



How to Create an Amazon-Like Experience with Fusion

The massive global success of Amazon can be tied to how the company employs search. Yes, of course, customers love free shipping, liberal return policies, and competitive prices. But first and foremost, customers know when they go to Amazon.com, they'll find exactly what they are looking for. Every part of Amazon.com is a big search UI powered by a powerful, scalable back end with several key features. You, too, can have these capabilities using the [Lucidworks Fusion suite of products](#).

Let's take an inventory of these features, from the top of the Amazon homepage all the way down to the footer.

The Search Bar

The Amazon search bar demonstrates two fundamental features: typeahead and faceting. Typeahead guesses what you're about to type before you finish typing the query. Faceting suggests departments you might limit your search to. For instance, once I type "blue sue," Amazon suggests I look for blue suede shoes in the men's shoes and music departments.



Once I select "blue suede shoes in men's shoes," I'm taken to a set of search results showing that type of shoe in the men's shoe category. I also see a dropdown box next to the search bar that lets me pick other departments. Fusion allows you to implement similar functionality to the typeahead (aka autocomplete) functionality. This is available by default.

Faceting is part of a [Fusion query pipeline](#). You can add facetting via a point-and-click interface to select the field you want to facet. You can also easily develop search UIs that incorporate typeahead and faceting features.

The screenshot shows the Fusion Query workbench interface. On the left, a sidebar titled "Add a field facet" lists various fields: absolute_resource_name_s, absolute_resource_name_t, body_t, charSet_s, Content_Encoding_s, Content_Encoding_t, Content_Length_i, Content_Location_s, Content_Location_t, and Content_Type_s. A "Filter" input field and a magnifying glass icon are at the top of this sidebar. On the right, search results are displayed:

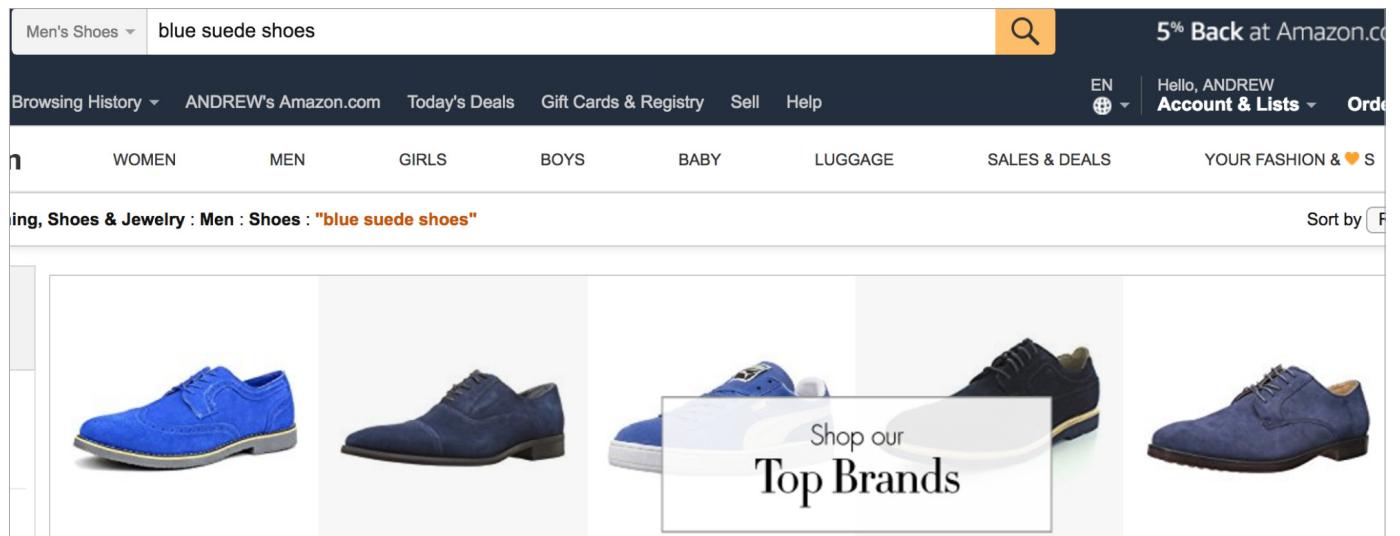
- Japanese Bobtail - Wikipedia**
https://en.wikipedia.org/wiki/Japanese_Bobtail
Score: 15.242611 [show fields](#)
- Cymric cat - Wikipedia**
https://en.wikipedia.org/wiki/Isle_of_Man_Longhair
Score: 13.888628 [show fields](#)
- Dragon Li - Wikipedia**
https://en.wikipedia.org/wiki/Dragon_Li
Score: 13.888628 [show fields](#)

At the bottom right, there are navigation links: 1, 2, 3, 4, 5, 6, 7, 8, ..., 15, Next. Below the navigation is the text: 1-10 of 141 docs (5 ms, max-score 15.242611).

Adding a faceted field via the Fusion Query workbench

Top Brands

After searching on “blue suede shoes” and going to the results page, Amazon boosts “top brands,” most likely based on merchandising policies or paid sponsorships.



Amazon Promoted Items

These top brand results can be “boosted” into the main set of search results or shown as a separate companion query. You can boost on a field in Fusion using the Query Fields stage of the Fusion query pipeline. This is all set up by default. You just specify the field (i.e. “sponsored”) and boost it.

You can also have a separate sub-query against a collection of brands and boost results matching those brands. Alternatively, this may just be another form of faceted search based on brand name.

Query Fields

Query Fields
Configure query parameters for Solr search

Label

Condition

Number of Results

Result Offset

Results Sort Order

Search Fields

<input type="button" value="+"/> Field Name	Field Boost
<input type="button" value="X"/> dc_title_t	<input type="text" value="40"/>

Boosting a field with Fusion “Query Fields” stage

Related to Items You've Viewed

On the Amazon homepage, you can see what are essentially the results of your most recent search or a few previous searches where you actually clicked on something.

Using [Fusion's signals](#) you can get a list of recently clicked items or just recent searches generally.



Amazon Promoted Items

params.user_name_s:whayes	Choose Sort Field ▾	Parameters (2)	URI
Add a field facet			
params.lw_url_s		https://en.wikipedia.org/wiki/Japanese_Bobtail	
params.page_s		1	
params.platform_s		MacIntel	
params.position_s		1	
params.user_agent_s		Mozilla/5.0 (Macintosh; Intel Mac OS X 10_11_6) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/56.0.2924.87 Safari/537.36	
params.user_name_s		whayes	
query_orig_s		*	
query_s		*	
query_t		[***]	
score		2.3025851	
timestamp_tdt		2017-03-14T17:26:33.720Z	
type_s		click	
tz_timestamp_txt		["Tue 2017-03-14 17:26:33.720 UTC"]	
version		1561866947858006000	

Signals for user clicks from Fusion Query Workbench

Based on your signals and the items you picked, you can use the [Item for Item](#)

[recommender](#) to find other things a user might be interested in purchasing.

More Items to Consider

Amazon also suggests even more items to consider based on your previous searches.

More top picks for you



Amazon top picks

These are “similar” to things that you’ve clicked on. For users that have history, Fusion can do this automatically with the [Recommend Items for User recommender](#). Previous user searches, clicks, shopping cart adds or purchases can be captured as [Fusion signals](#).

What about users without a lot of history? Using the “More Like This” feature you can provide the same functionality for your search UI. All you need is a single or list of item IDs.

MoreLikeThis takes the ID of an item and returns a list of similar items. It is accessed via a simple REST API for easy integration into your search application.

Query Fields

Query Fields

Configure query parameters for Solr search

Label

search-fields

OPEN EDITOR

Condition

1

Number of Results

10

Result Offset

0

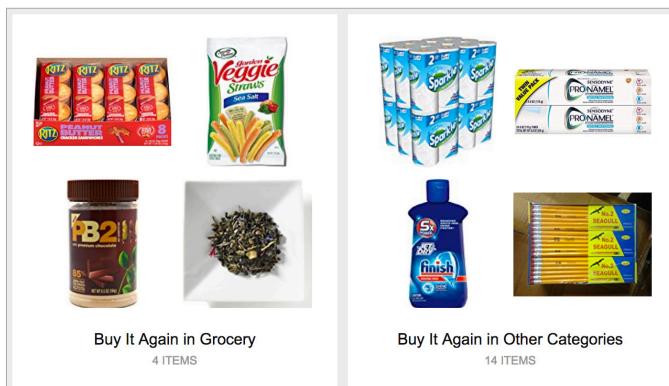
Results Sort Order

Search Fields

Field Name	Field Boost
dc_title_t	40

Your Amazon.com

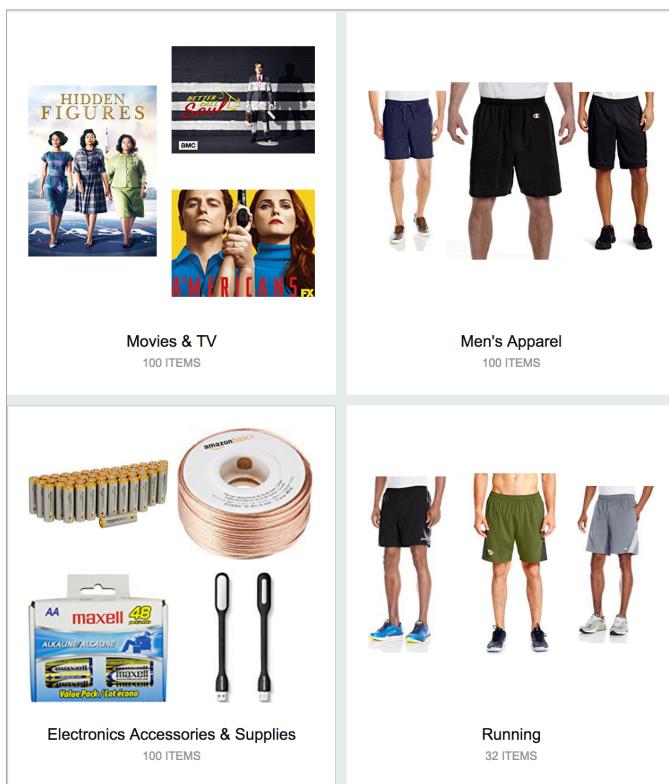
When you click on Your Amazon.com you get a much closer look at of how Amazon views you. First, they show you things that you've bought (which you can capture via signals). Next, they show you categories (Facets) which you've purchased items from with key items shown as representations (mainly Item or User recommendations or MoreLikeThis on a searched item or purchase) and finally a set of movies "inspired by your purchases." Which can just be accomplished by an Items for User recommendation based on a collection of your past purchases (signals).



Amazon "buy it again"

The Lucidworks Query Workbench interface shows a search for 'cats'. The results pane displays a list of documents related to Japanese Bobtail cats, including their Wikipedia page. The facets pane on the left shows categories like 'Recommendation Boosting', 'Query Fields', 'Field Facet', and 'Solr Query'. The parameters pane on the right shows various search parameters and their values.

Fusion signals in the Query Workbench



Amazon Recommended categories that you've made purchases in

Add a field facet

▼ absolute_resource_name_s x ≡

- https://en.wikipedia.org/wiki/1920_Nor... (1)
- <https://en.wikipedia.org/wiki/Abyssinian...> (1)
- https://en.wikipedia.org/wiki/American_... (1)
- https://en.wikipedia.org/wiki/American_... (1)
- https://en.wikipedia.org/wiki/American_... (1)

[View next 10](#)

Fusion Faceted Field in Query Workbench

Your recently viewed items and featured recommendations

Inspired by your purchases

Page 1 of 8



The Importance of
Being Earnest
› Oscar Wilde
★★★★★ 326
Mass Market Paperback
\$2.70



Ghosts
› Raina Telgemeier
★★★★★ 394
Paperback
\$6.55



Ghosts (Dover Thrift
Editions)
› Henrik Ibsen
★★★★★ 43
Paperback
\$1.54



Smile
› Raina Telgemeier
★★★★★ 1,040
Paperback
\$6.70



An Ideal Husband
(Dover Thrift Editions)
Oscar Wilde
★★★★★ 20
Paperback
\$3.00



Kristy's Great Idea:
Full-Color Edition...
› Ann M. Martin
★★★★★ 435
Paperback
\$6.70



Duduma Polarized
Designer Fashion Sports
Sunglasses for Baseball
Cycling Fishing Golf
Tr62...
★★★★★ 1,559
\$19.99

Amazon "Inspired by your purchases"

Recommend More Like This

Returns results similar to a given item using Solr's MoreLikeThis component. Provide parameters for the component here. For more information, see
<https://cwiki.apache.org/confluence/display/solr/MoreLikeThis>

Label

|

OPEN EDITOR

Condition

1

Use Query Parser

More Like This Fields

+ MoreLikeThisFields

Click the green plus icon above to add a row

DocId Field name

id

mindf

2

Fusion AI Recommend Similar Items stage (update to More Like This)

Improve your Recommendations

If you click on Improve your Recommendations you see a list of purchases and an opportunity to control how they affect Amazon's recommendations. You could do something similar with Fusion by capturing the feedback in another collection and aggregating it together in a subquery, boosting items with highly rated purchases first. You might also opt to simply remove purchase signals that users remove from this feedback or rate poorly. One might suspect that this "improve your recommendations" form is used rarely, but user ratings on other parts of the site are considered in a similar way.

Beyond Personalization

Normally at Lucidworks we'd put this first, but it isn't quite as obvious on Amazon.com as the more personalized features. When you search on a generic term like "green tea," what other users do affects your results as well. You can see this when you search while logged out and compare it to what you see when logged in.

The screenshot shows the Amazon search results for "green tea". The search bar at the top has "green tea" entered. Below the search bar, there are navigation links for "All", "Departments", "Your Amazon.com", "Today's Deals", "Gift Cards & Registry", "Sell", and "Help". On the right side, there are links for "EN", "Hello, Sign in", and "Account & Lists". The main search results area displays 1-24 of 47,580 results for "Grocery & Gourmet Food : Coffee, Tea & Beverages : Tea : green tea". A sidebar on the left provides filtering options: "Show results for Any Category", "Grocery & Gourmet Food", "Coffee, Tea & Beverages", "Tea" (with sub-options for Green, Tea Samplers, Herbal, Iced Tea), and "See more". Other filters include "Subscribe & Save", "Delivery Day", "Amazon Prime", "Eligible for Free Shipping", "Pantry", and "Tea Packaging". The main content area features a sponsored ad for VAHDAM TEAS, followed by several product cards for different types of green tea, including Matcha, Organic Green Tea Leaves, and Numi Organic Tea Gunpowder Green. Each product card includes an image, a brief description, a star rating, and purchase information.

Amazon "green tea" recommendations for an anonymous user

Users that have not logged in see sponsored content first and results boosted based on what other users clicked on or purchased based on that search. You see that Matcha brand is boosted to the first slot for the anonymous user. A more experienced tea drinker wouldn't have clicked on or purchased matcha powder when searching on green tea. So when a tea drinker, with a tea search, and purchase history is logged in, they see less sponsored content, and less content recommended by the broader user community.

Showing results in **Grocery & Gourmet Food**. Show instead results in [All Departments](#).

Product Image	Product Name	Rating	Price	Shipping	Actions
	Uncle Lee's Organic Green Tea -- 100 Tea Bags net wt...	★★★★★ 625	\$8.74	✓Prime FREE One-Day	Add to Cart
	Matcha Green Tea Powder Organic (Japanese Premium Culinary...)	★★★★★ 1,443	\$10.99	✓Prime	Add to Cart
	Stash Tea Premium Green Tea, 100 Count Box of Tea Bags...	★★★★★ 4,903	\$16.06 (\$0.16/Count)	Subscribe & Save	See Flavor Options

Amazon "green tea" Recommendations for a logged in user

The recommendations based on what other users do is a type of collaborative recommendation. This is a feature offered by the default Fusion query pipeline. A search application merely needs to capture the signals.

Query Workbench

cats-default

Recommendation Boosting

Recommendation Boosting
Uses recommended items for search time boosting

Label: recommendation

Condition

Number of Recommendations: 10

Number of Signals: 100

Aggregation Type: *

Solr Field to Boost On: id

```
/*
 * Helper method to document click events.
 * @param {string} docId The document id
 * @param {object} options An object containing parameter overrides and options.
 * - The object can be any parameter which will be passed through, including parameters.
 * - Ex: {type: 'custom', params: {filterQueries: ['something']}}
 * @return {promise}
 */
function postClickSignal(docId, options) {
  var date = new Date(),
    data = {};
  data = {
    docId: docId,
    head_field: ConfigService.config.head_field,
    language: ClientStatsService.getBrowserLanguage(),
    platform: ClientStatsService.getBrowserPlatform(),
    user_agent: ClientStatsService.getBrowserUserAgent(),
    user_name: ConfigService.getLoginCredentials().username || ConfigService.config.anonymous_access.username,
    query: QueryService.getQueryObject(),
    pipeline: ConfigService.config.signals_pipeline,
    timestamp: date.toISOString(),
    type: ConfigService.config.signal_type
  };
  _defaultsDeep(data, options);
  return postSignalData([data]);
}
```

A signal as written in JavaScript

Fusion Recommendation Stage

Another key feature of Amazon is the longer spread analytics. By capturing user purchase signals, they're able to see when two purchases are frequently made together and recommend purchasing them together. Sometimes they even offer deals for these complimentary purchases.

Frequently bought together

Total price: **\$35.03**

Add both to Cart

Add both to List

This item: Purina Tidy Cats LightWeight 24/7 Performance Clumping Cat Litter **\$16.14**

Purina ONE Indoor Advantage Adult Premium Cat Food **\$18.89**

Amazon Frequently Bought Together

Fusion captures not only signals but allows for aggregations and other kinds of derived information. This might include purchases that usually happen in the same cart or close together in time.

User Intent

As you can see, you can implement similar features to Amazon using Lucidworks Fusion. You can install Fusion, [ingest data and have basic collaborative recommendations with typeahead and faceting in less than an hour](#). When you get into deeper personalization, you'll need to measure and tweak your results. A key aspect of this is time. Just because someone purchased diapers 12.5 years ago, doesn't mean they're particularly likely to purchase them today. Quite a bit of [research has gone into the factors that make up measures of user intent](#). This includes derived results like demographics with time. Some research indicates that older users tend to be "more certain" and younger users "less certain." An older customer buying a particular type of soap is more likely to buy that same type for a longer time in the future than a younger customer. In concrete terms, this might mean you expire signals for older shoppers more quickly than for younger ones.

The screenshot shows the Lucidworks Fusion Query Workbench interface. At the top, there's a navigation bar with a file icon, the word 'cats', a dropdown arrow, the Lucidworks logo, 'Help', 'admin', and a notification icon. Below the header, the title 'Query Workbench' is displayed, along with 'New', 'Load...', 'Save', and a '+' button.

The main area is titled 'cats-default'. On the left, there's a sidebar with four items: 'Recommendation Boosting', 'Query Fields', 'Field Facet', and 'Solr Query', each with a three-dot menu icon. In the center, there's a search bar with 'cats' and a magnifying glass icon, followed by a 'Choose Sort Field' dropdown set to '*:*'. To the right of these are 'Parameters (2)' and a 'URI' field.

A large panel on the right contains a table with three rows. The columns are labeled '_lw_parser_resource_name_s', '_lw_parser_type_s', and '_version_'. The first row has a link to 'https://en.wikipedia.org/wiki/Japanese_Bobtail'. The second row has the value 'tika'. The third row has the value '1559492924171878400'. Below this table is a 'preview_t' section containing the text: "Japanese Bobtail From Wikipedia, the free encyclopedia. A jump to: navigation, search. This article is about the cat. For the squid, see Japanese Bobtail Squid. Japanese Bobtail is a breed of domestic cat with an unusual \"bobbed\" tail more closely related to the Cymric cat than to the domestic cat. It is a medium-sized cat with a thick, double coat and a short, rounded tail. The name 'Bobtail' comes from the fact that the tail is very short or non-existent. The Japanese Bobtail is a relatively new breed, having been developed in Japan in the 1960s. It is a popular breed in Japan and has gained popularity elsewhere in recent years. The Japanese Bobtail is a friendly and playful cat, often described as having a 'dog-like' personality. It is a good companion for children and other pets. The Japanese Bobtail is a healthy breed, with a long life expectancy of 15-20 years. It is a low-maintenance cat, requiring minimal grooming and care. The Japanese Bobtail is a unique and beautiful breed, making it a great choice for anyone looking for a special pet."

At the bottom of the preview panel, there's a link 'Cymric cat - Wikipedia' and a page navigation bar with links 1 through 15 and a 'Next' button. Below this, a note says '1-10 of 141 docs (6 ms, max-score 18.088642)'. At the very bottom of the preview panel are 'Format Results' and 'View As: Results' buttons.

Fusion's Query Workbench

The Fusion platform is perfect for everything that goes into measuring user intent. Fusion's Index and Query Workbench features let you test your personalization and recommendation algorithm tweaks before you deploy them or change your data. Fusion includes powerful recommender features as well as machine learning functionality powered by Apache Spark. Additionally, Fusion is especially adept at [managing time series data](#) which is key to handling the kind of event measures that go into user intent.

The screenshot shows the Lucidworks Fusion Index Workbench. At the top, there's a navigation bar with a file icon, the text "webserver", a dropdown arrow, the Lucidworks logo, "Help", "admin", and a notification bell icon.

The main area is titled "Index Workbench". On the left, there's a sidebar with a tree view showing two entries: "training-access-log" (Status: [\(Start Job\)](#), Last run: never) and "webserver-default" (with a "Replace" button). Below this are five stages: "Field Mapping", "Regex Field Extraction", "Date Parsing", "Solr Dynamic Field Name Mapping", and "Solr Indexer", each with a "Replace" button and a "Filter" input field.

The right side displays "Simulated Results [1]" for the "training-access.log#0" document. The document details are as follows:

```

id: "training-access.log#0"
context: {"batchId": "d2773e8a-1992-4134-b0d8-60a5290c5a8a", "simulate": true, "a...
fields: [35]
body_t: 139.162.108.53 -- [12/Feb/2017:13:45:52 +0000] "GET / HTTP/1.1" 40...
ip_addr_t: 139.162.108.53
ip_addr_s: 139.162.108.53
request_t: GET / HTTP/1.1
request_s: GET / HTTP/1.1
response_t: 401
response_s: 401
sent_s: 612
sent_t: 612
timestamp.tz_t: UTC
timestamp.tz_s: UTC
timestamp_dt: 2017-02-12T13:45:52.00Z
timestamp.tz_offset_s: +00:00
timestamp.tz_offset_t: +00:00

```

On the far right, there's a vertical sidebar with document counts: "1 Documents", "5 Documents", "10 Documents", and "20 Documents". Below that is a "View:" dropdown set to "1 Documents".

Fusion's Index Workbench simulating a change

Going Beyond Amazon

Automatic Groups (Clustering)

However, you can get one leg up on Amazon. Using Fusion's [clustering capabilities](#), you can automatically group items. This means that even though you might not have a facet or category identified for an item, you can still create a type of grouping with Fusion. This has several uses in retail. A user can find items that are in the 'same group' as other items and refine searches within that group in order to achieve more precise results. I can also use clustering to tune my promotions, matching up similar users with similar offers.

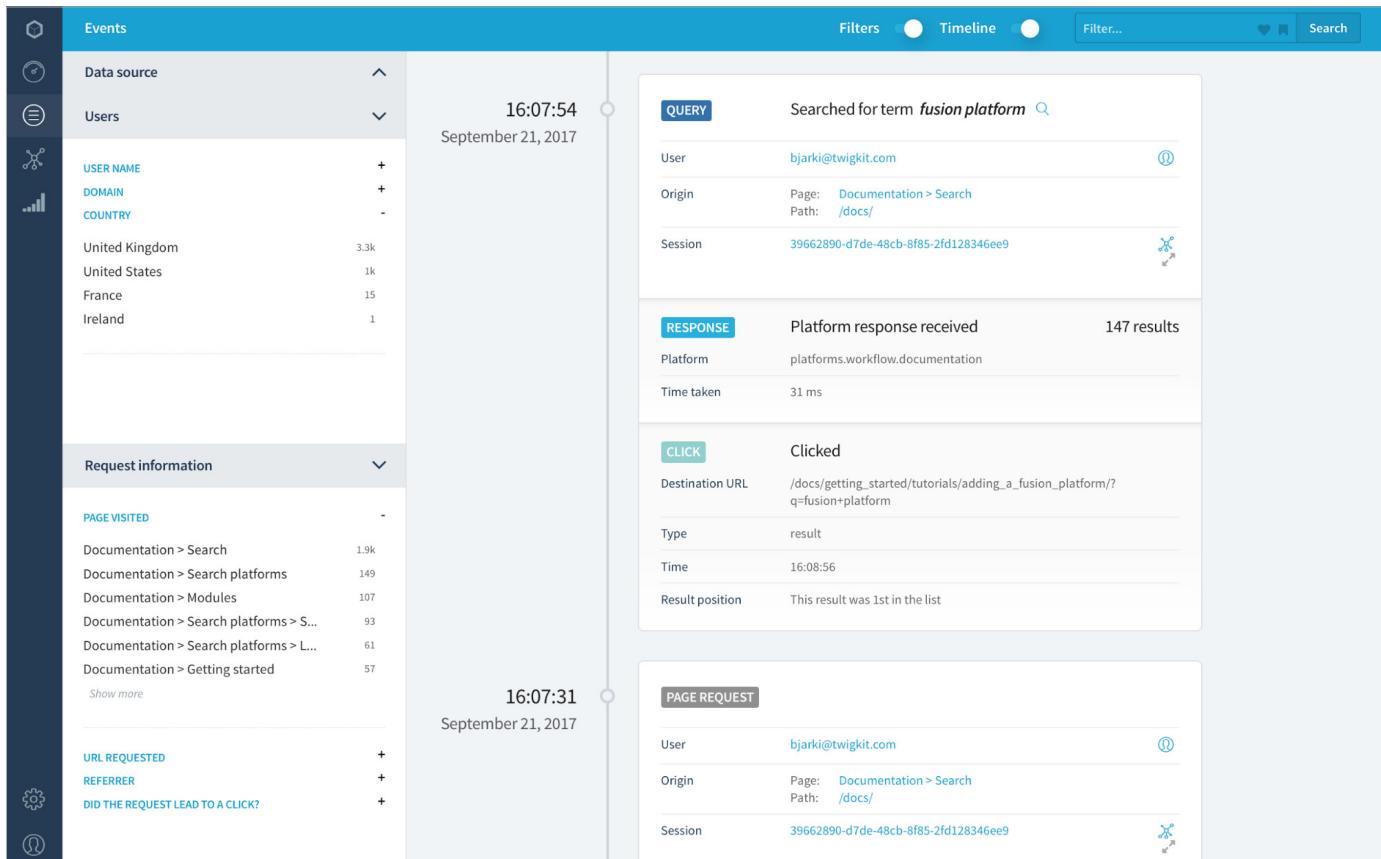
Classification

Where you do have categories, you can automatically determine which products fit into those categories via Fusion's [classification capabilities](#). This means you can avoid a manual process of sorting new inventory into categories. It also means you can discover what other categories an item might fall into. For example, just because the manufacturer called it a "4k smartphone" doesn't mean it isn't also a "android phone" or "phablet."

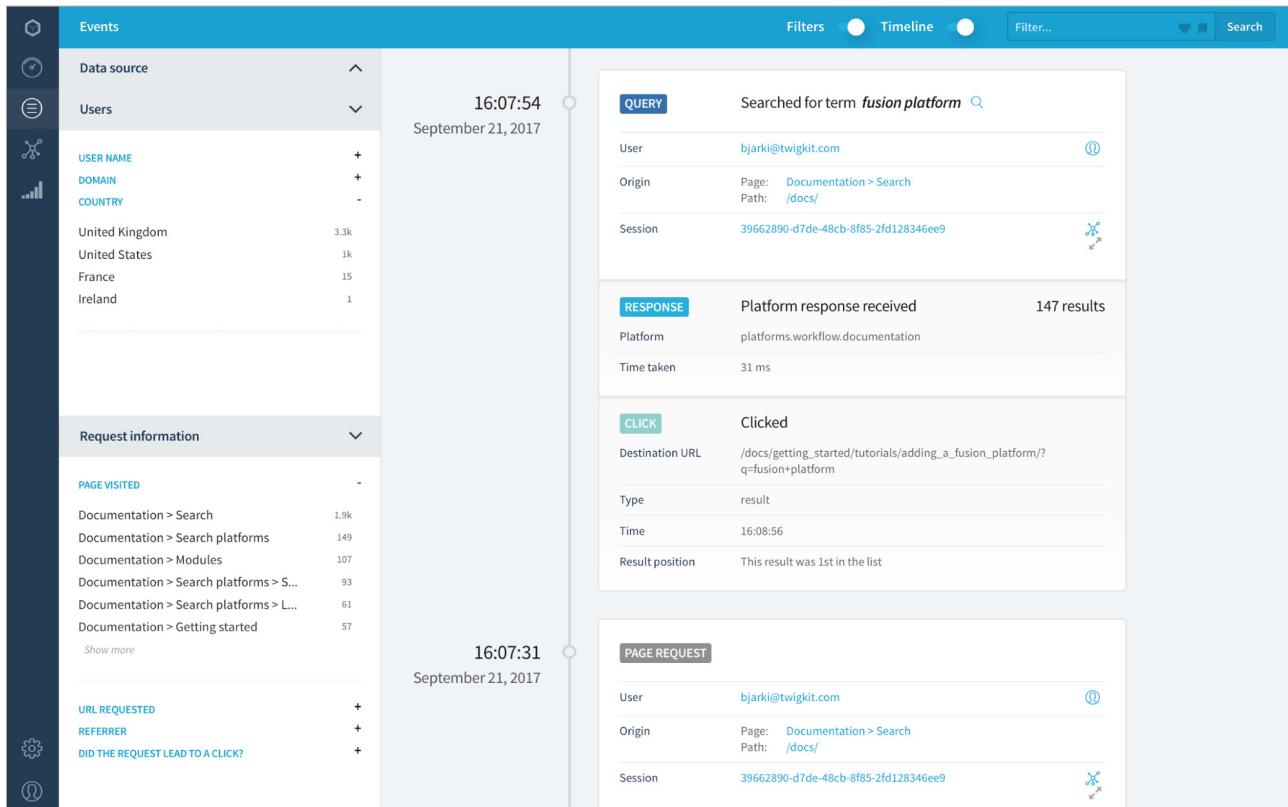
Effectiveness

One thing that is clear for longtime customers of Amazon is that they tweak their algorithms. Every ecommerce company has to do this. Customer behavior changes.

For example, some years back, customers rarely used their phones to make a purchase. Now, it is an everyday occurrence for many people.



It is critical to use analytics to measure the effectiveness of your site's personalization and general relevance tuning choices. Fusion comes with powerful analytics tools to do just that. You can use these capabilities to understand your top searches, your most frequent searches that didn't return a result, and other important metrics.



The Search UI

Using Fusion [App Studio](#), you can easily develop a basic search UI with collaborative recommendations. You'll still need some additional attention to UI to achieve the same kind of personalized recommendations as above. But the back end is relatively straightforward. Future Lucidworks releases will accelerate personalized search UI development just like Fusion has accelerated the back end data processing of search.

If you're selling stuff on the Internet, there is no reason you can't have all of the same features that make Amazon a success. All it takes is a little determination and the Lucidworks Fusion platform.

Getting Started

So you're thinking, "That's great! Where do I get this stuff?" You can get started with the Fusion platform with these simple steps:

- The first step is to download [Lucidworks Fusion](#).
- Check out our webinar on [Implementing Site Search in an Hour](#). Consider trying that out!
- Sign up for [training](#).
- [Talk to our experts](#) who have accelerated sales for companies like yours!

Amazon.com and Amazon are trademarks of Amazon.com, Inc. Screenshots and other representations are used under the principle of "fair use" for demonstrative and educational purposes. No association between Lucidworks, Lucidworks Fusion, and Amazon is intended to be represented or implied.