



# From climate data catalogues to dataset storage:

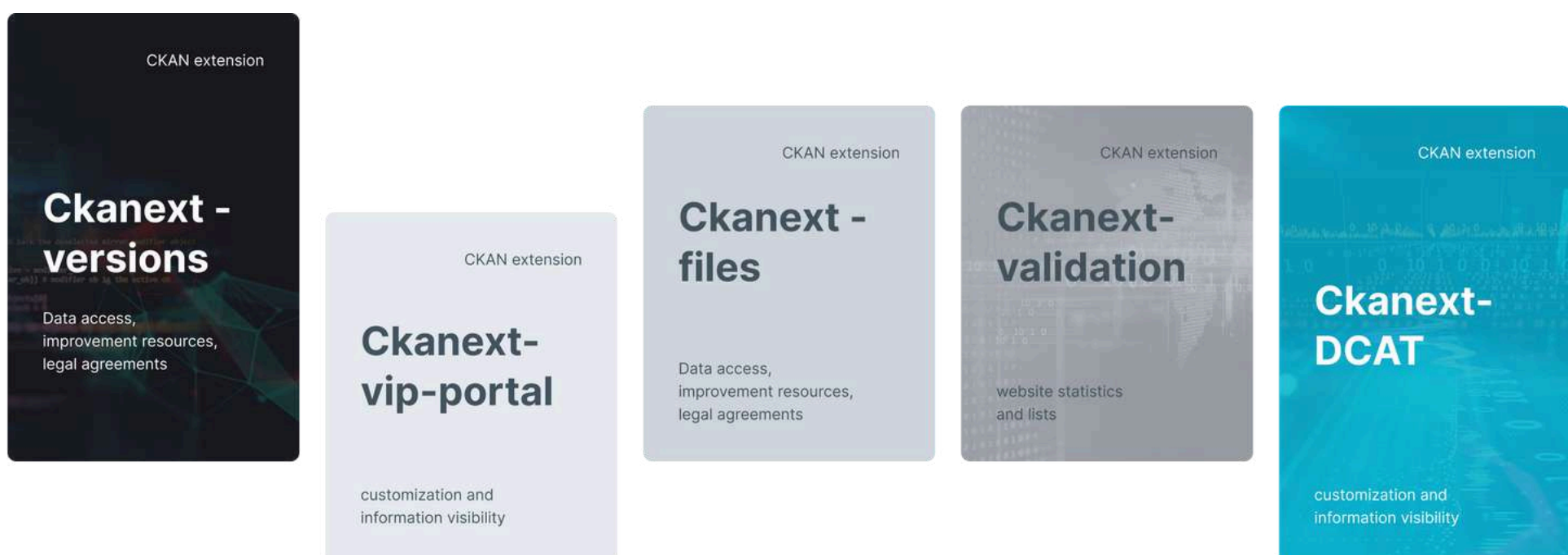
the CKAN extensions that make it possible

# Introduction

The [Comprehensive Knowledge Archive Network or CKAN](#) is what is referred to as 'out-of-the-box' software. This means it can be configured and set up to power an open or internal data platform once it is deployed and hosted on a web server, without requiring significant extra effort from the user.

This very basic or default functionality makes CKAN ideal for small users and those with significant software expertise or the resources to hire a specialist developer expertise, such as Link Digital. But CKAN's power in relation to data management applications can be significantly extended through a large range of extensions which add features and integrations to its core system.

There are literally hundreds of CKAN extensions, far too many to talk about in this piece. So, to give current and potential CKAN users a taste of what extensions are on offer, I asked some of Link Digital's staff, the people who work with CKAN every day, to recommend their most used, favourite or useful CKAN extension.



# What is a CKAN extension?

But, first, what is an extension? From a technical point of view extensions are the Python programming language package (which corresponds with a GitHub repository – a storage space or project folder where all the software files and their revision histories are kept) that one installs in a CKAN site.

Each extension provides one or more plugins that extend or modify CKAN's basic default behaviour. A developer installs an extension, and its features then become available. For instance, the extension [ckanext-spatial](#) provides various plugins, such as spatial-query, which are the ones that must be added to the configuration file. In a technical context the precise language that is used matters. For non-technical use, both the terms extension and plugin are often used interchangeably, although this article will use the former term throughout.

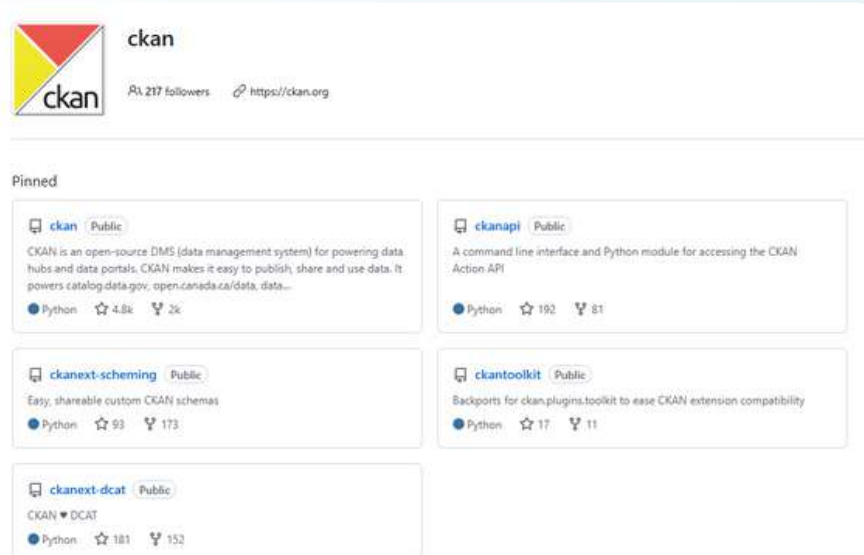
Extensions can vastly increase CKAN's basic functionality, integrate a CKAN application in a seamless manner with a user's existing data and digital operations or allow a user to up-scale their CKAN application as their data needs grow.

**CKAN is open source software, meaning the original source code has been made freely available, allowing anyone to redistribute and modify it.** One manifestation of this public ethos is the community of users and contributors around CKAN, who continuously contribute to the software's design, development and defect correction. This is in addition to the efforts of a core development team who work to keep the CKAN software in good shape. Many of these people are also involved in development and updating of extensions.

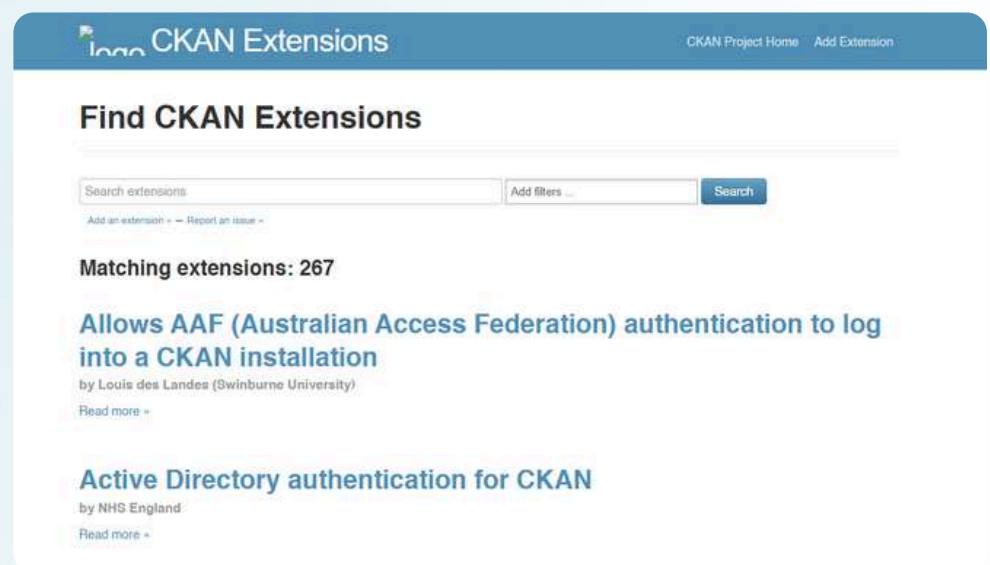
# How do extensions improve CKAN's functionality?

There is a wealth of CKAN extensions, and the exact number only grows over time as the community around the software, including specialist digital service providers such as Link Digital, contribute more. These extensions cover everything from data visualisation and data storage improvements to custom theming and metadata standards compliance.

You can get a sense of the additional functionality offered by CKAN extensions by exploring the [CKAN Extensions Repository on GitHub](#) or on [CKAN's extensions page](#). [The GitHub page gives you an idea of extensions and their 'forks'](#), a GitHub term for how many times a developer has made a copy of the extension and made extensions to the code, which is another way of saying how popular they are. Popular extensions include:



CKAN Extensions on GitHub



CKAN.org Extensions page



### **ckanext-geoview**

which contains plugins to display geospatial files and services in CKAN

### **ckanext-spatial**

an extension that provides some harvesters for importing ISO19139-based metadata into CKAN

### **ckanext-scheming**

a way to configure and share metadata schemas using a YAML (a data serialisation language used for configuring files) or JSON schema description (a file format and data exchange text to store and transmit data objects)

### **ckan-nextpages**

an easy way to add simple pages to CKAN

But while the GitHub page tells you about the most used extensions, a better way to grasp the truly innovative nature of many of the CKAN extensions and how they can be used to upscale the power of your CKAN application, is to ask the people who work on the software every day.



## **Sergey Motornyuk**

Full Stack Developer/ Architect

"A year ago, I created an extension called ckanext-collection. In CKAN we have a lot of datasets and resources, a lot of pages and text to discover. Some portals and data catalogues literally have thousands of different dataset titles.

And, of course, every file has a format. This extension helps in visualising those collections. You can describe what collection you are interested in, how you want it represented and the extension will build you a table on a page, or maybe a URL where you can download in a specific format all this data. Link Digital has used this extension when clients want their data to be shown in different formats.

For example, for an internal asset register, it enables you to build a single table that can display a workflow of the various data requests, those that are pending, granted, etc. The extension can also be customised to show maybe a certain type of request, for example pending data requests."

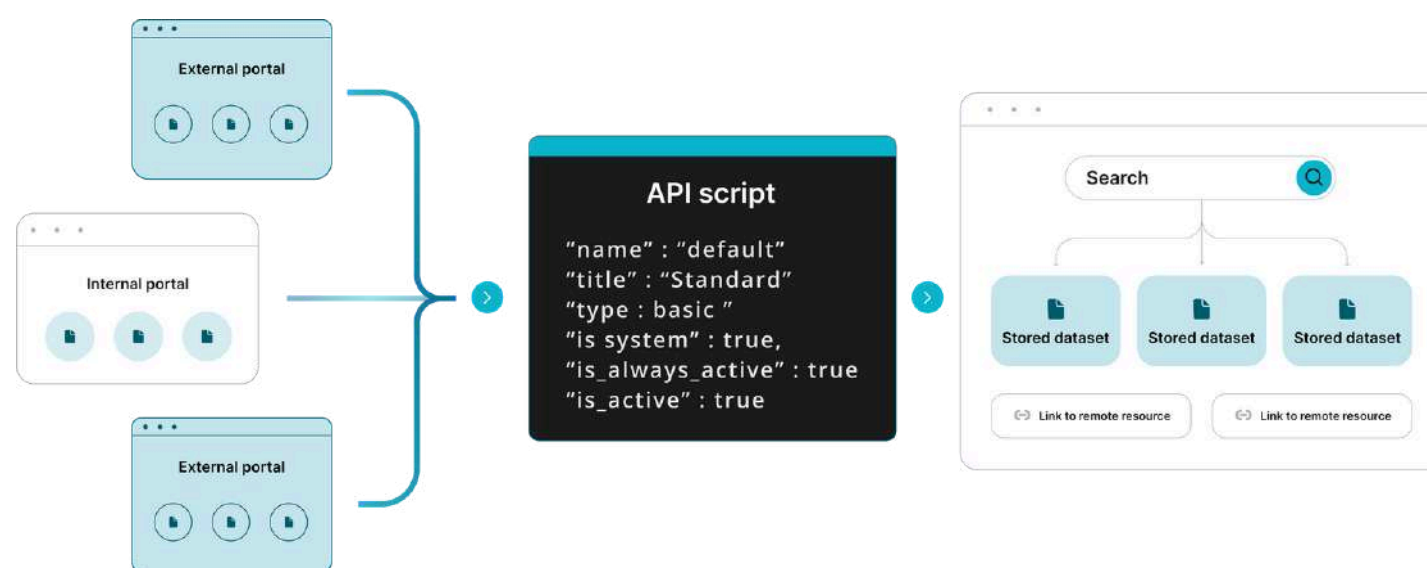


**Alex Gostev**

CKAN Contributor

“There has been an extensive rewrite of the CKAN extension for the Data Catalog Vocabulary called ckanext-dcat. DCAT is a standardised way to describe datasets in data catalogs and public facing data portals. It is designed to facilitate greater interoperability between data published on the web, for example, in open data catalogues or public facing data portals.

You can think of it as an enabler for many organisations in the European Union and United States as it's the foundation that should be covered to adhere to current governmental requirements. It's all about providing users with more comprehensive information about datasets. This means CKAN is staying up to date in terms of compatibility with the latest DCAT version, DCAT 3.”



*Ckanext-DCAT metadata mapping*



**Anastasiia Shevtsova**

UI/UX designer

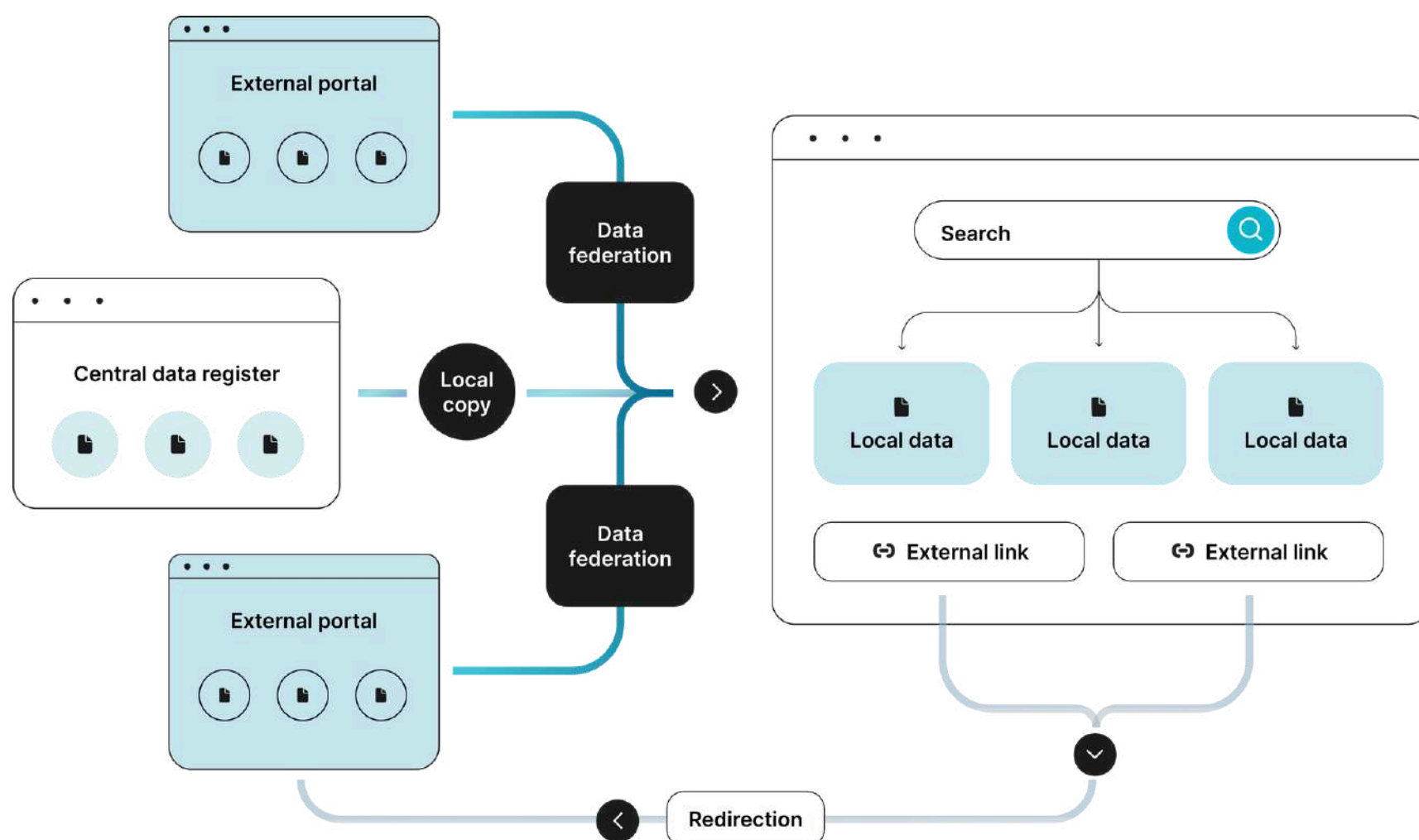
“I would like to recommend a CKAN extension that can be helpful for the aggregation of non-local data on the CKAN portal, [ckanext-federated-index](#). This extension searches datasets, saves the title, description and shows it in the relevant search results. The user is redirected to the original portal when the dataset details page is opened. I can see a lot of benefits in this extension. It doesn't require heavy resources for copying data to the portal. This is a lightweight solution.

This solution is fast. There is no delay so users can see data in real-time. Information in datasets is not modified, it is verified and updated. The CKAN portal shows the original owner of the data and redirects users to the original portal for further details. This scheme also allows the creation of informative solutions that contain and provide users with all necessary data without a long search and lots of effort.

I like this extension because as a designer I care about user experience, and this extension helps them to achieve their goals easily.



We used it in our work on [the CanWIN site](#) for the University of Manitoba project, an internal data management platform that collects and manages a large array of climate data, including spatial data, analyses and shares it with researchers.”



*ckanext-federated-index*



**Ian Ward**

Senior Solutions Architect

“There are many great extensions so I can’t choose just one. [ckanext-files](#) is a new extension designed to enhance CKAN’s file management capabilities and has potential to open up many new workflows. [ckanext-validation](#) can be used to nudge users to publish higher-quality data which is so important for data reuse. [ckanext-scheming](#) has become quite widely used for custom metadata schemas and is the basis for recent improvements to [ckanext-dcat](#), another really important extension for publishing metadata in an interoperable standard. The [ckanapi](#) tool is great for automating bulk tasks with CKAN and integrating CKAN with other python projects.”

[↗ ckanext-files](#)

[↗ ckanext-validation](#)

[↗ ckanext-scheming](#)

[↗ ckanext-dcat](#)

[↗ ckanapi](#)



## Lex du Plessis

Technical Delivery Manager

"I have a lot of extensions that I like but this one. It is called ckanext-gallery. It is from the London Natural History Museum Portal, which is powered by CKAN. For example, we can see it here being used to catalogue butterflies. It generates a preview of the image that is attached to the dataset and displays that on the webpage. And I just think this is such a really cool way of sharing the data, especially for a catalogue that is based on images. It is kind of the detail that you want to show."

The screenshot displays the 'Specimens' web application interface. At the top, there are three green buttons: 'Download', 'Open in image viewer', and 'Contact dataset curator'. Below these, the 'Specimen records' section features a navigation bar with 'Preview', 'Grid', 'Gallery' (selected), and 'Map' tabs. A search bar with the placeholder 'Enter search text' and a magnifying glass icon is present, along with a link to 'Advanced filters'. The left sidebar contains filters for 'Project' (set to 'papilionoidea new types ...'), 'Has image' and 'Has lat/long' checkboxes, and a 'Collection Code' dropdown menu. The main content area shows a grid of butterfly specimens, each with a thumbnail image and a label. The labels include the species name and a unique identifier, such as 'Heracles rumiko Shiraiwa & Gris...' and 'Anthene collinsi d'Abrera, 1980'.



## Steven De Costa

Chairman/Executive Director

“The [ckanext-hierarchy\\_extension](#) is a powerful yet simple tool for enhancing the flexibility of a CKAN portal by allowing the creation of hierarchical structures within organizations. This extension makes it possible to manage user privileges more efficiently by assigning different access and publishing roles to sub-organizations. This is particularly useful for data custodians and teams, as they can distribute responsibilities while maintaining clear oversight.

One of the key features I appreciate is its ability to preserve the organisational hierarchy when harvesting data from other portals. This means you can seamlessly integrate external data sources while keeping their structure intact, making it ideal for building a comprehensive metadata portal. For end users, the hierarchical access path offers an intuitive way to navigate and discover datasets, saving time and providing more focused search results.”



### ckanext-hierarchy

#### About

Organization hierarchy - CKAN extension

Readme

AGPL-3.0 license

Activity

Custom properties

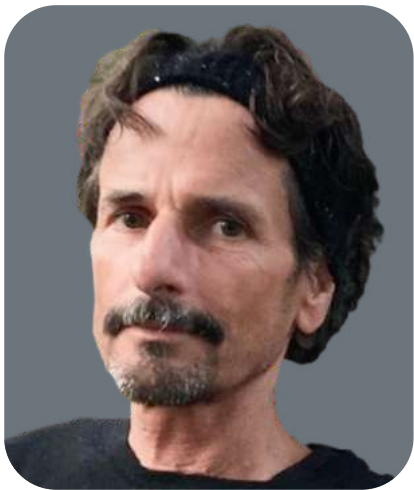
29 stars

17 watching

65 forks

[Report repository](#)



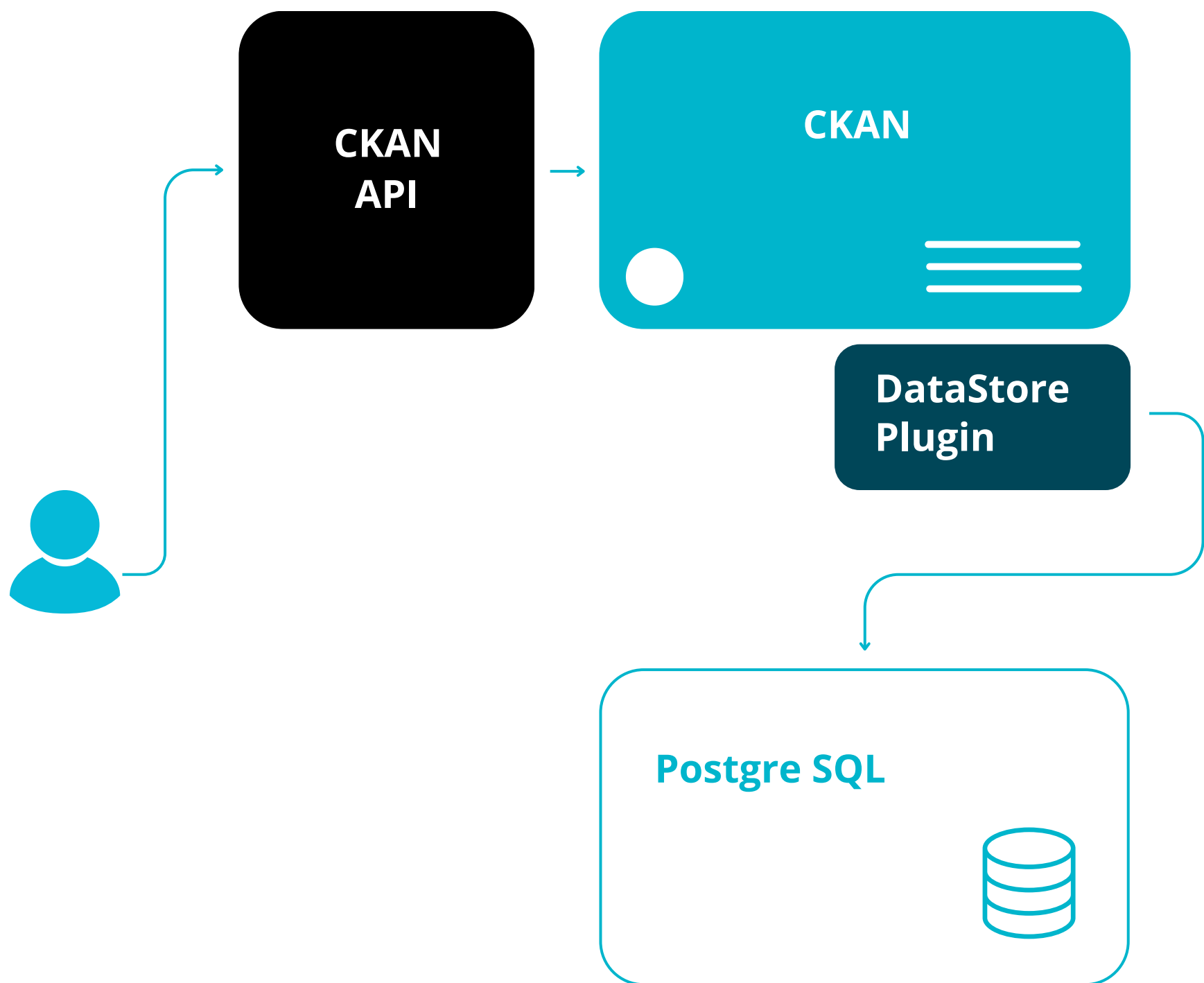


**Brett Jones**

DevOps Engineer

“The [CKAN DataStore plugin](#) is personally one of my favourites when it comes to working with data on the CKAN platform. It adds a powerful and flexible backend that allows for easy storage, querying, and filtering of datasets in a more structured and manageable way. This plugin makes it possible to store data in a [PostgreSQL database](#), which can then be queried directly through the CKAN API. This plugin is mature and easily installable.

In the Tech Team, we generally do not have many issues with this plugin. One of the key features that I really like is the ability to enable a data preview. Once the plugin is activated, users can interactively preview datasets right from the web interface. This is incredibly useful for exploring data before downloading it or running queries. You can see a sample of the dataset, which provides a clear and immediate understanding of the structure and content. Beyond that, the CKAN Datastore plugin offers support for filtering, sorting, and searching within the data, allowing users to extract just the pieces of information they need. For me, it’s an essential tool for making data more accessible and user-friendly, adding a lot of depth to CKAN’s already robust data management capabilities. It’s one plugin I always recommend to those looking to improve their CKAN experience.”



*Technical diagram of CKAN and PostgreSQL*

# Link Digital CKAN Extensions

<a href="#"><u>ckanext-selfinfo</u></a>	This extension is built to represent a basic information about the running CKAN Application accessible only to admins.
<a href="#"><u>ckanext-blocksmith</u></a>	Blocksmith is a CKAN extension that allows you to create and manage pages using a visual editor.
<a href="#"><u>ckanext-admin-panel</u></a>	Next generation admin interface for CKAN. See the extended documentation for more information.
<a href="#"><u>ckanext-event-audit</u></a>	This extension will capture and retain a comprehensive record of all changes within CKAN instances that Link Digital manages. The goal is to provide a robust mechanism for tracking modifications to entities such as users, datasets, resources, groups, and permissions. The need comes directly from clients with formal data handling and security requirements.
<a href="#"><u>ckanext-toolbelt/</u></a>	Collection of different tools for daily use.
<a href="#"><u>ckanext-resource-indexer/</u></a>	Discover more results in the Dataset search by searching through the content of resources.
<a href="#"><u>ckanext-federated-index/</u></a>	Lightweight solution for storing and searching remote datasets locally and redirecting to the original portal upon when dataset details page is opened.

# Link Digital CKAN Extensions

<a href="#"><u>ckanext-fpx/</u></a>	CKAN adapter for FPX service
<a href="#"><u>ckanext-ingest/</u></a>	Framework for transforming data stream into tasks. This extension provides tools for reading user-provided file-like object and producing tasks based on the input.
<a href="#"><u>ckanext-create-typed-package/</u></a>	Add dataset type selector to the "Add dataset" button.
<a href="#"><u>ckanext-editable-config/</u></a>	This plugin registers a set of API action for overriding config options and applying the changes to application without restarting the web server.
<a href="#"><u>ckanext-search-tweaks/</u></a>	Set of tools providing control over search results, sorting, etc.
<a href="#"><u>ckanext-nswdesignsystem/</u></a>	Collection of tools for styling CKAN using NSW Design System
<a href="#"><u>ckanext-collection/</u></a>	Tools for building interfaces for data collections using declarative style.
<a href="#"><u>ckanext-or-facet/</u></a>	Change logic of applying facets. With a bit of extra configuration, search for datasets, including any of applied facets, not necessary all of them
<a href="#"><u>ckanext-comments/</u></a>	This plugins provides comment threads linked to any of the main CKAN entities



# Link Digital CKAN Extensions

<a href="#"><u>ckanext-vip-portal/</u></a>	Generic access restrictions for anonymous user.
<a href="#"><u>ckanext-check-link/</u></a>	Link checker for CKAN. Provides API, CLI commands, and views for: checking availability of the file, refereed by resource checking availability of any arbitrary link. storing results of these checks visualizing stored results
<a href="#"><u>ckanext-flakes/</u></a>	Tools for creating and managing independent chunks of data.
<a href="#"><u>ckanext-fdt-sqlalchemy/</u></a>	[(debugging)] Enable Flask-DebugToolbar's SQLAlchemy pane.
<a href="#"><u>ckanext-iso19115/</u></a>	Export dataset into ISO 19115 XML.
<a href="#"><u>ckanext-composite-search/</u></a>	Complex search form for the dataset search page. Filter by field, search by multiple criteria, combine filter with the AND/OR operators and forget about Lucene queries
<a href="#"><u>ckanext-recaptcha/</u></a>	extension adds CAPTCHA verification to CKAN forms to help prevent spam submissions and automated abuse.
<a href="#"><u>ckanext-search-preset/</u></a>	Plugin that adds a number of API actions/helpers for storing search facets as a package fields and listing all the datasets that satisfies stored facets.
<a href="#"><u>ckanext-saml/</u></a>	This extension adds an ability to login from other source (known as IdP) using SAML2 standard. Your instance is presented as the SP.

# Link Digital CKAN Extensions

<a href="#"><u>ckanext-charts/</u></a>	This extension, ckanext-charts, provides additional functionality for working with charts in CKAN. It allows users to create, manage, and visualize charts based on data stored in CKAN datasets.
<a href="#"><u>ckanext-drupal-api/</u></a>	The extension helps synchronize the header and footer between CKAN and Drupal.
<a href="#"><u>ckanext-drupal-idp/</u></a>	When Drupal's session cookie is available use it for user authentication. Create missing users, using data from Drupal's DB and synchronize(conditionally) fields when user details changed on Drupal's side.
<a href="#"><u>ckanext-relationship/</u></a>	The extension adds an additional table to the database that stores relationships between entities in the form of triples (subject_id, object_id, relation_type). The relation_type parameter sets the type of relationship: peer-to-peer (related_to <=> related_to) and subordinate (child_of <=> parent_of). Adding, deleting and getting a list of relationships between entities is carried out using actions (relation_create, relation_delete, relations_list).
<a href="#"><u>ckanext-pygments</u></a>	Export dataset into ISO 19115 XML.
<a href="#"><u>ckanext-wysiwyg</u></a>	extension adds a "What You See Is What You Get" (WYSIWYG) rich text editor to CKAN, enhancing the content editing experience by allowing users to format text—such as bold, italics, headings, lists, and possibly embed media—directly within CKAN's interface, simplifying the creation and editing of resource descriptions or other text fields ckan-extensions.frictionlessdata.io+13

# Link Digital CKAN Extensions

<a href="#"><u>ckanext-auth</u></a>	The extension provides a 2FA authentication mechanism for CKAN.
<a href="#"><u>ckanext-unfold</u></a>	Enhance your CKAN experience with our extension that enables seamless previews of various archive formats, ensuring easy access and efficient data management.
<a href="#"><u>ckanext-mailcraft</u></a>	A custom mailer that can be conveniently expanded as needed
<a href="#"><u>ckanext-bulk</u></a>	An extension that helps to perform bulk operations over CKAN entities, such as datasets, organizations and groups.
<a href="#"><u>ckanext-clamav</u></a>	This is a basic example that helps to scan uploaded resources for malwares with clamd library.
<a href="#"><u>ckanext-harvest-basket</u></a>	The harvest_basket extension comes with a few custom harvesters for different data portals.
<a href="#"><u>ckanext-tour</u></a>	ckanext-tour is a CKAN extension that provides a guided tour feature for CKAN instances. It allows users to create interactive tours to showcase different features and functionalities of a CKAN instance.
<a href="#"><u>ckanext-transmute</u></a>	This extension helps to validate and convert data based on a specific schema.

# Link Digital CKAN Extensions

<a href="#"><u>ckanext-permissions</u></a>	The extension allows you to build a permission system within CKAN. For now, it uses 3 default roles: anonymous, user and sysadmin. We plan to expand the functionality to allow registering context roles, like organisation admin, editor or member.
<a href="#"><u>ckanext-let-me-in</u></a>	An extension that generates a one-time login link for a user by its id, name or email.
<a href="#"><u>ckanext-msal</u></a>	This extension allows you to sign in users with Microsoft identities (Azure AD, Microsoft Accounts and Azure AD B2C accounts). It uses <a href="#"><u>Microsoft MSAL</u></a> library.
<a href="#"><u>ckanext-latex</u></a>	Adds LaTeX scheming display and form snippets for CKAN forms.
<a href="#"><u>ckanext-alias</u></a>	The extension allow you to create a URL aliases for datasets. For example, if you have a dataset with my-dataset name, it can be accessed via URL like that https://your-domain/dataset/my-dataset. Now you could add multiple aliases that will redirect user to the dataset.
<a href="#"><u>ckanext-files</u></a>	Files as first-class citizens of CKAN. Upload, manage, remove files directly and attach them to datasets, resources, etc.
<a href="#"><u>ckanext-textcaptcha</u></a>	A text based captcha for the <a href="#"><u>CKAN open data platform</u></a> user registration form using <a href="http://textcaptcha.com"><u>http://textcaptcha.com</u></a> .
<a href="#"><u>ckanext-saml/</u></a>	This extension adds an ability to login from other source (known as IdP) using SAML2 standard. Your instance is presented as the SP.





## Let's talk CKAN extensions

Do you have a favourite CKAN extension? Or do you need a customised extension developed to meet a challenge you face in an open data portal or data catalogue? If so, contact us and tell us about your project.

Contact Us