Usability Testing

Refs: D. Petkovic

J. Rubin: “Handbook of Usability testing”, Wiley Technical Communication Library,

Testing Guru

Content

- Concept
- Types of usability tests
- Test environments
- Typical test plan format
- Task selection for evaluation
- Measurements and Questionnaires
- Summarizing performance results
- Analyzing Data

Usability Testing

- Process that employs representative target population participants to evaluate product usability using specific usability criteria
- Usability testing is not a guarantee for product success, but it should identify at least the key problems

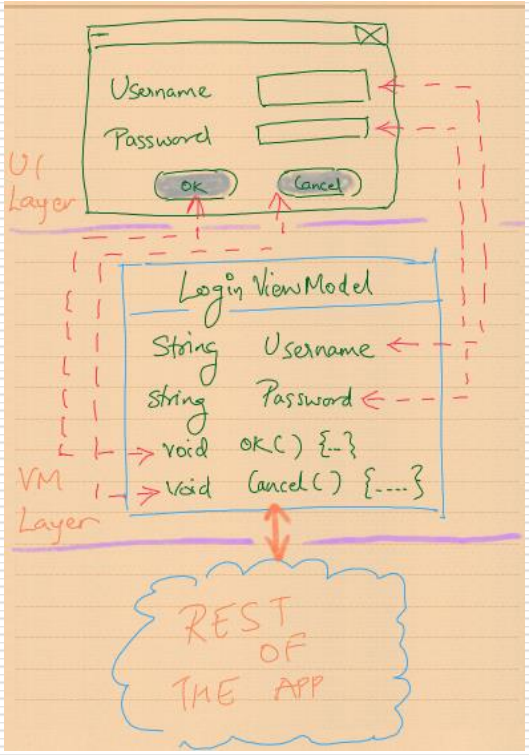
Basic components

1. Development of specific problems statements and test plans and objectives
2. Use of representative sample of end users
3. Representation of the actual work environment
4. Observation of end users during product use or review
5. Collection of quantitative and qualitative measurements
6. Analysis of results and recommendations

Types of usability tests

Exploratory:

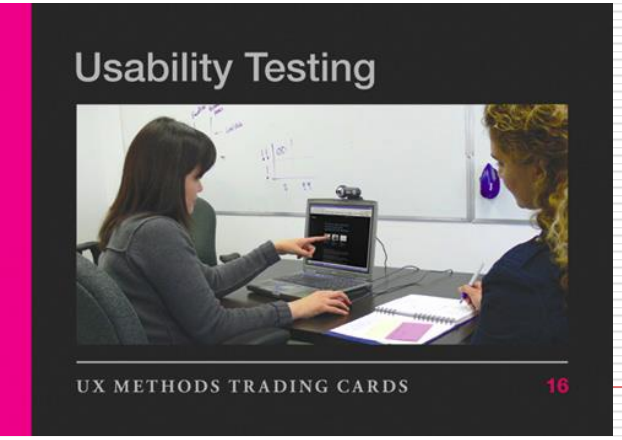
- Early in the process
- Can be based on any form of the GUI (sketch, wire diagrams etc.)
- Evaluate preliminary, basic design concept
- Perform representative tasks in a “shallow” mode
- Informal test methodology, a lot of interaction
- Discuss high level concepts

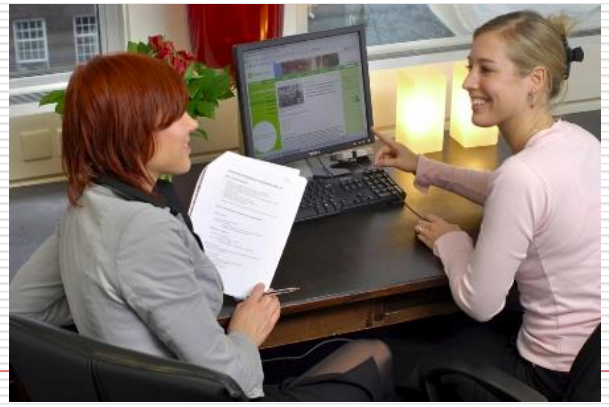


Types of usability tests

Assessment

- Done After fundamental concepts are done
- Evaluates usability of lower level operations
- The users actually perform set of well defined tasks
- Less interaction with test monitor
- Quantitative measurements are collected





Types of usability tests

Validation

- Done late in development cycle, close to release
- Goal is to certify product usability, "disaster insurance" against launching a poor product
- Often the first time when the whole product si tested (including help and docs)
- Evaluate wrt. some predetermined usability standard or benchmark
- Standards come from **previous testing, competitive information, marketing** etc.
- Very specific quantitative tests
- Can establish standards for future products
- Can be also done by beta customers

Types of usability tests

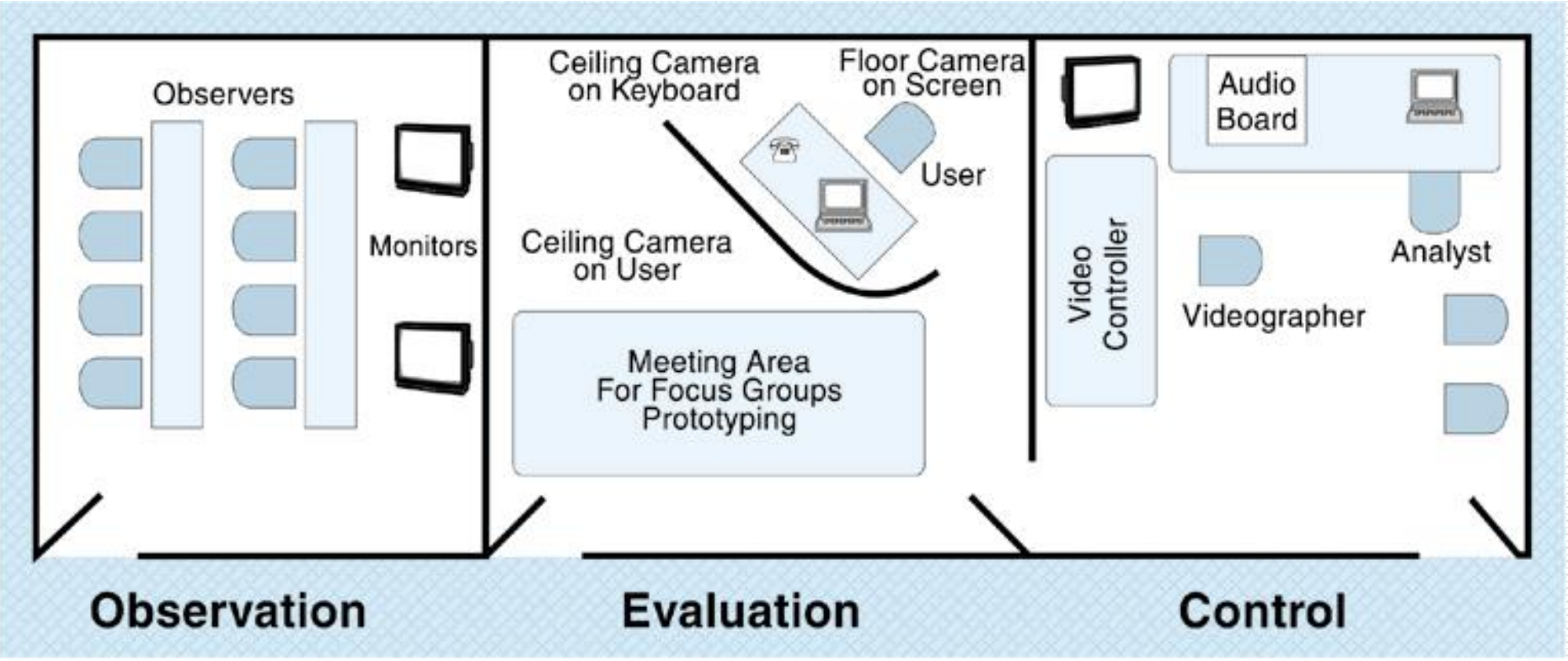
Comparison

- Can be done at any point in the development cycle
- Compare alternatives using objective measures
- Can be informal of formal, depending when it is done
- Often, best of alternative designs is combined

Test environments

- Simple single room setups
  - Observer/monitor close to evacuator
  - Observer removed from evaluator
- Electronic observation room
- Classic elaborate usability lab
- Mobile lab

Example


Observers
Monitors
Ceiling Camera on Keyboard
Floor Camera on Screen
User
Ceiling Camera on User
Meeting Area For Focus Groups Prototyping
Video Controller
Audio Board
Analyst
Videographer
Observation
Evaluation
Control

Typical test plan format

- Purpose: what is the main purpose of the test
- Problem statement: specific questions you want resolved
- Test plan and objectives: tasks the user will do
- User profile: who will be the users
- Method and test design: how will you observe it, how will you collect the data etc.
- Test environment and equipment
- Test monitor role
- Evaluation measures and data to be collected: how will you collect the feedback and how will you evaluate it
- Report: what will final report contain

Task selection for evaluation

- Tasks to be evaluated are functions users want to do with the product. **Focus is on user view of the tasks and NOT at the components and details** that you used to implement it. Examples:
  - Create and file document
  - Import several images
  - Find the right document
- Objective is to indirectly expose usability flaws by asking the user to perform typical tasks and **NOT telling them exactly how to do it.**
- Choose **key and most frequently** done tasks
- The **task has to be specific and measurable** (quantitatively or qualitatively)

Task components

TASK	DESCRIPTION
Task	Load paper into copier
Machine state	Paper tray empty
Successful completion criteria	Paper properly loaded
Benchmark	Completed in 1 minute

Selection of evaluators and test groups

- Evaluators should **be representative of the targeted users**
- Independent groups or within-subject design (but be careful to avoid exposing users to same tests since this will bias the results)
- Adequate numbers of testers
- Offer motivation and rewards

Measurements and Questionnaires

- Performance data: measures of user behavior such as **error rates, number of accesses to help, time** to perform the task etc.
  - Usually can and should be objectively and automatically measured
- Preference data: measures of **user opinion, thought process such as rankings, answers to questions, comments etc.**
  - Use questionnaires.

Some performance measures (measure what can be measured)

- Time to complete each task
- Number and percentage of tasks completed successfully/unsuccessfully
- Time required to access information
- Count of incorrect selections
- Count errors
- Time for system to respond
- ….

Data should be collected automatically or manually in an objective way.

Questionnaires (for preference data)

Likert scale

- I found GUI easy to use (check one)

  __ Strongly disagree __ Disagree

  __ Neither agree or disagree

  __ Agree __ Strongly agree

  (can also assign numbers from –2 to 2)

Semantic differentials

- I found File Open menu (circle one)

  Simple 3 2 1 0 1 2 3 Complex

Questionnaires

Fill in questions

- I found the following aspects of GUI particularly easy to use (list 0-4 aspects)

Questionnaires

Check-box

- Please check the statement that best describes your usage of spell check
  - --- I always use spell check
  - --- I use spell check only when I have to
  - --- I never use spell check

Questionnaires

Branching questions

- Would you rather use advanced search

  --- NO (skip to question 19)

  --- YES (continue)

- What kind of advanced search would you like? (check one)

  --- Boolean

  --- Relevance

Summarizing performance results

- Performance data
  - Mean time to complete
  - Median time to complete
  - Range (high and low)
  - Standard deviation (độ lệch) of completion times
  - System response time statistics
- Task accuracy
  - % of users completing the task within specified time
  - % of users completing the task regardless (không quan tâm) of time
  - Same as above, with assistance
  - Average error rate

Summarizing preference results

- For limited choice questions
  - Count how many participants selected each choice (number and %)
  - For Likert scale or semantic differentials provide average scores if there are enough evaluators
- For free form questions
  - List questions and group answers into categories, also into positive and negative answers
- For free comments
  - List and group them at the end of the report

Analyzing Data

- Identify and focus on tasks that **did not pass the test** or showed significant problems
- Identify user errors and difficulties
- Identify sources of errors
- Prioritize problems by criticality = severity AND probability of occurrence
- Analyze differences between groups (if applicable)

- Provide recommendations at the end

Problems statements and performance data to collect

Problem Statement	**Performance Data Collected**
How effective is the tutorial	Compare error rates of users who used and not used this
How easy is it to perform task X	Error rate OR Number of steps needed

Note: this is Performance data measurement only. You also need to asses user Preference data (see next slide)

Problems statements and preference data to collect

Problem Statement	**Preference Data Collected**
How effective is the tutorial	Ask user to rate it from very ineffective to very effective (Lickert scale or semantic differentials) + free comments
How easy is it to perform task X	Ask user to rate it from very easy to very difficult (Lickert scale or semantic differentials) + free comments

Relate problem statements with tasks

Problem Statement	**Task**
How effective is the tutorial	GroupA: Import image w/o using tutorial GroupB: Same but use tutorial first
How easy is it to Create Virtual Machine	Create Virtual machine with “this” properties using New VM Wizard

Task components

TASK	DESCRIPTION
Task	Create VM using New VM Wizard
Machine state	VMware WS SW just loaded
Successful completion criteria	Working VM created
Benchmark	Completed in 30 sec.

Some Suggested GUI issues to cover in preference data collection

Use GUI principles as general measures of quality to evaluate

- Screen layout matches tasks
- Amount of information is adequate
- Good use of text (headlines, help,messages, warnings…)
- Good use of colors
- Proper grouping of related info
- Navigational problems
- Users get lost in the system
- Organized by user tasks
- Icons are self-explanatory
- Consistency

Note: ask these questions in the context of concrete user tasks not in vacuum

How can we design GUI checklist ?

- UI elements
- User requirements
- Standards
- Trends

http://www.loop11.com/

- Loop^{11} is a web-based user-experience testing tool, allowing you to conduct remote, unmoderated usability testing on any kind of digital interface providing you with detailed quantitative usability metrics

Create New User Test

Hide website preview

Preview your website first to ensure it loads correctly.
Why do I need to do this?

http://www.fit.hcmus.edu.vn

Preview

Projects Under Construction

PROJECT	CREATED	TASKS	EDIT	DELETE

View full list

Launched Projects

PROJECT	LAUNCHED	RESPONSES	STATUS	REVIEW	ANALYSE	DELETE
SAMPLE PROJECT - Yahoo!	13 Apr 2009	224	Closed	View design	View report	Delete

View full list

COMMON QUESTIONS

If I require help and submit a query via email how long will it take to get a response?

Can I provide feedback about Loop[11]?

Can I speak to someone at Loop[11] if I need assistance?

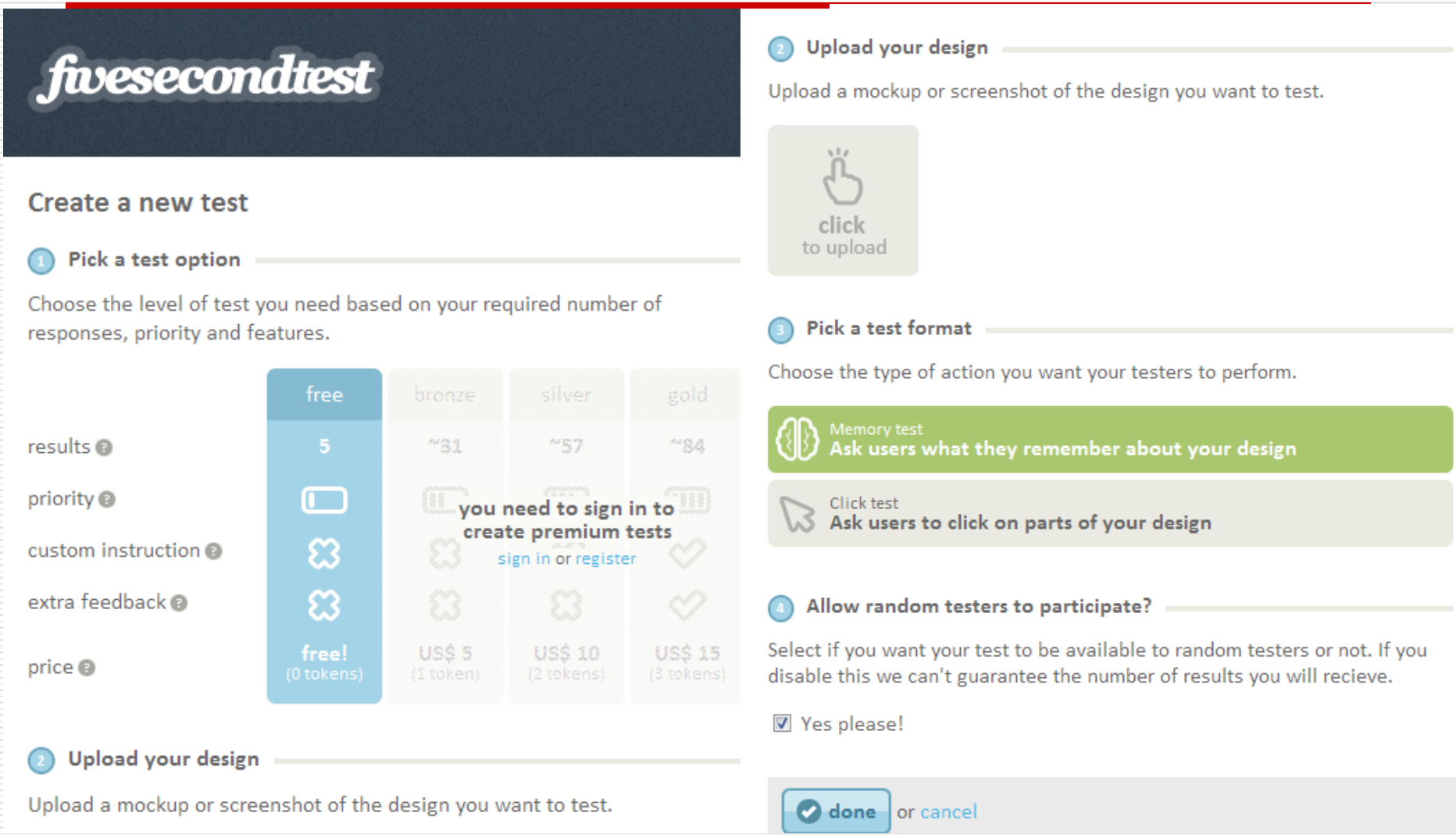
Is credit card payment the only method of purchase?

BUY TEST CREDITS

Pre-purchase credits instead of paying as you go.

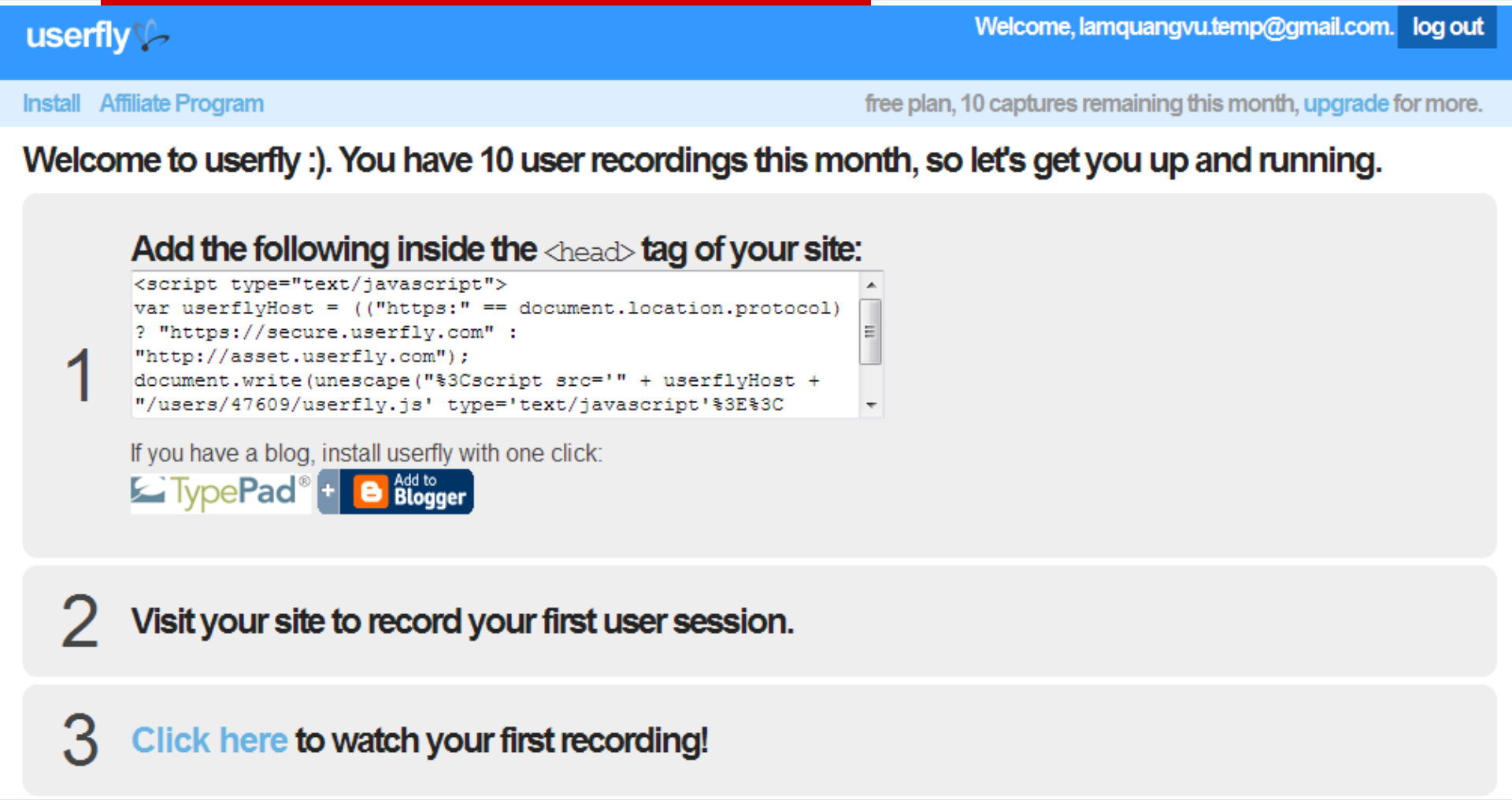
Go to purchase page

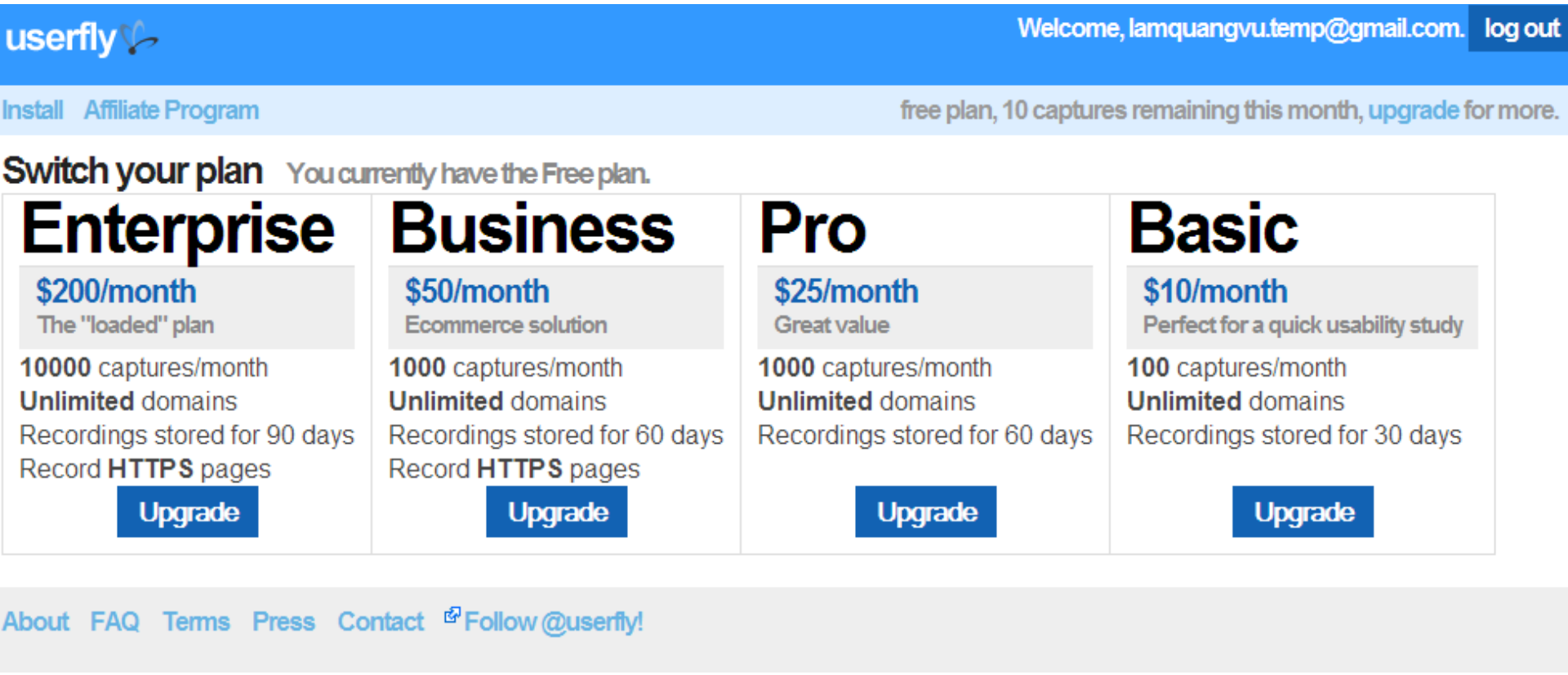
http://fivesecondtest.com/



http://userfly.com/

http://userfly.com/




userfly
Welcome, lamquangvu.temp@gmail.com. log out
Install Affiliate Program
free plan, 10 captures remaining this month, upgrade for more.
Switch your plan You currently have the Free plan.
Enterprise
$200/month
The "loaded" plan
10000 captures/month
Unlimited domains
Recordings stored for 90 days
Record HTTPS pages
Upgrade
Business
$50/month
Ecommerce solution
1000 captures/month
Unlimited domains
Recordings stored for 60 days
Record HTTPS pages
Upgrade
Pro
$25/month
Great value
1000 captures/month
Unlimited domains
Recordings stored for 60 days
Upgrade
Basic
$10/month
Perfect for a quick usability study
100 captures/month
Unlimited domains
Recordings stored for 30 days
Upgrade
About FAQ Terms Press Contact Follow @userfly!

http://www.usertesting.com/

http://www.clicktale.com/

- Similar to Userfly, ClickTale captures the actions of your website's visitors, including clicks, scrolling, and keystrokes. Other features offered include a scrolling heat map, form analytics, and individual link analysis (e.g. how many clicks or hovers did each link get?).
- ClickTale has a free plan that records 400 page views per month for one domain. Some of the features in the free plan are limited; for example, the scrolling heat map is available only for your most popular page.
- Paid plans start at $99 a month, which is still a competitive price for the features that this service offers.

CLICKTALE® Pricing Product Resources Company Contact Customer Login Search

Chat Now

Customer
Experience Analytics

Watch actual browsing behavior

Optimize website usability

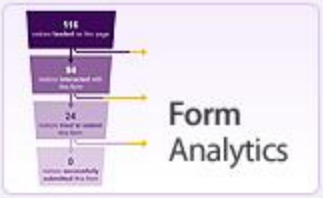
Boost conversion rates

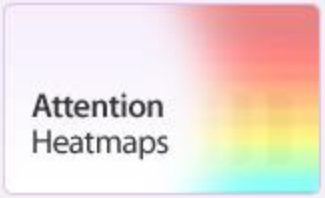
Plans & Pricing
3 months Free inside

Enterprise Solution

Request a Demo

Take Product Tour >>





40,000 Customers and Counting >>

Resources >>

http://smt.speedzinemedia.com/smt/

- Simple Mouse Tracking is another free tool to track a visitor's mouse movements and clicks on your site. You'll need PHP on your server to run it, although the actual pages don't have to be PHP-based; they just require a bit of JavaScript.

(smt) simple mouse tracking

preface abstract features system architecture references
demo examples downloads
comments contact

O_o

> The way to create something beautiful is often to make subtle tweaks to something that already exists, or to combine existing ideas in a slightly new way. This kind of work is hard to convey in a research paper.
>
> **Paul Graham - Hackers and Painters**

http://silverbackapp.com/

http://www.etre.com/tools/colourblindsimulator/

Colour Blindness Simulator

These are the results of your colour blindness simulation.

Original

This is the image you uploaded.

Simulation

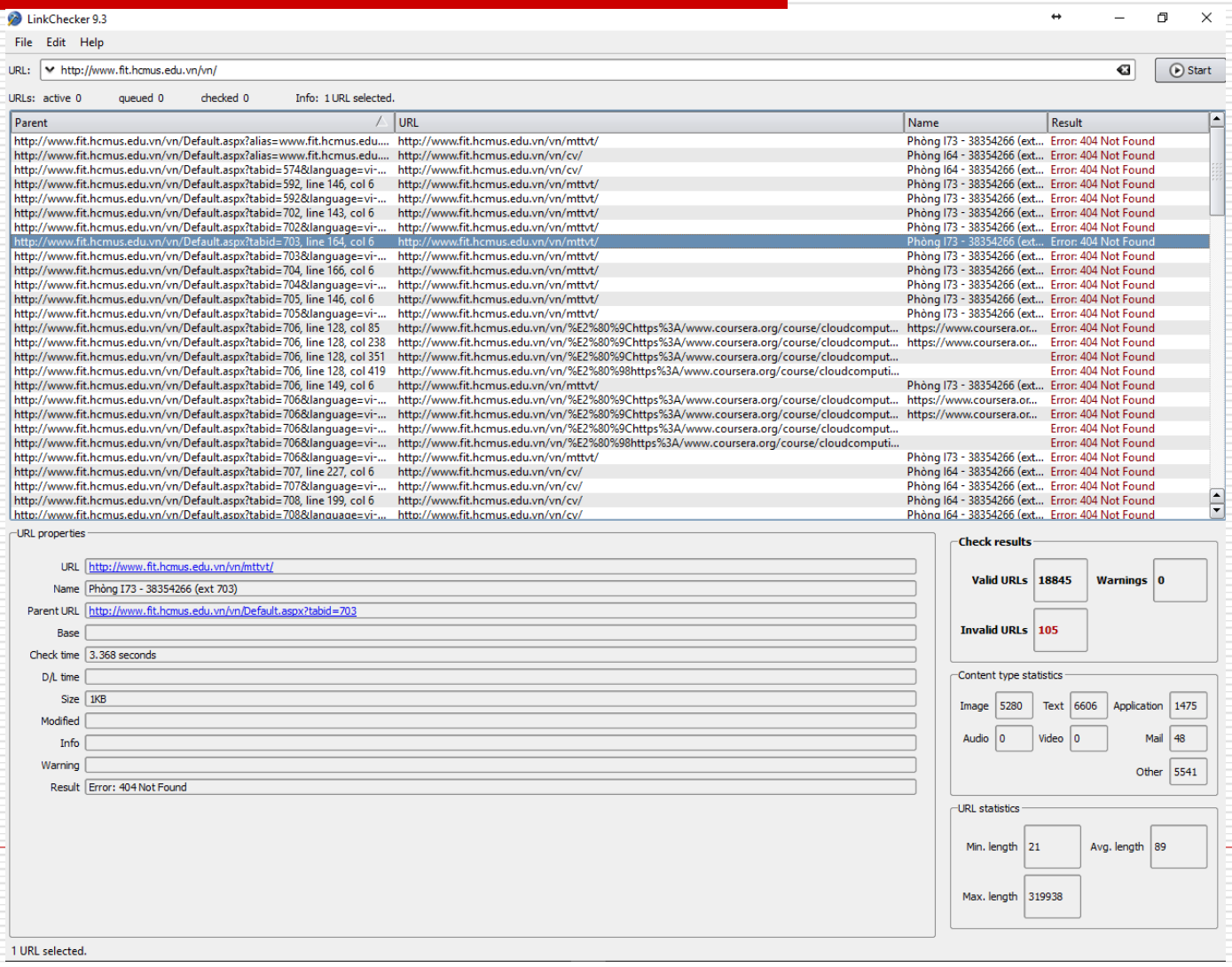
This is how your image may appear to a person with protanopia.

Protanopia is a more common form of colour blindness, occuring in between 1% and 5% of males (varying by race) and in approximately 0.1% of females. You may like to simulate how someone with deuteranopia or tritanopia views your image.

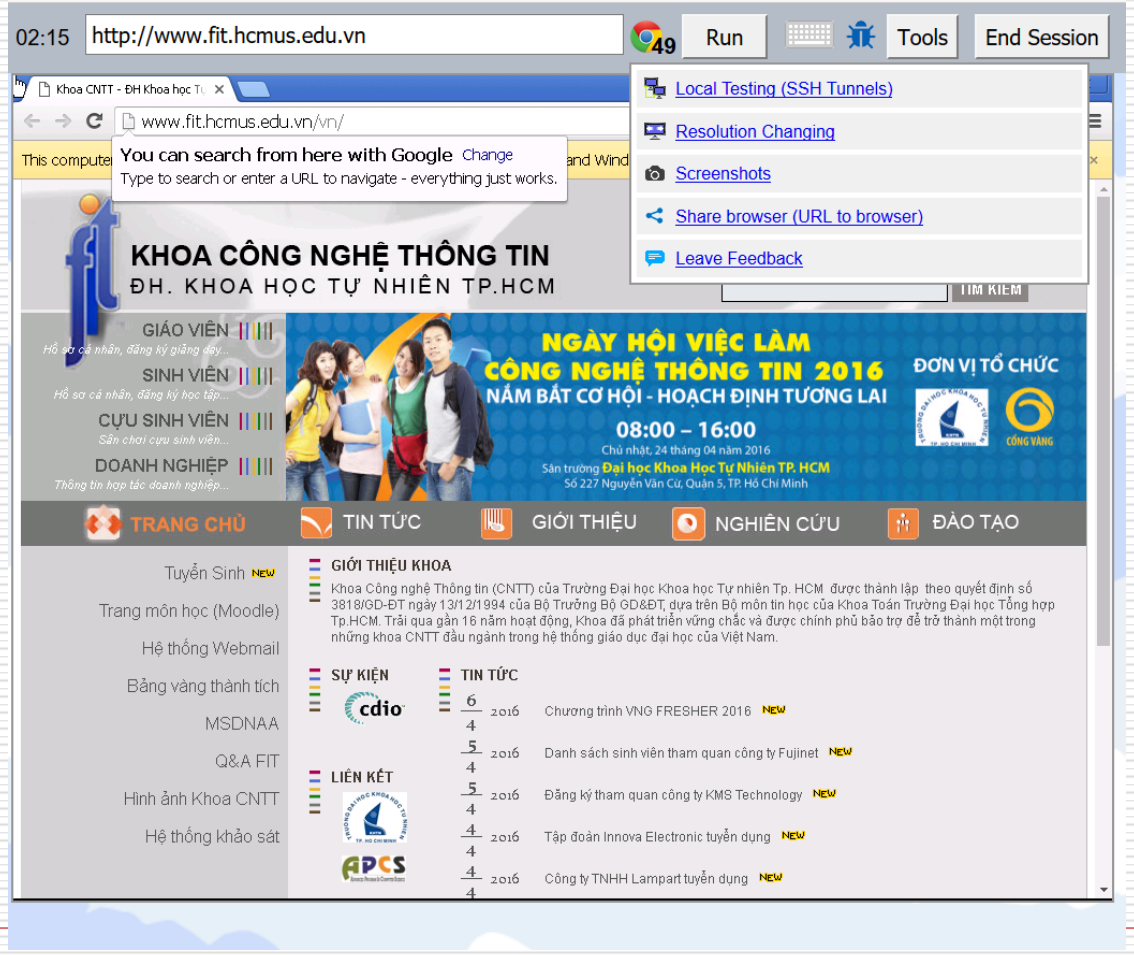
Other tools

- Link checker
- HTML validator
- CSS validator
- Browser compatibility

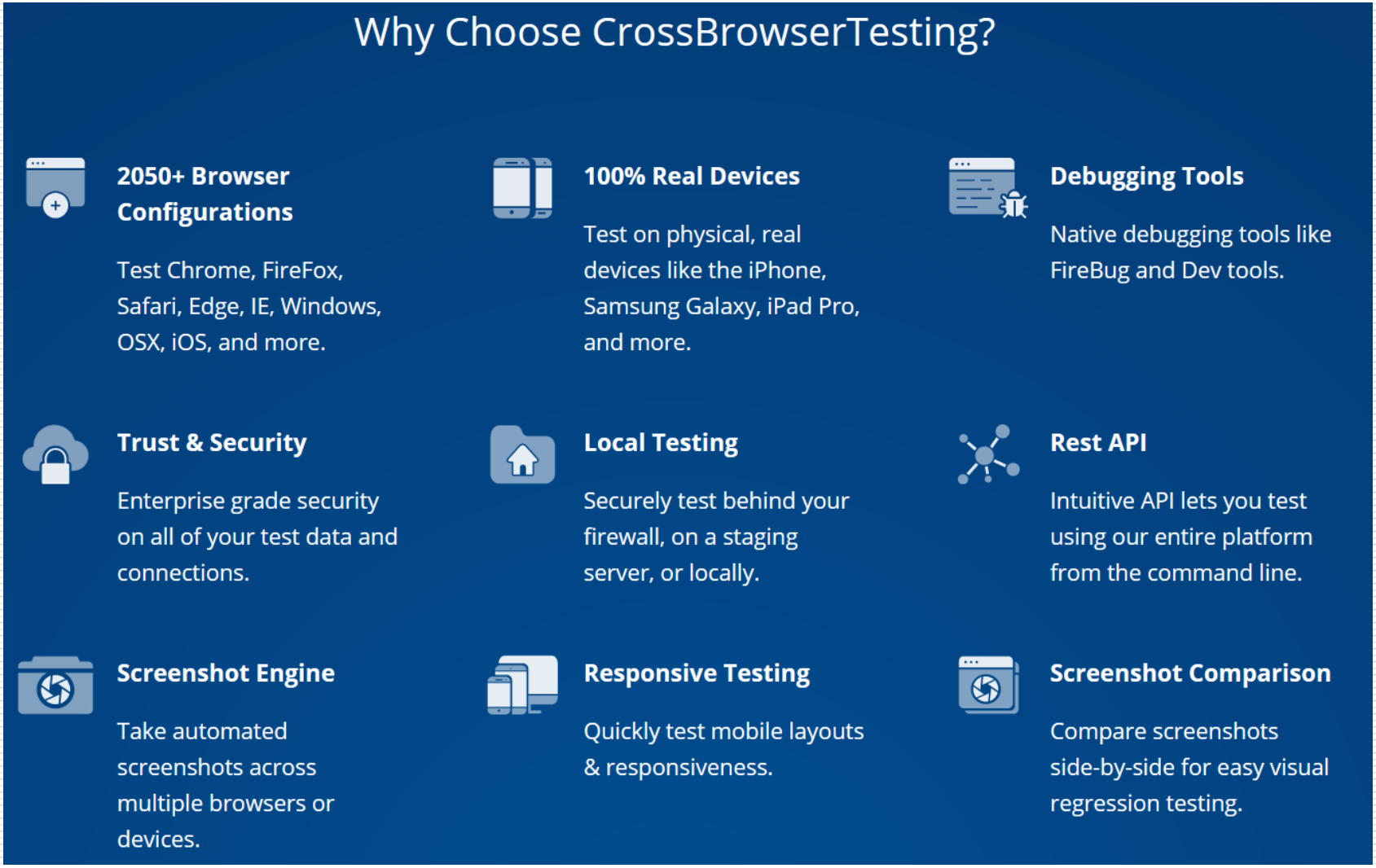
Link checker



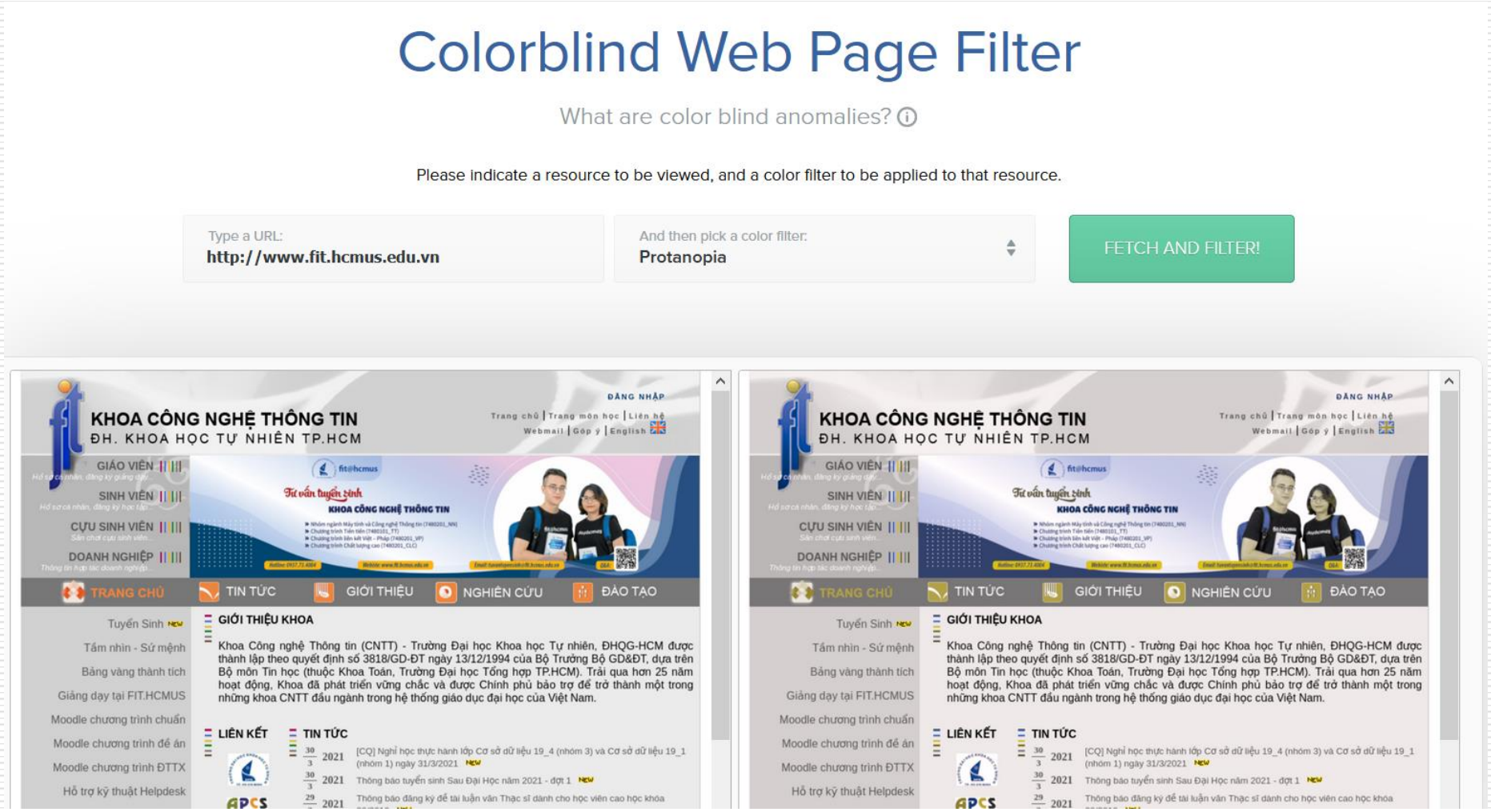
www.browserling.com

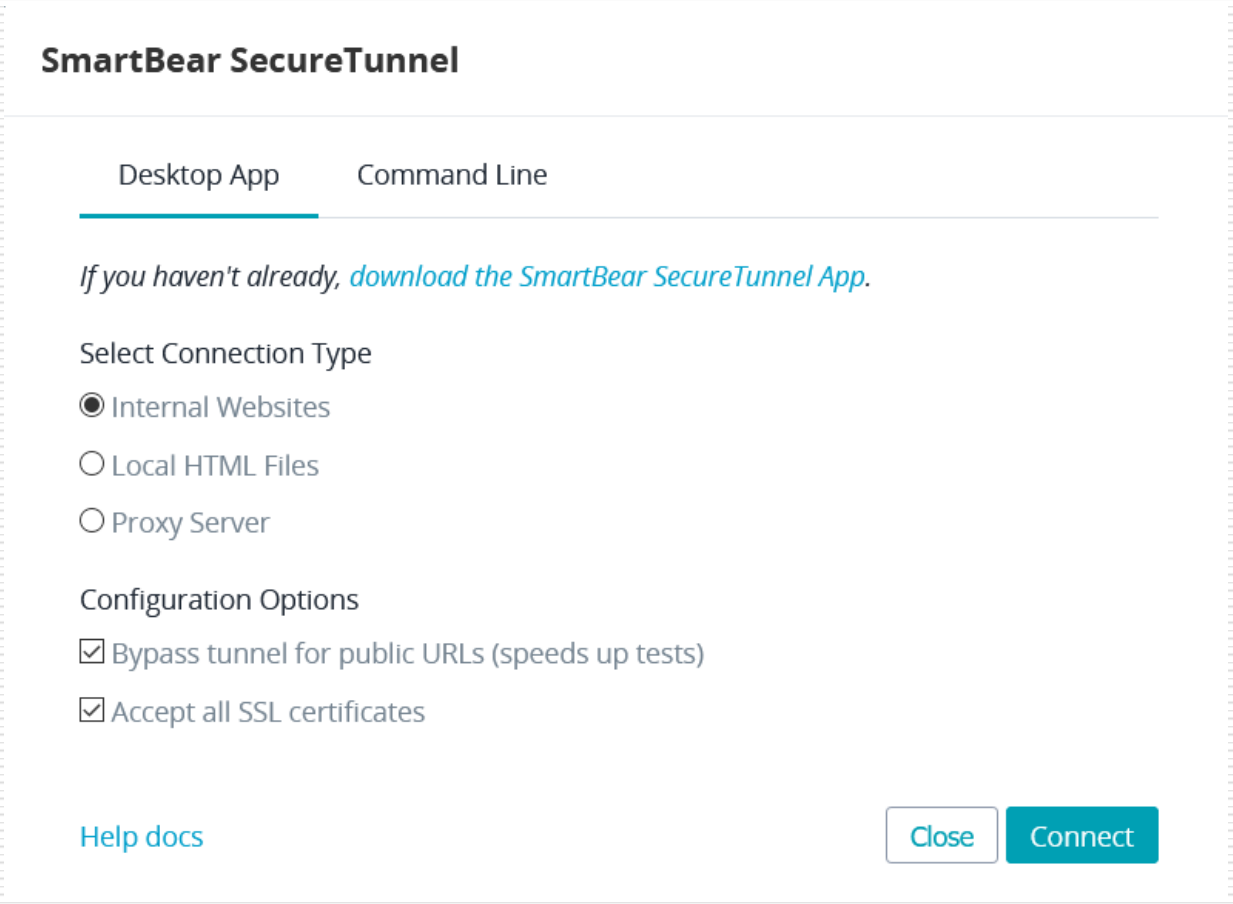


https://crossbrowsertesting.com



https://www.toptal.com/designers/colorfilter




SmartBear SecureTunnel
Desktop App
Command Line
If you haven't already, download the SmartBear SecureTunnel App.
Select Connection Type
Internal Websites
Local HTML Files
Proxy Server
Configuration Options
Bypass tunnel for public URLs (speeds up tests)
Accept all SSL certificates
Help docs
Close
Connect

Refs:

- https://uxchecklist.github.io/
- https://www.checklist.design/
- https://strongqa.com/qa-portal/testing-docs-templates/checklist
- https://www.hotjar.com/usability-testing/template-checklist/
- http://www.colorvisiontesting.com/