

Disclosure of AI Systems in Nested Knowledge

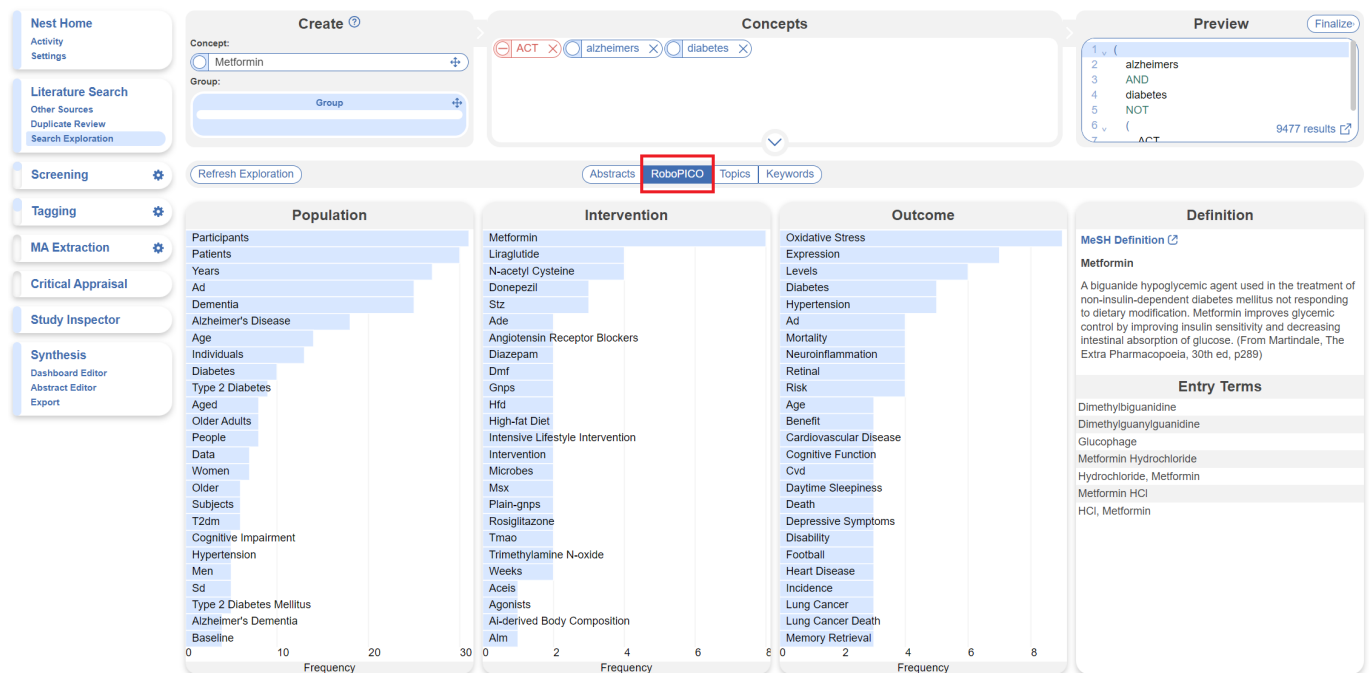
There are several tools in Nested Knowledge that utilize artificial intelligence and machine learning to make systematic reviews easier and more effective to conduct. This page provides technical details on what these features are, and how your data is used.

1. RoboPICO

- Used to highlight Populations, Interventions/Comparators, and Outcomes in abstracts during [Screening](#)
 - Optional:* by default this feature is toggled on, but can be toggled off and will remain off for all abstracts in the queue/ongoing modules and Study Inspector until toggled back on.



- Used to generate most commonly reported terms among the literature, to inform the build of search queries in [Search Exploration](#)
 - Optional:* Search Exploration is not a required step in AutoLit and simply offers assistance to build a search. However, when it is used, RoboPICO auto-generates terms when concepts are entered and “Refresh Exploration” is selected. This cannot be switched off when Search Exploration is used.



a. Which data does RoboPICO use?

The abstract of records is made available to the model.

b. How does the model work?

RoboPICO uses a fork of the machine learning system offered by [RobotReviewer](#). Specifically, [Named-entity Recognition models](#) extract Patient/Problem, Intervention, and Outcome entities from data in article abstracts. NK's modifications to RobotReviewer are open and General Public Licensed.

This model is not trained or updated from your data.

2. Bibliomine

[Bibliomine](#) extracts references from full text PDFs. Typically, previous systematic reviews or landmark studies are bibliomined, importing all cited references as records directly into your nest. This feature does not access your data unless you use it.

Optional: This feature is helpful if, for example, you are performing an update on an existing review, but is not required to successfully upload records to a nest.

a. Which data does Bibliomine use?

Bibliomine consumes any PDFs uploaded for its purposes. It writes records with full bibliographic data to your nest (pending user addition).

b. How does the model work?

Bibliomine uses [Cermine](#), an open source machine learning library for mining documents. Using DOI & title extracted, full bibliographic data will be retrieved from PubMed or CrossRef (in that order of preference).

This model is not trained or updated from your data.

3. Robot Screener

[Robot Screener](#) uses a model trained on human screening decisions to make reviewer-level screening decisions based on inclusion probabilities. In effect, it replaces one human reviewer when turned on in nests with a Dual Screening mode. This feature does not have access to your data unless you turn it on in Settings.

Optional: This feature is helpful to speed up the Screening process, but is not required to successfully Dual Screen all records in a nest.

Settings: Standard Modes: COVID-19

Screening

In Standard Screening, one user screens each record. Inclusion sends the record forward for gathering, such as tagging, meta-analytical extraction, and critical appraisal. Exclusion does not queue the record for gathering.

In Dual Screening, two users independently screen each record, and then all screening determinations are reviewed by an administrator. The administrator adjudicates any disagreement between the original screeners to set the final determination for each record.

In Two Pass Screening, all records are first rapidly screened using only title and abstract. Records may be advanced from title/abstract screening to more intensive full text screening, where final inclusion is determined.

In Dual Two Pass Screening, two users rapidly screen all records using only title/abstract and these determinations are reviewed and advanced by an administrator. Two users then screen all full texts and final inclusion is determined by the administrator.

Choose Mode:

☒ Standard

☐ Two Pass

Choose number of reviewers:

☐ Single

☒ Dual

Screening Model

Screening models predict the probability of individual records being included (Standard Screening) or abstract advanced (Two Pass Screening), using your past screening determinations. These probabilities help AutoLit determine which records to show first during screening and may be viewed and filtered in Inspector. They are also leveraged by Robot Screener in dual screening modes, if enabled.

The model may be trained manually or automatically (recommended). If the screening model is set to automatic, the model will be retrained after every 10 newly screened records. Otherwise, the model may be retrained manually throughout screening.

Choose mode:

☐ Automatic Training

☒ Robot Screener

[View Screening Model](#)

a. Which data does Robot Screener use?

The following data from your records are model inputs:

- Bibliographic data
 - Time since publication of the record
 - Page count
 - Keywords/Descriptors
- Abstract Content
 - N-grams
 - OpenAI text embedding (ada-002)
- Citation Counts from Scite, accessed using the DOI
 - Number of citing publications

- Number of supporting citation statements
- Number of contrasting citation statements

The model is trained on adjudicated AB or FT screening decisions, depending on your screening mode. This includes the exclusion reason. Similarly, the model outputs a screening decision for each record requiring a reviewer-level decision.

If enabled, Robot Screener will continuously screen new records as they imported into your nest.

b. How does this model work?

Learn more about the [screening model](#) that generates inclusion/advancement probabilities for Robot Screener. A probability threshold, optimized on a geometric mean of precision and recall in a cross validation, determines if records should be included or excluded.

This model is only available within your nest & its parameters are not shared with other nests or users.

4. Smart Tag Recommendations

Smart Tag Recommendations are context-aware recommendations of relevant concepts within full text PDFs. These recommendations include provenance via an annotation within the PDF, making review of the model's work possible.

Optional: This feature is helpful to speed up the data extraction process, but is not required to perform data extraction of all records in the nest. Standard Tag Recommendations should remain selected to keep this feature off.

Settings: Standard Modes: COVID-19

Tagging

In Standard tagging, the entire tagging hierarchy is made available as an open-ended list.

In Form-based tagging, tags can be turned into questions to be posed to the reviewer. There are three types of questions: Single Apply questions apply the tag selected, Single Select questions allow for only one of the child tags to be applied and Multiple Select questions allow for multiple child tags to be applied. All tags may have text content. Switching between these modes results in no loss of data.

Tag Recommendations

Smart tag recommendations use GPT 4, a large language model from OpenAI, to provide automatic highlighting of full texts based on your configured tags. Standard recommendations use keyword lookup.

Import

Import existing tagging data for any record already in your nest by uploading a spreadsheet.

Choose mode:

☒ Standard

☐ Form-based

Choose type:

☒ Standard

☐ Smart

Upload data:

[Begin](#)

a. Which data does Smart Tagging use?

Recommendations are generated for the full-text PDFs of all included records. You instruct the model through:

- Standard Tagging:

- Tag Names
- Hierarchical structure of tags
- Form Based Tagging:
 - Tag Names
 - Question Types
 - Question

b. How does the model work?

Smart Tagging uses [OpenAI's GPT-4](#), a large language model, to identify relevant concepts or answer questions about plain-text reductions of full-text PDFs. The model's recommendations back to the PDF as an annotation using fuzzy (edit distance) search.

5. Synthesis AI Disclosure

The records included in Synthesis, as well as the Tags applied to underlying reports, may have been collected by users with assistance of Artificial Intelligence. The Artificial Intelligence tools are generally integrated with direct researcher oversight, and the nest owners have final responsibility for the accuracy of all Screening and Tagging.

From:

<https://wiki.nested-knowledge.com/> - **Nested Knowledge**

Permanent link:

https://wiki.nested-knowledge.com/doku.php?id=wiki:support:ai_disclosure

Last update: **2023/11/16 15:19**