

For end users

- When viewing tags, underneath each contribution you will now see a highlight showing where the tag was used in the text rather than just the description of the contribution or its first few lines. This will also happen on group home pages and newsletters for contribution sections limited to a particular tag that show a description of the contribution. If the tag does not appear in the text then the behaviour will be as before.
- Clicking on a link to a source item will now open the URL in another tab for HTML pages and PDF documents.
- If you have permission to add tags but not to edit them then the associations part of the form will no longer be shown and a spurious error will no longer be given when submitting the form.
- A small number of minor issues have been fixed.

For group administrators

All groups

- If the system does not allow users to register and you select the *Invite user* action from the members index and try and invite an email address that is not one for a current user then you will now be given a message telling you this rather than a general error message.

Monitors

- When adding or editing a source that uses a spider, you can now set a list of schema.org or OpenGraph types to filter what pages are created as items using the new *Schema types* field. This allows you to, for example, restrict items from a site to just news articles.
- When adding or editing a source that uses a spider, you can now set a CSS selector for extracting the links from a page. This can be useful when there is only a specific set of links that you want to use on an update page.
- When adding or editing a source that uses a spider, you can now set a CSS selector for elements that should be removed from a page. This can be useful for sites that place promotional blocks or ads in the middle of article copy.
- If you change the interval at which an update page is checked for new items then this will now take effect immediately, rather than after the old interval has elapsed. You can also now set it to only carry out a single update, which can be useful for one-off retrieval of items from a page.
- If a source does not use a full spider (i.e. using an update page or a feed downloading full content) then it is now possible to redownload all its content by selecting the *Update content* button on the pop-up menu next to its name. This makes it easier if you change settings affecting content, such as the pattern for URLs to download or the selectors for content in a page.

- The source spider options that previously took a full regular expression now only take lists of simple wildcards in order to protect against expressions that could cause a denial-of-service attack. The *Article URL patterns* option has been renamed to *URL Patterns* and the *Page must contain pattern* and *Page must not contain pattern* options are now plural.
- There is a new *Source filter* group settings category that will apply a filter to all the sources within that monitor. Previously this would need to be configured for each source individually.
- Improvements have been made for handling non-compliant feeds, including those including a UTF-8 byte-order-mark at the start or missing the XML header.
- Non-HTML documents downloaded by the spider will now take their subject from the HTTP `Content-Disposition` header when available in preference to using a name taken from the URL.
- If a publication date from spidered content is later than the current date then the current date will be used instead.
- When you view the download status of a source, you will now be given more information about the last connection the feed or spider made and more specific counts for URLs that failed to be retrieved or were filtered out.
- Some other minor issues affecting sources, their editing and their retrieval of content have been resolved.

For system administrators

- When you add a new user that is a system administrator, you will now be given the option to set their system administrator password after creation.
- In the *External content* system settings category, you can now set the interval at which feeds are checked for new items.

For operations engineers

Upgrade

- As it can now be configured in the system settings, in the `bootstrap.properties` configuration file the `com.jiglu.external.feedCheckInterval` property has been removed.
- Because of an upgrade to the search system, search indexes will need to be rebuilt. See <https://support.jiglu.com/spaces/installation/knowledge/rebuilding-search-indexes> for instructions on how to do this if you host your own system.

Other changes

- Third-party libraries have all been updated to their latest recommended versions.

Jiglu 15.3

Changes in the August 2022 release



Security

- Because of the potential to use bad regular expressions that could cause a denial-of-service attack, we have replaced this use in sources with simpler wildcard patterns.
- The PostgreSQL database driver has been updated to resolve CVE-2022-31197, although this should not have been exploitable.