



#### Head Office

Level 4, 431 King William Street,  
Adelaide SA 5000

#### Phone

1300 769 809

#### Email

[contact@parashift.com.au](mailto:contact@parashift.com.au)

---

## Introducing AlfBI

### What is AlfBI

AlfBI is a turnkey solution for providing reporting on an Alfresco repository.

Utilising Alfresco Public API, it allows for data from the repository to be harvested and included in a separate data warehouse.

From this data warehouse, AlfBI provides preconfigured visual web based reports, while also allowing for custom configured reports.

Alongside this, there are a couple of other features that are included to help manage content at scale:

- Bulk Editing of Metadata via CSV Upload: allows sysadmins to make mass changes to documents
- Document and Image Deduplication: allows users to see whether the document exists in more than one place

### Why Would I need it?

There are a few use cases that AlfBI can provide a solution for:

- You want to provide reporting in Alfresco but not have to spend the time and resources in setting it up from scratch
- Your repository is becoming unwieldy to manage and you need some extra visualisation tools to assist in cleanup/maintenance
- You want to be notified when a certain action happens, such as a user accessing a confidential document

- Your managers want to align some KPIs to your users based upon how they use Alfresco or what content they are ingesting
- You have a custom content model that you want to be exposed as an SQL table but still updated/managed within Alfresco

### How does it work?

AlfBI is divided into 3 components:

- A module that is installed within Alfresco as an AMP. This module exposes some API endpoints to harvest information from Alfresco for storing in an external database
- An Apache Camel Component which reads all of these endpoints from an Alfresco instance, and stores it within a PostgreSQL database
- Ready made reports and dashboards for Metabase for visualisation and querying of information.

The Alfresco Module is simply installed as an AMP, and by itself exposes some endpoints to allow for an external process to collect and harvest information. These endpoints are read only restful APIs that provide information as JSON.



#### Head Office

Level 4, 431 King William Street,  
Adelaide SA 5000

#### Phone

1300 769 809

#### Email

[contact@parashift.com.au](mailto:contact@parashift.com.au)

---

The following endpoints are exposed by the module:

- nodes: Any metadata or information around documents/folders, including all custom properties
- audit: The audit log of Alfresco, including all user actions within the system
- system: System information including CPU, RAM, Disk Space
- workflow: Tasks and workflows that have taken place within Alfresco

These endpoints are polled by a ready-made Apache Camel Route, which is configured to store this information into a separate PostgreSQL database as a data warehouse. Periodically changes from the Alfresco instance are harvested, and stored within this data warehouse.

This camel route can run on the same server or a different one, providing there is access to the Alfresco instance via HTTP.

The PostgreSQL database schema is generated dynamically based upon the shape of the information sent. Each time a content model changes or is added, the database schema is updated to reflect and store the new fields and tables. The types of the fields are auto detected too, so if you have dates or numbers, they are stored within PostgreSQL as the native SQL type. It is also possible, however, to mandate a particular schema if necessary.

Utilising Metabase, AlfBI provides some standard reports and dashboards for Alfresco which are preconfigured based upon this schema. There are currently 25 Reports spread across 3 Dashboards that come with AlfBI, and range from things like Total Active users per day, to the Trashcan Size.

These default reports are purely client driven: more reports are added as we receive feedback.

These reports can easily be embedded into an Alfresco Share site's dashboard, or sent periodically via email using Metabase's pulse feature.

If you want to change or add extra reports, custom questions and dashboards can be created easily using the full features of Metabase, so that if you have any custom content model or requirements around reporting, you are not limited to the reports that are provided out of the box. Along with our existing documentation on how to do this, we can provide training and consultation if you need extra help here as well.

Since we are just using standard SQL, if you have an existing reporting tool, such as pentaho or jaspersoft, you can utilise these instead of metabase, and can be provided with the SQL of the default reports as a starting point for integrating into your own reporting workflow.