

Chapter 6 - Data Sourcing via Web

Part 3 - Data Parsing

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

import urllib
import urllib.request
import re
```

```
In [3]: with urllib.request.urlopen('https://raw.githubusercontent.com/BigDataGal/Data-Ma
html = response.read()
```

```
In [4]: soup = BeautifulSoup(html, "lxml")
type(soup)
```

```
Out[4]: bs4.BeautifulSoup
```

Parsing Your Data

```
In [6]: print(soup.prettify()[0:100])
```

```
<html>
<head>
  <title>
    IoT Articles
  </title>
</head>
<body>
  <p class="title">
    <b>
```

Getting data from a parse tree

```
In [7]: text_only = soup.get_text()
print(text_only)
```

IoT Articles

2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this article, I offer you a listing of new IoT device ideas that you can use...

It's almost 2018 and IoT is on the cusp of an explosive expansion. In this article, I offer you a listing of new IoT device ideas that you can use to get practice in designing your first IoT applications.

Looking Back at My Coolest IoT Find in 2017

Before going into detail about best new IoT device ideas, here's the backstory. Last month Ericsson Digital invited me to tour the Ericsson Studio in Kista, Sweden. Up until that visit, IoT had been largely theoretical to me. Of course, I know the usual mumbo-jumbo about wearables and IoT-connected fitness trackers. That stuff is all well and good, but it's somewhat old hat - plus I am not sure we are really benefiting so much from those, so I'm not that impressed.

It wasn't until I got to the Ericsson Studio that I became extremely impressed by how far IoT has really come. Relying on the promise of the 5g network expansion, IoT-powered smart devices are on the cusp of an explosive growth in adoption. It was Ericsson's Smart Car that sent me reeling:

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform that manages services for the Smart Cars to which it's connected. The Volvo pictured above acts as a drop-off location for groceries that have been ordered by its owner.

To understand how it works, imagine you're pulling your normal 9-to-5 and you know you need to grab some groceries on your way home. Well, since you're smart you've used Ericsson IoT platform to connect your car to the local grocery delivery service (Mat.se), so all you need to do is open the Mat.se app and make your usual order. Mat.se automatically handles the payment, grocery selection, delivery, and delivery scheduling. Since your car is IoT-enabled, Mat.se issues its trusted delivery agent a 1-time token to use for opening your car in order to place your groceries in your car for you at 4:40 pm (just before you get off from work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio, make sure to go watch the videos on this page.

Future Trends for IoT in 2018

New IoT device ideas won't do you much good unless you at least know the basic technology trends that are set to impact IoT over the next year(s). These include:

Big Data & Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it's going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.

Machine Learning Data Science: While a lot of IoT devices are still operated according to rules-based decision criteria, the age of artificial intelligence is upon us. IoT will increasingly depend on machine learning algorithms to

control device operations so that devices are able to autonomously respond to a complex set of overlapping stimuli.

Blockchain-Enabled Security: Above all else, IoT networks must be secure. Blockchain technology is primed to meet the security demands that come along with building and expanding the IoT.

Best New IoT Device Ideas

This listing of new IoT device ideas has been sub-divided according to the main technology upon which the IoT devices are built. Below I'm providing a list of new IoT device ideas, but for detailed instructions on how to build these IoT applications, I recommend the IoT courses on Udemy (Please note: if you purchase a Udemy course through this link, I may receive a small commission), or courses that are available at SkyFi and Coursera.

Raspberry Pi IoT Ideas

Using Raspberry Pi as open-source hardware, you can build IoT applications that offer any one of the following benefits:

- Enable built-in sensing to build a weather station that measures ambient temperature and humidity

- Build a system that detects discrepancies in electrical readings to identify electricity theft

- Use IoT to build a Servo that is controlled by motion detection readings

- Build a smart control switch that operates devices based on external stimuli.

- Use this for home automation.

- Build a music playing application that enables music for each room in your house

- Implement biometrics on IoT-connected devices

Arduino IoT Ideas

There are a number of new IoT device ideas that deploy Arduino as a microcontroller. These include:

- Integrate Arduino with Android to build a remote-control RGB LED device.

- Connect PIR sensors across the IoT to implement a smart building.

- Build a temperature and sunlight sensor system to remotely monitor and control the conditions of your garden.

- Deploy Arduino and IoT to automate your neighborhood streetlights.

- Build a smart irrigation system based on IoT-connected temperature and moisture sensors built-in to your agricultural plants.

[caption id="attachment_3807" align="aligncenter" width="300"] An IoT Chatbot Tree at the Ericsson Studio[/caption]

Wireless (GSM) IoT Ideas

Several new IoT device ideas are developed around the GSM wireless network. These are:

- Monitor soil moisture to automate agricultural irrigation cycles.

- Automate and control the conditions of a greenhouse.

- Enable bio-metrics to build a smart security system for your home or office building

- Build an autonomously operating fitness application that automatically makes recommendations based on motion detection and heart rate sensors that are embedded on wearable fitness trackers.

- Build a healthcare monitoring system that tracks, informs, and automatically alerts healthcare providers based on sensor readings that describe a patient's vital statistics (like temperature, pulse, blood pressure, etc).

IoT Automation Ideas

Almost all new IoT device ideas offer automation benefits, but to outline a few more ideas:

Build an IoT device that automatically locates and reports the closest nearby parking spot.

Build a motion detection system that automatically issues emails or sms messages to alert home owners of a likely home invasion.

Use temperature sensors connected across the IoT to automatically alert you if your home windows or doors have been left open.

Use bio-metric sensors to build a smart system that automate security for your home or office building

To learn more about IoT and what's happening on the leading edge, be sure to pop over to Ericsson's Studio Tour recap and watch these videos.

(I captured some of this content on behalf of DevMode Strategies during an invite-only tour of the Ericsson Studio in Kista. Rest assured, the text and opinions are my own)

...

Searching and retrieving data from a parse tree

Retrieving tags by filtering with name arguments

```
In [8]: soup.find_all("li")
```

```
Out[8]: [<li><strong>Big Data</strong> & Data Engineering: Sensors that are embedde
d within IoT devices spin off machine-generated data like it's going out of sty
le. For IoT to function, the platform must be solidly engineered to handle big
data. Be assured, that requires some serious data engineering.</li>,
<li><strong>Machine Learning</strong> Data Science: While a lot of IoT devices
are still operated according to rules-based decision criteria, the age of artif
icial intelligence is upon us. IoT will increasingly depend on machine learning
algorithms to control device operations so that devices are able to autonomousl
y respond to a complex set of overlapping stimuli.</li>,
<li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands th
at come along with building and expanding the IoT.</li>,
<li>Enable built-in sensing to build a weather station that measures ambient t
emperature and humidity</li>,
<li>Build a system that detects discrepancies in electrical readings to identi
fy electricity theft</li>,
<li>Use IoT to build a Servo that is controlled by motion detection readings</
li>,
<li>Build a smart control switch that operates devices based on external stimu
li. Use this for home automation.</li>,
<li>Build a music playing application that enables music for each room in your
house</li>,
<li>Implement biometrics on IoT-connected devices</li>,
<li>Integrate Arduino with Android to build a remote-control RGB LED device.</
li>,
<li>Connect PIR sensors across the IoT to implement a smart building.</li>,
<li>Build a temperature and sunlight sensor system to remotely monitor and con
trol the conditions of your garden.</li>,
<li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>,
<li>Build a smart irrigation system based on IoT-connected temperature and moi
sture sensors built-in to your agricultural plants.</li>,
<li>Monitor soil moisture to automate agricultural irrigation cycles.</li>,
<li>Automate and control the conditions of a greenhouse.</li>,
<li>Enable bio-metrics to build a smart security system for your home or offic
e building</li>,
<li>Build an autonomously operating fitness application that automatically mak
es recommendations based on motion detection and heart rate sensors that are em
bedded on wearable fitness trackers.</li>,
<li>Build a healthcare monitoring system that tracks, informs, and automatical
ly alerts healthcare providers based on sensor readings that describe a patient
s vital statistics (like temperature, pulse, blood pressure, etc).</li>,
<li>Build an IoT device that automatically locates and reports the closest nea
rby parking spot.</li>,
<li>Build a motion detection system that automatically issues emails or sms me
ssages to alert home owners of a likely home invasion.</li>,
<li>Use temperature sensors connected across the IoT to automatically alert yo
u if your home windows or doors have been left open.</li>,
<li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>]
```

Retrieving tags by filtering with keyword arguments

```
In [9]: soup.find_all(id="link 7")
```

```
Out[9]: [<a class="preview" href="http://www.skyfilabs.com/iot-online-courses" id="link 7">SkyFi</a>]
```

Retrieving tags by filtering with string arguments

```
In [10]: soup.find_all("ol")
```

```
Out[10]: [
  <ol>
    <li><strong>Big Data</strong> & Data Engineering: Sensors that are embed
    ded within IoT devices spin off machine-generated data like it's going out of
    style. For IoT to function, the platform must be solidly engineered to handle
    big data. Be assured, that requires some serious data engineering.</li>
    <li><strong>Machine Learning</strong> Data Science: While a lot of IoT devic
    es are still operated according to rules-based decision criteria, the age of
    artificial intelligence is upon us. IoT will increasingly depend on machine l
    earning algorithms to control device operations so that devices are able to a
    utonomously respond to a complex set of overlapping stimuli.</li>
    <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT network
    ks must be secure. Blockchain technology is primed to meet the security deman
    ds that come along with building and expanding the IoT.</li>
  </ol>,
  <ol>
    <li>Enable built-in sensing to build a weather station that measures ambient
    temperature and humidity</li>
    <li>Build a system that detects discrepancies in electrical readings to iden
    tify electricity theft</li>
    <li>Use IoT to build a Servo that is controlled by motion detection readings
    </li>
    <li>Build a smart control switch that operates devices based on external sti
    muli. Use this for home automation.</li>
    <li>Build a music playing application that enables music for each room in yo
    ur house</li>
    <li>Implement biometrics on IoT-connected devices</li>
  </ol>,
  <ol>
    <li>Integrate Arduino with Android to build a remote-control RGB LED device.
    </li>
    <li>Connect PIR sensors across the IoT to implement a smart building.</li>
    <li>Build a temperature and sunlight sensor system to remotely monitor and c
    ontrol the conditions of your garden.</li>
    <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
    <li>Build a smart irrigation system based on IoT-connected temperature and m
    oisture sensors built-in to your agricultural plants.</li>
  </ol>,
  <ol>
    <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
    <li>Automate and control the conditions of a greenhouse.</li>
    <li>Enable bio-metrics to build a smart security system for your home or off
    ice building</li>
    <li>Build an autonomously operating fitness application that automatically m
    akes recommendations based on motion detection and heart rate sensors that ar
    e embedded on wearable fitness trackers.</li>
    <li>Build a healthcare monitoring system that tracks, informs, and automatic
    ally alerts healthcare providers based on sensor readings that describe a pat
    ients vital statistics (like temperature, pulse, blood pressure, etc).</li>
  </ol>,
  <ol>
    <li>Build an IoT device that automatically locates and reports the closest n
    earby parking spot.</li>
    <li>Build a motion detection system that automatically issues emails or sms
    messages to alert home owners of a likely home invasion.</li>
  </ol>
]
```

```
<li>Use temperature sensors connected across the IoT to automatically alert  
you if your home windows or doors have been left open.</li>  
<li>Use bio-metric sensors to build a smart system that automate security fo  
r your home or office building</li>  
</ol>]
```

Retrieving tags by filtering with list objects


```
In [11]: soup.find_all(['ol', 'b'])
```

```
Out[11]: [<b>2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers
</b>,
<ol>
<li><strong>Big Data</strong> & Data Engineering: Sensors that are embed
ded within IoT devices spin off machine-generated data like it's going out of
style. For IoT to function, the platform must be solidly engineered to handle
big data. Be assured, that requires some serious data engineering.</li>
<li><strong>Machine Learning</strong> Data Science: While a lot of IoT devic
es are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine l
earning algorithms to control device operations so that devices are able to a
utonomously respond to a complex set of overlapping stimuli.</li>
<li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT network
ks must be secure. Blockchain technology is primed to meet the security deman
ds that come along with building and expanding the IoT.</li>
</ol>,
<ol>
<li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>
<li>Build a system that detects discrepancies in electrical readings to iden
tify electricity theft</li>
<li>Use IoT to build a Servo that is controlled by motion detection readings
</li>
<li>Build a smart control switch that operates devices based on external sti
mul. Use this for home automation.</li>
<li>Build a music playing application that enables music for each room in yo
ur house</li>
<li>Implement biometrics on IoT-connected devices</li>
</ol>,
<ol>
<li>Integrate Arduino with Android to build a remote-control RGB LED device.
</li>
<li>Connect PIR sensors across the IoT to implement a smart building.</li>
<li>Build a temperature and sunlight sensor system to remotely monitor and c
ontrol the conditions of your garden.</li>
<li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
<li>Build a smart irrigation system based on IoT-connected temperature and m
oisture sensors built-in to your agricultural plants.</li>
</ol>,
<ol>
<li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
<li>Automate and control the conditions of a greenhouse.</li>
<li>Enable bio-metrics to build a smart security system for your home or off
ice building</li>
<li>Build an autonomously operating fitness application that automatically m
akes recommendations based on motion detection and heart rate sensors that ar
e embedded on wearable fitness trackers.</li>
<li>Build a healthcare monitoring system that tracks, informs, and automatic
ally alerts healthcare providers based on sensor readings that describe a pat
ients vital statistics (like temperature, pulse, blood pressure, etc).</li>
</ol>,
<ol>
<li>Build an IoT device that automatically locates and reports the closest n
earby parking spot.</li>
```

```
<li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>
<li>Use temperature sensors connected across the IoT to automatically alert
you if your home windows or doors have been left open.</li>
<li>Use bio-metric sensors to build a smart system that automate security fo
r your home or office building</li>
</ol>]
```

Retrieving tags by filtering with regular expressions

```
In [12]: t = re.compile("t")
for tag in soup.find_all(t):
    print(tag.name)
```

```
html
title
strong
strong
strong
strong
strong
strong
```

Retrieving tags by filtering with a Boolean value

```
In [14]: for tag in soup.find_all(True):  
         print(tag.name)
```

```
html  
head  
title  
body  
p  
b  
p  
br  
br  
h1  
span  
strong  
a  
a  
a  
img  
a  
span  
strong  
a  
h1  
ol  
li  
strong  
li  
strong  
li  
strong  
h1  
a  
a  
a  
h2  
ol  
li  
li  
li  
li  
li  
li  
li  
h2  
ol  
li  
li  
li  
li  
li  
a  
img  
h2  
ol  
li  
li
```

```

li
li
li
h2
ol
li
li
li
li
span
strong
a
em
p

```

Retrieving weblinks by filtering with string objects

```
In [15]: for link in soup.find_all('a'):
          print(link.get('href'))
```

```

http://bit.ly/LP1NDJj (http://bit.ly/LP1NDJj)
http://www.data-mania.com/blog/m2m-vs-iot/ (http://www.data-mania.com/blog/m2m-
vs-iot/)
bit.ly/LP1NDJj
http://mat.se/ (http://mat.se/)
http://bit.ly/LP1NDJj (http://bit.ly/LP1NDJj)
https://click.linksynergy.com/deeplink?id=*JDLXjeE*wk&mid=39197&murl=https%3A%2
F%2Fwww.udemy.com%2Ftopic%2Finternet-of-things%2F%3Fsort%3Dhighest-rated (http
s://click.linksynergy.com/deeplink?id=*JDLXjeE*wk&mid=39197&murl=https%3A%2F%2F
www.udemy.com%2Ftopic%2Finternet-of-things%2F%3Fsort%3Dhighest-rated)
http://www.skyfilabs.com/iot-online-courses (http://www.skyfilabs.com/iot-onlin
e-courses)
https://www.coursera.org/specializations/iot (https://www.coursera.org/speciali
zations/iot)
bit.ly/LP1NDJj
http://bit.ly/LP1NDJj (http://bit.ly/LP1NDJj)

```

Retrieving strings by filtering with regular expressions

```
In [16]: soup.find_all(string=re.compile("data"))
```

```
Out[16]: [' & Data Engineering: Sensors that are embedded within IoT devices spin off ma
chine-generated data like it's going out of style. For IoT to function, the pla
tform must be solidly engineered to handle big data. Be assured, that requires
some serious data engineering.']
```

```
In [ ]:
```

