



南京大學  
NANJING UNIVERSITY



# 人机交互系统

## 评估的基础知识

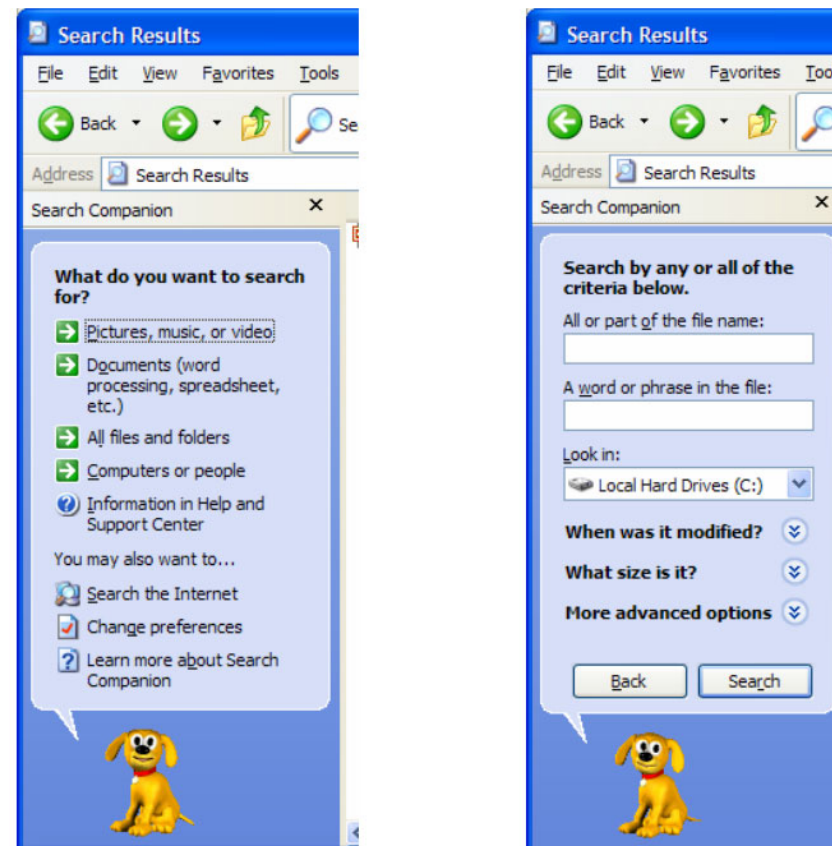
主讲教师：冯桂焕  
2022年秋季



# UI Hall of Shame or Hall of Fame?



南京大學  
NANJING UNIVERSITY



# Google™

Web

Images

Groups

Directory

News

Google Search

I'm Feeling Lucky

- [Advanced Search](#)
- [Preferences](#)
- [Language Tools](#)

[Advertise with Us](#) - [Business Solutions](#) - [Services & Tools](#) - [Jobs, Press, & Help](#)

©2003 Google - Searching 3,307,998,701 web pages

+你 搜索 图片 地图 Play YouTube 新闻 Gmail 更多

登录

# Google

谷歌

Google 搜索

手气不错

Google.com.hk 使用下列语言: [中文\(繁體\)](#) [English](#)

# Google

Google 搜索

手气不错

Gmail 图片

登录

Google 应用

Google 帐...

搜索

地图

YouTube

Play

新闻

Gmail

Meet

Chat

通讯录

云端硬盘

日历



 中国铁路客户服务中心 | 客运服务

意见反馈: 12306jfk@rails.com.cn 您好, 请 [登录](#) | [注册](#) 我的12306 [手机版](#)

[客运首页](#) [车票预订](#) [余票查询](#) [出行向导](#) [信息服务](#)

车票查询

☒ 单程 ☐ 往返

出发地

简拼/全拼/汉字

目的地

简拼/全拼/汉字

出发日

2014-03-12

返程日

2014-03-12

☒ 普通 ☐ 学生

查 询

出行向导

 车站查询

 中转查询

 票价查询

 正晚点查询

 客票代售点

最新动态

购票常见问题

使用须知

相关规章

1 关于2014年短途卧铺优惠有关事宜的公告 NEW (2014-03-05)

2 关于网站对用户身份进行核验的公告 NEW (2014-02-23)

3 关于调整部分车站互联网、电话订票起售时间的公告 NEW (2014-01-07)

4 广铁集团公司关于3月24日至3月31日长沙南开G6101次、广州南开G6118... (2014-03-04)

5 广铁集团公司关于3月7日至3月9日广州南--深圳北间临时加开部分动车的... (2014-03-04)

6 广铁集团公司关于自2014年3月1日起长沙开K9024/5次、3月2日吉首开K9... (2014-02-27)

7 广铁集团公司关于2014年2月28日、3月1日、2日广州东-深圳间加开D720... (2014-02-27)

8 关于杭深线厦深段开通持二代居民身份证直接检票进出站功能的公告 (2014-02-28)

9 太原局关于6834、6824次临时变更运行时刻的通知 (2014-02-28)

10 济南局关于莱阳站改造期间调整部分旅客列车车站的公告 (2014-02-26)

关于我们 | 网站声明

版权所有©2008-2015铁道部信息技术中心 中国铁道科学研究院

京ICP备10009636号



您现在的位置: 客运首页 > 我的12306 > 已完成订单

### 我的12306

#### 我的火车票

未完成订单

已完成订单(改/退)

#### 个人信息

查看个人信息

账号安全

#### 常用信息管理

常用联系人

车票快递地址

### 已完成订单

未出行订单

历史订单

按订票日期查询 查询日期 2014-12-23 - 2014-12-23 订单号/车次/乘客姓名 查询

全部 可改签 可退票

订单日期: 2014-12-23 冯桂焕 北京南—>南京南 乘车日期: 2015-01-05 17:55 订单详情						
序号	车次信息	席位信息	旅客信息	票款金额	车票状态	操作
1	2015-01-05 17:55开 G201 北京南-南京南	02车厢 05C号 二等座	二代身份证	成人票 443.5元	已支付	退票
订单号: E088032061						
全选 订单跟踪 退款详情 快递详情 改签						

订单日期: 2014-12-23 冯桂焕 南京—>北京南 乘车日期: 2015-01-04 21:57 订单详情

首页 上一页 1 下一页 末页

#### 温馨提示:

1. 订单信息在本网站保存期限不少于30日。
2. 本网站仅办理不晚于开车前2小时尚未换取纸质车票的退票、改签业务。
3. 在本网站办理退票, 只能逐次单张办理。
4. 车票只能改签一次, 已经改签的车票不能再次改签。
5. 退票、改签后, 如有应退票款, 请及时查询网上支付工具账户余额, 如有疑问请致电12306人工客服查询。
6. 如需取消或查看车票快递状态, 请点击“快递详情”。





您现在的位置: 客运首页 > 我的12306 > 已完成订单

我的12306

我的火车票

未完成订单

已完成订单 (改/退)

个人信息

查看个人信息

账号安全

常用信息管理

常用联系人

车票快递地址

已完成订单

未出行订单

历史订单

按订票日期查询 查询日期 2014-12-23 - 2014-12-23 订单号/车次/乘客姓名 查询

全部 可改签 可退票

订单日期: 2014-12-23 冯桂焕 北京南→南京南 乘车日期: 2015-01-05 17:55 订单详情

序号	车次信息	席位信息	旅客信息	票款金额	车票状态	操作
1	2015-01-05 17:55开 G201 北京南-南京南	02车厢 05C号 二等座	二代身份证	成人票 443.5元	已支付	退票

订单号: E088032061

全选

订单跟踪

退款详情

快递详情

改签

订单日期: 2014-12-23 冯桂焕 南京→北京南 乘车日期: 2015-01-04 21:57 订单详情

序号	车次信息	席位信息	旅客信息	票款金额	车票状态	操作
1	2015-01-04 21:57开 D312 南京-北京南	02车厢 08号上铺 软卧	二代身份证	成人票 492.5元	已支付	退票

订单号: E082203801

全选

订单跟踪

退款详情

快递详情

改签

首页

上一页

1

下一页

末页

南京大学  
NANJING UNIVERSITY



## ■评估总是需要的

## ■什么是评估？

- 系统化的数据搜集过程
- 目的是了解用户或用户组在特定环境中，使用产品执行特定任务的情况
  - 例如，用户能否找到特定的菜单项？图像是否有用，是否吸引人？产品是否引人入胜？
- 邀请用户进行评估的目的不是设法理解用户，而是评估特定用户在一个特定的环境背景中如何使用一个系统来执行一个特定的任务。



## 关于评估的错误认识

---

### ■只有在系统开发完成后才需要评估

- 如果开发逾期，但是交付日期不变，那么评估部分就会被取消

### ■只要一个系统做了开发者认为是正确的事，那就足够了

- 案件检索的例子

### ■设计人员能够准确了解用户工作的方式

- 12306的例子
- 短信系统最初是为了让工程师调试手机的通信系统而设计的





# 评估目标

---

## ■ 评估的优点

- 能够在交付产品之前（而不是之后）修复错误
- 设计小组能够专注于真实问题，而不是假想问题
- 工程师们能专心于编程而不是争论
- 能够大大缩短开发时间
- 销售部门可获得稳定的设计

## ■ 评估的目标

- 评估系统功能的范围和可达性
- 评估交互中用户的体验
- 确定系统的某些特定问题



# 评估原则

---

## ■评估应该依赖于产品的用户

- 与专业技术人员的水平和技术无关

## ■评估与设计应结合进行

- 仅靠用户最后对产品的一两次评估，是不能全面反映出软件可用性的

## ■评估应在用户的实际工作任务和操作环境下进行

- 根据用户完成任务的结果，进行客观的分析和评估

## ■要选择有广泛代表性的用户

- 参加测试的人必须具有代表性



# 评估范型和技术

---

## ■“范型”与“技术”

- 范型与具体学科相关，对如何评估有很大影响
  - 可用性测试是一种评估范型
- 每种范型有特定的技术
  - 可用性测试的技术有观察、问卷调查、访谈等

## ■评估范型

- 快速评估
- 可用性测试
- 实地研究
- 预测性评估



# 快速评估

---

■设计人员非正式地向用户或顾问了解反馈信息，以证实设计构思是否符合用户需要

- 可在任何阶段进行
- 强调 “快速了解”，而非仔细记录研究发现
  - 如在设计初期了解用户对新产品的意见、在设计末期了解用户对图标设计的看法等
- 得到的数据通常是非正式、叙述性的
  - 可以口语、书面笔记、草图、场景的形式反馈到设计过程
- 是设计网站时常用的方法

■基本特征：快速



# 可用性测试

---

- 20世纪80年代的主导方法
- 评测典型用户执行典型任务时的情况
  - 包括用户出错次数、完成任务的时间等
- 基本特征
  - 是在评估人员的密切控制之下实行的
- 主要任务
  - 量化表示用户的执行情况
- 缺点
  - 测试用户的数量通常较少
  - 不适合进行细致的统计分析



# 实地研究

---

- 基本特征：在自然工作环境中进行
- 目的：理解用户的实际工作情形以及技术对他们的影响
- 作用
  - 探索新技术的应用契机
  - 确定产品的需求
  - 促进技术的引入
  - 评估技术的应用
- 分类
  - 评测人员作为“局外人”
  - 评测人员也可作为“局内人”或测试用户



# 预测性评估

---

- 专家们根据自己对典型用户的了解（通常使用启发式过程）预测可用性问题
  - 也可使用理论模型
- 基本特征
  - 用户可以不在场
  - 使得整个过程快速、成本较低
- 启发式评估是典型的预测性评估方法
  - 注意：启发式原则应定制
  - 可能误导设计人员
    - 且有些结果可能并不准确



# 评估范型比较

评估范型	快速评估	可用性测试	实地研究	预测性评估
用户角色	自然行为	执行测试任务集	自然行为	用户通常不参与
控制权	评估人员实施最低限度控制	评估人员密切控制	评估人员与用户合作	评估人员为专家
评估地点	自然工作环境或实验室	实验室	自然工作环境	类似实验室的环境，通常在客户处进行
适用情形	快速了解设计反馈。可使用其他交互范型的技术，如专家评测	测试原型或产品	常用于设计初期，以检查设计是否满足用户需求，发现问题，发掘应用契机	专家（通常是开发顾问）检查原型，可在任何阶段进行。使用模型评测潜在设计的特定方面
数据类型	通常是定性的非正式描述	量化数据，有时是统计数据。可采用问卷调查或访谈搜集用户意见	应用草图、场景、例证等的定性描述	专家们列出问题清单，由模型导出量化数据，如两种设计的任务执行时间
反馈到设计	通过草图、例证、报告	通过性能评测、错误统计报告等为未来版本提供设计标准	通过描述性的例证、草图、场景和工作日志	专家列出一组问题，通常附带解决方案建议。为设计人员提供根据模型计算出的时间值
基本思想	以用户为中心，非常实用	基于试验的实用方法，即可用性工程	可以是客观观察或现场研究	专家检查以实用的启发式原则和实践经验为基础，采用基于理论的分析模型





# 评估技术

---

## ■观察用户

- 有助于确定新产品的需求
- 也可用于评估原型
- 挑战：如何在不干扰用户的前提下观察用户，以及如何分析大量数据

## ■询问用户意见

- 简单，调查用户数量从几个到几百不等

## ■询问专家意见

- “角色扮演”方式评估
- 同时专家会提出解决方案



## 评估技术-2

---

### ■测试用户的执行情况

- 可比较不同设计方案优劣
- 通常在受控环境中进行

### ■基于模型和理论，预测界面的有效性

- 常用技术如GOMS模型和KLM模型等



# 评估范型和技术的关系

评估技术	评估范型			
	快速评估	可用性测试	实地研究	预测性评估
观察用户	观察用户实际行为的重要方法	使用摄像和交互日志的记录方式，可做进一步分析，以找出问题，了解操作步骤，计算执行时间	实地研究的核心方法。在现场研究中，评测人员与测试环境相融合；在其他类型的研究中，评测人员只做客观观察	——
询问用户意见	与用户和潜在用户讨论，可采用个别会谈、集体会谈或专门小组的形式	通过问卷调查了解用户满意度，也可通过访谈了解更多详情	评测人员可采用访谈的形式，与用户讨论观察到的问题。现场研究可采用现场访谈	——
询问专家意见	专家评估原型的可用性(提供“评估报告”)	——	——	在设计初期，专家使用启发式原则预测界面的有效性
用户测试	——	在受控环境中，测试典型用户执行典型任务的情况，是可用性测试的基本方法	——	——
用户执行情况分析模型	——	——	——	使用分析模型预测界面的有效性，或比较用户使用不同设计方案的执行效率



# 区分评估技术的因素

---

## ■评估在周期中的位置

- 设计早期阶段的评估更快速、便宜

## ■评估的形式

- 实验室环境or工作环境

## ■技术的主客观程度

- 技术越主观，受评估人员知识的影响越大
  - 如认知走查等

## ■测量的类型

- 与技术的主客观性有关
  - 主观技术：定性数据
  - 客观技术：定量数据



## ■提供的信息

- 低层信息：这个图标是可理解的吗？
- 高层信息：这个系统是可用的吗？

## ■响应的及时性

- 边做边说法可及时记录用户行为
- 任务后的走查取决于对事件的回忆

## ■干扰程度

- 直接响应测量可能会影响用户表现

## ■所需资源

- 设备、时间、资金、参与者、评估人员的专业技术及环境等



# 评估技术比较

方法	生命周期阶段	用户人数	主要优点	主要缺点
启发式评估	早期设计，反复设计过程的“内循环”	无	能发现单个可用性问 题，能发现熟练用户 碰到的问题	没有涉及真实的用 户，故无法再用户需 求方面有“惊人发现”
绩效度量	竞争性分析，最终测 试	至少 10 人	硬性数据，对结果容 易进行比较	不能发现单个可用性 问题
边做边说	反复设计，形成性评 估	3~5 人	准确了解用户的错误 想法，测试费用低	用户感到不自然，熟 练用户感到很难用语 言表述
观察	任务分析，后续研究	3 人或以上	生态有效性；发现用 户的真实人物；建议 系统功能与特征	很难约定安排，实验 人员无法控制
问卷调查	任务分析，后续研究	至少 30 人	发现用户主观偏好， 容易重复进行	需要进行问卷预答 (避免出现误解)
访谈	任务分析，后续研究	5 人	灵活，可以深入了解 用户观点和用户体验	耗时，难以进行分析、 比较
焦点小组	任务分析，用于参与	每组 6~9 人		分析起来困难，有效 性低
使用过程记录 用户反馈	最终测试，后续研究 后续研究	至少 20 人 上百人	跟踪用户需求和想法 上的变化	需要专门部门来处理 回复



# 评估方法组合

---

■评估方法的组合取决于项目待评估的具体特性

■常用组合

- 启发式评估+边做边说等用户测试技术
  - 专家可通过启发性评估排除显而易见的可用性问题
  - 重新设计后，经用户测试，反复检查设计的效果
- 访谈+问卷调查
  - 先对小部分用户进行访谈，确定问卷中的具体问题

■启发式评估vs.用户测试

- 前者不需要用户参与
- 二者发现的可用性问题不同，可以互补

