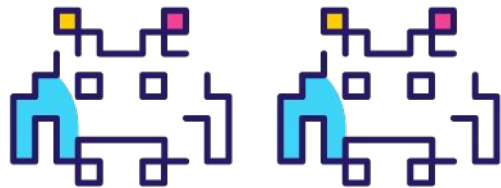




Webinar

**Creating
(your own) tools:
adapting to
enterprise needs**

MULTI PLAYER of the day



Mike King

Founder & Managing Director
at iPullRank



John Murch

Director of Software
Engineering at iPullRank



Colt Sliva

SEO Engineer at iPullRank

You work in Enterprise



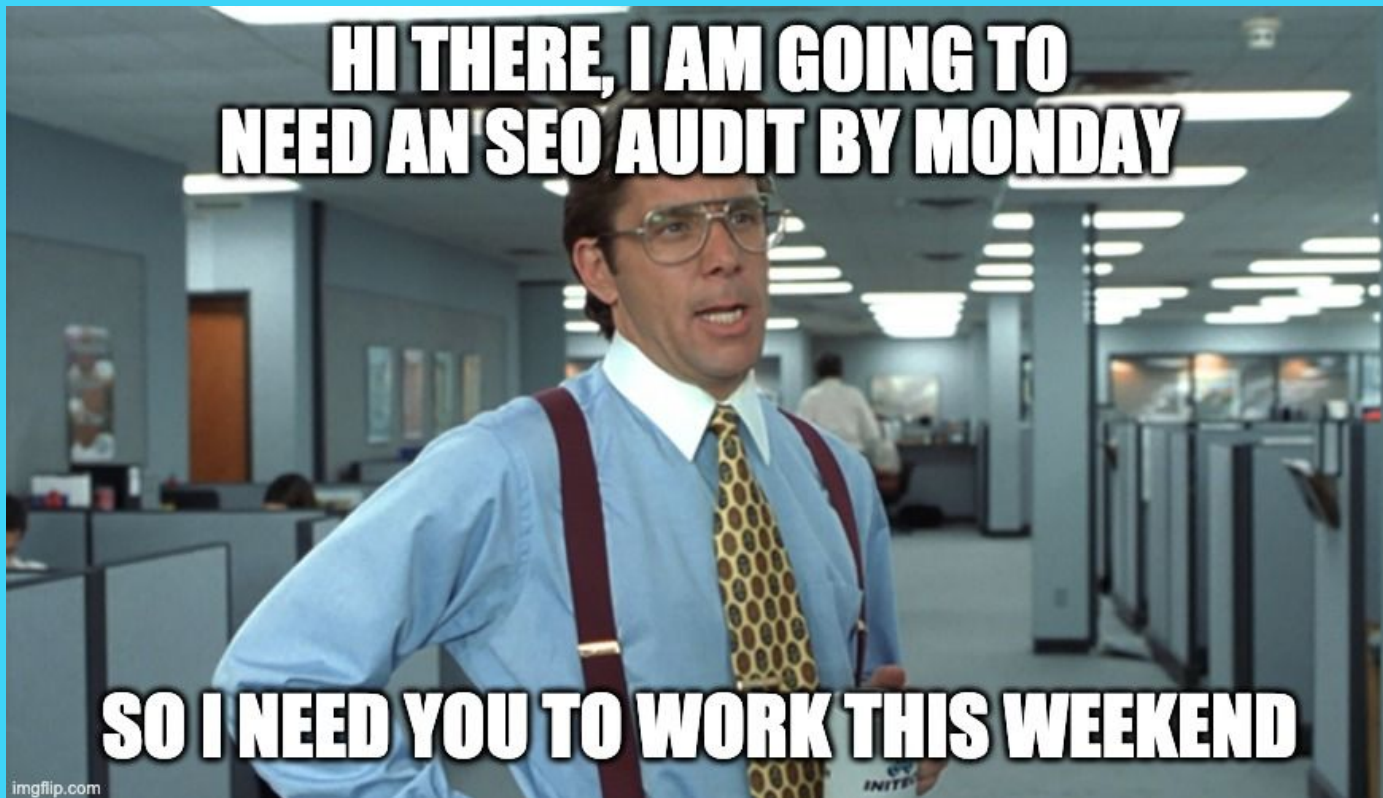
You have a lot of data



Where do you start?



How about an SEO Audit



What makes an enterprise site audit?

- **The same elements that make up all audits:**
 - **Content**
 - **Technical SEO**
 - **On-Page SEO**
 - **Off-Page SEO**
 - **Local SEO**
 - **Mobile vs Desktop**

Goal: Find Issues and Opportunities

We have tools available to make auditing easier, but sometimes they're not actionable for millions of URLs.

We end up with a lot of data from a lot of tools, that can be hard to digest



We have to bring the granularity to the tools.

**A site isn't just a domain. It's templates,
pages, and subfolders.**

**How do we segment and sample to create
insights from raw data?**

We Use Segments!



Focus on Segmentation

- **By Categories**
- **By Products**
- **By** 🤔
- **By URLs** 😍
 - **But first, we need a complete list of URLs.**

How can we Automate Segmentation?



- ~~Excel~~
- ~~Google Sheets~~
- Code



Coding Segments

- ~~URL Pathname~~
 - ~~But Not all sites have friendly URLs~~
- ~~If else Statements~~
- Define Rules and Logic



Awesome – Let's Rule!

- **Build Segmentation Rules and use RegEx**
 - **contains *blue***
- **Operators: And, Or, Not**
 - **contains /category/* but NOT /category/sitemap**
- **How about some simple syntax**

Let's not Reinvent the Wheel, but Expand it!

Build a **universal tool** that works for **any data set** and could be **expanded** to include future **use cases**

Expandable Rules and Logic

[segment: <type>]	// e.g. category
@us-products	// e.g. us only products
path /us/product/*	// e.g. Coffee Cup
not(// AND, OR, NOT
path /uk/products/*	
)	

Step 1. Download Sitemaps

Recursively download ALL the sitemaps!

Step 2. Extract all URLs from sitemaps

**Build a URL index with as many
URLs as possible**

Step 3. Split URLs on “/” and group

```
let folders = url.pathname.split('/')

```

```
www.asos.com/us/men/shoes-boots-sneakers/  
sneakers/cat/?cid=5775

```

```
[“us”, “men”, “shoes-boots-sneakers”, “sneakers”, “cat”]

```

Step 4. Rules to JSON Logic

Define and name your segments and use AND, OR, NOT operators to provide basic logic

Example Rules

Rules

```
[segment:category]
@customercare
path /customer-care/
@men
path /men/*
@jeans
path /jeans/*
not(
path /jeans/outlet*
)
```

=

```
1  {
2    "segments": {
3      "category": {
4        "customercare": {
5          "AND": [{
6            "path": "/customer-care/"
7          }]
8        },
9        "men": {
10         "AND": [{
11           "path": "/men/*"
12         }]
13       },
14       "jeans": {
15         "AND": [{
16           "path": "/jeans/*"
17         }],
18         "NOT": [{
19           "path": "/jeans/outlet*"
20         }]
21       }
22     }
23   }
24 }
```

Step 5. JSON Logic + URLs = Segments

```
"category":{  
  "jeans":{  
    AND:[{ "path":"/jeans/*"}]  
    NOT:[{ "path":"/jeans/outlet*"}]}
```

Example Logic

All URLs with
/jeans/* but not
any that have
/jeans/outlet*

```
1  {
2    "segments": {
3      "category": {
4        "customercare": {
5          "AND": [{
6            "path": "/customer-care/"
7          }]
8        },
9        "men": {
10         "AND": [{
11           "path": "/men/"
12         }]
13       },
14       "jeans": {
15         "AND": [{
16           "path": "/jeans/"
17         }],
18         "NOT": [{
19           "path": "/jeans/outlet*"
20         }]
21       }
22     }
23   }
24 }
```

Step 6. Show Stats & Repeat

```
{“category”:{  
  “mens”:1322,  
  “jeans”:11,  
  “customercare”: 10}  
}
```

Output your data

type	name	urls
category	customer care	https://www.asos.com/us/customer-care/
category	customer care	https://www.asos.com/us/customer-care/help/
category	customer care	https://www.asos.com/us/customer-care/get-in-touch/
category	customer care	https://www.asos.com/us/customer-care/delivery/
category	customer care	https://www.asos.com/us/customer-care/engagement/
category	customer care	https://www.asos.com/us/customer-care/order-issues/
category	customer care	https://www.asos.com/us/customer-care/payment-promos-gift-voucher/
category	customer care	https://www.asos.com/us/customer-care/returns-refunds/
category	customer care	https://www.asos.com/us/customer-care/product-stock/
category	customer care	https://www.asos.com/us/customer-care/technical/
category	men	https://www.asos.com/us/men/
category	men	https://www.asos.com/us/men/sale/a-to-z-of-brands/cat/?cid=8420
category	men	https://www.asos.com/us/men/a-to-z-of-brands/cat/?cid=1361
category	men	https://www.asos.com/us/men/sale/a-to-z-of-brands/one-above-another/
category	men	https://www.asos.com/us/men/gifts/cat/?cid=16091
category	men	https://www.asos.com/us/men/sale/cat/?cid=8409
category	men	https://www.asos.com/us/men/outlet/cat/?cid=27396
category	men	https://www.asos.com/us/men/overalls-boiler-suits/cat/?cid=50009
category	men	https://www.asos.com/us/men/two-piece/cat/?cid=28291
category	men	https://www.asos.com/us/men/loungewear/cat/?cid=18797
category	men	https://www.asos.com/us/men/sweaters-cardigans/cat/?cid=7617
category	men	https://www.asos.com/us/men/shoes-boots-sneakers/cat/?cid=4209

Segment Type
Name
URL

Ways you can leverage this data

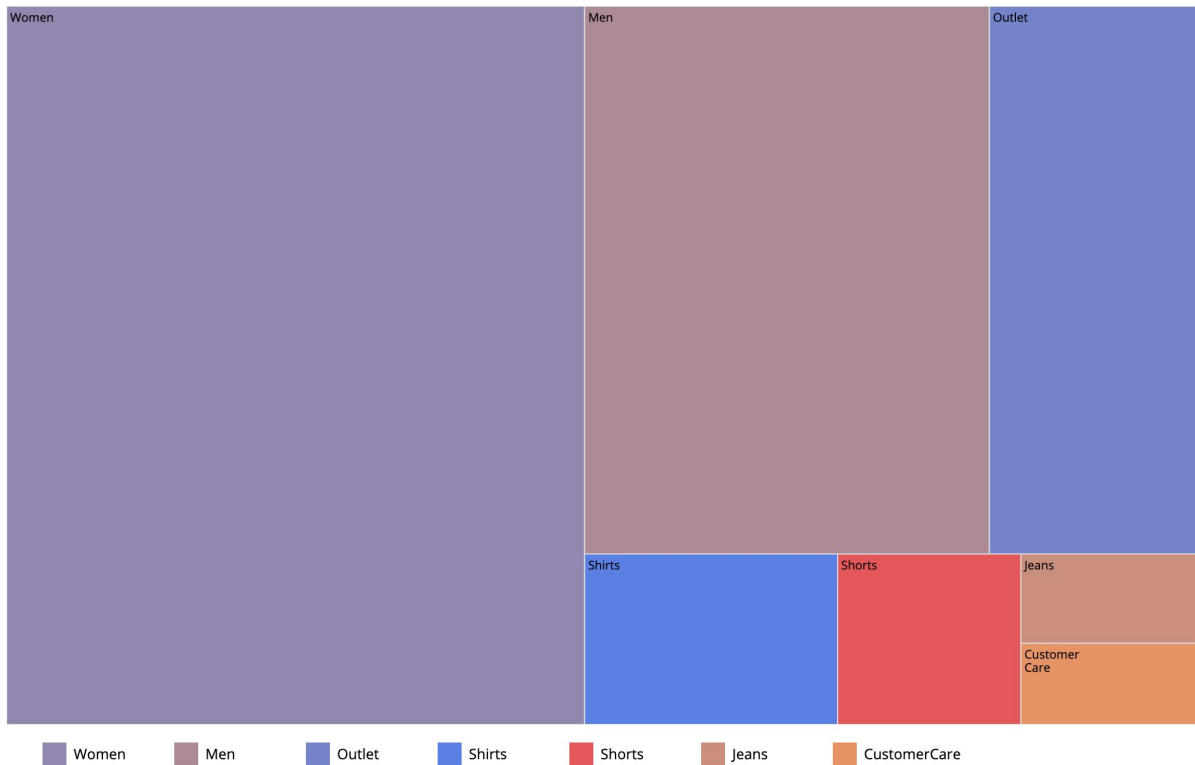
- **Redirect Mapping for expiring content**
- **Keyword Opportunities**
- **Internal Linking**



Segments  but client wants more...

- **Data Visualization**
- **Data Enrichment**

Treemap



Data Enrichment



- **Use URLs and Segments to lookup data on**
 - **ahrefs**
 - **Semrush**
 - **Oncrawl**
 - **Google Search Console**

What does the client get?

- **Keyword universe size by segment**
 - Discover thin content
- **Backlink volume by segment**
 - Find low authority clusters
- **Average internal link count within segment**
 - Flag pagerank sculpting opportunities
- **Average Click-through-rate within segments**
 - Prioritize conversion rate optimization
- **Keyword winners and losers with greater detail**
 - Understand performance across lines of business



🔥 Terminal Commands 🔥

- `sort -nk1 urls.txt`
- `cat urls.txt | less`
- `diff -y jan.txt feb.txt`
 - Compare sitemap data versus archive.org
- `wc -l urls.txt`



🤔 Thoughts on Tools

- **Keep it Simple**
- **Make Building Blocks**
- **Edge Cases**
- **Git First, API Second**
- **Solve for today, plan for tomorrow**

Thank you!



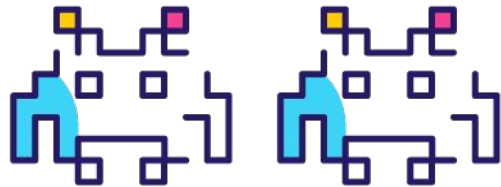
Mike King
@ipullrank



John Murch
@johnmurch



Colt Sliva
@signorcolt





**Grab your controller,
it's time for Q&A!**