UNC UNIVERSITY LIBRARIES

ICASR

September 2024

Automation for Screening Search Results at an R1 Research University

Michelle Cawley

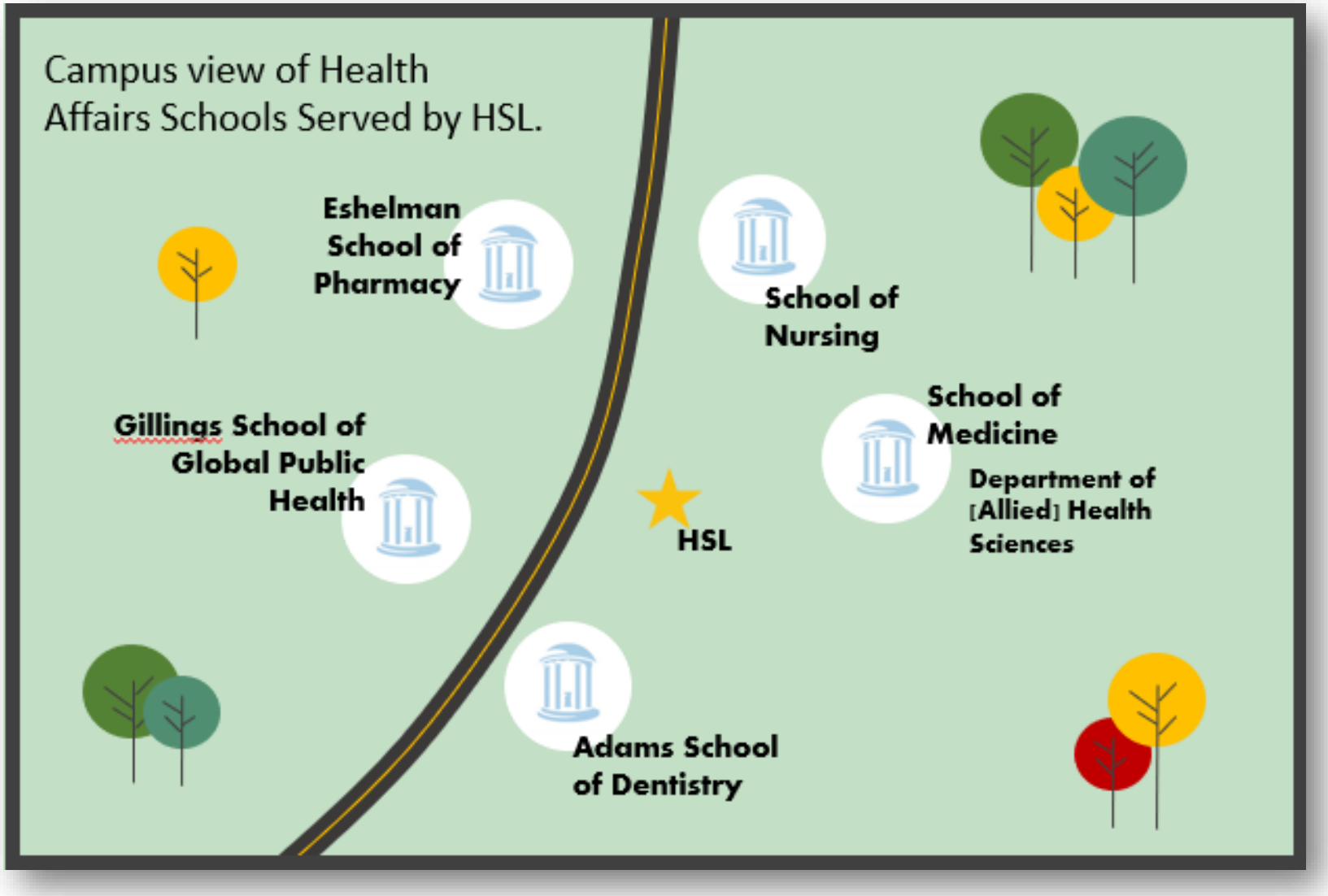
mcawley@unc.edu

Photo Credit: Rebecca Carlson

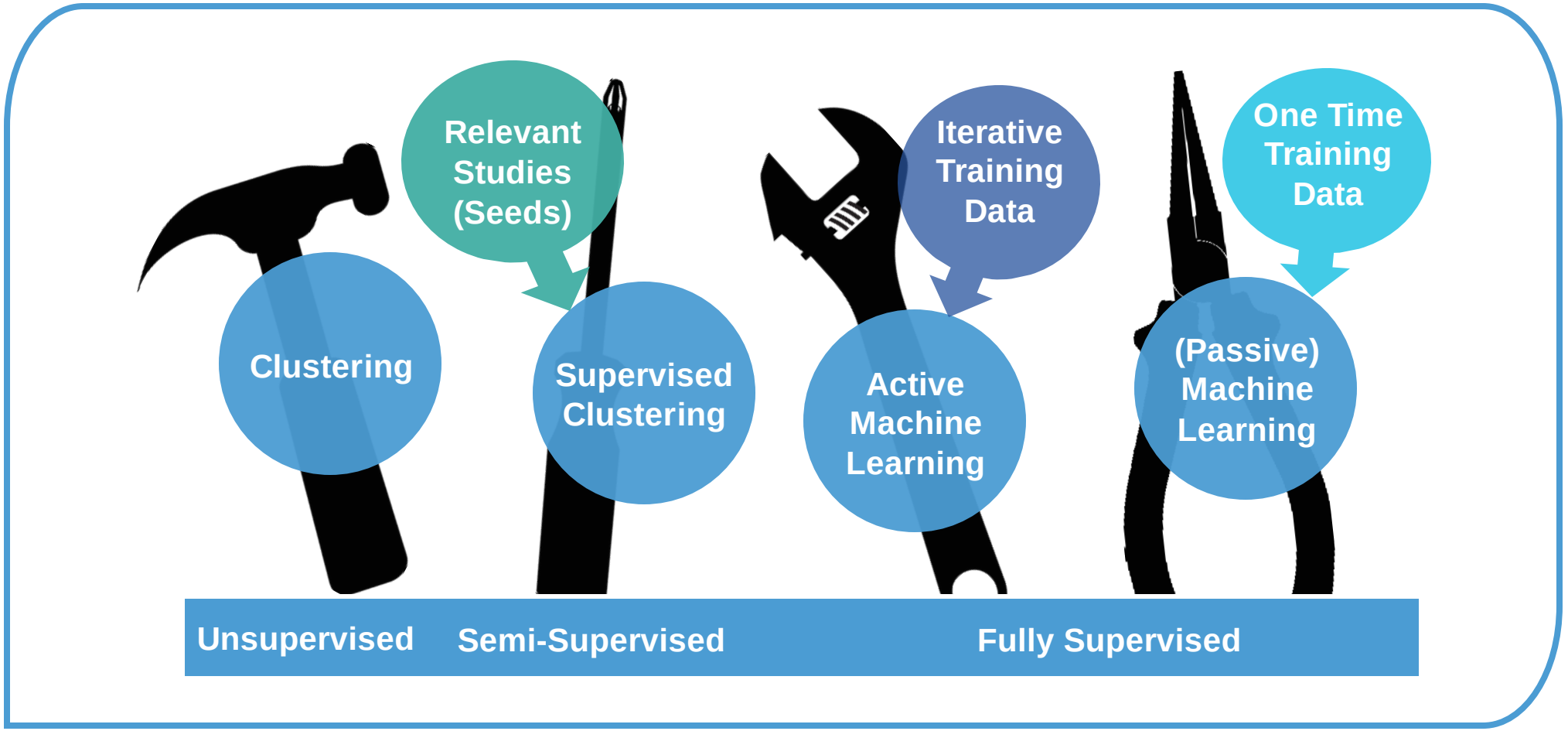
Session Roadmap

- UNC Chapel Hill
  - Context
  - Current state
- Screening and Tagging Process
- Advancing automation going forward

UNC Health Affairs Landscape



Text Analytics Approaches



ML for Bibliographic Data

Most often: Large, comprehensive literature searches, including systematic reviews (SRs).

Search Updates

Image by Tom Majric from Pixabay

Needle in the Haystack

Image by Евгения from Pixabay

Wring Out the Towel

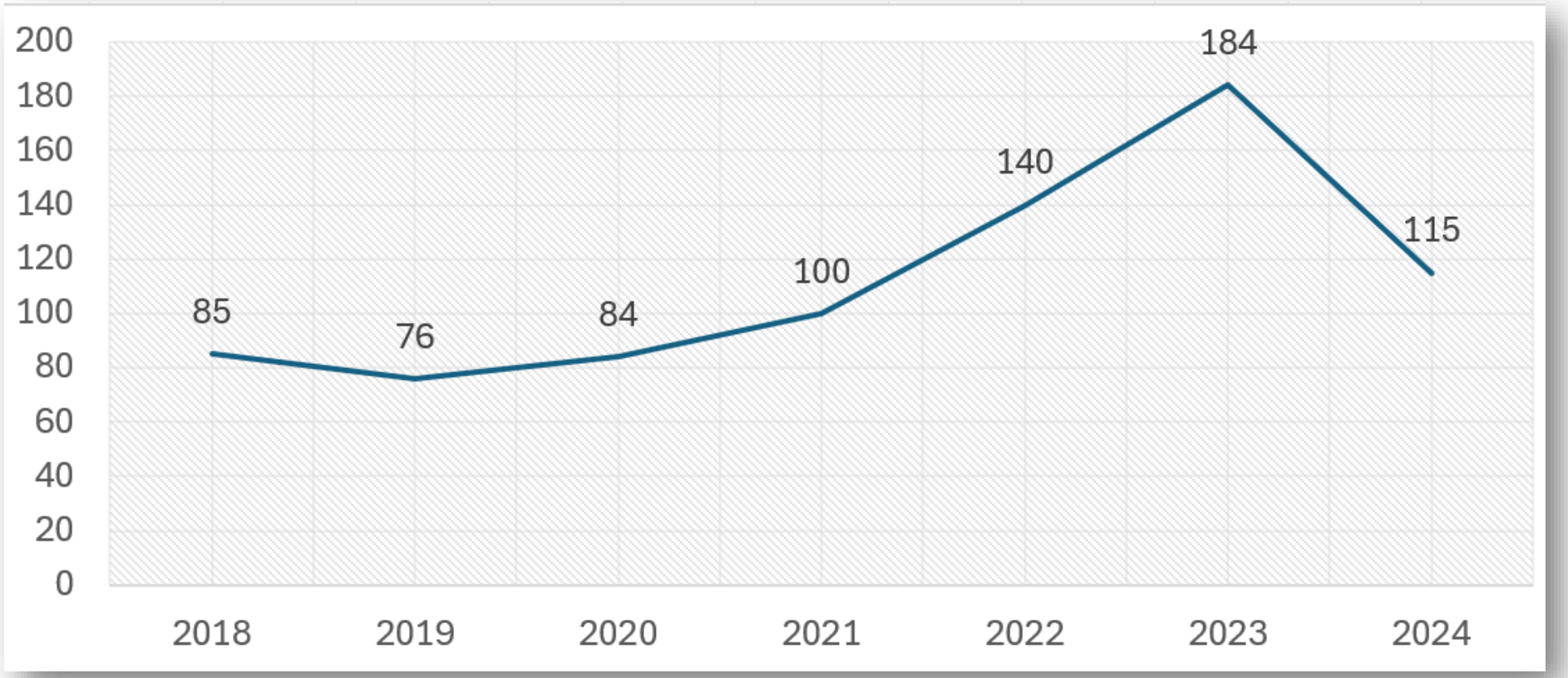
Image by Siala from Pixabay

More ML Success Stories @ UNC

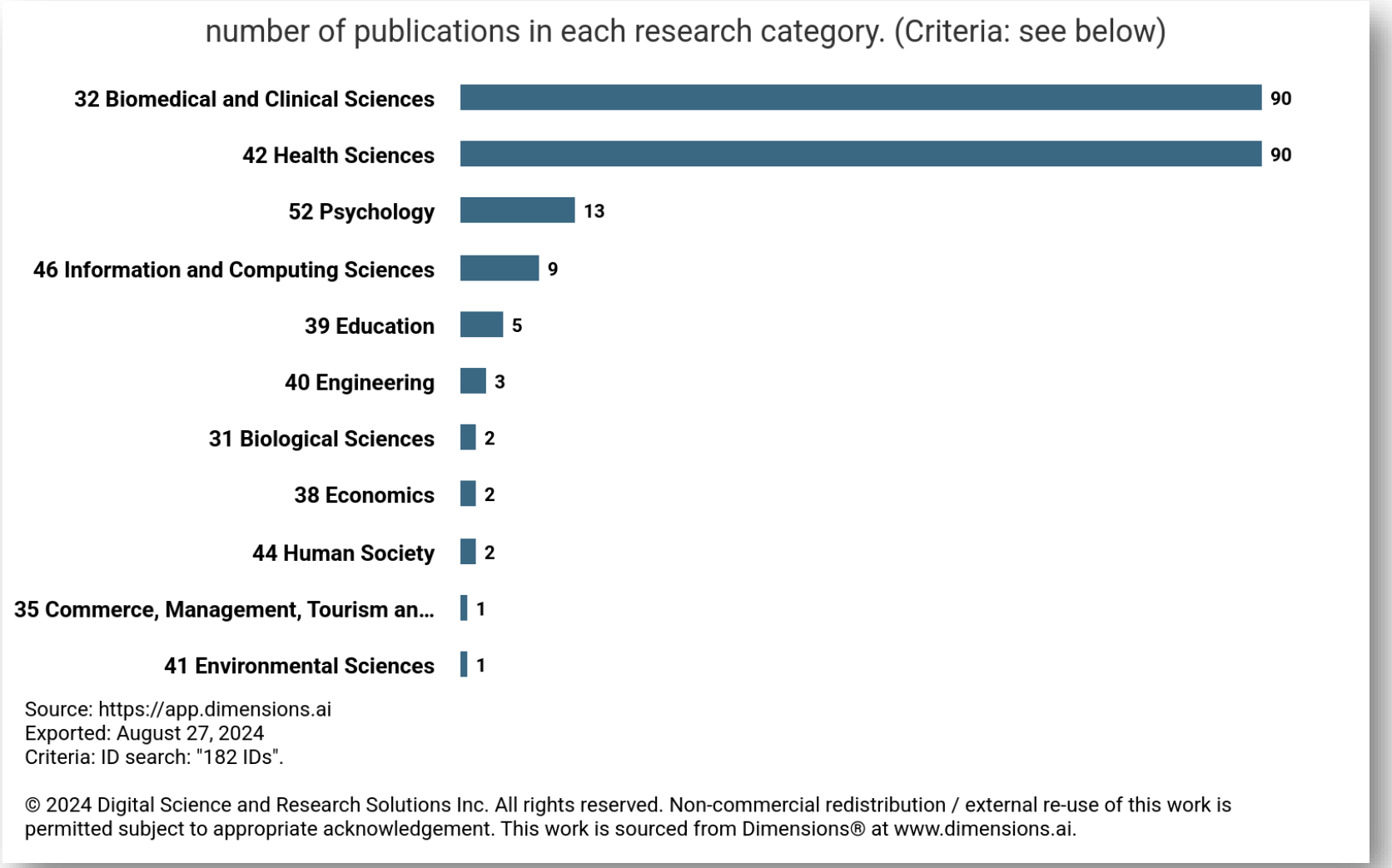
- Clustering for quick answers.
- Stratifying search results.
- Low precision dataset: Finding the ‘needle in the haystack’.
- Evidence for two-phase approach: Supervised clustering → Machine Learning.
- Externally derived training data.

Cawley, M. (2022). Supporting efficiencies in locating evidence using machine learning and other automation approaches. In Mani, NS; Cawley, M. (Eds.), Handbook of Research on Academic Libraries as Partners in Data Science Ecosystems. IGI Global, Hershey, PA.

UNC SR Requests since 2018



By Research Category



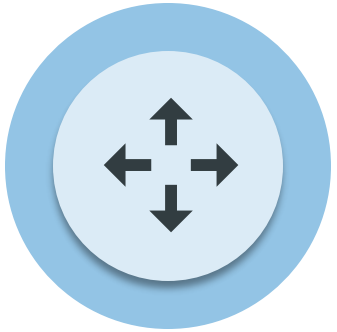
Growth and Evaluation

30 Projects

12 Staff Trained

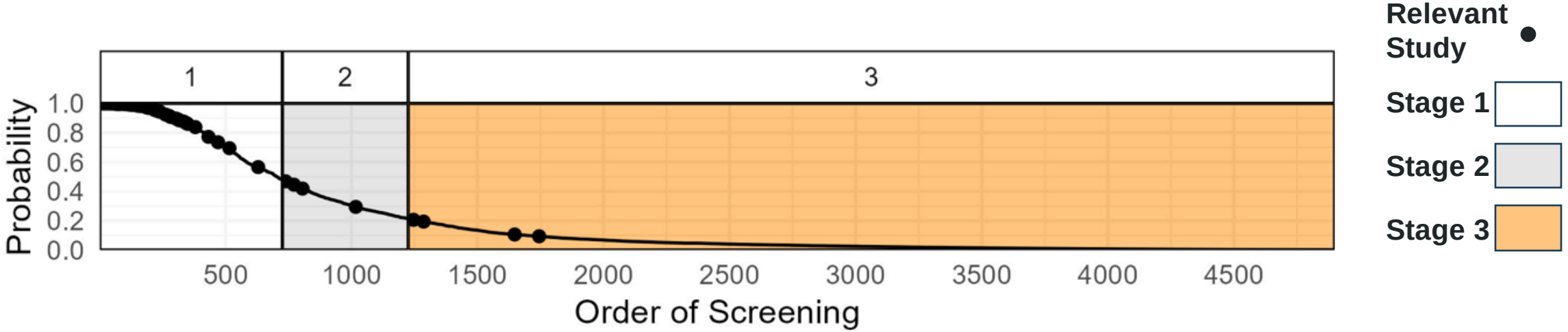
Save Time

Expand Scope



Simulation 1

Distribution of Relevant Studies



Simulation 1

Confusion Matrix

Predicted Classification	n=4,899	Actual Classification: Positive	Actual Classification: Negative
	Positive	TP (75)	FP (1,150)
	Negative	FN (4)	TN (3,670)

Precision = 6% | Recall = 95%

Classified Correctly

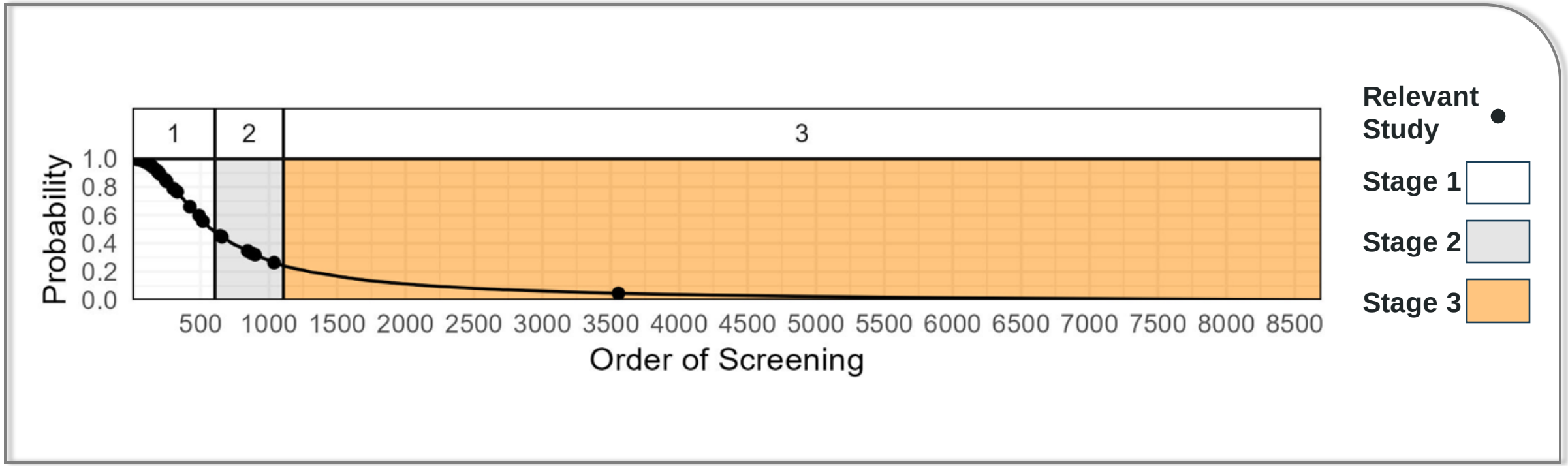
TP: True Positives
TN: True Negatives

Misclassifications

FP: False Positives (Type 1)
FN: False Negatives (Type 2)

Simulation 2

Distribution of Relevant Studies



Simulation 2

Confusion Matrix

Predicted Classification	n=8,694	Actual Classification: Positive	Actual Classification: Negative
	Positive	TP (48)	FP (1,057)
	Negative	FN (1)	TN (7,588)

Precision = 5%

Recall = 98%

Classified Correctly

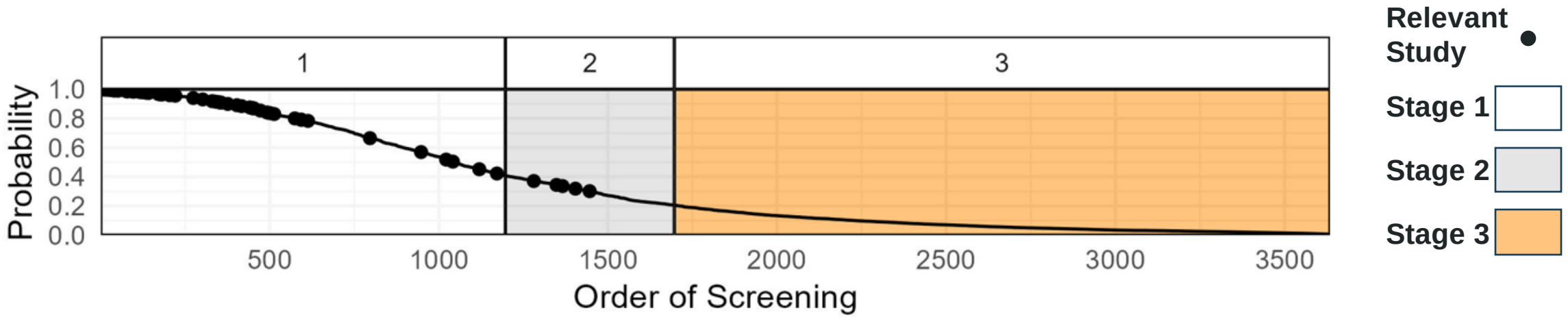
TP: True Positives
TN: True Negatives

Misclassifications

FP: False Positives (Type 1)
FN: False Negatives (Type 2)

Simulation 3

Distribution of Relevant Studies



Simulation 1

Confusion Matrix

Predicted Classification	n=3,634	Actual Classification: Positive	Actual Classification: Negative
	Positive	TP (64)	FP (1,632)
	Negative	FN (0)	TN (1,938)

Precision = 4% | Recall = 100%

Classified Correctly

TP: True Positives
TN: True Negatives

Misclassifications

FP: False Positives (Type 1)
FN: False Negatives (Type 2)

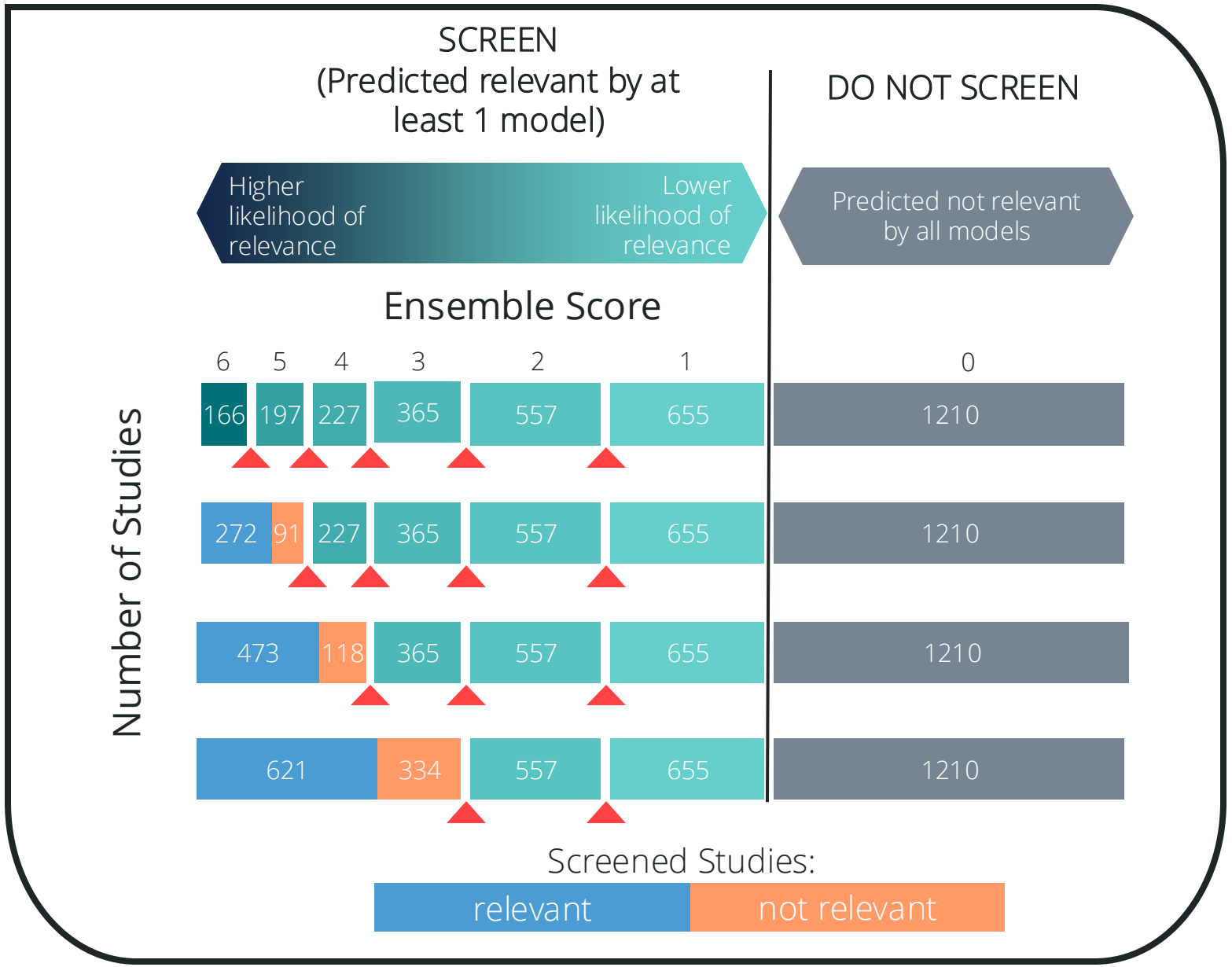
Screening & Tagging Process

- Enterprise license for Covidence
- Librarian recommends automation using DoCTER
- Typical process
- 2 Phase Supervised Clustering + Machine Learning
  - Research team screens 250 studies for “Seeds”
  - Librarian runs SC
  - Research team screens in Covidence
  - Switch to ML
  - Research completes screening in Covidence
  - Librarian recommends when to stop

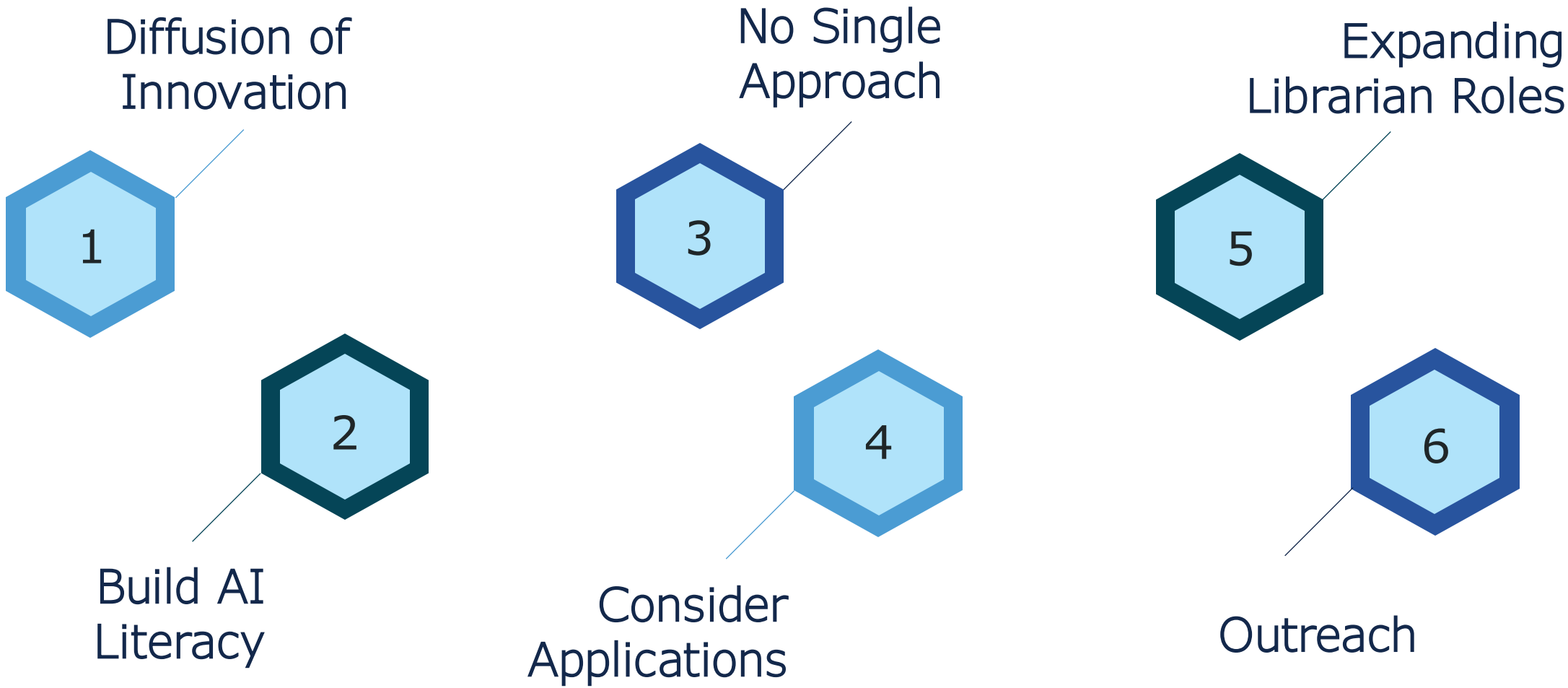
Supervised Clustering with an Ensemble Approach

Options at each ▲ checkpoint:

1. Keep screening
2. Stop screening
3. Modify Approach (e.g., switch to machine learning)



Advancing Automation



Selected Publications

Anderson, D. M., R. Cronk, D. Fejfar, E. Pak, M. Cawley and J. Bartram (2021). Safe Healthcare Facilities: A Systematic Review on the Costs of Establishing and Maintaining Environmental Health in Facilities in Low- and Middle-Income Countries. Int J Environ Res Public Health **18**(2).

Cawley, M. (2022). Supporting efficiencies in locating evidence using machine learning and other automation approaches. In Mani, NS; Cawley, M. (Eds.), Handbook of Research on Academic Libraries as Partners in Data Science Ecosystems. IGI Global, Hershey, PA.

Cawley, M., R. Beardslee, B. Beverly, A. Hotchkiss, E. Kirrane, R. Sams, A. Varghese, J. Wignall and J. Cowden (2020). Novel text analytics approach to identify relevant literature for human health risk assessments: A pilot study with health effects of in utero exposures. Environment International 134: 105228.

Cohen, A. M., W. R. Hersh, K. Peterson and P.-Y. Yen (2006). Reducing workload in systematic review preparation using automated citation classification. Journal of the American Medical Informatics Association : JAMIA 13(2): 206-219.

Mostafa, J. and W. Lam (2000). Automatic classification using supervised learning in a medical document filtering application. Information Processing & Management 36(3): 415-444.

O'Connor, A. M., Tsafnat, G., Thomas, J., Glasziou, P., Gilbert, S. B., & Hutton, B. (2019). A question of trust: Can we build an evidence base to gain trust in systematic review automation technologies? *Systematic Reviews*, *8*(1), 143. https://doi.org/10.1186/s13643-019-1062-0

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews*, *4*, 5. https://doi.org/10.1186/2046-4053-4-5

Thomas, J., McDonald, S., Noel-Storr, A., Shemilt, I., Elliott, J., Mavergames, C., & Marshall, I. J. (2020). Machine learning reduced workload with minimal risk of missing studies: development and evaluation of a randomized controlled trial classifier for Cochrane Reviews. *Journal of Clinical Epidemiology*. https://doi.org/10.1016/j.jclinepi.2020.11.003

Tsafnat, G., Glasziou, P., Karystianis, G., & Coiera, E. (2018). Automated screening of research studies for systematic reviews using study characteristics. *Systematic Reviews*, *7*(1), 64. https://doi.org/10.1186/s13643-018-0724-7

Wallace, B. C., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2010). Modeling annotation time to reduce workload in comparative effectiveness reviews. *Proceedings of the ACM International Conference on Health Informatics - IHI '10*, 28. https://doi.org/10.1145/1882992.1882999

Wallace, B. C., Trikalinos, T. A., Lau, J., Brodley, C., & Schmid, C. H. (2010). Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics*, *11*, 55. https://doi.org/10.1186/1471-2105-11-55

Varghese, A., M. Cawley and T. Hong (2018). "Supervised clustering for automated document classification and prioritization: a case study using toxicological abstracts." Environment Systems and Decisions 38(3): 398-414.

UNIVERSITY
LIBRARIES

The University of North Carolina at Chapel Hill