来的这种失败。给自己设定合理的期望，也给会收到结果的人设定合理期望，利用问题来制定进一步的计划，从访谈对象那里获得反馈，并检查研究进行情况，即使在研究进行过程中也可以。

### 尽早邀约利益相关者

启发式评估——或者称为睿智的批评(Thoughtful Criticism)——常常无法让合作伙伴相信有这样那样的问题。而实况事件——要有面包圈——在单向镜子后面的实时观察，就能让合作伙伴相信有这样那样的问题。很不幸，但的确是事实。

——大卫·亨德利，助理教授，华盛顿大学信息学院，个人电子邮件

如果客户能改变设计师对工作的初始理解，这个过程才是真正以客户为中心。

——拜尔和霍尔茨布拉特，《情境设计》，第54页

请人们观察并参与研究是向他们推销产品价值和效力的最有效方式之一。观察人们无法使用产品是一种令人难以置信的强大体验，它向观察者传达了他们的看法和用户的想法可能完全不同。

开发团队和关键利益相关者需要参与研究，但参与方式不同。

#### 利益相关者

做出产品决定的人要能直接看见研究过程的价值，让他们看到价值的





### 忠告：

要尊重开发团队。就像开发人员会把不理解“软件”的用户当成笨蛋一样。如果开发人员不理解以用户为中心的设计过程，用户体验研究员也会把他们当成“笨蛋”或者最“一无所知的人”。要积极参与开发过程，解释研究过程，并倾听建议。如果刚开始时开发团队看起来没有头绪，当他们作为合作伙伴参与到研究项目中时，你会惊奇地发现他们很快就变得非常聪明。

最佳方式是让他们首先直接观察研究。他们不必积极参加研究，但如果他们能在同一间屋子里观察人们艰难的使用情况，并听听这些情况给开发团队带来的启示，他们就更会支持将来的这种研究。

### 开发团队

直接参与开发的人需要看到研究过程和结果。对他们而言，参与研究就要直接参与开发研究目标，创建研究原型，分析研究结果。一旦参与进来，他们就会更自然地支持研究过程，并把研究集成到将来的计划中。

但每段研究都要包含每个人常常很困难，而且开发人员需要被引导到对他们最有用的研究上。例如，营销和业务开发团队的人更关心产品的广泛战略问题。这些问题常常包含在焦点小组或者情境调查中，而在可用性测试中则较少。技术文档人员和培训人员要知道人们需要哪些概念进行更好解释，以及在任务分析或者可用性测试访谈中揭示的问题，从而受益。工程和交互设计团队绝对需要参与可用性测试，但参与情境研究可能对他们就没有那么迫切(也就是说，工程和交互设计团队很可能要广泛参与研究，因为他们所处的位置最适合用自下而上的用户体验知识)。

### 展示结果

以利益相关者能看到价值的方式来展示研究结果，这和让人们参与研究几乎同样重要。以最直接有用的方式展现结果最重要，但理想研究也具有说服力。每份报告都能作为讨论以用户为中心的开发的另一个平台。

如第17章所讨论的，要根据特定受众定制展示。例如，一家媒体公司在线业务的副总裁没有时间参加任何用户研究，甚至都没有坐下来看过完整展示结果。但他是确认研究价值的重要人物。因此，研究团队请他从臭名昭著的快速午餐时间抽了十分钟。没有时间详细叙述发现结果，但有充足的时间叙述重要概念。展示关注研究过程，并给他看几个仿真叙述，显示产品中的改进情况。这样就能告诉这位副总裁他的职员在以原始方式思考客户，并向他介绍开发过程中的新概念。

即使要为开发人员编写报告，也要思考如何做好报告为改变企业文化



做好铺垫。要向小团队展示变化来鼓励快速迭代和讨论。通过每份报告，突出未被回答的问题以及将来研究的可能性，来鼓励更多研究。要讨论用户、广告商和公司形象需求之间的互动关系，来了解人们如何从三者综合的角度来考虑产品。

### 保持耐心，但要坚持

要顽强、坚定而且机灵，但也要知道何时让步，何时闯进高级副总裁的办公室。

——昌西·威尔逊，本特利学院设计和可用性测试中心主管，个人交流

软件开发文化会进行它们自己的革命。在这种环境下，很难说服人们相信值得采取的革命方法，但在组织内部引入以用户为中心的方法中，肯定会发生这种事情。仔细选择步骤能保持持续创新，但同时又不会对开发过程或者参加开发的人增加压力。

但有时候，对研究的反应会非常冷淡。报告放在经理办公桌上没人看，遭受工程师的质疑，这些工程师能发现报告缺乏科学严密性。这些情况下，继续研究显得特别重要，以表明结果真实、有价值而且一致。例如，如果和副总裁聊了 10 分钟，就可以在每次研究后定期和副总裁聊聊，每个月一次向其展示研究重点。即使他们不关心每项研究的具体情况，信息的花絮也能说明研究过程和价值。安排午餐时间进行关于以用户为中心的设计座谈会，并邀请公司之外的人来谈谈成功项目。要买些披萨，请外面的人谈谈你的团队碰到的相似情况会非常具有说服力。

### 鼓励具体用户体验角色

如果公司意识到用户体验研究和以用户为中心的设计的价值，接下来就要开始推动创建能管控流程的职位。如果有人负责研究过程，他们就即是发言人又是传播者。

按照职位高低的倒序，有以下几个职位描述样本。





### 用户体验专家

用户体验专家负责在开发团队内部评估、跟踪和设计有效的用户体验。专家会监控对需要执行研究以及向开发团队解释结果的开发过程。

要求具备可用性、营销研究、实验心理学或者真实世界的文化人类学方面的知识。最好有设计和软件开发或者项目管理经验。理想的候选人除了主持研究分析结果之外，还要能管理研究项目，并向设计、工程和技术文档受众解释研究。

### 用户体验主管

用户体验主管负责公司内的所有用户体验研究。职责包括管理所有内部研究项目，向管理层传达结果，培训初级研究人员基本方法，在所有项目之间恰当传递知识。

要求熟悉常见用户体验研究方法(可用性测试，任务分析，情境调查)，并且所有候选人至少都要有3年进行定期完整研究的经验。具体经验包括为研究和开发建立用户背景资料，招募参加者，主持研究，分析和展现结果。要求有和交互设计师和软件开发人员密切工作的经验。最好有信息架构经验。

另外，候选人要能培训他人来进行用户体验研究，并能调整研究方法来适应产品或者开发团队的需求。

### 首席体验官(CXO)

首席体验官属于高级管理职位，负责所有情况下所有级别用户的完整用户体验。该职位包含任何会影响人们与公司交互方式的设计、有效性以及客户支持工作。要求熟悉用户体验以及有效研究的所有方面。另外，无论是从视觉创意方面、交互角度、还有作为客户关系管理系统的一部分的经验，首席体验官要有设计最终用户交互的广泛经验。

为了保证尽最大可能提供最佳体验给公司产品用户，首席体验官要



负责引入并维护“在开发过程所有关键点能够集成用户输入和用户研究”的文化。最后，首席体验官还要负责平衡用户需求和公司需求，既要向用户，也要向利益相关者和公司董事会报告首要、最重要的东西。

## 理由

如果没有理由，就没有集成的动力。提高用户体验质量是长期投资。在受到财季和不稳定的价值驱动的世界中，很容易放弃长期战略，而且如果必须要在财务上进行妥协，用户研究极少会被认为与当前危机相关。用户研究会被放下，和其他仍善努力放在一起，如保护雨林，收养流浪猫。

以用户为中心的开发需要定位成一种能产生直接财务和公司意义的解决方案。它需要显示出它是“必须要有东西”，而不是“锦上添花的东西”。如何才能显示出来需要有能证明它应该存在的证据和方法。

## 以用户为中心的过程的原因

如果把以用户为中心的方法引入到开发过程，常常会有争论和辩论。这些方法通常都是新概念和新的工作方式，因此会遭遇抵抗，此时需要澄清它们的优点。

虽然实际重点会根据情况变化，但有几个基本争论过去已取得成功。

- **有效性。**人们实际想要的产品并不需要重新制造。不需要改变围绕人们的工作方式而产生的产品。如果有了人们如何使用产品的模型，就会少一些异议，少一些模糊，少一些开发延迟，而且你也能知道在哪里增加功能(在哪里不增加功能)。总而言之，这个过程能更有效利用公司资源：它能在发布之前创建更清楚的路线图，而在发布后又能减少服务器和支持职员承担的压力。
- **名声。**有正面用户体验的用户会更有可能继续使用你的产品，并通过口碑相传告诉其他人使用。符合人们需求、愿望和能力的产





品创造的满足感会超越简单的满足，而且这种满足感能获得超越功能的氛围。人们会把自己和产品从情感上关联起来，创造出一种比仅仅建立在理性功能平衡上更强烈的联系。

- **竞争性优势。**用户模型越详细，产品就越容易满足他们的需求。这样就有可能确定竞争对手产品没有满足的需求，并驱动创新，这种创新并非基于技术能力(这些技术能力可能有，也可能没有实际应用)，而基于真实需求(真实需求当然能发挥作用)。不是仅仅对用户行为做出反应，而是能参与并且驱动用户行为。
- **信任。**如果产品能根据人们的期望和能力进行活动，人们就会更信任产品。而信任感反过来会导致忠诚、满意和耐心。
- **利润。**最终，如果产品制造成本更低，维护费用更低，能吸引更多客户，能给商业伙伴提供更好的价值，产品就能赚更多钱。详情可参见后面的“可用性定价”。

这些一般改进反过来意味着直接影响开发的具体短期优点。

- 减少预发布重新设计的次数。
- 减少部门之间、开发团队个体成员之间的争论。
- 更可能与进度保持一致，因为减少了意外之外的事情，能更准确预测开发时间。
- 更容易从整体上和开发人员及公司分享产品理念。
- 更容易向最终用户传达产品优点。
- 减轻客户服务负担，因为必须回答的问题减少了。
- 更容易预测设备负担，因为能估计到特性的受欢迎程度。
- 减少质量保证需求，因为典型使用模式决定了应该关注哪里。

所有这些情况综合在一起，就能节省公司费用，扩展公司品牌价值，并对市场情况作出更快反应。这些情况都是保证创建用户研究项目的好理由，并能对公司首席执行官“客户至上”的宣言和以用户为中心的开发的严密广泛计划提供中间步骤。

### 忠告：

用户体验的很多重要方面都不大容易度量。根据很多定量指标来衡量良好用户体验设计的效果很容易，但把数字加到设计上常常会产生误解，因为看起来有客观(数字)度量，却基本没有应该度量的现象。例如，满意和信任只能从调查看出来，但调查数字常常是表现为精确的“科学”度量。日志文件中可度量的回访只是忠诚度的一部分，但它们常常被用来代替忠诚度。





### 度量有效性

最终，如果争论能赢得赞许，这些方法就能开始集成到开发过程中。理想情况下，结果很明显，并且所有产品和步骤都能得到改善，没有必要说服每个人都相信这些想法的有效性。否则，演示这些步骤的有效性就很有必要。即使变化很明显，也总是需要一些度量方法来确定发生了多少变化。

这时就需要指标。指标抽象化了用户体验能度量和比较的方面。在推测和观点的环境下，指标能澄清灰色区域并验证假设。指标对理解问题级别、评估解决方案的有效性发挥着重要作用。

#### 选择指标

就像指向月亮的手指，不要关注手指，否则会错过所有壮阔景象。

——李小龙，《猛龙过江》

指标是可以进行度量的东西，并与正在研究的现象(希望能)相关。不要把所有精力都集中在度量和影响指标上，而忽略了指标所代表的东西。就像人们说的，地图不是领土。

指标可以从目标着手。最终目标是要评估产品所提供的用户体验质量，要目标是要了解新方法带来了多少变化。编写研究计划时，定义的相同目标要有基础。目标是产品存在的问题，定义指标是了解体验让人失望的过程的一部分。就像有些目标会影响最终用户的体验，而其他目标能解决公司关心的问题一样，某些指标在跟踪变化的有效性显得更重要，而同时其他某些指标能说服利益相关者。

例如在第 6 章，针对购物车的使用，有一个问题是“放弃购物车和完成交易的人比例是多少？”这就是指标。虽然可以度量若干情况(人们的善变程度，愿意体验产品的程度，等等)，这个指标主要用来衡量人们的困惑程度。





## 一些典型的网站指标

最终访问“详情”页(信息页)的比例。

打开购物车的人数和到达“感谢购买”页面人数的比例。

在网站上的停留时长。

搜索特性在单次会话中的使用次数。

顶层导航页载入的次数(即人们后退的次数)。

每天关键评论的数量。

每位访问者每月的访问次数。

当然,所有这些指标都必须要用来理解用户行为的原因。例如,“详情”页访问数量的增加可能代表了人们更关心材料,或者说明人们查找信息时发生了更多错误。

更系统性的方法是把当前项目的目标列出来,并根据用户体验类型进行交叉,每个交叉的地方就创建出一个指标。这样能确保能从重要角度检查主要问题。因此,第5章中提到的 Sport-i.com 例子看起来如下所示。

SPORT-i.COM 的一些指标

	效率高低	有效性	满意度
访问者到购物者的转换率	导致购买的点击轨迹的长度	访问者与购买者的比例	调查中正面评论和负面评论的比例
改进过的导航	哪些导航页被访问的频率	不同页面被访问的次数	调查中正面评论和负面评论的比例
信息的时间性	日常用户的数量	偶尔用户和日常用户的比例	调查中“时效性”评论和“非时效性”评论的比例

由此可见,有些指标能对网站情境产生很多意义,并且产生可供收集的数据,而其他数据还有待查证。不管是哪种情况,和利益相关者确定指标能迫使每个人都考虑产品目标和用户体验方面之间的关系,但并不意味着必须要完全根据表格并度量表格中的每样东西。

创建的精确指标是需要跟踪的特性的混合体,为了理解变化的有效性,同时要结合其他一些东西,这些东西对于说服公司此过程的有效性是



**提示:**

指标存在隐藏性问题,如果页面浏览量代表成功,但若干页面没有放入典型任务中,即使产品更容易使用、更让人满意,此指标实际还是可能会下降。

必须的,此外,指标不一定要基于内部收集的数据,可以采用外部度量。例如,Google 根据关键词被提到的频率来确定关键词费用,因此可以采用 Google 的定价作为品牌名称知名度的指标。

**收集和比较指标**

根据指标度量表现有两种主要方法,本书其他章节都提到了这两种方法。点击流能揭示大量当前用户的聚合行为。获得新指标的结果也就意味着采用不同方法处理日志文件。

但通过可用性测试收集到的指标是另一回事。正常情况下,以某种方式做出反应的参加者准确数量对可用性测试并不重要,可用性测试关心的是广泛趋势。但如果要回答具体比例问题(能找到新网站地图的用户比例和重新设计前的 47%相比),必须要采用不同方式来进行用户测试。

除非要访谈几百人,否则用户测试就没有统计学意义,测试结果也不能推广到整个用户人口。但测试结果能和其他测试的结果进行比较。在变化前后仔细进行两个用户测试,可以进行比较得到一些定量数据,来看看是否发生了变化。

对于任何试验设计,都要尽量减少要变化的变量。测试之间有三种情况会发生变化。

- **人口。**第二次测试中会有不同群体。他们应该是可进行比较的人,要和第一次测试的人非常吻合而且能招募到的,但他们不能一样(即使一样,他们也应该曾经看到过产品,这样他们对产品不会有相同反应)。
- **测试程序。**原型的展现方式发生变化能导致人们对原型的体验发生变化。即使措辞或者顺序有微小变化也会影响认知。
- **原型。**界面变化明显会改变人们对产品的体验。通常就是可比较的用户测试之间的变化。

如果这些因素变化超过一个,测试过程中通常就会引入更多变化,无法进行有效理解和管理,如果没有意义,还会使得收集到数据非常值得怀疑。为了最大化过程的有效性,也可以对更多人进行测试。更多人能让你





了解答复的广度和一致性。雅克布·尼尔森(Jakob Nielsen)建议每场测试最少有 20 人,而没有定量组成部分的测试只需 6~8 人,这些数字很可能是好的参考数字。

如果偶尔会枯燥,和测试管理相比较,用户测试中的度量非常简单。只要数数就可以了。用来了解界面效率的计时度量可以用视频面板上的计数器或者手动码表来完成。错误计数(返回按钮被使用了多少次)可以采用 & 来完成。

对于日志文件或者调查,和第 11 章及第 13 章提到的程序一样。需要提前定义指标,需要知道体验变化的准确日期和时间。对于大型产品,会同时发生很多变化,如果知道同时还发生了什么变化,有助于识别观察到的变化是否一样或者是由于检查的界面改变造成的。例如,如果首页进行了重新设计,并且闪亮登场,度量首页的访问量既可以是度量营销,也可以是度量首页的可用性。这种情况下,衡量已完成任务也许更能准确说明交互变化的效果。

### 可用性定价

利用指标变化来计算投资回报率(ROI)非常具有说服力,但非常困难。很多因素会同时影响产品在财务上的成功,但常常无法从其他变化的集合效果中去除用户体验变化的效果。

也就是说,如果有确凿的案例,底线财务优点对于实施和继续以用户为中心的设计过程最具有说服力。在互联网上,这种情况要比套装软件容易,发布黄金正式版(Golden Master)的时候,套装软件中会同时发生很多变化。电子商务网站最容易产生底线,这些网站有能迅速转化成收入的指标。

- 访问者到购物者的转换率直接度量有多少访问者最终进行了购买(“最终”可以代表首次访问的三个月内,也可以代表某个时间窗口)。
- 购物车大小指平均购买数量。
- 购物车放弃率度量有多少人开始了购物过程,但从未完成购买。

#### 提示:

本章提到的想法深受《可用性成本合理化》(Cost-Justifying Usability)一书的影响,该书 1994 年由 Academic Press 出版,作者为伦道夫·G·拜厄斯和德博拉·梅休。该书比本书更全面详细地叙述了对用户体验定价的问题。



乘上购物车中的物品数量，这项指标能度量失去的收入。

每项指标都能非常合理，并直接显示出网站中发生的变化使网站多赚钱或者少赚钱。

对于其他类型的服务，例如有项服务的功能是向雇员出售广告或者推销信息，就需要找到不同指标。因为拿薪水的雇员要答复客户支持电话和电子邮件，因此减少支持电话和电子邮件的数量能转换成职员的数量减少、但支持成本相对较小，减少支持通常不会显著提高收入底线。重要的是要找到度量提高收入的方法，因为产品是以用户为中心的方式开发出来的。

例如，有一家新闻网站经过了重新设计，人们更容易找到内容。平均点击流从 1.2 页上升到 1.5 页，代表页面浏览量上升了 25%，可以转换成广告收入的对应比例增长。看起来非常枯燥无味，但假设同时在进行营销活动。可用性测试团队和营销团队都声称是他们的努力造成了收入增加。为了评判用户体验并将其与营销分离开，可以估计可用性的作用并计算投资回报率。这样能创建一个方程，它可激发对广告和可用性效果的讨论，但至少讨论可以在相对平等的水平上进行。

最近，网站经过重新设计，页面浏览量和广告收入都有增加。同时还有营销活动鼓励人们来访问网站。

经过对使用日志的分析，我们发现，在重新设计前 8 周，对话的平均时间是 1.2 页。这就是说，人们主要看的是首页，大约有 20% 的人会看两页或者更多页(极少有人会看超过四页)。

可用性测试表明，如果内容不在首页上出现，用户找到这些内容就会碰到很多麻烦。重新设计的目标之一就是让用户更容易找到这种内容。

重新设计 4 周后，平均点击流是 1.5 页，每次会话的页面和页面浏览量都增加了 25%。营销活动当然对增加做出了贡献，但多少贡献是归功于网站可用性的提高？如果我们假设增加的数量中有 30% 要归功于人们能更容易找到内容，这就表明网站更可用直接导致页面浏览量增加了 7.5%。





参考去年每月页面浏览量的平均数字(150 万)，还有标准按展示付费的广告是 10 美元，就表明月收入增加 1125 美元。如果营销工作造成了人们来到网站，但新设计导致了所有额外的页面浏览量，增加的访问量应该导致月收入增加 3750 美元。

但对网站的深层使用和仅仅访问首页不同，而且对额外收入有影响。子板块按展示显示广告的价格是 16 美元。如果 30%有效，就意味着收入增加 1800 美元，每年 21600 美元，或者说总收入增加约 10%。

成本是每小时 50 美元，共 80 小时，或者说 4000 美元加上 1000 美元奖励和设备费用。因此，公司总成本约 5000 美元。和年收入相比，第一年年底时的投入回报率从 270%提高到了 330%。

在有些情况下，投资回报率完全内部化。如果通过减少发布后的修正次数，良好的用户研究能使开发过程流线化，用户研究实际上就为公司节省了无法用直接收入来度量的大量费用。采用和没有采用以用户为中心的方法进行开发循环中的修正或者延迟的成本，两者比较能够有效度量内部投资回报率。

## 如果太困难了怎么办？

这些事情都不容易，但如果企业文化不允许集成可用性概念该怎么办？该做些什么？

首先，确认抵抗的性质。抵抗主要有两种形式：动力和敌意。

人们很容易回到旧习惯中。以用户为中心的设计需要时间、精力和奉献。以“老办法”做事情总是会比较容易，而且暂时看起来是效率更高的方式。“如果我们用老办法，3 周就能完成。你想让我们先进行一个月的研究？我们会严重延期！”利用短期效率作为放弃正确程序的托词会错失这种方法的要点。虽然靠“老派”方式做出的直觉决定最终是合适的，但它们更可能造成长期问题，最终导致更多工作。

对抗动力的方法就是把它引入不同方向。如前所述，如果与内部“营



销活动”结合起来,把用户体验研究过程对产品及公司的影响记录下来(见第17章),逐渐引入用户体验研究方法,这样是有效的。同样,如果有一位执行官承诺改造整个开发过程,这样也有效,不过所花资源和付出比大多数公司更多。

有些情况下,虽然公司能听取这些想法,但短期内都不可能改变任何主要过程。这种情况下,可能每次都需要证明用户体验研究工作的好处。

虽然在某些方式下,比较激进挑战的允许在思维方式中有更多理性而快速的变化,但敌意更难对付。再次强调,如果出错的话,真的会非常混乱。

杰夫·维恩描述了他和一家公司的职员讨论用户研究时碰到的反应:“对工程师而言,可用性是约束我的镣铐,对设计师而言,可用性是营销。”杰夫的解决方案是把可用性变成事件,他清楚说明有些特别的事情将要发生到公司和开发人员身上,管理层会听到并仔细看他们的产品。杰夫首先安排日程表,给开发人员更多空间,并缓和他们对可用性过程会增加额外工作而无法完成所造成的恐惧。然后他邀请所有开发人员去舒适的房间观看可用性研究,并提供食物。开发人员一起看到了访谈,杰夫向他们解释测试,强调某些行为的重要性,并把其他行为放到情境中。他还鼓励开发人员讨论产品。过了一段时间,开发人员注意到他们会根据特定用户及其陈述来辩论功能,而不是根据原则或者观点。

在某些情况下,用户会威胁人们。开发人员把这种人称为“讨厌的计算机用户”(Luser),把开发可用产品的过程描述成让他们变成“笨蛋”的过程,并会带着嘲弄的心态看待用户的要求或者错误理解。开发人员这种陈述背叛了对产品介于愤怒和不安之间的态度。如果敌意显得特别缺乏理性,并且顽固,那事情就会相当棘手。但直接接触用户有助于说服有疑问的人。首先,看见用户使用东西发生错误能确认怀疑者最害怕的东西:人们基本不能使用产品而且不友好。这的确让人感觉不舒服,但能带来很多核心想法。与用户广泛接触几乎总能提示出用户的问题和疑虑比开发人员所害怕的东西来更容易消除,不过用户之间也有令人困惑的区别。

不要让研究分散注意力而忘记产品。没有首先询问人们就做出决定也

#### 忠告:

研究麻烦。如果以用户为中心的过程开始受到欢迎,开发团队也将其纳入基本过程,就会有一种倾向想把所有变化都基于用户研究。这种想法让人仰慕,但一般都不可行,而且还会导致开发停滞。





没有关系。只不过不要用这种方式做所有决定。

### 唯一的方向

即使很难在公司实施这种过程，变化也会造成困难和混乱，也并不意味着永远不会发生。以用户为中心的设计背后的理念已经存在 20 多年，慢慢渗透到即使是最害怕用户、最关注内部的公司。“人无我有，人有我优”的压力挥之不去，以用户为中心、用户研究为基础的方法是实现这一目标的最佳方法之一(如果不是最佳方法)。以用户为中心的公司能更好传达用户理念，猜测更少，解决问题更多，并能更好地应对变化。简而言之，公司运行中的摩擦更少，理念更清晰。

这种公司现在还很少，但以后会越来越多，因为必须越来越多。随着信息革命——刚刚开始于 20 世纪 80 年代末及 90 年代初期——渗透到全世界的社会，工业革命的习惯已经融化而消失了。世界不再需要批量生产、批量营销、批量分销的产品和想法，没有道理再去考虑批量市场，不再需要根据对一些人的了解来为所有人创建解决方案。随着经济环境变得更艰难，竞争更激烈，每个地方的公司都会意识到好的商业并非止于而是始于产品或者服务的最终用户。





## 附录 A 用户研究实验室的预算

用户研究实验室不一定要很大。其主要目标是为研究提供一致、安静和舒适的空间。除此之外，投资实验室是一场回报不断减少的游戏。价值百万美元的实验室日常进行的工作不会远远超出在会议室中完成的工作。尽管如此，拥有可用性研究实验室仍然有很大的优势。从实用角度而言，实验室可以统一设备设置，显著减少对音频和视频文件质量的担心。

第 10 章描述的可用性测试实验室的物理布局如图 A.1 所示(第 9 章描述了焦点小组的物理布局)。

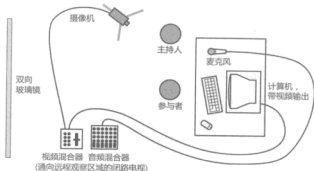


图 A.1 典型的可用性测试实验室

以下是根据预算购置的设备清单，包括闭路观察设备(可用来代替或者辅助双向玻璃镜)和参加者能看到的画中画成像系统。

### 视频设备

- 消费者级别的计算机：1500 美元。研究中使用的计算机设备要类似于用户可能使用的计算机，通常都不会是最先进的计算机。另外，保持计算机中性设置的最佳方法是从普通设置开始。





- 扫描转换器：300 美元。扫描转换器能把测试计算机的显卡信号转换成视频信号，然后可以进行处理和录制。确保挑选测试会采用的典型分辨率(很多低端转换器只能转换 640x480，大多数可用性测试都无法接受这么低的分辨率)。
- 两台微型摄像机，IEEE1394(Firewire/i.LINK)接口，并带外部视频输入，每台 800~2000 美元。要准备两台摄像机：一台作为摄像机，另一台作为视频录制机来录制视频混合器画中画的输出。摄像机不一定要特别好，因为基本现代摄像机都能满足要求。摄像机不一定要一样，因为在特定时间只会有一台进行视频录制，这台摄像机可以是较高质量的摄像机(实际上，录制用的摄像机都不一定是摄像机；盒式录放像机就可以了)。但如果要从头建立实验室，可以买两台相同摄像机备用。IEEE1394 接口输出更容易把视频导入到桌面视频编辑器。
- 视频转换器：1000 美元。如果能把用户画面和用户所看的内容录制到同一盘磁带上，可以采用低端视频转换器。
- 专业视频监视器：700~1000 美元。可以在回放时不损失质量，对于分析非常重要。
- 两台大电视机：700 美元。对于闭路观察，要有两台大电视机。一台显示测试者画面，另一台显示测试者所看的东西。两台电视机也可以用于向较大的群体回放视频重点。
- 视频三脚架：150 美元。
- 互联线缆：50 美元。
- 闭路电缆：200 美元。这些电缆都是长漆包线，能通到另一个房间，不会受到不同计算机设备的干扰。可以购买无线“电视机加倍器”单元，但这些单元不能总是在办公室环境下工作。
- 无线电波调幅器：30 美元。调幅器能把高端视频设备使用的组成信号转换成电视机天线信号，但有些电视机型号不必使用。
- 微型摄像机磁带：每盒 8~12 美元。
- 混合器、控制台和显示器运输车：200 美元。



## 可选设备

- 领夹式麦克风：150 美元，音频质量更好。
- 全向麦克风：150 美元。
- 全向麦克风支架：80 美元。
- 麦克风混合器：200 美元。
- 麦克风线：50 美元。

## 研究实验室作为目的地

但创建实验室的驱动因素并不仅仅是为了获得实用的视频，这些用户研究设备对于在公司中建立以用户为中心的开发流程发挥着重要作用。它能为公司内部的人建立观察用户使用产品的场所。这样能把用户的抽象概念更真实地集成到开发中，把对“用户”的想法转变成实际人群，减少解决方案的不稳定性，并强化来自于这个过程的好处。



## 附录 B 常见调查问题

以下问题类型清单可以用于调查，并提供了建议答案。这些问题不完整，也不完美，并不适用于所有情况，因为调查目标决定了问题以及答复的实际措辞。

### 人口统计问题

1. 您属于哪个年龄阶段？
  - 小于 18 岁
  - 18~24 岁
  - 25~34 岁
  - 35~49 岁
  - 50~64 岁
  - 65 或者超过 65 岁
  - 保密
2. 您的性别是什么？
  - 女性
  - 男性
  - 保密
3. 请选择您的教育水平。
  - 小学
  - 高中
  - 高中毕业
  - 职业/技术学校
  - 大学肄业



- 大学本科毕业
  - 硕士学位
  - 教授学位(法律硕士, MD 等)
  - 高级学位(博士, 教育心理学博士等)
  - 其他
- 4 您的家庭年收入属于下列哪个级别?
- 低于 20,000 美元
  - 20,001~29,999 美元
  - 30,000~39,999 美元
  - 40,000~49,999 美元
  - 50,000~59,999 美元
  - 60,000~69,999 美元
  - 70,000~79,999 美元
  - 80,000~99,999 美元
  - 100,000~119,999 美元
  - \$120,000~\$149,999 美元
  - 150,000 美元
5. 您现在的就业状态?
- 全职
  - 兼职
  - 无业
  - 自己创业
6. 您的工作职位是什么?
7. 您的公司或者组织的主要产品或者服务是什么?

## 互联网/互联网使用问题

1. 您使用计算机多长时间了?
- 少于 1 个月
  - 1~6 个月





- 6个月~1年
  - 1~2年
  - 2~3年
  - 3~5年
  - 5年或者更长
2. 您多久使用一次计算机？
- 每月少于一次
  - 每月一次
  - 每月几次
  - 每周一次
  - 每周几次
  - 每天
  - 每天几次或者每天大部分时间
3. 您最近常在哪里使用计算机？
- 家中
  - 单位
  - 学校
  - 图书馆
  - 公共计算机实验室
  - 任何地方(笔记本/PDA)
  - 其他
4. 您使用互联网电子邮件多长时间了？
- 少于一年
  - 1~2年
  - 2~3年
  - 3~4年
  - 4~5年
  - 超过5年



5. 您使用互联网多长时间了?
- 少于一个月
  - 1~6 个月
  - 6 个月~1 年
  - 1~2 年
  - 2~3 年
  - 3~5 年
  - 5 年或者更长
6. 您多长时间使用一次互联网?
- 每月少于一次
  - 每月一次
  - 每月几次
  - 每周一次
  - 每周几次
  - 每天
  - 每天几次
7. 在一次互联网访问会话中, 您平均花多少时间?
- 5 分钟或者更少
  - 5~15 分钟
  - 15~30 分钟
  - 30 分钟~1 小时
  - 1~3 小时
  - 3~6 小时
  - 6 小时或者更多

### 产品使用问题

1. 您使用[产品名称]已经多少时间了?
- 首次使用
  - 少于一周





- 1周~1个月
- 1~6个月
- 6个月~1年
- 1~2年
- 2~3年
- 3年或者更多
- 不知道

2. 下列特性清单中,您在最近三次使用[产品名称]时用过哪些?(选择所有用过的特性。)

- 功能特性 1
- 功能特性 2
- 没有
- 不知道

### 技术能力问题

提示:

连接速度和技术能力的变化非常频繁,因此到本书出版时,本书中提到的连接选项肯定已发生变化。

1. 您最常用哪种操作系统?
  - Windows 2000
  - Windows NT
  - Macintosh OSX
  - Linux
  - 其他
  - 不知道
2. 您的主要计算机 CPU 速度是多少?
  - 低于 500MHz
  - 500~1000MHz
  - 1000~2000MHz
  - 高于 2000MHz
  - 其他
  - 不知道





3. 您的主要显示器分辨率是多少?
- 640×480 像素
  - 800×600 像素
  - 1024×768 像素
  - 1600×1200 像素
  - 高于 1600×1200 像素
  - 其他
  - 不知道
4. 您连接互联网的速度是多少?
- 低于 56K
  - 56K
  - DSL 或者有线调制解调器(144K 到 1.5M)
  - T1
  - T3
  - 其他
  - 不知道



## 附录 C 观察员指南

### 焦点小组

(1) 仔细聆听。很容易立即讨论观察到的情况，但首先一定要听听人们真正说的内容，可以随时讨论你观察到的东西，但别忘了要仔细听。

(2) 不要仓促下结论。利用人们的陈述来了解他们思考问题的方式和他们的价值观，但不要把陈述的具体情况当信条。如果小组中每个人都表示喜欢或者讨厌什么东西，并不意味着整个世界都会这么认为，但它清楚表明有足够多的人会这么认为，应该足够引起注意。

(3) 焦点小组没有统计意义。如果五个人里面有四个人说了一些情况，并不意味着 80%的人都会这么认为。这说明有很多人可能会这么认为，但并不代表整个用户人口比例。什么也不是。就只是数字而已。

(4) 焦点小组的参加者都是专家。参加焦点小组的人知道他们要做什么，也知道他们当前是怎么做的。倾听他们的需求和体验，把他们当成顾问一样，他们能告诉你客户需要什么，而不要把他们当成产品消费者或者销售目标。

(5) 焦点小组不是灵丹妙药。如果每个焦点小组都有几个好主意，焦点小组就足够有价值，但并不是参加者的每个陈述都应该详细研究。

(6) 随时把问题传递给主持人，但不要过分传递。可以偶尔向小组提出问题，但一个环节中不要超过两个。应该简洁明了地写出问题，并且应该就好像要告诉主持人一样进行措辞。然后把问题给助理主持人，助理主持人会给主持人。在对话过程适当的时候，主持人会引入问题。但如果时间或者主题不合适，主持人可以决定不引入问题。

(7) 人们是矛盾的。听听人们如何思考主题、他们用什么标准来得出



结论，不一定要听他们说出来的具体愿望。人们可能没有意识到两个愿望不可能同时实现，或者人们根本就并不在乎。要做好准备偶尔会觉得无聊或者受到混淆。人们的行为并不总是有意义或者有见地。

(8) 不要贬低人们。参加者有时候会说些情况表明他们没搞清楚。不要假设人们是因为从一开始就没兴趣或没想法才没说什么重要的东西。了解为什么有人“没搞清楚”，这是了解那些“搞清楚”的人的观点的关键。

(9) 给主持人留点比萨。

## 可用性测试

(1) 仔细聆听。立即讨论观察结果很诱人，但首先一定要听听人们真正在说什么。可以随时讨论你看到的東西，但别忘了要仔细听。

(2) 可用性测试在统计学上没有代表意义。如果4个人里有3个人说了什么，并不代表75%的人都会这么认为。它的确说明有人会这么认为，但没有数量上的意义。

(3) 不要字字当真。只是几个人的观点。如果他们有好想法，那很好，但除非有重要证据能证明，否则请凭你的直觉来判断重要性。因此，如果有人说，“我讨厌绿色”，并不是说要改变颜色（但如果每个人都说“我讨厌绿色”，就说明需要进一步研究）。

(4) 人是矛盾的。听听人们如何思考主题、他们用什么标准得出结论，不一定要听他们所说的具体愿望。人们可能没有意识到两个愿望不可能同时实现，或者人们根本就并不在乎。要准备迎接偶尔的无聊或混淆。人们的行为并不总是有意义或者有见地。

(5) 不要期望通过可用性测试带来革命。如果得到一两个好想法，可用性测试就算是成功的。

(6) 除了注意观察人们做的事情、注意到的事情以外，同样要注意观察没有做的事情、没有注意到的事情。

如果是室内观察员，应该增加以下指南。

(7) 如果主持人提供明确机会，应该随时提出问题。提出的问题不要





以任何方式对产品产生价值暗示。因此，不要问“这是同类产品中最好的吗？”，而应该问“是否还有其他产品能这么做？您对这些产品有什么看法吗？”

(8) 不要说你和产品有直接关系。如果人们觉得同一间屋子中的人和产品没有太多关系，他们更容易对产品效果发表意见。



## 参 考 书 目

本书不是科学教材，而是一本基本的指南手册，因此我有意没有去找一大堆很难找到或者对于日常使用太技术化的论文和文献。以下常见书籍就能构成用户体验和以用户为中心的开发材料的良好图书馆。如果您有兴趣了解每个主题的更多技术层面或者历史层面，可以从每本书的参考书目中找到更多信息。

### 定性研究

Beyer, H. and Holzblatt, K. *Contextual Design*. San Francisco, CA: Morgan Kaufmann, 1998.

Kirwan, B. and Ainsworth, L. K. *A Guide to Task Analysis*. London, UK: Taylor and Francis, 1992.

Krueger, R. A. *Focus Groups: A Practical Guide for Applied Research (3d edition)*. Thousand Oaks, CA: Sage Publications, 1988.

Rubin, J. *Handbook of Usability Testing*. New York, NY: John Wiley & Sons, 1994.

### 定量研究

Babbie, E. R. *Survey Research Methods (Second Edition)*. Belmont, CA.: Wadsworth Publishing, 1990.

Dillman, D. A. *Mail and Internet Surveys (Second Edition)*. New York, NY: John Wiley & Sons, 1999.

Grossnickle, J. and Raskin, O. *The Handbook of Online Marketing Research*. New York, NY: McGraw-Hill, 2000.

Kirk, R. E. *Statistics: An Introduction*. Belmont, CA.: Wadsworth Publishing,





1998.

Moore, D. S. *Statistics: Concepts and Controversies (Fifth Edition)*. New York, NY: W.H. Freeman, 2001.

Rosenthal, R. and Rosnow, R. L. *Essentials of Behavioral Research: Methods and Data Analysis (Second Edition)*. New York, NY: McGraw-Hill, 1991.

### 设计哲学

Brinck, T., Gergle, D., Wood, S. D. *Usability for the Web*. San Francisco, CA: Morgan Kaufmann, 2002.

Burroughs, W. S. *Naked Lunch (Reissue)*. New York, NY: Grove Press, 1992.

Cooper, A. *The Inmates are Running the Asylum*. Indianapolis, IN: SAMS, 1999.

Garrett, J. J. *Elements of User Experience*. Indianapolis, IN: New Riders, 2002.

Nielsen, J. *Usability Engineering*. San Francisco, CA: Morgan Kaufmann, 1994.

Norman, D. A. *The Design of Everyday Things (Revised Edition)*. New York, NY: Currency/Doubleday, 1990.

Petroski, H. *The Evolution of Useful Things*. New York, NY: Vintage Books, 1994.

Veen, J. *The Art and Science of Web Design*. Indianapolis, IN: New Riders, 2000.

### 可用性商业

Bias, R. G. and Mayhew, D. J. (eds.). *Cost-Justifying Usability*. Boston, MA: Academic Press, 1994.

Donoghue, K. *Built for Use: Driving Profitability Through the User Experience*. New York: McGraw-Hill, 2002.



软件项目管理

Demarco, T. and Boehm, B.W. *Controlling Software Projects*. Englewood Cliffs, NJ: Prentice Hall PTR/Sun Microsystems Press, 1998.

Highsmith, J. *Agile Software Development Ecosystems*. Reading, MA: Addison-Wesley, 2002.

Mayhew, D. J. *The Usability Engineering Lifecycle*. San Francisco, CA: Morgan Kaufmann, 1999.

McConnell, S. C. *Rapid Development*. Redmond, WA: Microsoft Press, 1996.

编者注:

请参见清华大学出版社出版发行的《快速软件开发(珍藏版)》。



分享说明

如果你看到这段文字，那么通常来说你已经看完了本书，恭喜你！

马云曾说：“昨天很残酷，今天更残酷，明天很美好... 绝大部分人死在今天晚上。”也许坚持到明天，我们未必能成功。但是更多时候，是我们根本没有坚持！

在这个充满诱惑的时代，**我们花了多少时间在玩手机游戏、看电影、抑或是刷微博看小说上面，又花了多少时间静下心来读几本书？**你有多久没有完整地看完一本书了？更多时候是匆匆翻几页吧。这也是我整理这个系列的原因之一，让自己花更多时间系统化地阅读，更深入的思考，不让自己的生活被碎片化。

有时候看到一本好书，会心跳加速，觉得世间最大的享受莫过于此。我在 2011 年的第一天买了《三双鞋》送给自己，最开始是被美捷步创始人的商业传奇吸引买的，可是读到后来，发现自己深受启发的远非他的商业创奇，更多的是被他对生活的热情所鼓舞。当然，别指望看一本书就能改变很多，读书是一个过程。

关于我

非著名产品经理，产品观察家，痴迷交互设计、用户心理、产品体验、原型工具，给极客公园、虎嗅网、钛媒体、创业邦、i 黑马、互联网的一些事、搜狐新闻等码过几个字，扯扯产品设计。微信公众号：「idesign123」。 <http://pmithink.com> 信仰好设计，更有爱。哦，对了，小子尚未婚配，有女友，且固定，您来晚了→\_→

联系方式

如果你有一些好的资源（别想歪咯）、好的想法或者就是单纯地想和我交流下，欢迎之至，下面是我的联系方式：

-  514359876
-  @zchening
-  zchening@gmail.com
-   扫一扫可加载更多.....



扫一扫或搜索：*idesign123*