Web Automation with Selenium Webdriver

Andrew Malone
Harvard abcd-www
November 13, 2013

WALL·E

SELENIUM

Selenium

Selenium is a suite of tools to automate web browsers across many platforms.

Confusion!

Selenium

Selenium IDE

Selenium RC

Selenium 1

Selenium 2

Selenium Webdriver

Selenium Server

Selenium Grid

Webdriver

Selenium Webdriver makes direct calls to the browser using each browser’s native support for automation.

Current browser support

Google Chrome 12+

Internet Explorer 6, 7, 8, 9

Firefox 3+

Opera 11.5+

HtmlUnit 2.9

Android – 2.3+

iOS 3.2+

Safari 5.1+

Language bindings

Java

C#

Ruby

Python

Javascript (Node)

Perl

PHP

Haskell

Objective-C

Examples in this presentation will be in Python

GETTING UP AND RUNNING

Assumptions

Python is installed – version 2.6, 2.7, 3.2 or 3.3

Selenium Webdriver bindings are installed (available as a Python package)

Starting a browser session

```
# import webdriver
from selenium import webdriver

# start a browser session
browser = webdriver.Chrome()

# Load a web page
browser.get("http://w3.abcd.harvard.edu")
```

LOCATING ELEMENTS

Elements

Interactions are with specific elements on a page.

An element is an individual DOM node.

Webdriver location strategies

Id

Name

Tag name

Link text

Partial link text

Class name

Css selector

Xpath

Webdriver location strategies

```
<input type="text" name="textInput" id="textInput" class="input">
```

id	`browser.find_element_by_id("textInput")`
name	`browser.find_element_by_name("inputName")`
tag name	`browser.find_element_by_tag_name("input")`
class name	`browser.find_element_by_class_name("input")`
css	`browser.find_element_by_css_selector("input.input")`
xpath	`browser.find_element_by_xpath("/input[@name=textInput]")`

```
<a href="/index.html">Web page</a>
```

link text	`browser.find_element_by_link_text("Web page")`
partial link text	`browser.find_element_by_partial_link_text("page")`

Singular and Plural

Singular

```
element = browser.find_element_by_css_selector("a.link")
element.click()
```

Finds the first element that matches the selector

Returns the element

Throws a Webdriver exception if no match is found

Singular and Plural

Plural

```
elements = browser.find_elements_by_css_selector("a.link")
elements[0].click()
```

Finds all elements that match the selector

Returns a list of elements

Returns an empty list if no match is found

Alternate locator syntax

These two statements are equivalent

```
cheese = browser.find_element_by_name("cheese")
```

```
from selenium.webdriver.common.by import By
cheese = browser.find_element(By.NAME, "cheese")
```

Choosing a location strategy

Use id or name whenever possible

CSS attribute selectors and pseudoclasses can be extremely powerful

Use xpath as a last resort

CSS attribute selectors

```
<a href="index.html?form=registerForm&user=newUser">Register</a>
```

Attribute contains:

```
a[href*=registerForm]
```

```
<a href="someExternalDocument.pdf">Read this!</a>
```

Attribute ends with:

```
a[href?=pdf]
```

CSS pseudoclasses

```
<table id="datatable">
    <tr>
        <th>1</th>
        <td>Num 1</td>
        <td>Num 2</td>
    </tr>
    <tr>
        <th>2</th>
        <td>Num 3</td>
        <td>Num 4</td>
    </tr>
</table>
```

```
#datatable tr:nth-child(2) td:nth-of-type(1)
```

XPATH alternatives

```
<table id="datatable">
    <tr>
        <th>1</th>
        <td>Num 1</td>
        <td>Num 2</td>
    </tr>
    <tr>
        <th>2</th>
        <td>Num 3</td>
        <td>Num 4</td>
    </tr>
</table>
```

```
//table[@id='datatable']/tr[2]/td[1]
```

XPATH alternatives

```
<table id="datatable">
    <tr>
        <th>1</th>
        <td>Num 1</td>
        <td>Num 2</td>
    </tr>
    <tr>
        <th>2</th>
        <td>Num 3</td>
        <td>Num 4</td>
    </tr>
</table>
```

```
//th[text()='2']/following-sibling::td[1]
```

BUILDING A SIMPLE SCRIPT

Interaction methods

```
# click the element
element.click()

# type into the element
element.send_keys("some text")

# is the element visible to a user
element.is_displayed()

# get an attribute of the element
element.get_attribute("name")

# get the text of the element
element.text
```

```
from selenium import webdriver

browser = webdriver.Chrome()
browser.get("https://www.directory.harvard.edu/")
browser.find_element_by_name("lastName").send_keys("Malone")
browser.find_element_by_id("search").click()
browser.find_element_by_link_text("Andrew P. Malone").click()
xpath = "//font[text()='Appointment Title:']/../following-sibling::td"
title = browser.find_element_by_xpath(xpath).text

if title == "Senior Systems Operations Analyst":
    print "Title is correct"
else:
    print "Title is not correct"
```

HARVARD Web Directory

Tips

Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.

Search Help

Need help with your search?

Search the Public Harvard Directory

To look up Harvard students, faculty, and staff members, fill in one or more of the fields and click Search.

Last Name:
First Name:
Email:
Phone:

Search Reset

Search the Internal Directory

Authenticate and view additional information.

Authenticate Via PIN

Privacy

The compilation or redistribution of information from Harvard University directories is forbidden.

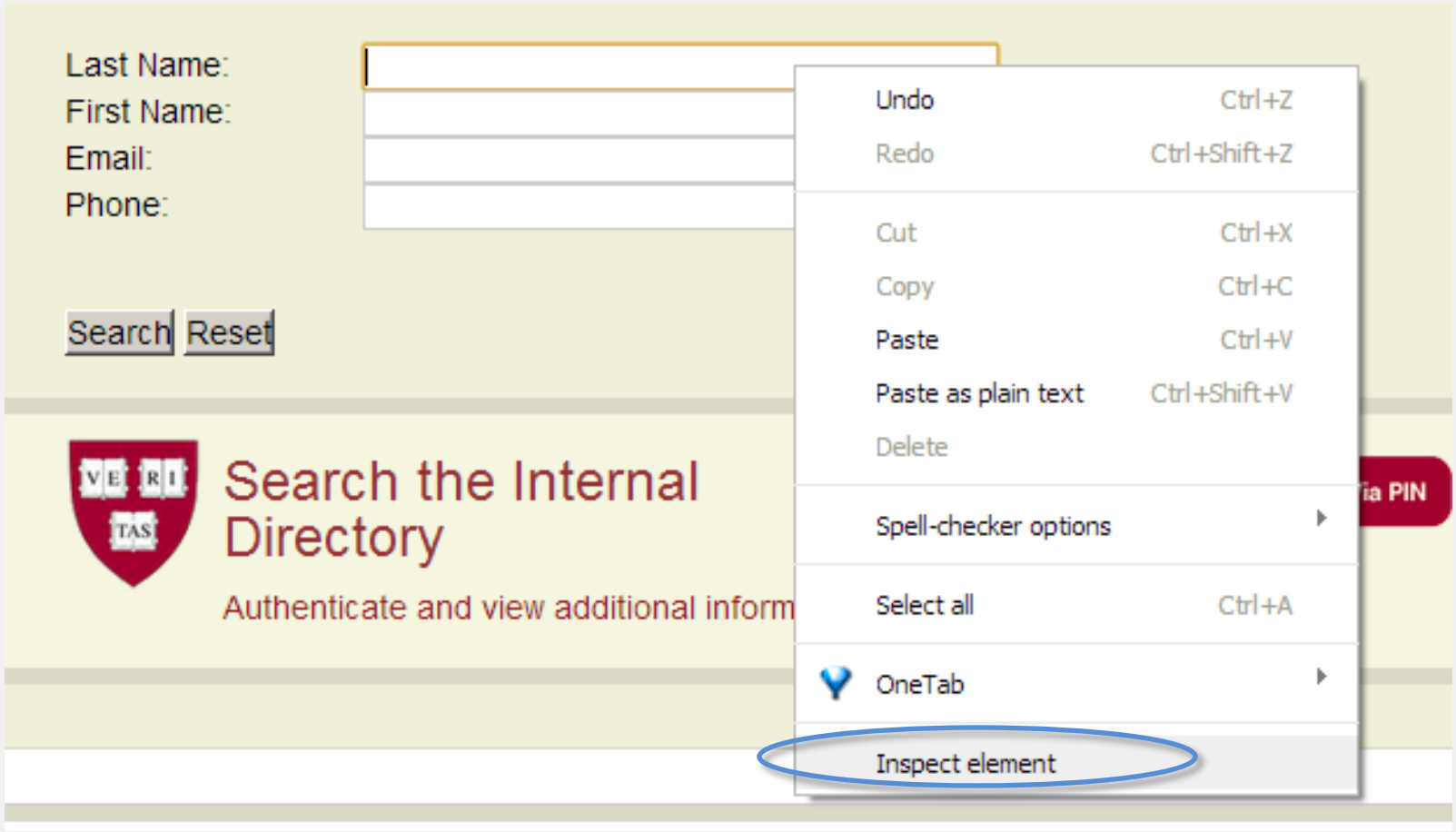
Privacy Statement

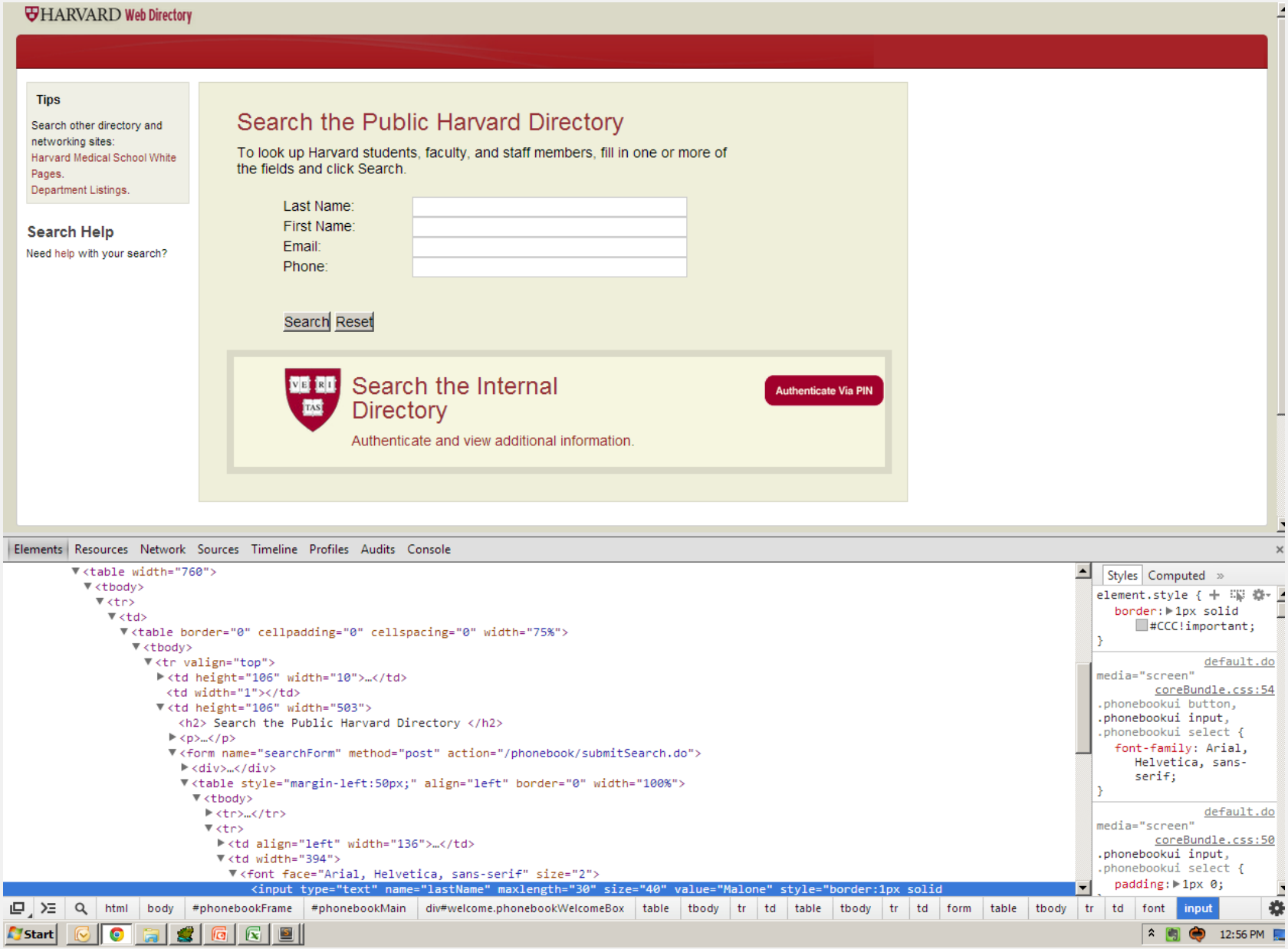
Help

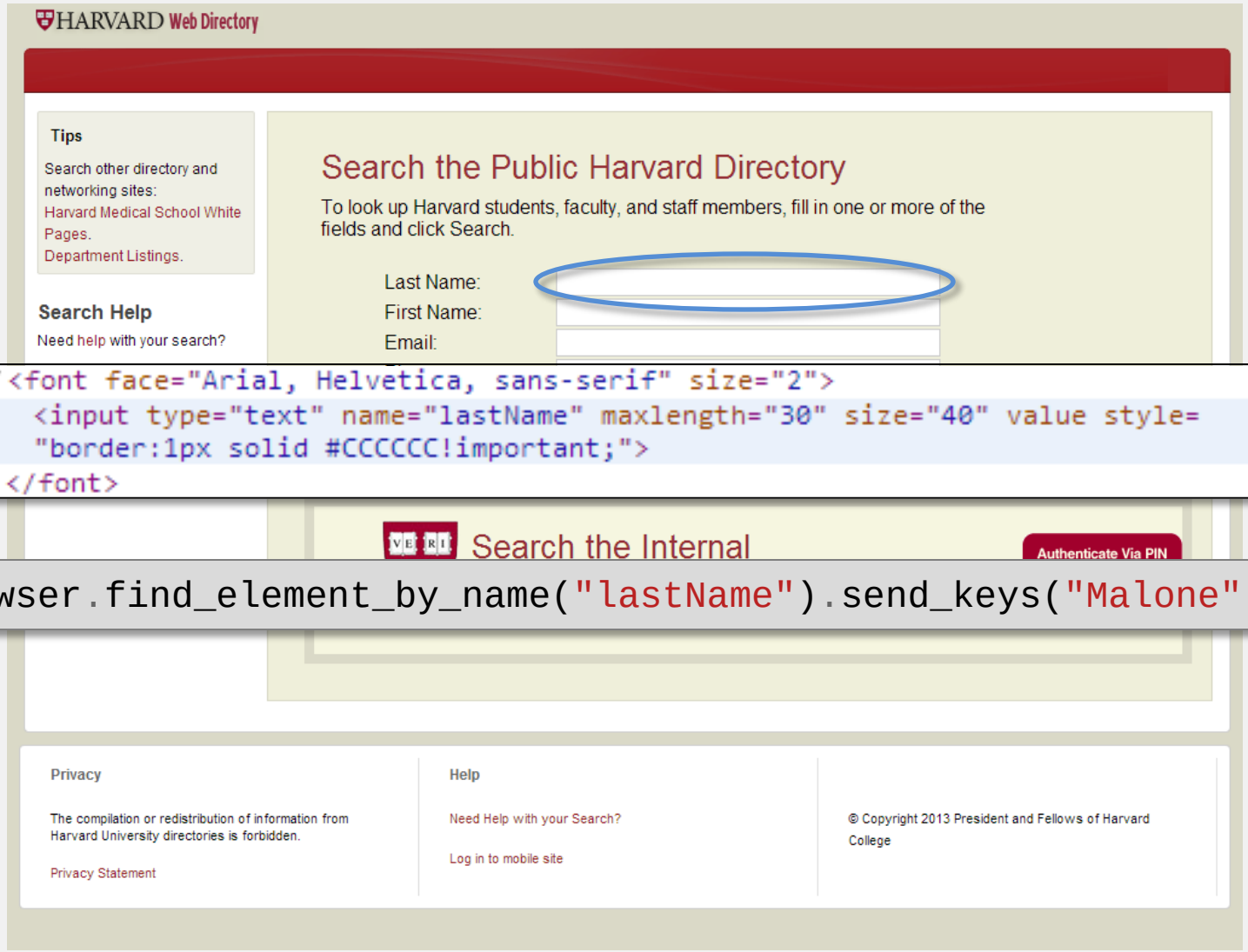
Need Help with your Search?

Log in to mobile site

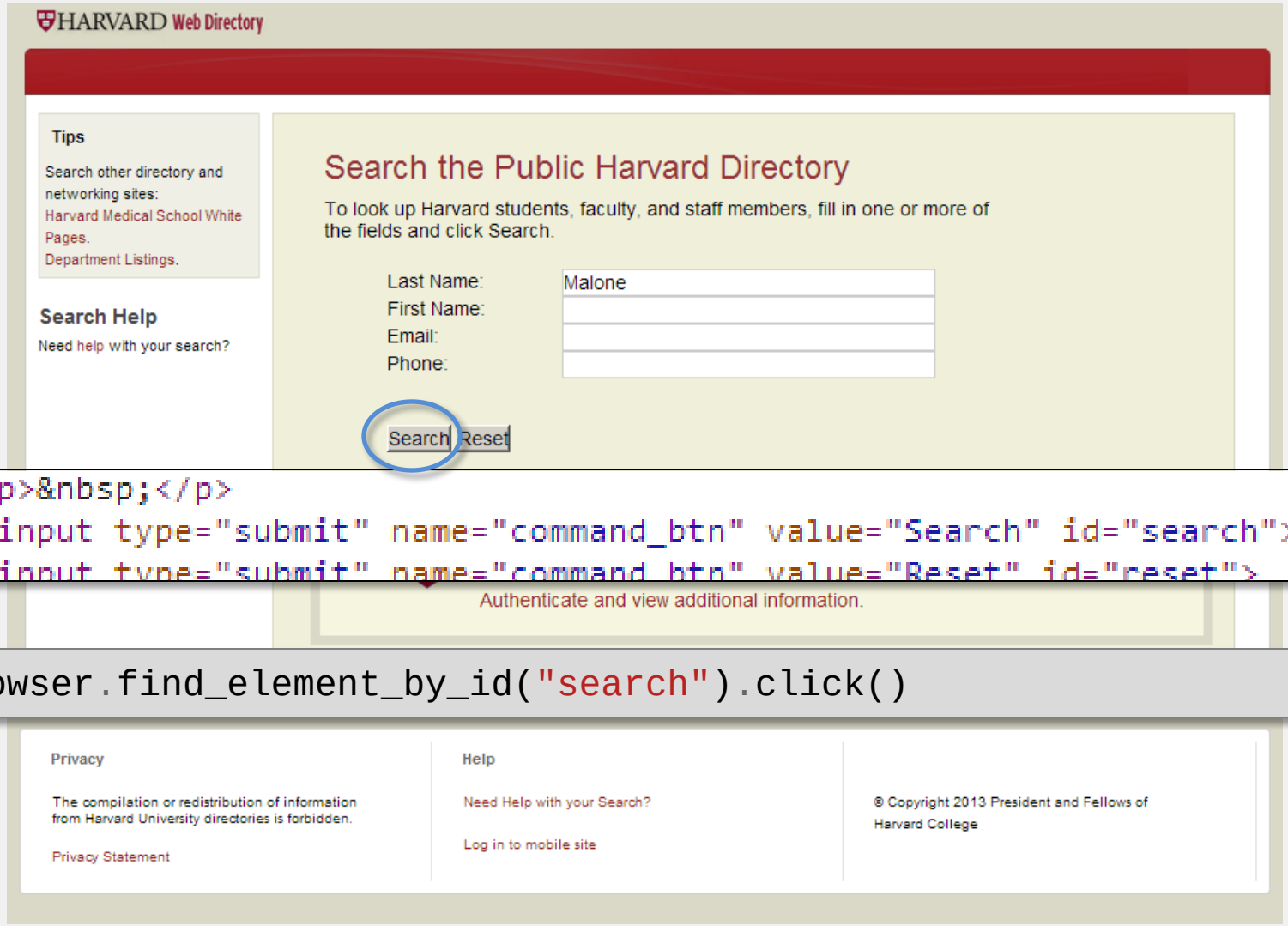
© Copyright 2013 President and Fellows of Harvard College


Last Name:
First Name:
Email:
Phone:
Search
Reset
Search the Internal Directory
Authenticate and view additional inform
Undo Ctrl+Z
Redo Ctrl+Shift+Z
Cut Ctrl+X
Copy Ctrl+C
Paste Ctrl+V
Paste as plain text Ctrl+Shift+V
Delete
Spell-checker options
Select all Ctrl+A
OneTab
Inspect element

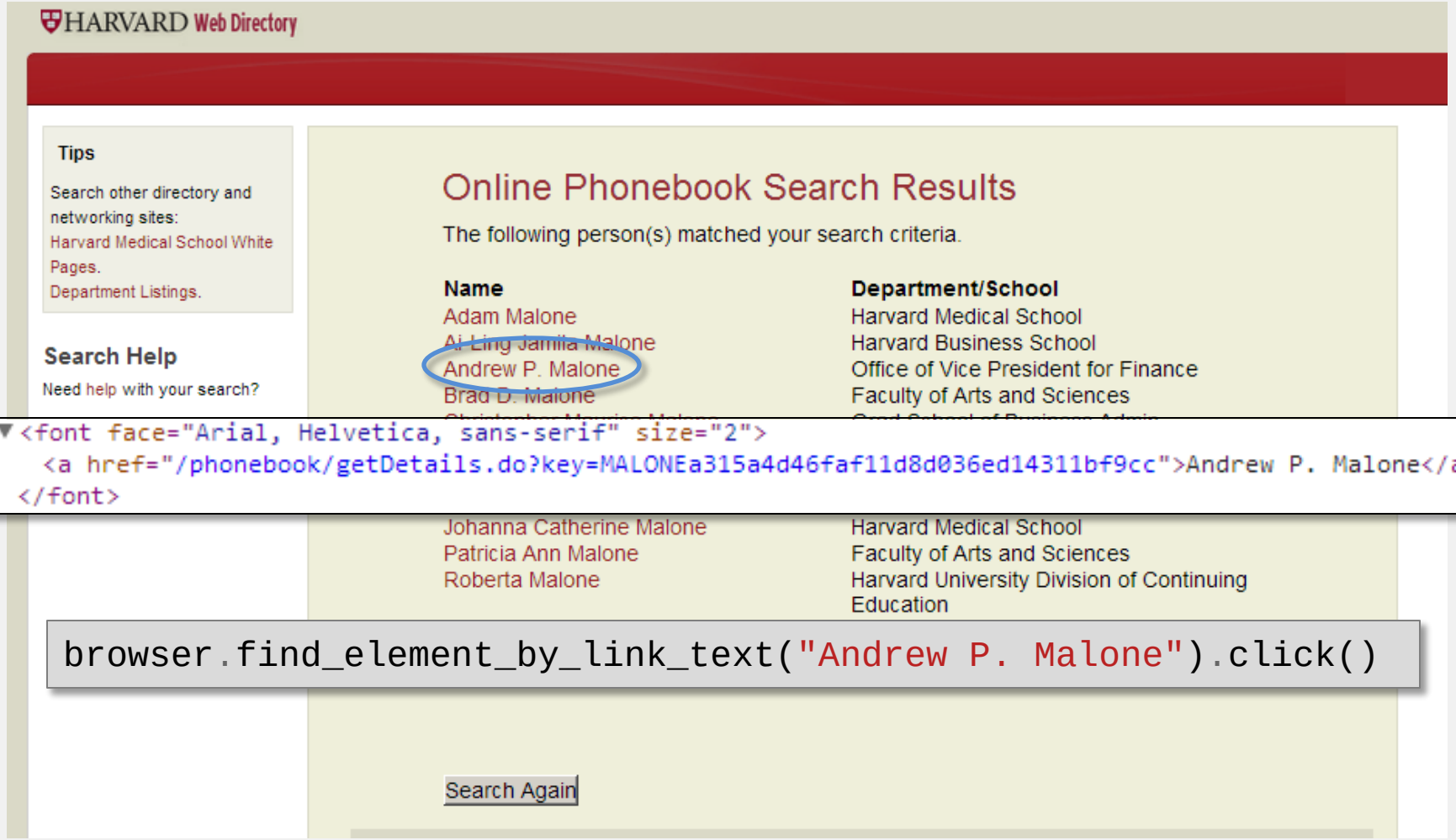

HARVARD Web Directory
Tips
Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.
Search Help
Need help with your search?
Search the Public Harvard Directory
To look up Harvard students, faculty, and staff members, fill in one or more of the fields and click Search.
Last Name:
First Name:
Email:
Phone:
Search
Reset
Search the Internal Directory
Authenticate and view additional information.
Authenticate Via PIN
Elements
Resources
Network
Sources
Timeline
Profiles
Audits
Console
<table width="760">
<tbody>
<tr>
<td>
<table border="0" cellpadding="0" cellspacing="0" width="75%">
<tbody>
<tr valign="top">
<td height="106" width="10">…</td>
<td width="1"></td>
<td height="106" width="503">
<h2> Search the Public Harvard Directory </h2>
…
<form name="searchForm" method="post" action="/phonebook/submitSearch.do">
…
<table style="margin-left:50px;" align="left" border="0" width="100%">
<tbody>
<tr>…</tr>
<tr>
<td align="left" width="136">…</td>
<td width="394">
<font face="Arial, Helvetica, sans-serif" size="2">
<input type="text" name="lastName" maxlength="30" size="40" value="Malone" style="border:1px solid
Styles
Computed
element.style {
border: 1px solid #CCC!important;
}
default.do
media="screen"
coreBundle.css:54
.phonebookui button,
.phonebookui input,
.phonebookui select {
font-family: Arial, Helvetica, sans-serif;
}
default.do
media="screen"
coreBundle.css:50
.phonebookui input,
.phonebookui select {
padding: 1px 0;
html
body
#phonebookFrame
#phonebookMain
div#welcome.phonebookWelcomeBox
table
tbody
tr
td
table
tbody
tr
td
form
table
tbody
tr
td
font
input
Start
12:56 PM




```
browser.find_element_by_name("lastName").send_keys("Malone")
```

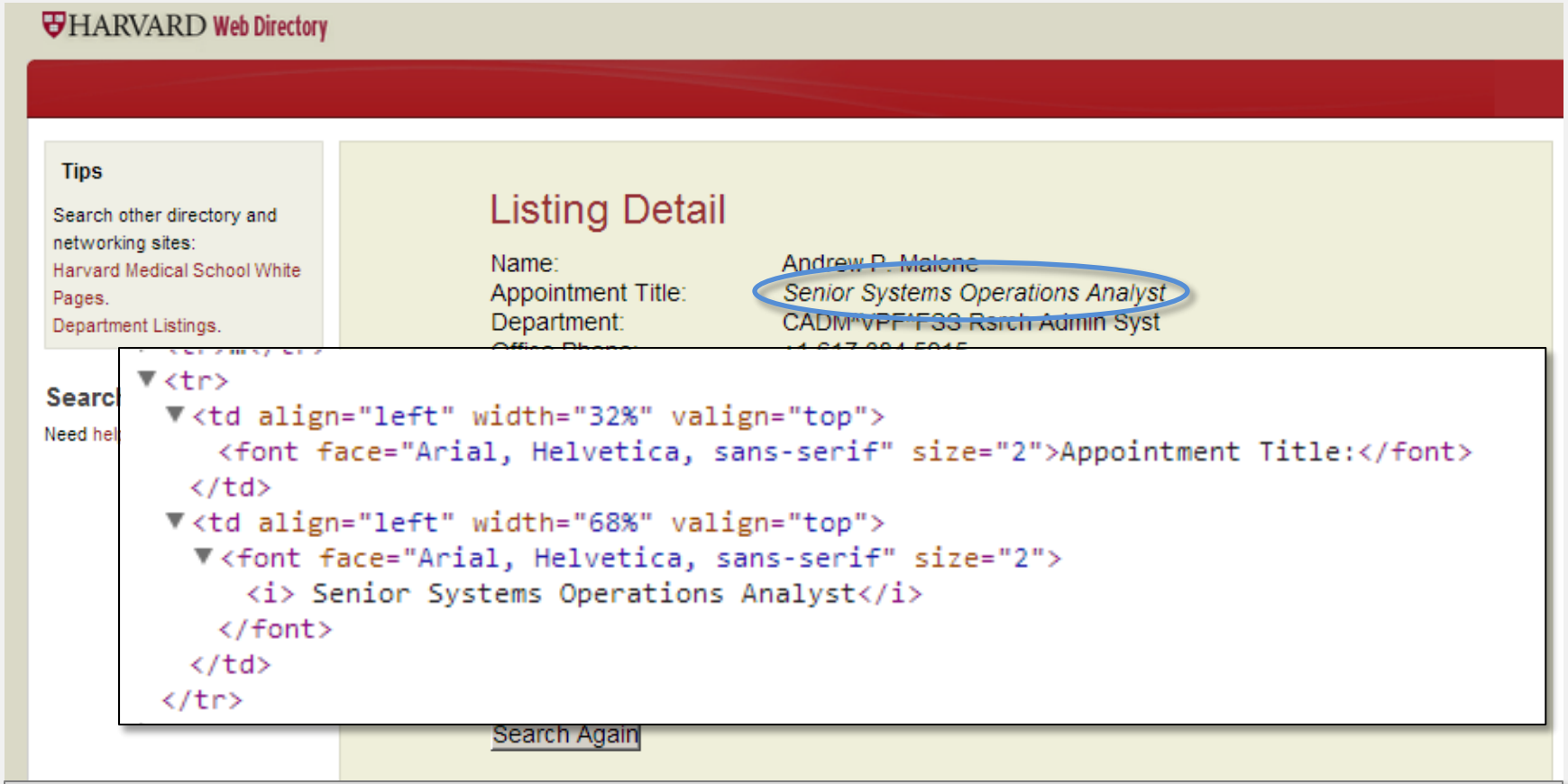

HARVARD Web Directory
Tips
Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.
Search Help
Need help with your search?
Search the Public Harvard Directory
To look up Harvard students, faculty, and staff members, fill in one or more of the fields and click Search.
Last Name:
Malone
First Name:
Email:
Phone:
Search
Reset

<input type="submit" name="command_btn" value="Search" id="search">
<input type="submit" name="command_btn" value="Reset" id="reset">
Authenticate and view additional information.
browser.find_element_by_id("search").click()
Privacy
The compilation or redistribution of information from Harvard University directories is forbidden.
Privacy Statement
Help
Need Help with your Search?
Log in to mobile site
© Copyright 2013 President and Fellows of Harvard College


HARVARD Web Directory
Tips
Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.
Search Help
Need help with your search?
Online Phonebook Search Results
The following person(s) matched your search criteria.
Name
Department/School
Adam Malone
Harvard Medical School
Ai-Ling Jamila Malone
Harvard Business School
Andrew P. Malone
Office of Vice President for Finance
Brad D. Malone
Faculty of Arts and Sciences

Andrew P. Malone

Johanna Catherine Malone
Harvard Medical School
Patricia Ann Malone
Faculty of Arts and Sciences
Roberta Malone
Harvard University Division of Continuing Education
browser.find_element_by_link_text("Andrew P. Malone").click()
Search Again




```
xpath = "//font[text()='Appointment Title:']/../following-sibling::td"
title = browser.find_element_by_xpath(xpath).text
```

```
from selenium import webdriver

browser = webdriver.Chrome()
browser.get("https://www.directory.harvard.edu/")
browser.find_element_by_name("lastName").send_keys("Malone")
browser.find_element_by_id("search").click()
browser.find_element_by_link_text("Andrew P. Malone").click()
xpath = "//font[text()='Appointment Title:']/../following-sibling::td"
title = browser.find_element_by_xpath(xpath).text

if title == "Senior Systems Operations Analyst":
    print "Title is correct"
else:
    print "Title is not correct"
```

Tips

Learn and use the web inspector and browser console

$(“css selector”)

$x(“xpath selector”)

PAGE OBJECTS

Page object pattern

Keep all the webdriver interactions in one place

Adds a layer of abstraction between webdriver implementation and website/application functionality

HARVARD Web Directory

Tips

Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.

Search Help

Need help with your search?

Search the Public Harvard Directory

To look up Harvard students, faculty, and staff members, fill in one or more of the fields and click Search.

Last Name:
First Name:
Email:
Phone:

Search Reset

Search the Internal Directory

Authenticate and view additional information.

Authenticate Via PIN

Privacy

The compilation or redistribution of information from Harvard University directories is forbidden.

Privacy Statement

Help

Need Help with your Search?

Log in to mobile site

© Copyright 2013 President and Fellows of Harvard College

Search page

```
class Search_Page(Page):
    locators = {
        "last_name": "name=lastName",
        "search": "id=search"
    }

    def go(self):
        self.browser.get("https://www.directory.harvard.edu/")
        return self

    def set_last_name(self, last_name):
        self.find(self.locators["last_name"]).send_keys(last_name)

    def search(self):
        self.find(self.locators["search"]).click()
        return Results_List(self.browser)
```

Base page object

```
class Page(object):
    def __init__(self, browser):
        self.browser = browser

    def find(self, locator):
        # convenience method for writing simpler locators
        from selenium.webdriver.common.by import By
        locator_map = {
            "css": By.CSS_SELECTOR,
            "id": By.ID,
            "name": By.NAME,
            "link": By.LINK_TEXT,
            "xpath": By.XPATH
        }
        locator_type = locator[:locator.find("=")]
        locator_value = locator[locator.find("=") + 1:]
        return self.browser.find_element(locator_map[locator_type],
                        locator_value)
```

Search page

```
class Search_Page(Page):
    locators = {
        "last_name": "name=lastName",
        "search": "id=search"
    }

    def go(self):
        self.browser.get("https://www.directory.harvard.edu/")
        return self

    def set_last_name(self, last_name):
        self.find(self.locators["last_name"]).send_keys(last_name)

    def search(self):
        self.find(self.locators["search"]).click()
        return Results_List(self.browser)
```

HARVARD Web Directory

Tips

Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.

Search Help

Need help with your search?

Online Phonebook Search Results

The following person(s) matched your search criteria.

Name	Department/School
Adam Malone	Harvard Medical School
Ai-Ling Jamila Malone	Harvard Business School
Andrew P. Malone	Office of Vice President for Finance
Brad D. Malone	Faculty of Arts and Sciences
Christopher Maurice Malone	Grad School of Business Admin
Clare Felicity Malone	Grad School of Arts and Sci
Daniel James Malone	Graduate School of Design
David Malone	Campus Services
Johanna Catherine Malone	Harvard Medical School
Patricia Ann Malone	Faculty of Arts and Sciences
Roberta Malone	Harvard University Division of Continuing Education
Ryan Malone	DCE - Harvard Extension School
Sean Ryne Malone	Harvard College

Search Again

Results list

```
class Results_List(Page):
    locators = {
        "name": "link=REPLACE"
    }

    def click_name(self, name):
        self.find(self.locators["name"]
              .replace("REPLACE", name)).click()
        return Results_Page(self.browser)
```

HARVARD Web Directory

Tips

Search other directory and networking sites:
Harvard Medical School White Pages.
Department Listings.

Search Help

Need help with your search?

Listing Detail

Name:	Andrew P. Malone
Appointment Title:	*Senior Systems Operations Analyst*
Department:	CADM^VPF^FSS Rsrch Admin Syst
Office Phone:	+1 617 384 5915
	8 Story St. 209
Office Fax:	+1 617 384 8324
University Mailing Address:	Harvard Univ Suite 209 8 Story St Cambridge MA 02138
Email:	andrew_malone@harvard.edu
Unit:	Office of Vice President for Finance

Return to Query Results
Search Again

Results page

```
class Results_Page(Page):
    locators = {
        "title":"xpath=//font[text()='Appointment
                    Title:']/../following-sibling::td"
    }

    def title(self):
        return self.find(self.locators["title"]).text
```

```
from selenium import webdriver
browser = webdriver.Chrome()
search_page = Search_Page(browser).go()
search_page.set_last_name("Malone")
results_list = search_page.search()
results_page = results_list.click_name("Andrew P. Malone")
title = results_page.title()
```

Why use page objects?

When underlying source of pages change, updates only need to be made in one place

Allows script writers to focus on tests, not on details of webdriver interactions

WHAT CAN YOU USE THIS FOR?

Streamline testing efficiency

Data staging

Creating specific application state for testing

Other uses

Automated test execution

Data driven test

Cross browser testing

Running Selenium on a server

Integration with a build process

RISKS

“Tools can extend or enhance or enable fabulous testing. But by accelerating some action, tools can enable us to do bad testing faster than ever...”

http://www.developsense.com/blog/2013/02/manual-and-automated-testing/

Automation is software development

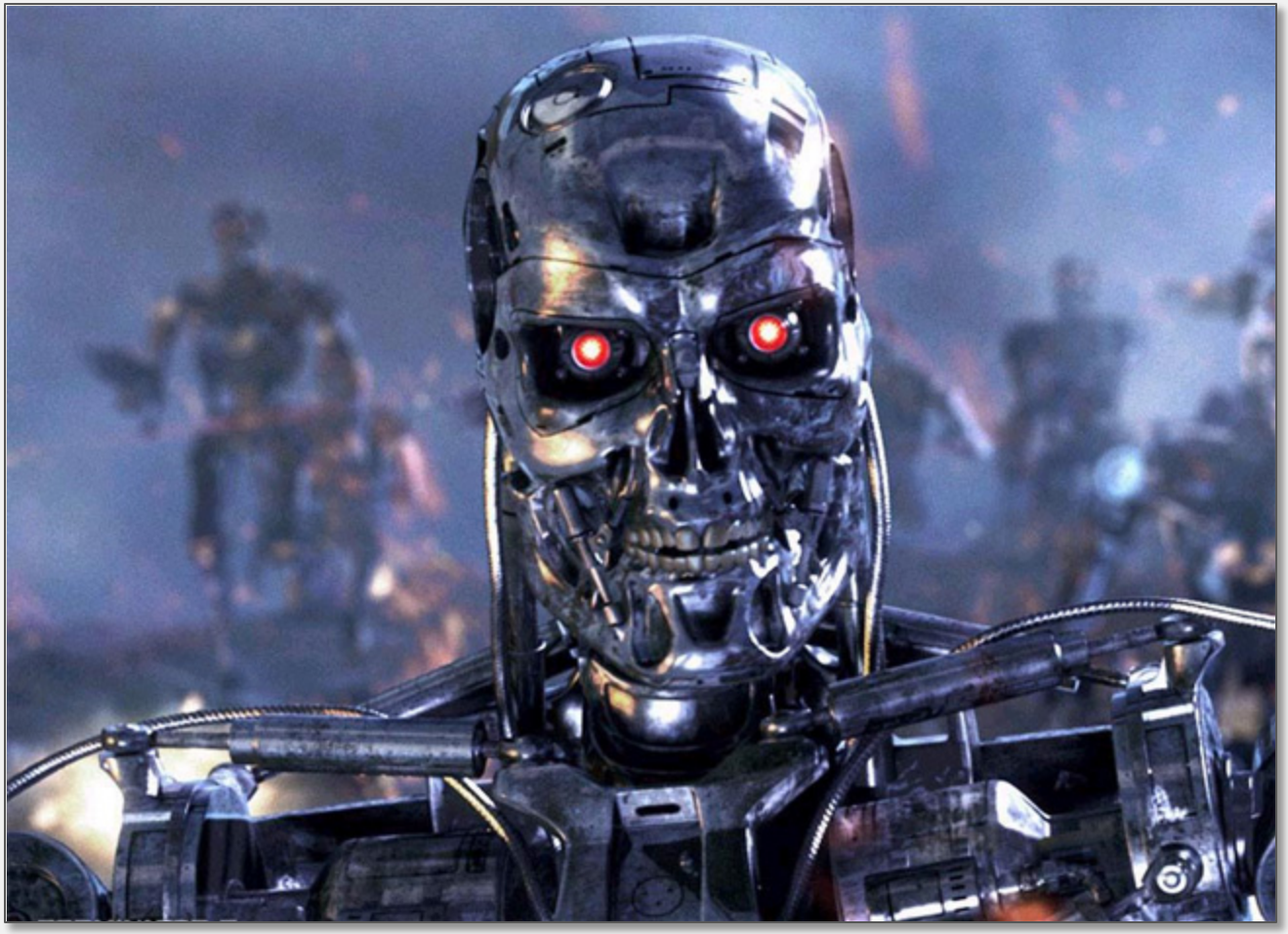
Does your team have the right skills to create maintainable, extendable, reliable code?

Scope creep

Aiming for 100% test coverage is not the right goal

Focus instead on reducing risk or increasing efficiency

Don’t be like this…



Be like this!

THANK YOU!

Andrew Malone

Financial Systems Solutions (FSS)
andrew_malone@harvard.edu

External resources

Selenium project homepage

http://www.seleniumhq.org/

Python webdriver bindings

http://selenium-python.readthedocs.org/en/latest/

GMAS automation project

https://git.huit.harvard.edu/gmas_webdriver