



Federated Search with Drupal, Solr, and React



Matt Carmichael & Dan Montgomery • MidCamp 2019

Today's Session

- + Introductions
- + Project Overview
- + Live Demo
- + Application Components
- + Getting Started
- + Question & Answer



About Palantir.net

Matt Carmichael

+ Engineer, Palantir.net



#midcamp @palantir

Dan Montgomery

+ Senior Engineer and
Technical Architect,
Palantir.net



Project Overview

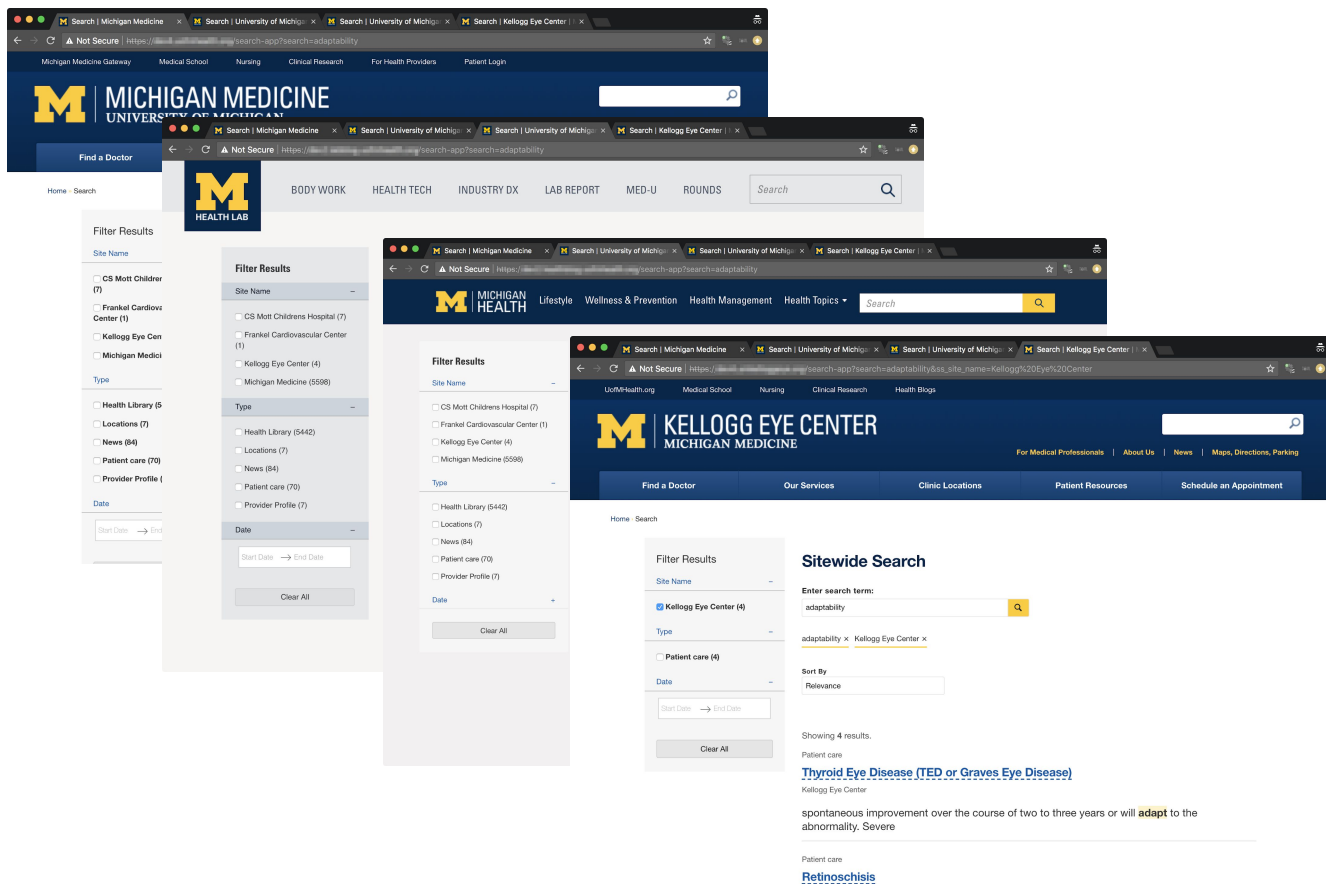
Our Use Case

Existing Drupal 7 sites.

New Drupal 8 sites.

Domain Access.

No guarantees of common architecture.



 palantir.net

#midcamp

Make it Simple!

Make it work like this.

The Google logo is centered at the top of the page, featuring its characteristic multi-colored letters: 'G' in blue, 'o' in red, 'o' in yellow, 'g' in green, and 'l' in blue.A simple, minimalist search bar with a thin grey border and a small microphone icon on the right side.

Google Search

I'm Feeling Lucky

 **palantir.net**

#midcamp

Search Appliance

Google Search Appliance

End of Life product.



 palantir.net

#midcamp

Unlock your data with the power and relevance of Google Search, customized and secured in an appliance for use inside your business.

Site Search

End of Life product.

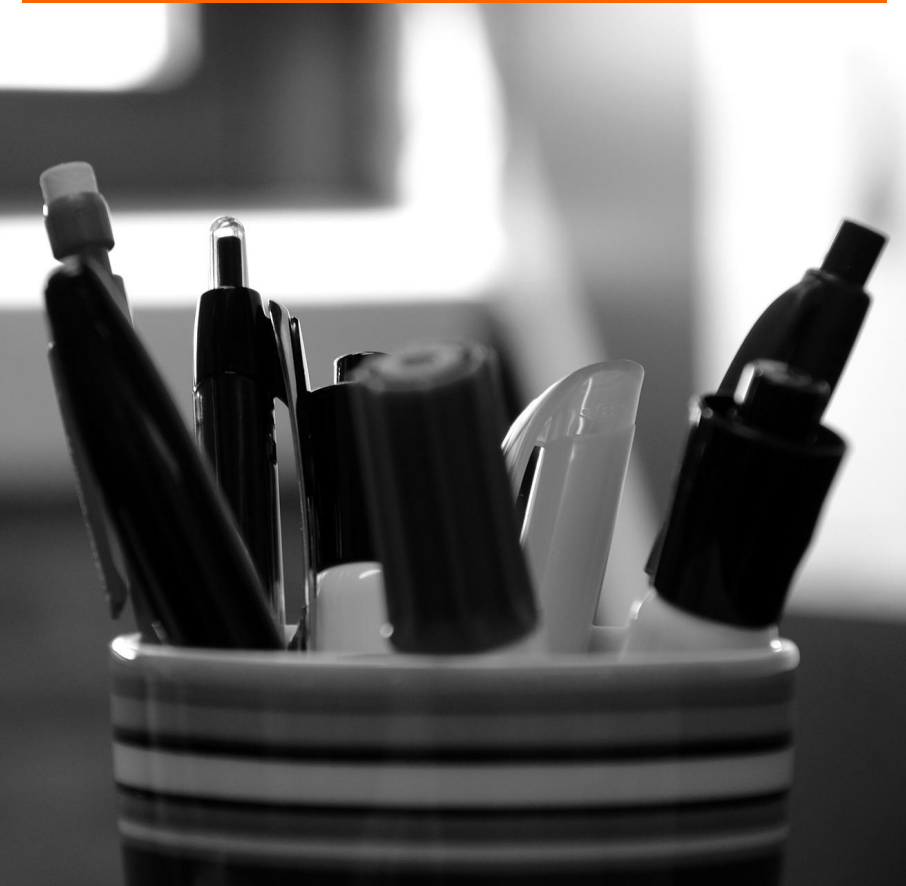
 palantir.net

#midcamp

Google Site Search



Add the power of Google.com technology to your website's search box, helping customers quickly find what they're looking for.



Unified Search

- + A simple way to store, retrieve, and parse content.
- + A cross-platform search application.
- + A speedy, usable, responsive front-end.
- + A flexible, extensible, reusable model.
- + A drop-in replacement for deprecated Google Products

Let's Do It Live!

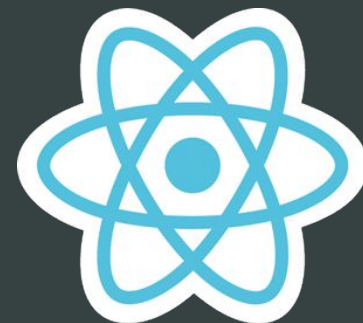
Three Part Solution



Solr



Drupal



React

Solr components

Solr expects a **uniform schema** for proper results across an index.

An index can have multiple sources.



#midcamp

Label	Machine Name	Type	Required?	Single/ Multi- value	Description
Federated Title	federated_title	string	Yes	single	The title of the item. Displayed as the title of each search result.
Federated Date	federated_date	date	No	single	Usually the date the content was created. Used to provide date-based filtering.
Federated Type	federated_type	string	Yes	single	The shared type label for faceting. Also used to label each result.
Federated Terms	federated_terms	string	No	multi	Terms for additional, topic-based facets, mapped to shared topic terms if necessary.
Federated Image	federated_image	string	No	single	An absolute url to an image which, if it exists, will be displayed with the text. Recommended image size: ___ x ___
Rendered HTML output	rendered_item	fulltext	Yes	multi	The full text of the item, with HTML stripped.
URI	url	string	Yes	single	The absolute path to the item, used to provide a link to each result.
Site Name	site_name	string	Yes	single	The descriptive name of the source site. Used to provide site-based filtering.
Site	site	string	Automatic	single	The base url of the source site, like <code>https://labblog.uofmhealth.edu</code> . This will be sent automatically by Drupal, and is required for external content sources.

Solr API

Solr already knows how to return JSON objects in response to a search request.



#midcamp

The screenshot displays the Apache Solr Admin interface. On the left is a navigation sidebar with options like Dashboard, Logging, Core Admin, Java Properties, Thread Dump, and a dropdown menu for 'drupal8'. The main area is titled 'Request-Handler (qt)' and shows a search query: `/select`. Below the query, there are fields for `q` (containing `*:*`), `fq`, `sort`, `start, rows` (set to 0 and 10), `fl`, and `df`. The 'Raw Query Parameters' section shows `key1=val1&key2=val2`. The 'wt' dropdown is set to 'json'. There are checkboxes for `indent` (checked), `debugQuery`, `dismax`, `edismax`, `hl`, `facet`, `spatial`, and `spellcheck`. A blue 'Execute Query' button is at the bottom of this section.

The right side of the screenshot shows the JSON response from the search. The URL bar indicates the request: `http://federated-search-demo.local:8983/solr/drupal8/select?q=%3A*&wt=json&indent=true`. The JSON response is as follows:

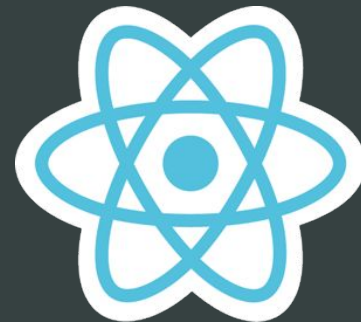
```
{
  "response": {
    "numFound": 15,
    "start": 0,
    "docs": [
      {
        "id": "ijs7c5-federated_search-entity:node/10:en",
        "index_id": "federated_search",
        "hash": "ijs7c5",
        "site": "http://federated-search-demo.local/",
        "ss_federated_title": "Victoria sponge cake",
        "ss_federated_type": "Recipe",
        "tm_rendered_item": [
          "Victoria sponge cake Cooking time 20 minutes Difficulty Easy Ingredients 225g butter or margarine 225g"
        ],
        "spell": [
          "Victoria sponge cake Cooking time 20 minutes Difficulty Easy Ingredients 225g butter or margarine 225g"
        ],
        "ss_site_name": "Umami Demo",
        "ds_federated_date": "2018-08-13T20:23:25Z",
        "ss_federated_image": "http://federated-search-demo.local/sites/default/files/styles/search_api_federated",
        "ss_url": "http://federated-search-demo.local/recipes/victoria-sponge-cake",
        "ss_search_api_id": "entity:node/10:en",
        "ss_search_api_datasource": "entity:node",
        "ss_search_api_language": "en",
        "v_version": "1608717676293652500",
        "timestamp": "2018-08-13T20:38:50.23Z"
      },
      {
        "id": "ijs7c5-federated_search-entity:node/11:en",
        "index_id": "federated_search",
        "hash": "ijs7c5",
        "site": "http://federated-search-demo.local/",
        "ss_federated_title": "Gluten free pizza",
        "ss_federated_type": "Recipe",
        "tm_rendered_item": [
          "Gluten free pizza Cooking time 15 minutes Difficulty Medium Ingredients 400g gluten free flour 125mL w"
        ],
        "spell": [
          "Gluten free pizza Cooking time 15 minutes Difficulty Medium Ingredients 400g gluten free flour 125mL w"
        ],
        "ss_site_name": "Umami Demo"
      }
    ]
  }
}
```




✓ Solr



Drupal



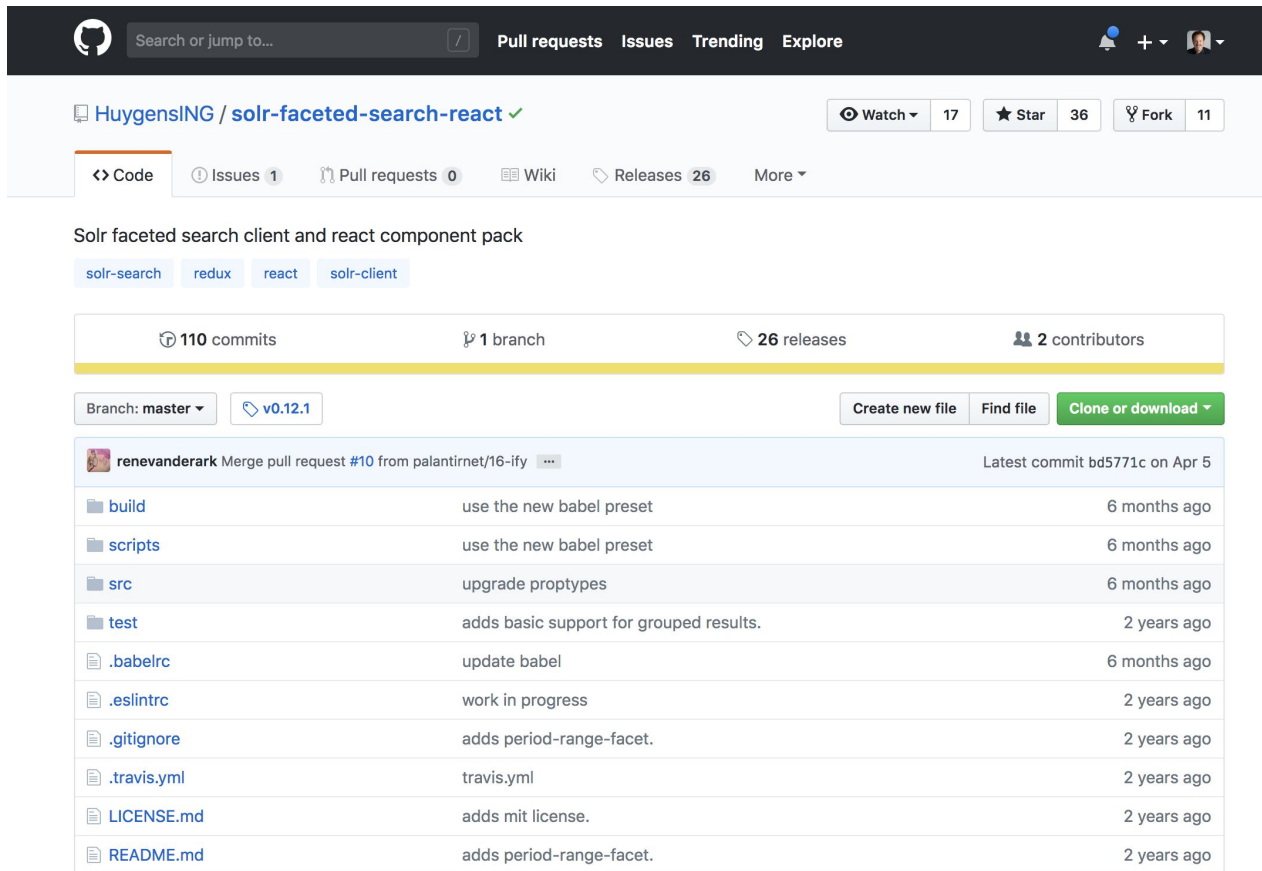
React

React Library

React applications love to consume JSON and format its results.

 palantir.net

#midcamp



The screenshot shows the GitHub repository page for 'solr-faceted-search-react' by HuygensING. The repository has 17 watches, 36 stars, and 11 forks. It contains 110 commits, 1 branch, 26 releases, and 2 contributors. The current branch is 'master' and the version is 'v0.12.1'. A recent pull request by 'renevanderark' is merged, with the latest commit 'bd5771c' on Apr 5. The commit message is 'Merge pull request #10 from palantirnet/16-ify'. The commit includes changes to several files:

File	Change	Time
build	use the new babel preset	6 months ago
scripts	use the new babel preset	6 months ago
src	upgrade proptypes	6 months ago
test	adds basic support for grouped results.	2 years ago
.babelrc	update babel	6 months ago
.eslintrc	work in progress	2 years ago
.gitignore	adds period-range-facet.	2 years ago
.travis.yml	travis.yml	2 years ago
LICENSE.md	adds mit license.	2 years ago
README.md	adds period-range-facet.	2 years ago

Open Source FTW

The existing library has the components, but not the design we need.

It's also MIT Licensed for reuse.



#midcamp

A screenshot of a web application's search interface. On the left, there are filter panels for "All text fields", "Name", "Characteristics", and "Date of birth". The "Characteristics" filter is expanded, showing a list of categories with checkboxes and counts: "Publicist (177)", "Bestuurslid vakvereniging (106)", "Redacteur (100)", and "Schrijver (91)". The "Date of birth" filter shows a range from 1855 to 1902. On the right, the search results are displayed as a list of cards. Each card shows a person's name, their characteristics, and their dates of birth and death. The results are sorted by name in ascending order. The interface includes a search bar, a "Found 449 results" header, and sorting options (Name, asc, desc).

Search	
All text fields	<input type="text"/>
Name	<input type="text"/>
Characteristics	<input type="text"/> Filter... a-z 0-9
<input type="checkbox"/> Publicist (177)	
<input type="checkbox"/> Bestuurslid vakvereniging (106)	
<input type="checkbox"/> Redacteur (100)	
<input type="checkbox"/> Schrijver (91)	
Date of birth	<input type="text"/> 1855 1902
Date of death	<input type="text"/>

Found 449 results	
Name	Johannes Addens
Characteristics	Lid diplomatieke dienst, Redacteur, Scheikundige
Date of birth	1829
Date of death	1911
Name	Louis Aert
Characteristics	(leeg)
Date of birth	
Date of death	
Name	Franciscus Lambertus Theodorus Alberdingk Thijm
Characteristics	Boekhouder, Vertaler
Date of birth	1854
Date of death	1925
Name	Johanna Catharina Alphonsina Alberdingk Thijm
Characteristics	Publicist
Date of birth	1893
Date of death	1953
Name	Joseph Willem Lambertus Alberdingk Thijm
Characteristics	Editeur
Date of birth	1888
Date of death	1970
Name	Josephus Albertus Alberdingk Thijm
Characteristics	Bedrijfsleider, Redacteur, Aquarellist, Schrijver, Boekhandelaar, Criticus, Drukker, Editeur, Bestuurslid vakvereniging, Dichter, Hoogleraar, Publicist, Uitgever, Tekenaar
Date of birth	1820
Date of death	1889
Name	Karel Joan Lodewijk Alberdingk Thijm
Characteristics	Publicist, Redacteur, Journalist, Schrijver, Criticus, Bestuurslid vakvereniging, Polemist
Date of birth	1864
Date of death	

Design Direction

We did usability testing around search components and how people use facets.

 palantir.net

#midcamp

Filter Results

Search site —

Healthblog

Labblog

MCancer

UofMHealth

Type +

Patient Care +

Specialty +

Date +

Clear All

Sitewide Search

Enter search term:

[Healthblog](#) × [03/01/2018 - 03/25/2018](#) × [Events](#) × [Bronchitis](#) ×

Showing **10** results out of **158**



Conditions & Treatments

[Eye Cancer](#)

Kellogg Eye Center · December 7, 2017

Lorem ipsum **dolor sit** amet, consectetur adipiscing sit amet pharetra faucibus. Fusce quis nibendum metus.

Maps & Directions

[Taubman Health Center- Floor 1](#)

Michigan Medicine · October 23, 2016

Lorem ipsum dolor sit amet. consectetur adipiscing elit. Donec **dolor sit** vo

Application Design

Our forked application
was designed to be
extended and skinned.

 **palantir.net**

#midcamp

```
export default {
  searchFields: {
    text: TextSearch,
    "list-facet": ListFacet,
    "range-facet": RangeFacet,
    "period-range-facet": RangeFacet,
    container: SearchFieldContainer,
    currentQuery: CurrentQuery
  },
  results: {
    result: Result,
    resultCount: CountLabel,
    header: ResultHeader,
    list: ResultList,
    container: ResultContainer,
    pending: ResultPending,
    preloadIndicator: PreloadIndicator,
    csvExport: CsvExport,
    paginate: ResultPagination
  },
  sortFields: {
    menu: SortMenu
  }
}
```

Templates Design

Templates use common markup that can be enhanced through a single CSS file.



#midcamp

```
render() {
  const { doc, highlight } = this.props;

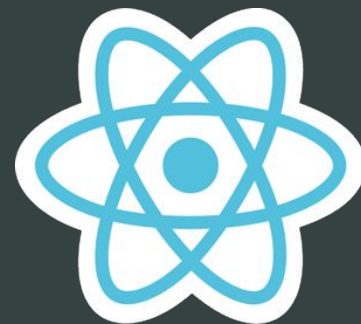
  return (
    <li onClick={() => this.props.onSelect(doc)}>
      {doc.ss_federated_image &&
        <div className="search-results__container--left">
          <img src={doc.ss_federated_image} alt="" />
        </div>
        <div className="search-results__container--right">
          <span className="search-results__label">{doc.ss_federated_type}</span>
          <h3 className="search-results__heading"><a href={doc.ss_url}>{doc.ss_federated_title}</a></h3>
          <div className="search-results__meta">
            <cite className="search-results__citation">{doc.ss_site_name}</cite>
            {this.dateFormat(doc.ds_federated_date)}
          </div>
          <p className="search-results__teaser" dangerouslySetInnerHTML={{__html:
            highlight.tm_rendered_item}} />
        </div>
      </li>
    )
  }
```



✓ Solr



Drupal



✓ React

Solr Indexing

We have a working application.

We merely have to tell Solr how to index the content consistently.



#midcamp

```
{
  "response": {
    "numFound": 15,
    "start": 0,
    "docs": [
      {
        "id": "ijs7c5-federated_search-entity:node/10:en",
        "index_id": "federated_search",
        "hash": "ijs7c5",
        "site": "http://federated-search-demo.local/",
        "ss_federated_title": "Victoria sponge cake",
        "ss_federated_type": "Recipe",
        "tm_rendered_item": [
          "Victoria sponge cake Cooking time 20 minutes Difficulty Easy In
        ],
        "spell": [
          "Victoria sponge cake Cooking time 20 minutes Difficulty Easy In
        ],
        "ss_site_name": "Umami Demo",
        "ds_federated_date": "2018-08-13T20:23:25Z",
        "ss_federated_image": "http://federated-search-demo.local/sites/de
        "ss_url": "http://federated-search-demo.local/recipes/victoria-spo
        "ss_search_api_id": "entity:node/10:en",
        "ss_search_api_datasource": "entity:node",
        "ss_search_api_language": "en",
        "_version_": 1608717676293652500,
        "timestamp": "2018-08-13T20:38:50.23Z"
      },
    ]
  }
}
```

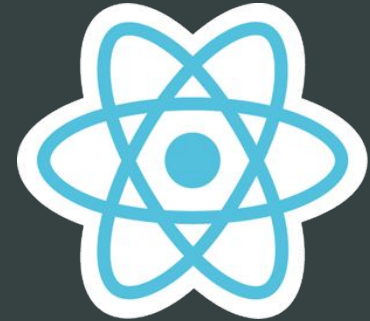

**Do you want to know a
secret?**



✓ Solr

Drupal is not
a hard
requirement

CMS of choice



✓ React

Why is that?

Drupal search templates simply call the React application.

Any HTML can.

[search_api_federated_solr / templates / search-app.html.twig](#)

```
{% set app_config = federated_search_app_config|json_encode(constant('JSON_UNESCAPED_SLASHES'), constant('JSON_NUMERIC_CHECK')) %}
```

```
<noscript>This search page requires Javascript in order to function. <a href="https://www.whatismybrowser.com/guides/how-to-enable-javascript/auto">Learn how to enable Javascript in your browser</a>.</noscript>
```

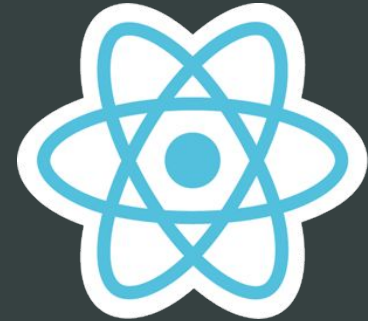
```
<div id="root" data-federated-search-app-config="{ { app_config } }">
  <p class="element-invisible" aria-hidden="true">Error message...</p>
</div>
```



✓ **Solr**



**Drupal does
make it easier**



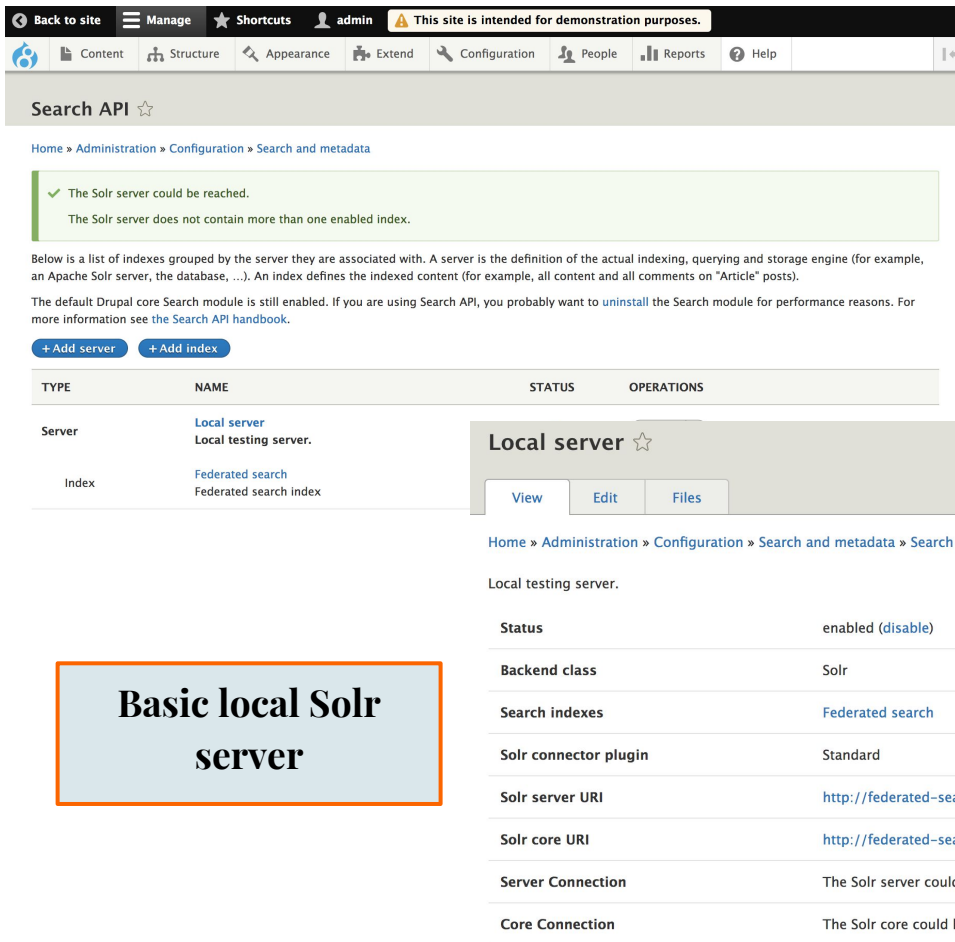
✓ **React**

Drupal components

Existing Search API and Search API Solr modules for Drupal 8 and Drupal 7.

 palantir.net

#midcamp



The screenshot shows the Drupal Search API configuration interface. At the top, there's a navigation bar with 'Back to site', 'Manage', 'Shortcuts', and 'admin'. A warning message states 'This site is intended for demonstration purposes.' Below the navigation bar, there's a search bar and a breadcrumb trail: 'Home » Administration » Configuration » Search and metadata'. A green message box indicates that the Solr server could be reached and does not contain more than one enabled index. Below this, there's a table listing indexes grouped by server. The table has columns for 'TYPE', 'NAME', 'STATUS', and 'OPERATIONS'. The 'Local server' is listed with a 'Local testing server' index. The 'Federated search' is listed with a 'Federated search index'.

TYPE	NAME	STATUS	OPERATIONS
Server	Local server Local testing server.		
Index	Federated search Federated search index		

Below the table, there's a 'Local server' configuration page. It has tabs for 'View', 'Edit', and 'Files'. The breadcrumb trail is 'Home » Administration » Configuration » Search and metadata » Search API'. The configuration page shows the 'Local testing server' configuration. The status is 'enabled (disable)'. The backend class is 'Solr'. The search indexes are 'Federated search'. The Solr connector plugin is 'Standard'. The Solr server URI is 'http://federated-search-demo.local:8983/solr/'. The Solr core URI is 'http://federated-search-demo.local:8983/solr/#/drupal8'. The server connection is 'The Solr server could be reached.' The core connection is 'The Solr core could be accessed (latency: 4.3788419494629 ms).'

**Standard Search
API config**

**Basic local Solr
server**

Federated field mapping

We use conversions to map disparate data to a standard schema.



#midcamp

Content Structure Appearance Extend Configuration People Reports Help

Manage fields for search index *Federated search* ☆

View Edit Fields Processors

Home » Administration » Configuration » Search and metadata » Search API » Federated search

[+ Add fields](#)

The data type of a field determines how it can be used for searching and filtering. The boost is used to give additional weight to certain fields, for example titles or tags. For information about the data types available for indexing, see the [data types table](#) at the bottom of the page.

▼ GENERAL

LABEL	MACHINE NAME	PROPERTY PATH	TYPE	BOOST
<input type="text" value="Federated Date"/>	<input type="text" value="federated_date"/>	federated_field	Date	
<input type="text" value="Federated Image"/>	<input type="text" value="federated_image"/>	federated_field	String	
<input type="text" value="Federated terms"/>	<input type="text" value="federated_terms"/>	federated_terms	String	
<input type="text" value="Federated Title"/>	<input type="text" value="federated_title"/>	federated_field	String	
<input type="text" value="Federated Type"/>	<input type="text" value="federated_type"/>	federated_field	String	
<input type="text" value="Rendered HTML output"/>	<input type="text" value="rendered_item"/>	rendered_item	Fulltext	1.0
<input type="text" value="Site Name"/>	<input type="text" value="site_name"/>	site_name	String	
<input type="text" value="URI"/>	<input type="text" value="url"/>	federated_field	String	

▼ CONTENT

LABEL	MACHINE NAME	PROPERTY PATH	TYPE	BOOST	OPERATIONS
-------	--------------	---------------	------	-------	------------

Federated field mapping

Simple string mappings.

 palantir.net

#midcamp

Edit field Federated Type ✕

Federated data

Field data for *Content* » *Article*

Field data for *Content* » *Author*

Field data for *Content* » *Listacle*

Field data for *Content* » *Page*

Field data for *Content* » *Podcast*

Field data for *Content* » *Video*

Set the data to be sent to the index for each bundle in the data sources set in your index configuration. Use static values or choose tokens using the picker below.

[Browse available tokens.](#)
[Cancel](#)

Federated field mapping

Token support.

 palantir.net

#midcamp

Edit field Federated Date ✕

Federated data

Field data for *Content* » *Article*

Field data for *Content* » *Author*

Field data for *Content* » *Listacle*

Field data for *Content* » *Page*

Field data for *Content* » *Podcast*

Field data for *Content* » *Video*

Set the data to be sent to the index for each bundle in the data sources set in your index configuration. Use static values or choose tokens using the picker below.

[Browse available tokens.](#)
[Cancel](#)

Federated field mapping

Preview images can pull from existing or custom fields and support multiple media types.



#midcamp

Edit field Federated Image ✕

Federated data

Field data for *Content* » *Article*

Field data for *Content* » *Author*

Field data for *Content* » *Listacle*

Field data for *Content* » *Page*

Field data for *Content* » *Podcast*

Field data for *Content* » *Video*

Set the data to be sent to the index for each bundle in the data sources set in your index configuration. Use static values or choose tokens using the picker below.

[Browse available tokens.](#)
[Cancel](#)

[Save](#)

Federated field mapping

We can't really use Drupal's site name since it's not user-friendly, also we need the site name for special things.

Edit field Site Name ✕

Site Name *

The name of the site from which this content originated. This can be useful if indexing multiple sites with a single search index.

[Cancel](#)

[Save](#)

Federated field mapping

Site names also integration with Domain Access for those who need it.

 palantir.net

#midcamp

Callback settings

Bundle filter Enabled	Michigan Medicine Domain Label * Michigan Medicine Map the Domain to a custom label for search.
Complete entity view Enabled	
Site Name Enabled	CS Mott Children's Hospital Michigan Medicine Domain Label * CS Mott Childrens Hospital Map the Domain to a custom label for search.
Federated Field Enabled	
Re-map Field Names Enabled	Frankel Cardiovascular Center Michigan Medicine Domain Label * Frankel Cardiovascular Center Map the Domain to a custom label for search.
	Von Voigtlander Women's Hospital Michigan Medicine Domain Label * Von Voigtlander Womens Hospital Map the Domain to a custom label for search.
	Kellogg Eye Center Michigan Medicine Domain Label * Kellogg Eye Center Map the Domain to a custom label for search.
	Victors Care Domain Label * Victors Care Map the Domain to a custom label for search.

Taxonomy fields

Some fields are uniform across sites and easy to map.

Taxonomy is more complex.



#midcamp

Filter Results

Search site	+
Type	-
<input type="checkbox"/> News	
<input type="checkbox"/> Events	
<input type="checkbox"/> Multimedia	
<input type="checkbox"/> Locations	
<input type="checkbox"/> People	
Patient Care	+
Specialty	+
Date	+

Clear All

Filter Results

Search site	+
Type	+
Patient Care	+
Specialty	-
<input type="checkbox"/> Dermatology	
<input type="checkbox"/> Endocrinology	
<input type="checkbox"/> Gynecology	
<input type="checkbox"/> Oncology	
Date	+

Clear All

Taxonomy fields

We can't use the token system to map taxonomies and their hierarchies.

Edit field Federated terms ✕

Federated terms

By adding this field to your search index configuration, you have enabled the federated terms processor to run when new items are indexed. Next, add a "Federated Terms" field to any taxonomy vocabulary whose terms should be mapped to a "federated" term (this helps map terms across vocabularies and sites to a single "federated" term). Then, edit terms in those vocabularies to add the federated term destination value (i.e. "Conditions>Blood Disorders"). Once that tagged content gets indexed, it will have "federated_terms" populated with any matching federated term destination values.

[Cancel](#)

[Save](#)

Taxonomy fields

Hierarchies are mapped on the term itself, providing aliases that can be used across sites with unique vocabularies and terms.

Name *

Dermatology



FEDERATED TERMS

Specialty>Dermatology



This mapped term is used as a facet value in your search application. It should consist of a hierarchy made up of a term (i.e. "Diabetes"), separated by ">". For example: Condition>Diabetes.

Specialty>Skin Care



This mapped term is used as a facet value in your search application. It should consist of a hierarchy made up of a term (i.e. "Diabetes"), separated by ">". For example: Condition>Diabetes.



This mapped term is used as a facet value in your search application. It should consist of a hierarchy made up of a term (i.e. "Diabetes"), separated by ">". For example: Condition>Diabetes.



This mapped term is used as a facet value in your search application. It should consist of a hierarchy made up of a term (i.e. "Diabetes"), separated by ">". For example: Condition>Diabetes.

Add another item

Putting it all together



Drupal provides indexing, a search page callback, and a search box block.

Filter Results

Search site -

Healthblog

Labblog

MCancer

UofMHealth

Type +

Patient Care +

Specialty +

Date +

Clear All

Sitewide Search

Enter search term:

Healthblog x 03/01/2018 - 03/25/2018 x Events x Bronchitis x

Showing 10 results out of 158



Conditions & Treatments

[Eye Cancer](#)

Kellogg Eye Center · December 7, 2017

Lorem ipsum **dolor sit** amet, consectetur adipiscing elit. Donec volutpat augue sit amet pharetra faucibus. Fusce quis malesuada turpis, ac bibendum metus.

 palantir.net

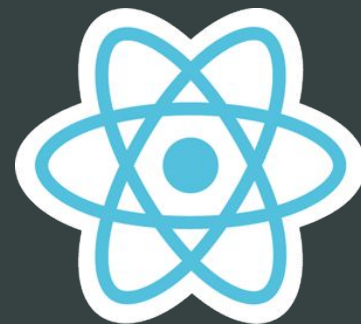
#midcamp



✓ Solr



✓ Drupal



✓ React

How to Get Started

Open Source components

We provide a starting point that supports Drupal 8 and Drupal 7.

+ Federated Search Demo

+ <https://github.com/palantirnet/federated-search-demo>

+ macOS development environment

+ Composer

+ VirtualBox

+ Ansible

+ Vagrant

+ vagrant-hostmaster

+ vagrant-auto-network

 palantir.net

#midcamp

Build Your Own

Items installed by the demo box.

 palantir.net

#midcamp

- + Search API Federated Solr
 - + https://www.drupal.org/project/search_api_federated_solr
- + Search API Field Map
 - + https://www.drupal.org/project/search_api_field_map (d8)
- + Federated Search React
 - + <https://github.com/palantirnet/federated-search-react>
- + Drupal 8 and Drupal 7
- + Ubuntu 64bit version 16.04.4 LTS
- + Apache HTTP version 2.x
- + MySQL version 5.7.21
- + PHP version 7.1.16
- + Apache Solr version 4.5.1 (Acquia version)

Questions?



Thank You.

622 Davis Street, Suite 400
Evanston, IL 60201

773.645.4100



Ken Rickard

Avi Schwab

Jes Constantine

Michelle Jackson

Kelsey Bentham

Malak Desai

Luke Wertz

Kristen Mayer

Adrian Muresan

Bec White

James Sansbury

The University of Michigan



Thank You.

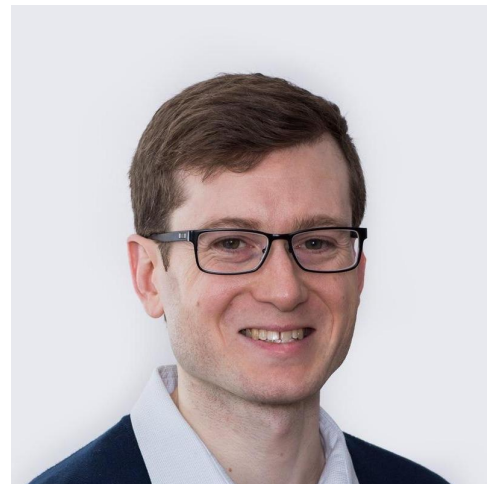
622 Davis Street, Suite 400
Evanston, IL 60201

773.645.4100



Matt Carmichael

Engineer



Dan Montgomery

**Senior Engineer and
Technical Architect**

PLEASE PROVIDE YOUR FEEDBACK!

mid.camp/272

The top rated sessions will be captioned, courtesy
of Clarity Partners



CONTRIBUTION DAY

Saturday 10am to 4pm

You don't have to know code to give back!

New Contributor training 10am to Noon
with **AmyJune Hineline** of Kanopi Studios