

Intro to Web Scraping 1

April 1, 2019

0.1 Tutorial contents

Part 1: Introduction * Section ?? * Section ??

Part 2: Basic web scraping * Section ?? * Section ?? * Section ?? * Section ??

Part 3: Advanced topics * Section ??

Part 4: Other advice * Section ?? * Section ??

1 Part I: Introduction

1.1 What is web scraping?

Collecting information from web pages. There are multiple ways of doing this, such as writing your own scraper, using APIs, using boxed software and tools. Today we will focus on the first type, but I will provide resources that cover the other two broad categories.

1.2 Libraries

Multiple useful libraries. The main ones that we'll be using today are [Requests](#), [BeautifulSoup](#), and [Selenium WebDriver](#).

```
In [2]: import requests
        from bs4 import BeautifulSoup
        import selenium
```

2 Part II: Basic web scraping

2.1 Making web requests

When you open a web page, you send instructions to a web server, requesting for the information to be displayed in your browser. The server in return sends the code (usually HTML) to your browser. The browser translates the code using a Document Object Model(DOM).

Let's make a request for the IOC webpage.

```
In [3]: #Specify the url
        url = "https://www.iocsummerschoolexeter.org/"
        # Query the website and return the html.
        # Save the html as a new variable page.
        page = requests.get(url)
```

We can inspect the content of the page object, which is what we'll be working on later.

```
In [4]: page.content[0:3000]
```

```
Out [4]: b'<!DOCTYPE html>\n<html lang="en">\n\t<head><script src="/gdpr/gdprscript.js?buildTime
```

2.2 Intro to HTML

The response looks intimidating, but once you understand basic HTTP things look a lot easier. Let's have a look at what this code would look like for a very simple page. Go to www.example.com, and right click anywhere in the page, selecting "View page source". You need to be aware of the simple elements of syntax for HTML pages. Every tag serves a block inside the webpage: 1.

: HTML documents must start with a type declaration. 2. The HTML document is contained between `html` and `/html`. 3. The meta and script declaration of the HTML document is between `head` and `/head`. 4. The visible part of the HTML document is between `body` and `/body` tags. 5. Title headings are defined with the `h1` through `hn` tags. 6. Paragraphs are defined with the `p` tag.

Other useful tags include `a` for hyperlinks, `table` for tables, `tr` for table rows, and `td` for table columns.

HTML tags sometimes come with `id` or `class` attributes. The `id` attribute specifies a unique id for an HTML tag and the value must be unique within the HTML document. The `class` attribute is used to define equal styles for HTML tags with the same class. We can make use of these ids and classes to help us locate the data we want.

2.3 Parsing HTML

The next step is to parse the HTTP and get the info that is relevant to us.

```
In [5]: #Parse the html in the 'page' variable, and store it in BeautifulSoup format
        soup = BeautifulSoup(page.text, 'html.parser')
        soup
```

```
Out [5]: <!DOCTYPE html>
```

```
<html lang="en">
<head><script src="/gdpr/gdprscript.js?buildTime=1553821729&amp;hasRemindMe=true&amp;s
<title>The Institute of Coding Summer School at the University of Exeter - Home</title>
<meta content="The Institute of Coding Summer School at the University of Exeter" prop
<meta content="The Institute of Coding (IoC) Summer School at the University of Exeter
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/publish
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/publish
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/publish
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/publish
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/editor/3
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/ks0a2107
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/c57a7516
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/c57a7563
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/c57a7522
<meta content="https://www.iocsummerschoolexeter.org/uploads/2/3/9/8/23981353/c57a7514
```

```

<meta content="https://www.iocsummerschoolexeter.org/" property="og:url"/>
<meta content="The Institute of Coding (IoC) Summer School at the University of Exeter"
<meta content="coding, institute, hefce, university, exeter, data, algorithm, analytics"
<meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
<meta content="width=device-width, initial-scale=1.0" name="viewport"/>

```

512917575486359

```

<link href="//cdn2.editmysite.com/css/sites.css?buildTime=1550172868"
<link href="//cdn2.editmysite.com/css/old/fancybox.css?1550172868" rel="stylesheet" type="text/css"
<link href="//cdn2.editmysite.com/css/social-icons.css?buildtime=1550172868" media="screen"
<link href="/files/main_style.css?1550225720" rel="stylesheet" title="wsite-theme-css"
<link href="//fonts.googleapis.com/css?family=Roboto+Slab:400,300,700&subset=latin"
<link href="//fonts.googleapis.com/css?family=Fjalla+One&subset=latin,latin-ext"
<link href="//fonts.googleapis.com/css?family=Montserrat:400,700&subset=latin,latin-ext"
<link href="//fonts.googleapis.com/css?family=Lato:400,300,300italic,700,400italic,700italic"
<style type="text/css">
.wsite-elements.wsite-not-footer:not(.wsite-header-elements) div.paragraph, .wsite-elements.wsite-not-footer
#wsite-content div.paragraph, #wsite-content p, #wsite-content .product-block .product-title, #wsite-content
.wsite-elements.wsite-footer div.paragraph, .wsite-elements.wsite-footer p, .wsite-elements.wsite-footer
.wsite-elements.wsite-not-footer:not(.wsite-header-elements) h2, .wsite-elements.wsite-not-footer
#wsite-content h2, #wsite-content .product-long .product-title, #wsite-content .product-long .product-title,
.wsite-elements.wsite-footer h2, .wsite-elements.wsite-footer .product-long .product-title,
#wsite-title {}
.wsite-not-footer h2.wsite-content-title a, .wsite-not-footer .paragraph a, .wsite-not-footer
.wsite-footer h2 a, .wsite-footer .paragraph a, .wsite-footer blockquote a {color:#c2a5c0}
.wsite-menu-default a {font-family:"Arial" !important;text-transform: none !important}
.wsite-menu a {}
.wsite-image div, .wsite-caption {}
.galleryCaptionInnerText {}
.fancybox-title {}
.wslide-caption-text {}
.wsite-phone {}
.wsite-headline,.wsite-header-section .wsite-content-title {font-family:"Arial" !important}
.wsite-headline-paragraph,.wsite-header-section .paragraph {font-family:"Arial" !important}
.wsite-button-inner {font-family:"Arial" !important;font-weight:700 !important;text-transform: none !important}
.wsite-not-footer blockquote {}
.wsite-footer blockquote {}
.blog-header h2 a {}
#wsite-content h2.wsite-product-title {}
.wsite-product .wsite-product-price a {}
.wsite-not-footer h2.wsite-content-title a:hover, .wsite-not-footer .paragraph a:hover
@media screen and (min-width: 767px) {.wsite-elements.wsite-not-footer:not(.wsite-header-elements)
#wsite-content div.paragraph, #wsite-content p, #wsite-content .product-block .product-title, #wsite-content
.wsite-elements.wsite-footer div.paragraph, .wsite-elements.wsite-footer p, .wsite-elements.wsite-footer
.wsite-elements.wsite-not-footer:not(.wsite-header-elements) h2, .wsite-elements.wsite-not-footer
#wsite-content h2, #wsite-content .product-long .product-title, #wsite-content .product-long .product-title,
.wsite-elements.wsite-footer h2, .wsite-elements.wsite-footer .product-long .product-title,
#wsite-title {}

```

```

.wsite-menu-default a {}
.wsite-menu a {}
.wsite-image div, .wsite-caption {}
.galleryCaptionInnerText {}
.fancybox-title {}
.wslide-caption-text {}
.wsite-phone {}
.wsite-headline, .wsite-header-section .wsite-content-title {}
.wsite-headline-paragraph, .wsite-header-section .paragraph {}
.wsite-button-inner {}
.wsite-not-footer blockquote {}
.wsite-footer blockquote {}
.blog-header h2 a {}
#wsite-content h2.wsite-product-title {}
.wsite-product .wsite-product-price a {}
}</style>
<style type="text/css">.wsite-com-category-product-group .wsite-com-column { width: 33
<script src="/files/templateArtifacts.js?1550225720"></script>
<script>
var STATIC_BASE = '//cdn1.editmysite.com/';
var ASSETS_BASE = '//cdn2.editmysite.com/';
var STYLE_PREFIX = 'wsite';
</script>
<script src="https://ajax.googleapis.com/ajax/libs/jquery/1.8.3/jquery.min.js"></script>
<script src="//cdn2.editmysite.com/js/lang/en/stl.js?buildTime=1550172868&" type="
<script src="//cdn2.editmysite.com/js/site/main.js?buildTime=1550172868"></script><scri
        function initCustomerAccountsModels() {
                                (function(){_W.setup_rpc({"url":"/ajax/api/
_W.setup_model_rpc({"rpc_namespace": "_W.CustomerAccounts.RPC", "model_namespace": "_W.Cus
})();
        }
        if(document.createEvent && document.addEventListener) {
                var initEvt = document.createEvent('Event');
                initEvt.initEvent('customerAccountsModelsInitialized', true, fa
                document.dispatchEvent(initEvt);
        } else if(document.documentElement.initCustomerAccountsModels === 0){
                document.documentElement.initCustomerAccountsModels++
        }
</script>
<script type="text/javascript"> _W = _W || {}; _W.securePrefix='www.iocsummerschoolexe
        _W.customerLocale = "en_GB";
        _W.storeName = null;
        _W.storeCountry = "GB";
        _W.storeCurrency = "GBP";
        _W.storeEuPrivacyPolicyUrl = "";
        com_currentSite = "299388504590429306";
        com_userID = "23981353";</script><script type="text/javascript"
                _W = _W || {};

```

```

_W.Commerce = _W.Commerce;
_W.Commerce.hasCart = true;
</script><script> base_context=
<script>function initCommerceModels() { (function(){_W.setup_rpc({"url":"\ajax\api\api\_W.setup_model_rpc({"rpc_namespace":"_W.Commerce.RPC","model_namespace":"_W.Commerce.B
})(); }</script>
<script src="//cdn2.editmysite.com/js/site/commerce-core.js?buildTime=1550172868"></script>
<script src="//cdn2.editmysite.com/js/site/main-commerce-browse.js?buildTime=1550172868"></script>
<script src="//cdn2.editmysite.com/js/lang/en/stl.js?buildTime=1550172868&" type="text/javascript"></script>

function initFlyouts(){
    initPublishedFlyoutMenus(
        [{"id":"114533134437475288","title":"Home","url":"index.html",
        "114533134437475288",
        '',
        'active',
        false,
        {"navigation\item":"<li {{#id}}id=\"{{id}}\"{{/\id}} class=\"w
        {"hasCustomMinicart":true}
    )
}
//-->
</script>
</head>
<body class="header-page wsite-page-index collapsed-menu-off banner-bg-on wsite-theme-1">
<div class="wrapper">
<div class="paris-header">
<div class="container">
<div class="nav dummy-nav wsite-menu-default"><ul class="site-menu">
<li class="wsite-menu-item-wrap" id="active">
<a class="wsite-menu-item" href="/">
        Home
    </a>
</li>
<li class="wsite-menu-item-wrap" id="pg537531721210351813">
<a class="wsite-menu-item" href="/about.html">
        About
    </a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-447875943371678234">
<a class="wsite-menu-subitem" href="/instructors.html">
<span class="wsite-menu-title">
        Instructors
    </span>
</a>
</li>

```

```

<li class="wsite-menu-subitem-wrap " id="wsite-nav-428449932915725692">
<a class="wsite-menu-subitem" href="/contact.html">
<span class="wsite-menu-title">
    Contact
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-115774652140471228">
<a class="wsite-menu-subitem" href="/privacy-statement.html">
<span class="wsite-menu-title">
    Privacy Statement
</span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg750005411943018735">
<a class="wsite-menu-item" href="/structure.html">
    Structure
</a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-281688027203735335">
<a class="wsite-menu-subitem" href="/learning-environment.html">
<span class="wsite-menu-title">
    Learning Environment
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-868244784447957774">
<a class="wsite-menu-subitem" href="/block-1---programme-outline.html">
<span class="wsite-menu-title">
    Block 1 - Programme Outline
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-181475333474170941">
<a class="wsite-menu-subitem" href="/block-2---programme-outline.html">
<span class="wsite-menu-title">
    Block 2 - Programme Outline
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-761909471185475700">
<a class="wsite-menu-subitem" href="/block-3---programme-outline.html">
<span class="wsite-menu-title">
    Block 3 - Programme Outline

```

```

        </span>
    </a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg524658878774308606">
    <a class="wsite-menu-item" href="/eligibility.html">
        Eligibility
    </a>
    <div class="wsite-menu-wrap" style="display:none">
        <ul class="wsite-menu">
            <li class="wsite-menu-subitem-wrap " id="wsite-nav-163829561948474038">
                <a class="wsite-menu-subitem" href="/q-step-students.html">
                    <span class="wsite-menu-title">
                        Q-Step Students
                    </span>
                </a>
            </li>
            </ul>
        </div>
    </li>
    <li class="wsite-menu-item-wrap" id="pg971730385455396268">
        <a class="wsite-menu-item" href="/bursaries.html">
            Bursaries
        </a>
    </li>
    <li class="wsite-menu-item-wrap" id="pg775168331197737129">
        <a class="wsite-menu-item" href="/accommodation.html">
            Accommodation
        </a>
    </li>
    <li class="wsite-menu-item-wrap" id="pg251883312953949911">
        <a class="wsite-menu-item" href="/registration-costs.html">
            Registration Costs
        </a>
    </li>
    <li class="wsite-menu-item-wrap" id="pg253676446592865942">
        <a class="wsite-menu-item" href="/faq.html">
            FAQ
        </a>
    </li>
    <li class="wsite-menu-item-wrap" id="pg905718591334408166">
        <a class="wsite-menu-item" href="/terms-and-conditions.html">
            Terms and Conditions
        </a>
    </li>
</ul>

```

```

</div>
<div class="nav desktop-nav wsite-menu-default"><ul class="site-menu">
<li class="wsite-menu-item-wrap" id="active">
<a class="wsite-menu-item" href="/">
    Home
</a>
</li>
<li class="wsite-menu-item-wrap" id="pg537531721210351813">
<a class="wsite-menu-item" href="/about.html">
    About
</a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-447875943371678234">
<a class="wsite-menu-subitem" href="/instructors.html">
<span class="wsite-menu-title">
    Instructors
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-428449932915725692">
<a class="wsite-menu-subitem" href="/contact.html">
<span class="wsite-menu-title">
    Contact
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-115774652140471228">
<a class="wsite-menu-subitem" href="/privacy-statement.html">
<span class="wsite-menu-title">
    Privacy Statement
</span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg750005411943018735">
<a class="wsite-menu-item" href="/structure.html">
    Structure
</a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-281688027203735335">
<a class="wsite-menu-subitem" href="/learning-environment.html">
<span class="wsite-menu-title">
    Learning Environment
</span>

```



```

</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-868244784447957774">
<a class="wsite-menu-subitem" href="/block-1---programme-outline.html">
<span class="wsite-menu-title">
Block 1 - Programme Outline
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-181475333474170941">
<a class="wsite-menu-subitem" href="/block-2---programme-outline.html">
<span class="wsite-menu-title">
Block 2 - Programme Outline
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-761909471185475700">
<a class="wsite-menu-subitem" href="/block-3---programme-outline.html">
<span class="wsite-menu-title">
Block 3 - Programme Outline
</span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg524658878774308606">
<a class="wsite-menu-item" href="/eligibility.html">
Eligibility
</a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-163829561948474038">
<a class="wsite-menu-subitem" href="/q-step-students.html">
<span class="wsite-menu-title">
Q-Step Students
</span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg971730385455396268">
<a class="wsite-menu-item" href="/bursaries.html">
Bursaries
</a>
</li>
<li class="wsite-menu-item-wrap" id="pg775168331197737129">

```

```
<a class="wsite-menu-item" href="/accommodation.html">
    Accommodation
</a>
</li>
<li class="wsite-menu-item-wrap" id="pg251883312953949911">
<a class="wsite-menu-item" href="/registration-costs.html">
    Registration Costs
</a>
</li>
<li class="wsite-menu-item-wrap" id="pg253676446592865942">
<a class="wsite-menu-item" href="/faq.html">
    FAQ
</a>
</li>
<li class="wsite-menu-item-wrap" id="pg905718591334408166">
<a class="wsite-menu-item" href="/terms-and-conditions.html">
    Terms and Conditions
</a>
</li>
</ul>
</div>
<label class="hamburger">
<span class="hamburger-icon">
<div class="hamburger-line"></div>
<div class="hamburger-line"></div>
<div class="hamburger-line"></div>
</span>
<span class="menu-text menu-open-text">menu</span>
<span class="menu-text menu-close-text">close</span>
</label>
<div class="logo"><span class="wsite-logo">
<a href="/">

</a>
</span></div>
<div class="site-utils">
<span class="wsite-search">
<form action="/apps/search" id="wsite-header-search-form" method="get">
<input autocomplete="off" class="wsite-search-input" name="q" placeholder="Search" type="text">
</form>
</span>
<button class="search-icon">
<svg height="13px" viewBox="0 0 12 13" width="12px">
<path d="M7.77381641,9.19956292 L10.9920719,12.4178184 L12.414847,10.9950432 L9.209804,8.5851568 L7.77381641,9.19956292 Z" />
</svg>
</button>
<button class="search-close">
<span class="search-close-text">close</span>
```

```

<span class="search-close-icon">
<svg height="11px" viewBox="51 4 14 11" width="14px">
<path d="M58,8.57142857 L61.5714286,5 L63,6.42857143 L59.4285714,10 L63,13.5714286 L61
</svg>
</span>
</button>
<div class="mini-cart-toggle">
</div>
</div>
</div><!-- end .container -->
<div class="mini-cart-overlay">
<button class="mini-cart-close">
<span class="mini-cart-close-text">close</span>
<span class="mini-cart-close-icon">
<svg height="11px" viewBox="51 4 14 11" width="14px">
<path d="M58,8.57142857 L61.5714286,5 L63,6.42857143 L59.4285714,10 L63,13.5714286 L61
</svg>
</span>
</button>
</div>
<div class="nav collapsed-nav wsite-menu-default"><ul class="site-menu">
<li class="wsite-menu-item-wrap" id="active">
<a class="wsite-menu-item" href="/">
Home
</a>
</li>
<li class="wsite-menu-item-wrap" id="pg537531721210351813">
<a class="wsite-menu-item" href="/about.html">
About
</a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-447875943371678234">
<a class="wsite-menu-subitem" href="/instructors.html">
<span class="wsite-menu-title">
Instructors
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-428449932915725692">
<a class="wsite-menu-subitem" href="/contact.html">
<span class="wsite-menu-title">
Contact
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-115774652140471228">
<a class="wsite-menu-subitem" href="/privacy-statement.html">

```

```

<span class="wsite-menu-title">
    Privacy Statement
</span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg750005411943018735">
<a class="wsite-menu-item" href="/structure.html">
    Structure
</a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-281688027203735335">
<a class="wsite-menu-subitem" href="/learning-environment.html">
<span class="wsite-menu-title">
    Learning Environment
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-868244784447957774">
<a class="wsite-menu-subitem" href="/block-1---programme-outline.html">
<span class="wsite-menu-title">
    Block 1 - Programme Outline
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-181475333474170941">
<a class="wsite-menu-subitem" href="/block-2---programme-outline.html">
<span class="wsite-menu-title">
    Block 2 - Programme Outline
</span>
</a>
</li>
<li class="wsite-menu-subitem-wrap " id="wsite-nav-761909471185475700">
<a class="wsite-menu-subitem" href="/block-3---programme-outline.html">
<span class="wsite-menu-title">
    Block 3 - Programme Outline
</span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg524658878774308606">
<a class="wsite-menu-item" href="/eligibility.html">
    Eligibility

```

```

        </a>
<div class="wsite-menu-wrap" style="display:none">
<ul class="wsite-menu">
<li class="wsite-menu-subitem-wrap " id="wsite-nav-163829561948474038">
<a class="wsite-menu-subitem" href="/q-step-students.html">
<span class="wsite-menu-title">
        Q-Step Students
    </span>
</a>
</li>
</ul>
</div>
</li>
<li class="wsite-menu-item-wrap" id="pg971730385455396268">
<a class="wsite-menu-item" href="/bursaries.html">
        Bursaries
    </a>
</li>
<li class="wsite-menu-item-wrap" id="pg775168331197737129">
<a class="wsite-menu-item" href="/accommodation.html">
        Accommodation
    </a>
</li>
<li class="wsite-menu-item-wrap" id="pg251883312953949911">
<a class="wsite-menu-item" href="/registration-costs.html">
        Registration Costs
    </a>
</li>
<li class="wsite-menu-item-wrap" id="pg253676446592865942">
<a class="wsite-menu-item" href="/faq.html">
        FAQ
    </a>
</li>
<li class="wsite-menu-item-wrap" id="pg905718591334408166">
<a class="wsite-menu-item" href="/terms-and-conditions.html">
        Terms and Conditions
    </a>
</li>
</ul>
</div>
</div><!-- end .header -->
<div class="banner-wrap">
<div class="wsite-elements wsite-not-footer wsite-header-elements">
<div class="wsite-section-wrap">
<div class="wsite-section wsite-header-section wsite-section-bg-image" style='height: 400px;'>
<div class="wsite-section-content">
<div class="container">
<div class="banner">

```

```

<div class="wsite-section-elements">
<div><div class="wsite-image wsite-image-border-none " style="padding-top:10px;padding:
<a>
</div>
</div></div>
<h2 class="wsite-content-title" style="text-align:right;"><font color="#003865" size="
<div><div class="wsite-multicol"><div class="wsite-multicol-table-wrap" style="margin:
<table class="wsite-multicol-table">
<tbody class="wsite-multicol-tbody">
<tr class="wsite-multicol-tr">
<td class="wsite-multicol-col" style="width:33.33333333333%; padding:0 15px;">
<div><div class="wsite-image wsite-image-border-none " style="padding-top:10px;padding:
<a>

<div style="display:block;font-size:90%"></div>
</div></div>
</td> <td class="wsite-multicol-col" style="width:33.33333333333%; padding:0 15px;">
<div><div class="wsite-image wsite-image-border-none " style="padding-top:10px;padding:
<a>

<div style="display:block;font-size:90%"></div>
</div></div>
</td> <td class="wsite-multicol-col" style="width:33.33333333333%; padding:0 15px;">
<div><div class="wsite-image wsite-image-border-none " style="padding-top:10px;padding:
<a>
</div>
</div></div>
</td> </tr>
</tbody>
</table>
</div></div></div>
</div>
</div>
</div>
</div>
<div class=""></div>
</div>
</div>
</div>
<div class="main-wrap">
<div class="wsite-elements wsite-not-footer" id="wsite-content">

```

```

<div class="wsite-section-wrap">
<div class="wsite-section wsite-body-section wsite-background-13" style="height: auto;
<div class="wsite-section-content">
<div class="container">
<div class="wsite-section-elements">
<div style="text-align:center;"><div style="height: 10px; overflow: hidden;"></div>
<a class="wsite-button wsite-button-small wsite-button-highlight" href="/instructors.h
<span class="wsite-button-inner">Meet our team of world-leading researchers</span>
</a>
<div style="height: 10px; overflow: hidden;"></div></div>
</div>
</div>
</div>
</div>
</div>
<div class="wsite-section-wrap">
<div class="wsite-section wsite-body-section wsite-section-bg-color wsite-background-1
<div class="wsite-section-content">
<div class="container">
<div class="wsite-section-elements">
<h2 class="wsite-content-title" style="text-align:center;">Applications Open for 2019!
<div style="text-align:center;"><div style="height: 10px; overflow: hidden;"></div>
<a class="wsite-button wsite-button-small wsite-button-highlight" href="https://mycaree
<span class="wsite-button-inner">APPLY NOW</span>
</a>
<div style="height: 10px; overflow: hidden;"></div></div>
<div><div class="wsite-image wsite-image-border-none " style="padding-top:10px;padding
<a href="https://twitter.com/IoCExeter" target="_blank">
</div>
</div></div>
<h2 class="wsite-content-title" style="text-align:center;"><font color="#ffb500" size=
<div class="paragraph" style="text-align:left;">The Institute of Coding (IoC) Summer S
<div><div style="height: 20px; overflow: hidden;"></div>
<div class="imageGallery" id="402275968959232433-gallery" style="line-height: 0px; pad
<div id="402275968959232433-imageContainer0" style="float:left;width:33.28%;margin:0;"
</div>
<div style="height: 20px; overflow: hidden;"></div></div>
</div>
</div>
</div>
</div>
</div>
<div class="wsite-section-wrap">
<div class="wsite-section wsite-body-section wsite-section-bg-color wsite-background-1
<div class="wsite-section-content">
<div class="container">

```

```

<div class="wsite-section-elements">
<div><div class="wsite-multicol"><div class="wsite-multicol-table-wrap" style="margin:
<table class="wsite-multicol-table">
<tbody class="wsite-multicol-tbody">
<tr class="wsite-multicol-tr">
<td class="wsite-multicol-col" style="width:33.333333333333%; padding:0 5px;">
<div id="244036296635446800"><div><style type="text/css">
    #element-49782be4-1d10-4cef-bf68-dee3830f8af6 .colored-box-content {
        clear: both;
        float: left;
        width: 100%;
        -moz-box-sizing: border-box;
        -webkit-box-sizing: border-box;
        -ms-box-sizing: border-box;
        box-sizing: border-box;
        background-color: #ff6a13;
        padding-top: 20px;
        padding-bottom: 20px;
        padding-left: 20px;
        padding-right: 20px;
        -webkit-border-top-left-radius: 0px;
        -moz-border-top-left-radius: 0px;
        border-top-left-radius: 0px;
        -webkit-border-top-right-radius: 0px;
        -moz-border-top-right-radius: 0px;
        border-top-right-radius: 0px;
        -webkit-border-bottom-left-radius: 0px;
        -moz-border-bottom-left-radius: 0px;
        border-bottom-left-radius: 0px;
        -webkit-border-bottom-right-radius: 0px;
        -moz-border-bottom-right-radius: 0px;
        border-bottom-right-radius: 0px;
    }

</style><div class="platform-element-contents" data-platform-element-id="8488572479797979"
<div class="colored-box">
<div class="colored-box-content">
<div style="width: auto"><div></div><h2 class="wsite-content-title" style="text-align:
</div>
</div>
</div>
<div style="clear:both;"></div><script class="element-script" type="text/javascript">f
    var requireFunc = window.platformElementRequire || window.require;

    // Relies on a global require, specific to platform elements
    requireFunc([
        'w-global',
        'underscore',

```



```

        'jquery',
        'backbone',
        'util/platform/elements/PlatformElement',
        'util/platform/elements/PlatformElementSettings'
    ], function(
        _W,
        _,
        $,
        Backbone,
        PlatformElement,
        PlatformElementSettings
    ) {
        var dependencies = null || [];
        var platform_element_id = "848857247979793891-1.0.1";

        if (typeof _W.loadedPlatformDependencies === 'undefined') {
            _W.loadedPlatformDependencies = [];
        }

        if (typeof _W.platformElements === 'undefined') {
            _W.platformElements = [];
        }

        if (typeof _W.platformElements[platform_element_id] === 'undefined') {
            _W.platformElements[platform_element_id] = {};
            _W.platformElements[platform_element_id].deferredObject = new S
            _W.platformElements[platform_element_id].deferredPromise = _W.p
        }

        if(_.intersection(_W.loadedPlatformDependencies, dependencies).length
            _.reduce(dependencies, function(promise, nextScript){
                _W.loadedPlatformDependencies.push(nextScript);
                return promise.then(function(){
                    return $.getScript(nextScript);
                });
            }, $.promise()).then(function(){
                _W.platformElements[platform_element_id].deferredObject
            }));

        if (dependencies.length === 0){
            _W.platformElements[platform_element_id].deferredObject.resolve
        }

        _W.platformElements[platform_element_id].deferredPromise.done(function
            var _ElementDefinition = /**
* This is required for element rendering to be possible
* @type {PlatformElement}

```

```

    */
(function() {
    var ColoredBox = PlatformElement.extend({

        initialize: function() {
        }

    });

    return ColoredBox;
})();

if (typeof _ElementDefinition == 'undefined' || typeof _Element
    var _ElementDefinition = PlatformElement.extend({});
}

var _Element = _ElementDefinition.extend({
    initialize: function() {
        // we still want to call the initialize function
        // however, we don't want to call it until placeholderInterval
        this.placeholderInterval = setInterval(function() {
            // so use setInterval to check for placeholderInterval
            if (this.$('.platform-element-child-placeholder')
                clearInterval(this.placeholderInterval);
                this.constructor.__super__.initialize.call(this);
            }
        }.bind(this), 100);
    }
});

_Element.prototype.settings = new PlatformElementSettings({"top":
_Element.prototype.settings.page_element_id = "2440362966354468";

_Element.prototype.element_id = "49782be4-1d10-4cef-bf68-dee3830f8af6";
_Element.prototype.user_id = "23981353";
_Element.prototype.site_id = "299388504590429306";
_Element.prototype.assets_path = "//marketplace.editmysite.com";
new _Element({
    el: '#element-49782be4-1d10-4cef-bf68-dee3830f8af6'
});

});

});

}

if (typeof document.documentElement.appReady == 'undefined') {
    document.documentElement.appReady = 0;
}

```

```

if (document.documentElement.appReady || (window.inEditor && window.inEditor())) {
    setupElement244036296635446800();
} else if (document.createEvent && document.addEventListener) {
    document.addEventListener('appReady', setupElement244036296635446800, false);
} else {
    document.documentElement.attachEvent('onpropertychange', function(event){
        if (event.propertyName == 'appReady') {
            setupElement244036296635446800();
        }
    });
}

</script></div></div>
</td> <td class="wsite-multicol-col" style="width:33.333333333333%; padding:0 5px;">
<div id="937001622770875045"><div><style type="text/css">
    #element-45115d2c-55fb-40ee-8140-ef0e8336799c .colored-box-content {
        clear: both;
        float: left;
        width: 100%;
        -moz-box-sizing: border-box;
        -webkit-box-sizing: border-box;
        -ms-box-sizing: border-box;
        box-sizing: border-box;
        background-color: #003865;
        padding-top: 20px;
        padding-bottom: 20px;
        padding-left: 20px;
        padding-right: 20px;
        -webkit-border-top-left-radius: 0px;
        -moz-border-top-left-radius: 0px;
        border-top-left-radius: 0px;
        -webkit-border-top-right-radius: 0px;
        -moz-border-top-right-radius: 0px;
        border-top-right-radius: 0px;
        -webkit-border-bottom-left-radius: 0px;
        -moz-border-bottom-left-radius: 0px;
        border-bottom-left-radius: 0px;
        -webkit-border-bottom-right-radius: 0px;
        -moz-border-bottom-right-radius: 0px;
        border-bottom-right-radius: 0px;
    }

</style><div class="platform-element-contents" data-platform-element-id="8488572479797979">
<div class="colored-box">
<div class="colored-box-content">
<div style="width: auto"><div></div><h2 class="wsite-content-title" style="text-align: center;">
</div>

```

```

</div>
</div>
<div style="clear:both;"></div><script class="element-script" type="text/javascript">f
    var requireFunc = window.platformElementRequire || window.require;

    // Relies on a global require, specific to platform elements
    requireFunc([
        'w-global',
        'underscore',
        'jquery',
        'backbone',
        'util/platform/elements/PlatformElement',
        'util/platform/elements/PlatformElementSettings'
    ], function(
        _W,
        _,
        $,
        Backbone,
        PlatformElement,
        PlatformElementSettings
    ) {
        var dependencies = null || [];
        var platform_element_id = "848857247979793891-1.0.1";

        if (typeof _W.loadedPlatformDependencies === 'undefined') {
            _W.loadedPlatformDependencies = [];
        }

        if (typeof _W.platformElements === 'undefined') {
            _W.platformElements = [];
        }

        if (typeof _W.platformElements[platform_element_id] === 'undefined') {
            _W.platformElements[platform_element_id] = {};
            _W.platformElements[platform_element_id].deferredObject = new $
            _W.platformElements[platform_element_id].deferredPromise = _W.
        }

        if(_.intersection(_W.loadedPlatformDependencies, dependencies).length
            _.reduce(dependencies, function(promise, nextScript){
                _W.loadedPlatformDependencies.push(nextScript);
                return promise.then(function(){
                    return $.getScript(nextScript);
                });
            }, $.promise()).then(function(){
                _W.platformElements[platform_element_id].deferredObject
            });
    }
}

```



```

        });
    });
}

if (typeof document.documentElement.appReady == 'undefined') {
    document.documentElement.appReady = 0;
}

if (document.documentElement.appReady || (window.inEditor && window.inEditor())) {
    setupElement937001622770875045();
} else if (document.createElement && document.addEventListener) {
    document.addEventListener('appReady', setupElement937001622770875045, false);
} else {
    document.documentElement.attachEvent('onpropertychange', function(event){
        if (event.propertyName == 'appReady') {
            setupElement937001622770875045();
        }
    });
}

</script></div></div>
</td> <td class="wsite-multicol-col" style="width:33.333333333333%; padding:0 5px;">
<div id="455599425849470879"><div><style type="text/css">
    #element-06bfb7dc-e75e-419e-a8bb-41db63abc44c .colored-box-content {
        clear: both;
        float: left;
        width: 100%;
        -moz-box-sizing: border-box;
        -webkit-box-sizing: border-box;
        -ms-box-sizing: border-box;
        box-sizing: border-box;
        background-color: #00b5e2;
        padding-top: 20px;
        padding-bottom: 20px;
        padding-left: 20px;
        padding-right: 20px;
        -webkit-border-top-left-radius: 0px;
        -moz-border-top-left-radius: 0px;
        border-top-left-radius: 0px;
        -webkit-border-top-right-radius: 0px;
        -moz-border-top-right-radius: 0px;
        border-top-right-radius: 0px;
        -webkit-border-bottom-left-radius: 0px;
        -moz-border-bottom-left-radius: 0px;
        border-bottom-left-radius: 0px;
        -webkit-border-bottom-right-radius: 0px;

```



```

        _W.loadedPlatformDependencies.push(nextScript);
        return promise.then(function(){
            return $.getScript(nextScript);
        });
    }, $().promise()).then(function(){
        _W.platformElements[platform_element_id].deferredObject.resolve();
    });
}

if (dependencies.length === 0){
    _W.platformElements[platform_element_id].deferredObject.resolve();
}

_W.platformElements[platform_element_id].deferredPromise.done(function(){
    var _ElementDefinition = /**
     * This is required for element rendering to be possible
     * @type {PlatformElement}
     */
    (function() {
        var ColoredBox = PlatformElement.extend({

            initialize: function() {

            }

        });

        return ColoredBox;
    })();

    if (typeof _ElementDefinition == 'undefined' || typeof _ElementDefinition == 'function')
        var _ElementDefinition = PlatformElement.extend({});

    var _Element = _ElementDefinition.extend({
        initialize: function() {
            // we still want to call the initialize function
            // however, we don't want to call it until platform elements are loaded
            this.placeholderInterval = setInterval(function(){
                // so use setInterval to check for platform elements
                if (this.$('.platform-element-child-placeholder').length > 0)
                    clearInterval(this.placeholderInterval);
                this.constructor.__super__.initialize.call(this);
            }, 100);
        }
    });
}

```



```

        _Element.prototype.settings = new PlatformElementSettings({"top
        _Element.prototype.settings.page_element_id = "4555994258494708

        _Element.prototype.element_id = "06bfb7dc-e75e-419e-a8bb-41db63
        _Element.prototype.user_id = "23981353";
        _Element.prototype.site_id = "299388504590429306";
        _Element.prototype.assets_path = "//marketplace.editmysite.com/
        new _Element({
            el: '#element-06bfb7dc-e75e-419e-a8bb-41db63abc44c'
        });
    });
});

}

if (typeof document.documentElement.appReady == 'undefined') {
    document.documentElement.appReady = 0;
}

if (document.documentElement.appReady || (window.inEditor && window.inEditor())) {
    setupElement455599425849470879();
} else if (document.createEvent && document.addEventListener) {
    document.addEventListener('appReady', setupElement455599425849470879, false);
} else {
    document.documentElement.attachEvent('onpropertychange', function(event){
        if (event.propertyName == 'appReady') {
            setupElement455599425849470879();
        }
    });
}

</script></div></div>
</td> </tr>
</tbody>
</table>
</div></div></div>
</div>
</div>
</div>
</div>
</div>
</div>
</div>
</div>
<div class="footer-wrap">
<div class="container">
<div class="footer"><div class="wsite-elements wsite-footer">
<div><div class="wsite-multicol"><div class="wsite-multicol-table-wrap" style="margin:
<table class="wsite-multicol-table">

```

```

<tbody class="wsite-multicol-tbody">
<tr class="wsite-multicol-tr">
<td class="wsite-multicol-col" style="width:42.056074766355%; padding:0 25px;">
<h2 class="wsite-content-title" style="text-align:left;"><font size="2">Exeter Q-Step (
</td> <td class="wsite-multicol-col" style="width:57.943925233645%; padding:0 25px;">
<h2 class="wsite-content-title" style="text-align:left;"><font size="4"><a href="mailto:
<div style="text-align:left;"><div style="height:10px;overflow:hidden"></div>
<span class="wsite-social wsite-social-default"><a alt="Facebook" class="first-child w
<div style="height:10px;overflow:hidden"></div></div>
</td> </tr>
</tbody>
</table>
</div></div></div></div></div>
</div><!-- end container -->
</div><!-- end footer-wrap -->
</div>
<script src="/files/theme/plugins.js?1536789990" type="text/javascript"></script>
<script src="/files/theme/jquery.pxuMenu.js?1536789990" type="text/javascript"></script>
<script src="/files/theme/jquery.trend.js?1536789990" type="text/javascript"></script>
<script src="/files/theme/jquery.revealer.js?1536789990" type="text/javascript"></script>
<script src="/files/theme/debounce.js?1536789990" type="text/javascript"></script>
<script src="/files/theme/custom.js?1536789990" type="text/javascript"></script>
<div id="customer-accounts-app"></div>
<script src="//cdn2.editmysite.com/js/site/main-customer-accounts-site.js?buildTime=15
<!-- Global site tag (gtag.js) - Google Analytics -->
<script async="" src="https://www.googletagmanager.com/gtag/js?id=UA-117961586-1"></script>
<script>
  window.dataLayer = window.dataLayer || [];
  function gtag(){dataLayer.push(arguments);}
  gtag('js', new Date());

  gtag('config', 'UA-117961586-1');
</script><script type="text/javascript">
  var _gaq = _gaq || [];
  _gaq.push(['_setAccount', 'UA-7870337-1']);
  _gaq.push(['_setDomainName', 'none']);
  _gaq.push(['_setAllowLinker', true]);

  (function() {
    var ga = document.createElement('script'); ga.type = 'text/javascript';
    ga.src = ('https:' == document.location.protocol ? 'https://ssl' : 'http://www') +
    var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
  })();

  _W.Analytics = _W.Analytics || {'trackers': {}};
  _W.Analytics.trackers.wGA = '_gaq';
</script>
<script async="1" type="text/javascript">

```

```

;(function(p,l,o,w,i,n,g){if(!p[i]){p.GlobalSnowplowNamespace=p.GlobalSnowplow
    p.GlobalSnowplowNamespace.push(i);p[i]=function(){(p[i].q=p[i]
    };p[i].q=p[i].q||[];n=l.createElement(o);g=l.getElementsByTagName
    n.src=w;g.parentNode.insertBefore(n,g)}}(window,document,'scrip

var r = [99, 104, 101, 99, 107, 111, 117, 116, 46, 40, 119, 101, 101, 98, 108,
var snPlObR = function(arr) {
    var s = '';
    for (var i = 0 ; i < arr.length ; i++){
        s = s + String.fromCharCode(arr[i]);
    }
    return s;
};
var s = snPlObR(r);

var regEx = new RegExp(s);

_W.Analytics = _W.Analytics || {'trackers': {}};
_W.Analytics.trackers.wSP = 'snowday';
_W.Analytics.user_id = '23981353';
_W.Analytics.site_id = '299388504590429306';

// Setting do not track if the GDPR cookie is not present. This is then checked
// to set tracking decisions. https://github.com/snowplow/snowplow-javascript-
window.doNotTrack = document.cookie.indexOf('gdpr-kb') === -1 ? 'yes' : null;

(function(app_id, ec_hostname, discover_root_domain) {
    var track = window[_W.Analytics.trackers.wSP];
    if (!track) return;
    track('newTracker', app_id, ec_hostname, {
        appId: app_id,
        post: true,
        platform: 'web',
        discoverRootDomain: discover_root_domain,
        cookieName: '_snow_',
        contexts: {
            webPage: true,
            performanceTiming: true,
            gaCookies: true
        },
        crossDomainLinker: function (linkElement) {
            return regEx.test(linkElement.href);
        },
        respectDoNotTrack: document.cookie.indexOf('gdpr-kb') === -1
    });
    track('trackPageView', _W.Analytics.user_id+':'+_W.Analytics.site_id);
    track('crossDomainLinker', function (linkElement) {

```

```

        return regEx.test(linkElement.href);
    });
    })(
        '_wn',
        'ec.editmysite.com',
        true
    );
</script>
<script>
    (function(jQuery) {
        try {
            if (jQuery) {
                jQuery('div.blog-social div.fb-like').attr('class', 'b
                var $commentFrame = jQuery('#commentArea iframe');
                if ($commentFrame.length > 0) {
                    var frameHeight = jQuery($commentFrame[0].cont
                    $commentFrame.css('min-height', frameHeight +
                }
                if (jQuery('.product-button').length > 0){
                    jQuery(document).ready(function(){
                        jQuery('.product-button').parent().each
                        if(jQuery(product).attr('target')
                        if (!jQuery(product).f
                            jQuery('<input:
                                type:
                                name:
                                value:
                            }).appendTo(pr
                        }
                    }
                });
            });
        }
        else {
            // Prototype
            $$('div.blog-social div.fb-like').each(function(div) {
                div.className = 'blog-social-item blog-fb-like
            });
            $$('#commentArea iframe').each(function(iframe) {
                iframe.style.minHeight = '410px';
            });
        }
    }
    catch(ex) {}
})(window._W && _W.jQuery);

window._W.isEUUser = true;

```

```

        window._W.showCookieToAll = "";
    </script>
</body>
</html>

```

```

In [6]: # Get the page title
        soup.title

```

```

Out[6]: <title>The Institute of Coding Summer School at the University of Exeter - Home</title>

```

```

In [7]: soup.title.string
        # Returns the content inside a title tag as a string

```

```

Out[7]: 'The Institute of Coding Summer School at the University of Exeter - Home'

```

```

In [8]: soup.title.parent.name

```

```

Out[8]: 'head'

```

```

In [9]: soup.a
        # Return the first matching anchor tag</a>

```

```

Out[9]: <a class="wsite-menu-item" href="/">
        Home
        </a>

```

```

In [10]: # Or, we could use the find all, and return all the matching anchor tags
         soup.find_all('a')

```

```

Out[10]: [<a class="wsite-menu-item" href="/">
        Home
        </a>, <a class="wsite-menu-item" href="/about.html">
        About
        </a>, <a class="wsite-menu-subitem" href="/instructors.html">
        <span class="wsite-menu-title">
        Instructors
        </span>
        </a>, <a class="wsite-menu-subitem" href="/contact.html">
        <span class="wsite-menu-title">
        Contact
        </span>
        </a>, <a class="wsite-menu-subitem" href="/privacy-statement.html">
        <span class="wsite-menu-title">
        Privacy Statement
        </span>
        </a>, <a class="wsite-menu-item" href="/structure.html">
        Structure
        </a>, <a class="wsite-menu-subitem" href="/learning-environment.html">
        <span class="wsite-menu-title">

```

```

Learning Environment
</span>
</a>, <a class="wsite-menu-subitem" href="/block-1---programme-outline.html">
<span class="wsite-menu-title">
    Block 1 - Programme Outline
</span>
</a>, <a class="wsite-menu-subitem" href="/block-2---programme-outline.html">
<span class="wsite-menu-title">
    Block 2 - Programme Outline
</span>
</a>, <a class="wsite-menu-subitem" href="/block-3---programme-outline.html">
<span class="wsite-menu-title">
    Block 3 - Programme Outline
</span>
</a>, <a class="wsite-menu-item" href="/eligibility.html">
    Eligibility
    </a>, <a class="wsite-menu-subitem" href="/q-step-students.html">
<span class="wsite-menu-title">
    Q-Step Students
</span>
</a>, <a class="wsite-menu-item" href="/bursaries.html">
    Bursaries
    </a>, <a class="wsite-menu-item" href="/accommodation.html">
    Accommodation
    </a>, <a class="wsite-menu-item" href="/registration-costs.html">
    Registration Costs
    </a>, <a class="wsite-menu-item" href="/faq.html">
    FAQ
    </a>, <a class="wsite-menu-item" href="/terms-and-conditions.html">
    Terms and Conditions
    </a>, <a class="wsite-menu-item" href="/">
    Home
    </a>, <a class="wsite-menu-item" href="/about.html">
    About
    </a>, <a class="wsite-menu-subitem" href="/instructors.html">
<span class="wsite-menu-title">
    Instructors
</span>
</a>, <a class="wsite-menu-subitem" href="/contact.html">
<span class="wsite-menu-title">
    Contact
</span>
</a>, <a class="wsite-menu-subitem" href="/privacy-statement.html">
<span class="wsite-menu-title">
    Privacy Statement
</span>
</a>, <a class="wsite-menu-item" href="/structure.html">
    Structure

```

```

        </a>, <a class="wsite-menu-subitem" href="/learning-environment.html">
<span class="wsite-menu-title">
    Learning Environment
</span>
</a>, <a class="wsite-menu-subitem" href="/block-1---programme-outline.html">
<span class="wsite-menu-title">
    Block 1 - Programme Outline
</span>
</a>, <a class="wsite-menu-subitem" href="/block-2---programme-outline.html">
<span class="wsite-menu-title">
    Block 2 - Programme Outline
</span>
</a>, <a class="wsite-menu-subitem" href="/block-3---programme-outline.html">
<span class="wsite-menu-title">
    Block 3 - Programme Outline
</span>
</a>, <a class="wsite-menu-item" href="/eligibility.html">
    Eligibility
    <a>, <a class="wsite-menu-subitem" href="/q-step-students.html">
<span class="wsite-menu-title">
    Q-Step Students
</span>
</a>, <a class="wsite-menu-item" href="/bursaries.html">
    Bursaries
    <a>, <a class="wsite-menu-item" href="/accommodation.html">
    Accommodation
    <a>, <a class="wsite-menu-item" href="/registration-costs.html">
    Registration Costs
    <a>, <a class="wsite-menu-item" href="/faq.html">
    FAQ
    <a>, <a class="wsite-menu-item" href="/terms-and-conditions.html">
    Terms and Conditions
    <a>, <a href="/">

    Home
    <a>, <a class="wsite-menu-item" href="/about.html">
    About
    <a>, <a class="wsite-menu-subitem" href="/instructors.html">
<span class="wsite-menu-title">
    Instructors
</span>
</a>, <a class="wsite-menu-subitem" href="/contact.html">
<span class="wsite-menu-title">
    Contact
</span>
</a>, <a class="wsite-menu-subitem" href="/privacy-statement.html">
<span class="wsite-menu-title">

```



```

</a>, <a href="https://twitter.com/IoCExeter" target="_blank">
<img alt="Picture" src="/uploads/2/3/9/8/23981353/editor/3ufy0wpd-400x400.jpg?155022
</a>, <a href="/uploads/2/3/9/8/23981353/ks0a2107_orig.jpg" rel="lightbox[gallery402

```

In [11]: *#Or to to get them in a nicer format:*

```

for link in soup.find_all('a'):
    print(link.get('href'))

```

```

/
/about.html
/instructors.html
/contact.html
/privacy-statement.html
/structure.html
/learning-environment.html
/block-1---programme-outline.html
/block-2---programme-outline.html
/block-3---programme-outline.html
/eligibility.html
/q-step-students.html
/bursaries.html
/accommodation.html
/registration-costs.html
/faq.html
/terms-and-conditions.html
/
/about.html
/instructors.html
/contact.html
/privacy-statement.html
/structure.html
/learning-environment.html
/block-1---programme-outline.html
/block-2---programme-outline.html
/block-3---programme-outline.html
/eligibility.html
/q-step-students.html
/bursaries.html
/accommodation.html
/registration-costs.html
/faq.html
/terms-and-conditions.html
/
/
/about.html
/instructors.html
/contact.html
/privacy-statement.html

```

```

/structure.html
/learning-environment.html
/block-1---programme-outline.html
/block-2---programme-outline.html
/block-3---programme-outline.html
/eligibility.html
/q-step-students.html
/bursaries.html
/accommodation.html
/registration-costs.html
/faq.html
/terms-and-conditions.html
None
None
None
None
/instructors.html
https://mycareerzone.exeter.ac.uk/Form.aspx?id=592139
https://twitter.com/IoCExeter
/uploads/2/3/9/8/23981353/ks0a2107_orig.jpg
/uploads/2/3/9/8/23981353/c57a7516_orig.jpg
/uploads/2/3/9/8/23981353/c57a7563_orig.jpg
/uploads/2/3/9/8/23981353/c57a7522_orig.jpg
/uploads/2/3/9/8/23981353/c57a7514_orig.jpg
/uploads/2/3/9/8/23981353/c57a7552_orig.jpg
/uploads/2/3/9/8/23981353/c57a7529_orig.jpg
/uploads/2/3/9/8/23981353/ks0a2125_1_orig.jpg
/uploads/2/3/9/8/23981353/ks0a2704_1_orig.jpg
/about.html
/structure.html
/bursaries.html
mailto:ioc@exeter.ac.uk
https://www.facebook.com/qstepexeter
https://twitter.com/IoCExeter
https://www.youtube.com/channel/UCWA-QbolpCJS3zZQzDSKXGg
mailto:ioc@exeter.ac.uk

```

```

In [12]: #Or to to get only links that contain a certain text
import re
for link in soup.find_all('a', href=re.compile("mailto")):
    print(link.get('href'))

```

```

mailto:ioc@exeter.ac.uk
mailto:ioc@exeter.ac.uk

```