

# Optimized Merchandising - Unleashing Search

I don't think companies can spend too much time on thinking about search and making search great. That is how customers are engaging with you.

Paul Miller,  
Digital Acceleration Leader.



Product findability may be the single most important function of your commerce solution. Search engine users are around 7-10 times more likely to convert than regular website visitors<sup>1</sup>, despite the fact that most online shoppers report their onsite search experience to be mediocre at best. The good news is that technology has made incredible advancements in supporting areas such as natural language processing, machine learning, cloud computing, and big data, which when combined are creating much more powerful and concise search tools for today's online shoppers. In this whitepaper, we explore the key challenges online shoppers and organizations face and the core elements of a search solution which can meet them.

Imagine searching for "dress shoes" on a fashion retailer's site and the search engine returned a list of dresses and athletic shoes. What would cross your mind? You would not be alone if you simply abandoned the site to find one that does a better job of respecting your time by delivering quality results. The Baymard Institute found in a 2020 study performed that 61% of sites require their users to search by the same product type jargon the site uses, e.g. failing to return all relevant products for a search such as "blow dryer" if "hairdryer" is used on the site, or "multifunction printer" vs "all-in-one printer"<sup>2</sup>.

Shoppers are free to enter any combination of characters they want. They will misspell words, they will use slang and colloquialisms. They will use the measurement system of their choice, they will be hyper-specific and hyper-general. They will paste in long phrases that they copied from somewhere on the web. They will search in their native language and use accent marks. If they can do it, they will do it and they still expect you to return high quality results. Why? Because others have figured it out and they have raised the bar of expectations. The following illustration is a case in point:

Google is the world leader in intelligent search. Over 80% of all web searches in the world flows through them. Why? Because they have figured out how to deliver relevant results to a world of busy people on the move, typing only with their thumb! Despite all of the necessary underlying technology being available, B2C and B2B brands are still struggling to provide high-quality results for even the most basic of searches.

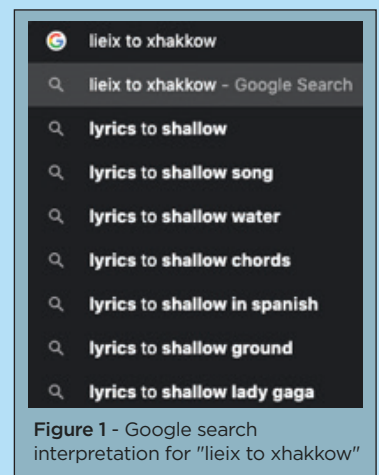


Figure 1 - Google search interpretation for "lieix to xhakkow"

<sup>1</sup> "Top 25 Best Practices for eCommerce Search", Paige Tyrell, 201

<sup>2</sup> "Deconstructing E-Commerce Search: The 8 Most Common Query Types", Baymard Institute, 2020

## Did you know...

**46% of sites**  
don't support thematic search queries such as "spring jacket" or "office chair".<sup>2</sup>

**32% of sites**  
don't support symbols and abbreviations for even the most basic units, resulting in users missing out on perfectly relevant products if searching for inch when the site has used " or in in their product data.<sup>2</sup>

**27% of sites**  
won't yield useful results if users misspell just a single character in a product title.<sup>2</sup>

**25% of sites**  
don't support non-product search queries, like "returns" or "order tracking".<sup>2</sup>

## Why Elasticsearch



From the creators of the ELK stack, Elasticsearch is an open-source search and analytics solution that combines an unmatched breadth and depth of capabilities with extreme scalability and performance. Adopted by mega-companies such as Netflix, eBay, and Facebook, Elasticsearch is a revolutionary update to previous search solutions in several important ways:

### The ability to ask your data questions of all kinds.

- **Search your way:** Elasticsearch lets you perform and combine many types of searches — structured, unstructured, geo, metric — any way you want. Start simple with one question and see where it takes you.
- **Analyze at scale:** It's one thing to find the 10 best documents to match your query. But how do you make sense of, say, a billion log lines? Elasticsearch aggregations let you zoom out to explore trends and patterns in your data.

### Elasticsearch is fast. Really, really fast.

- **Rapid results:** When you get relevant answers instantly, your customers have a better experience.
- **Powerful design:** Elasticsearch implemented inverted indices with finite state transducers for full-text querying, BKD trees for storing numeric and geo data, and a column store for analytics. That's tech speak for "a stable platform to take on any search challenge and execute it fast".
- **All-inclusive:** Since everything is indexed you can leverage and access all of your data at ludicrously awesome speeds.

### Scale isn't a problem

- Elasticsearch doesn't care if its running on a single node or a 300-node cluster - the experience is the same either way...awesome.
- It scales horizontally to handle millions or more of events per second, while automatically managing how indices and queries are distributed across the cluster for smooth operations.

## Search across everything and get the specific results you are looking for.

- Rank your search results based on a variety of factors — from term frequency or recency to popularity and beyond. Mix and match these along with functions to fine tune how your results show up to your users.
- And because most of our users are human, Elasticsearch is equipped to handle human mistakes including complexities like typos.
- Augment your data with data from other data sources.
- Search for retail store locations, customer service terms like “returns”, “shipping” and other non-product content.

## Elasticsearch has your back and is constantly maintaining and improving the platform so you can focus on your business.

- Elasticsearch detects failures to keep your cluster (and your data) safe and available.
- With cross-cluster replication, a secondary cluster can spring into action as a hot backup.
- Elasticsearch operates in a distributed environment designed from the ground up for perpetual peace of mind.

## Natural Language Processing

If you're like most people and have used a digital assistant like Siri, Alexa or Cortana, you are already at least subconsciously familiar with natural language processing or NLP. For the uninitiated, natural language processing is the technology at the heart of understanding the human meaning behind spoken or written phrases. In the case of a digital voice assistant, there is a three-step process - converting sounds to text, processing the text to understand it, then delivering relevant results based on the meaning of what was asked. When you ask Siri “What will the weather be like tomorrow?”, the NLP might break down your question as follows:

1. ~~What will the~~ **weather** ~~be like~~ **tomorrow**
2. **weather** (noun), **tomorrow** (time)

Similarly, in a search solution, an NLP tool can break down a search such as “This year’s fall look for women” into:

1. **This year’s fall** ~~look for~~ **women**
2. **This year** (product attribute: current year), **fall** (product attribute: current season), **women** (gender: female)

Assuming your product catalog data organizes products by year, season and gender, a search like this would hit on all of these results.

Another significant benefit of NLP is the ability to understand natural language across many different languages - a big boost to any vendors who are global or going global as it adds a level of assurance that you will be able to deliver a great search experience in the searchers native language.

Beyond digital assistants, virtually every industry in the world relies on NLP in an attempt to understand information expressed in one of the hundreds of human languages and act on it programmatically. NLP does pretty much exactly what it sounds like it does - it processes natural language. It can determine parts of speech like nouns, verbs, and adjectives. It can determine if the first period in “That will be \$15.99 please.” is a decimal point and not the end of the sentence. It can understand if a sentence has been negated, or double negated, or even triple negated. It deciphers meaning so that you can ask questions in natural language and get a reasonable response from a computer system.

## Apache NiFi

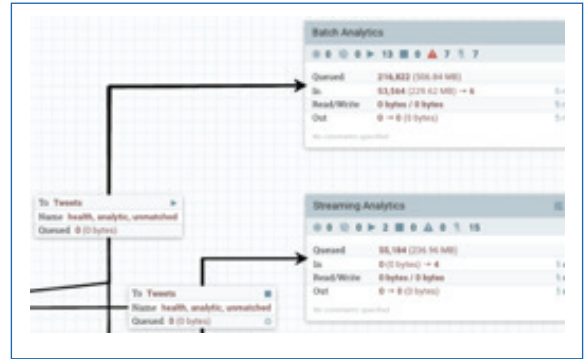


Data pipelining is nothing new in the world of data science and analysis, however, it can also be used to make your site search extremely intelligent and save your business users time. Traditional site searches index information in the commerce database - usually product data--but it also indexes related information such as ratings and reviews.

If you sell groceries online, the information you index would generally come from your suppliers and merchandisers by way of a PIM tool of some sort that feeds its information into the commerce tables. What if you wanted to have your products appear when users entered the name of a recipe though? What if a customer searched for “brownie swirl cheesecake” and your search engine could return all of the ingredients AND the recipe for this cake instead of an irrelevant listing of other boxed brownie mixes, and plain cheesecakes? How do you execute well on that search result without manually curating every single recipe? Data pipelines! You simply ingest recipe data into your Elasticsearch and index your products with the ingredients in each recipe.

## Here are a couple more interesting pipelining use cases:

- Pipelining an Instagram feed of fashion trends, models, and celebrities who have worn certain clothing items into a fashion retailer's Elasticsearch. When a customer searches "Gal Gadot jeans", all of the jeans that retailer sells that Gal has worn appear alongside a feed of inspiring images.
- Pipelining all of the Tour De France historical rider data with the bikes they ride, the shoes, helmets, and kits into a cycling retailer's Elasticsearch. When a customer searches for "Mark Cavendish's 2018 shoes", the exact pair of orange Nike Superfly 360's show up along with cross-selling recommendations of the rest of his gear.



Pipelines are a powerful way to expand the data flow into your search. Managing your data pipelines, though, will quickly become a problem without a visual tool. Enter Apache NiFi. NiFi puts the power of data pipeline manipulation in the hands of less-technical business users who can easily make modifications to the ingestion.

NiFi is a highly configurable and intuitive web interface. It allows for a seamless experience between design, control, feedback, and monitoring of your pipelines. It handles backpressure very well and flows can be modified at runtime.

## Deep Customizations for eCommerce

HCL Commerce has ensured that existing search merchandising tools will work with the new search solution so that your business users don't need to familiarize themselves with completely new tooling to perform basic search administration. The new search solution was built to have inherent B2C and B2B object-awareness. For example, in a B2B context, search results are aware of the current user's entitlements. In both B2B and B2C, searches involving colors and sizes will leverage these defining attributes in the catalog.

One of the big limitations of the previous generation search solution was its inability to match on key defining attributes if they weren't entered exactly as they were stored. For example, if a product was listed as an Adidas Ultraboost and it came in "granite" and "indigo" colors and the US sizes 7-13, but a user searched for "blue adidas ultraboost size 44", it would not necessarily be returned. The new HCL Software search solution now contains proprietary MatchMaker™ technology which will translate all known units of measurement and attempt to fuzzy-match on all known color names. Color name thresholds can be configured so that search results can be fine-tuned based on the needs of the organization.

## End-Customer Focused

Customers don't always care about technology, but they do care about results. When customers interact with your brand's digital property, they have at least one goal in mind – at least one objective that you must meet as quickly as possible. So, how does all of this new technology help them? It's like the brilliant philosopher Jerry Maguire once said, "Help me help you". Elasticsearch, NLP and NiFi helps you help your customers by better being able to understand the words they are using to express their objectives and deliver relevant and rich results one search at a time. Don't make your customers figure out your business language. If they want to search for "Five ½" plastic fitting", let them! If they want to search for that in Spanish using the metric system, let them! These new capabilities make it incredibly friendly for users to express themselves in their own way.

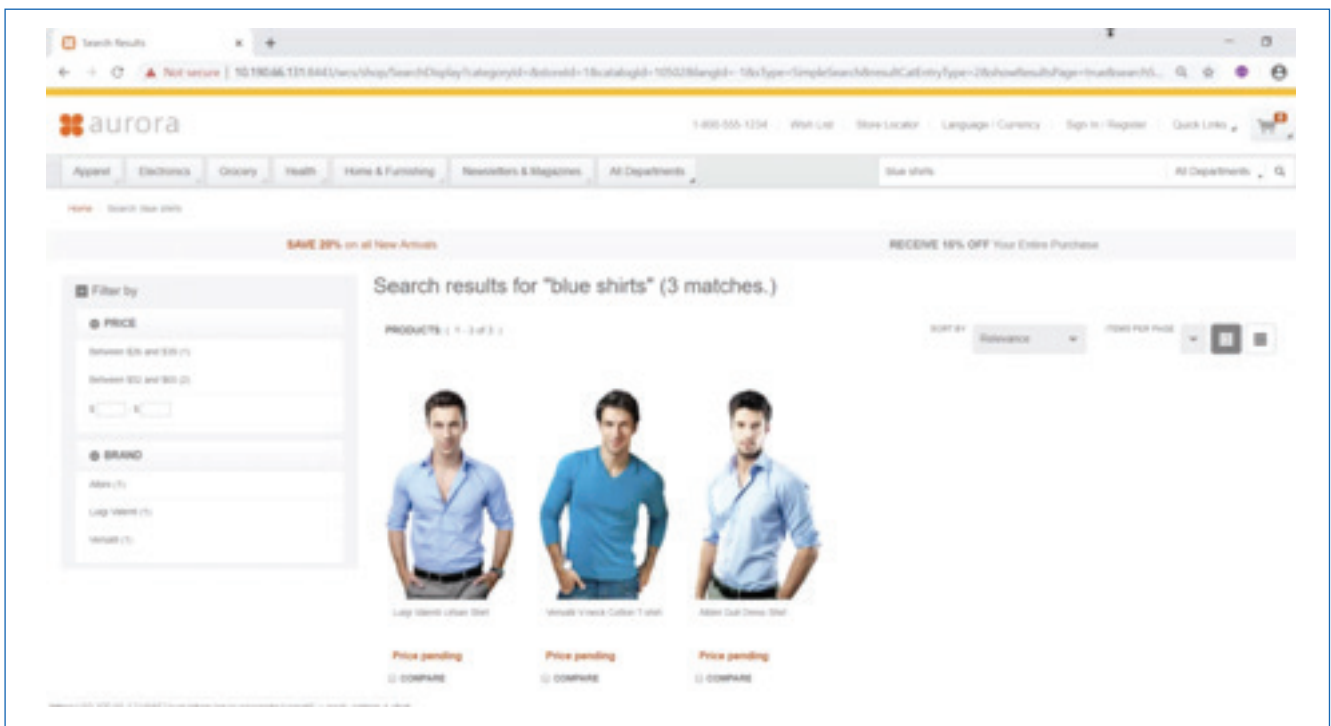
## Technical User Benefits

Extreme flexibility at the hands of a developer allows them to build amazing things. With NiFi, CoreNLP and Elasticsearch, technical users will not only have this level of flexibility, but they will be using cutting edge tech along the way. Built on top of Lucene, many developers and admins will already be very familiar with Elasticsearch's core underpinnings. Elasticsearch is also a Kubernetes-ready search solution, which means that your devops team will be able to build, deploy and scale your search solution with ease. Developers will be able to leverage JSON as the serialization for document formats which is simple, easy to read and a standard across many modern technology sets. Elasticsearch is a schema-free solution and as such does not require field-level definitions to be created before indexing. When an object has a new property, it will automatically be added to the mapping definitions. Dev teams can also fine tune queries with JSON-based DSL, allowing them to construct very complex queries as needed. They can also use JSON to rank and group results.

NiFi offers layers of flexibility and capability for data manipulation on top of the already-powerful Elasticsearch tool. When moving data around, it is important that you guarantee delivery, and this is a key tenant of the NiFi solution. This is achieved with a persistent write-ahead log and content repository. They are designed to work together in high transaction rate environments with effective load-spreading. NiFi also handles backpressure well by supporting buffered queues and aging out data as necessary. Like Elasticsearch, NiFi is designed to scale-out through node clustering. The new Ingest Service leveraging NiFi provides tremendous flexibility and a visual flow programming tool so that the technical user can more easily create, test and manage data ingestion pipelines and can quickly determine where there may be issues with the transformation processing of the data comprising the indexes.

## Business User Advantages

Merchandisers will appreciate the time-savings that an intelligent search solution affords. No longer will they need to laboriously build and manage exhaustive lists of synonyms and search rules. While there will still be capability to manage synonyms, many will largely be automatic and with the use of NiFi, they will now have the advantage of being able to specify data ingestion rules which can take massive manual load off of their plates. These tools heavily mitigate the need to jump through hoops just to deliver basic relevancy. NLP will have the ability to intelligently parse search input for a vast range of potential color names and units of measure. And as we mentioned above, NLP can easily handle multiple languages so as you add new sites in new geographies you can rest assured that your search will still deliver a great experience for all of your expansion plans. For example, searching for “6” and “six inches” will return the same results without any manual search rules needed. This applies to equivalencies in the imperial and metric systems as well. If your merchandise is loaded in metric, searching for items in imperial will also be automatically returned. Likewise, when searching for “blue shirts”, any shirt that has been listed with a color in the blue family or even slightly outside of the blue family can be automatically returned:



## Pushing the limits of Search in HCL Commerce

When 40 to 80%<sup>3</sup> of your revenue can be driven by search, your approach to search sets the tone for your customers' entire onsite experience and ultimately, the success of your business. That is why HCL Commerce has made Search one of the three pillars for our 9.1 release. We will empower our customers to immediately improve the onsite search experience by embedding Elasticsearch right into the platform. And we will continue to expand on those embedded capabilities in future releases as well, diving deeper into the rich features Elasticsearch offers – all to the benefit of our customers.

This continued focus on making the lives of our customers better is what HCL Commerce is all about.

Beyond the Transaction.

<sup>3</sup> <https://www.prefixbox.com/blog/ecommerce-site-search/>

HCL Commerce makes agile commerce @ scale a reality for enterprise companies. Leveraging a modern, cloud-ready, extensible architecture our clients can deliver more innovation in less time to unlock growth, achieve higher ROI, accelerate innovation and reduce their business risk. HCL combines the trusted scale, security and performance you expect, with the speed and flexibility of a modern cloud-ready, extensible commerce architecture – all built on 20 years of unmatched, omni-channel innovation that’s been co-created with our customers and partners. Start delivering world-class experiences that are easy to customize on a platform that enables non-stop innovation. **Visit the website to learn more.**



**About HCL Software**

HCL Software is a division of HCL Technologies (HCL) that operates its primary software business. It develops, markets, sells, and supports over 20 product families in the areas of DevSecOps, Automation, Digital Solutions, Data Management, Marketing and Commerce, and Mainframes. HCL Software has offices and labs around the world to serve thousands of customers. Its mission is to drive ultimate customer success with their IT investments through relentless innovation of its products. For more information, please visit [www.hcltechsw.com](https://www.hcltechsw.com).

**Learn more: <https://www.hcltechsw.com>**