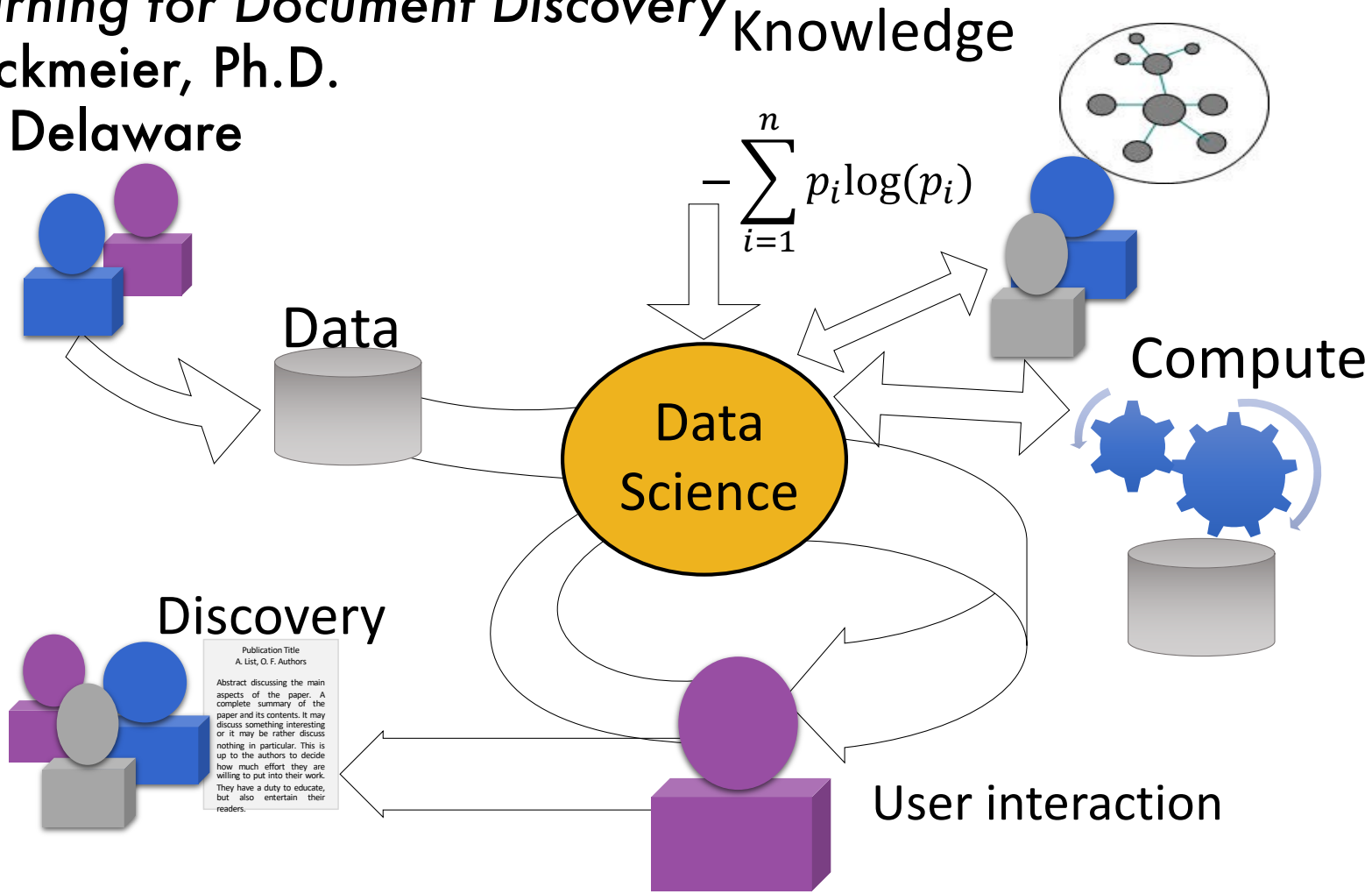
Machine Learning for Document Discovery

Austin J. Brockmeier, Ph.D.
University of Delaware



Systematic reviews: Evidence based research

?

P Patient Population

I Intervention or Issue

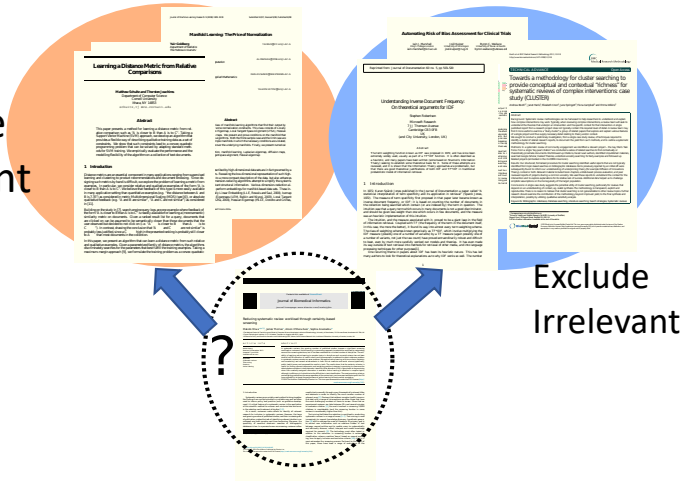
C Comparison intervention (optional)

O Outcome of interest

https://library.triton.edu/nursing

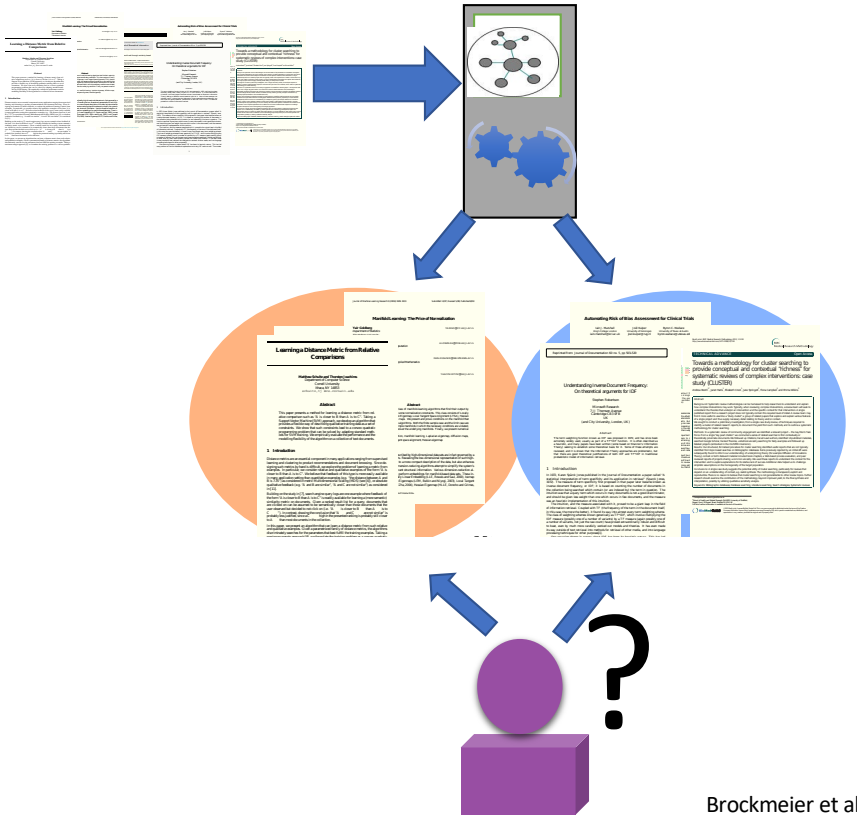
Query

The University of Edinburgh

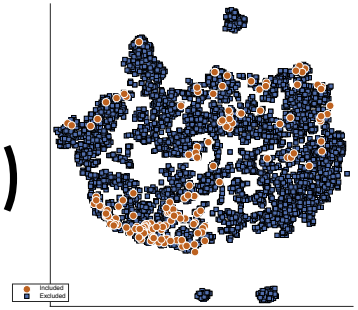


Hits

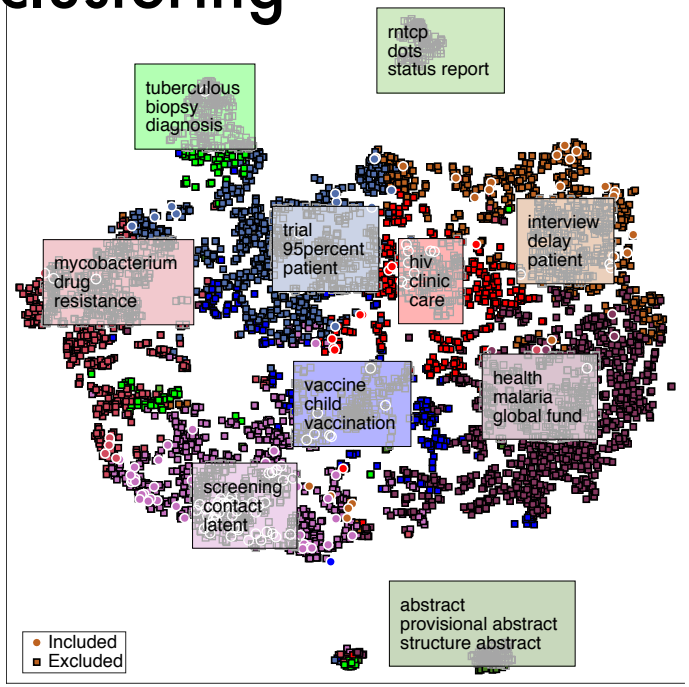
Clustering



Visualization with dimension reduction / embedding (t-SNE)

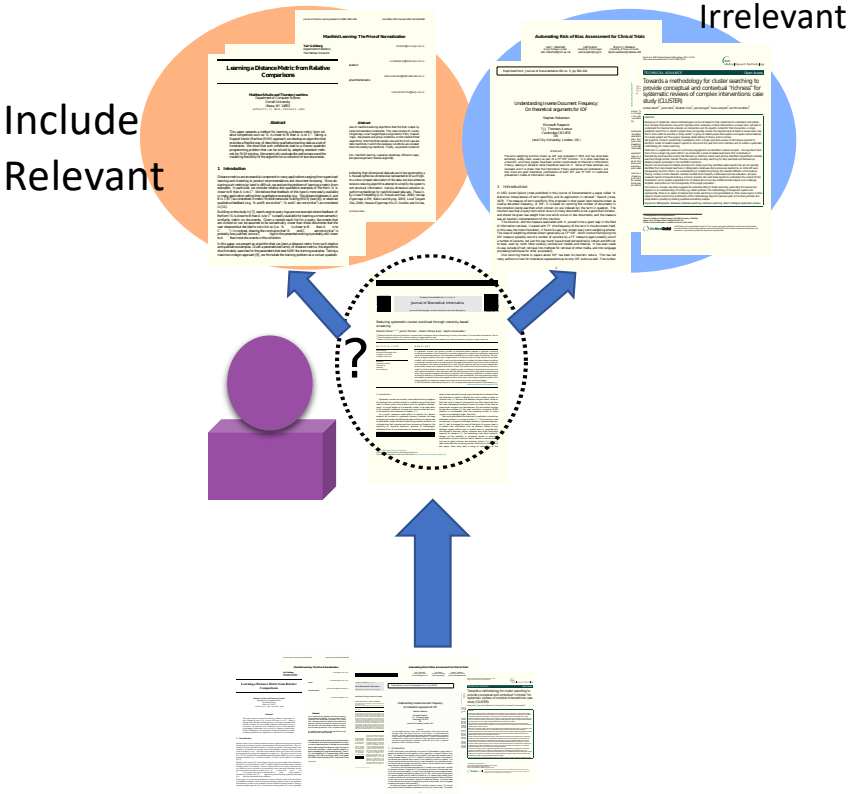


Descriptive clustering

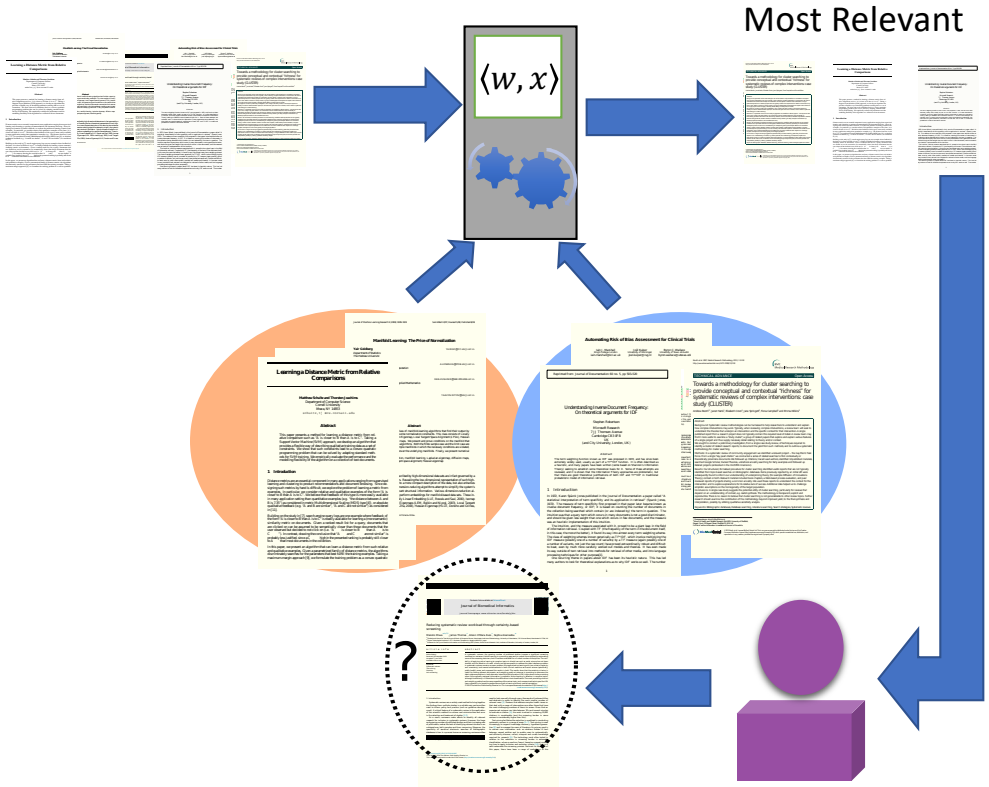


Brockmeier et al., "Self-Tuned Descriptive Document Clustering Using a Predictive Network," in *IEEE TKDE*
doi: 10.1109/TKDE.2017.2781721.

Initial training

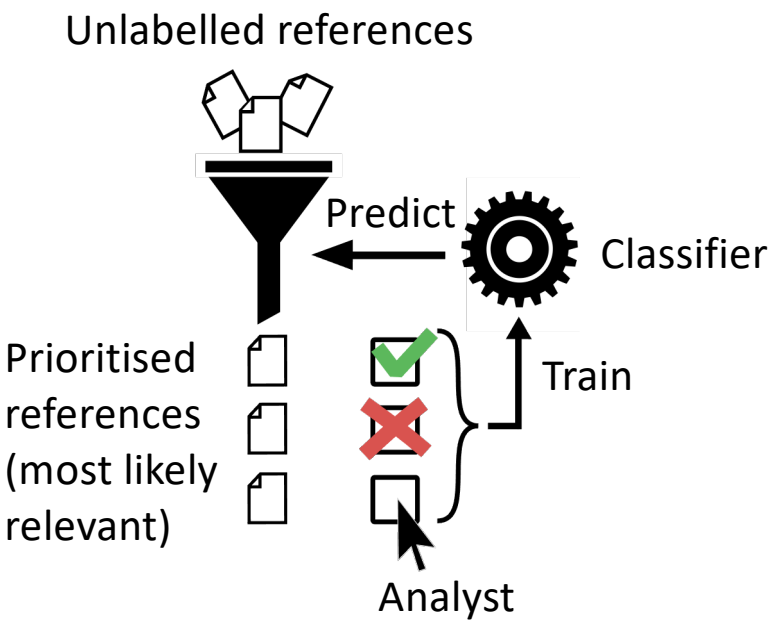


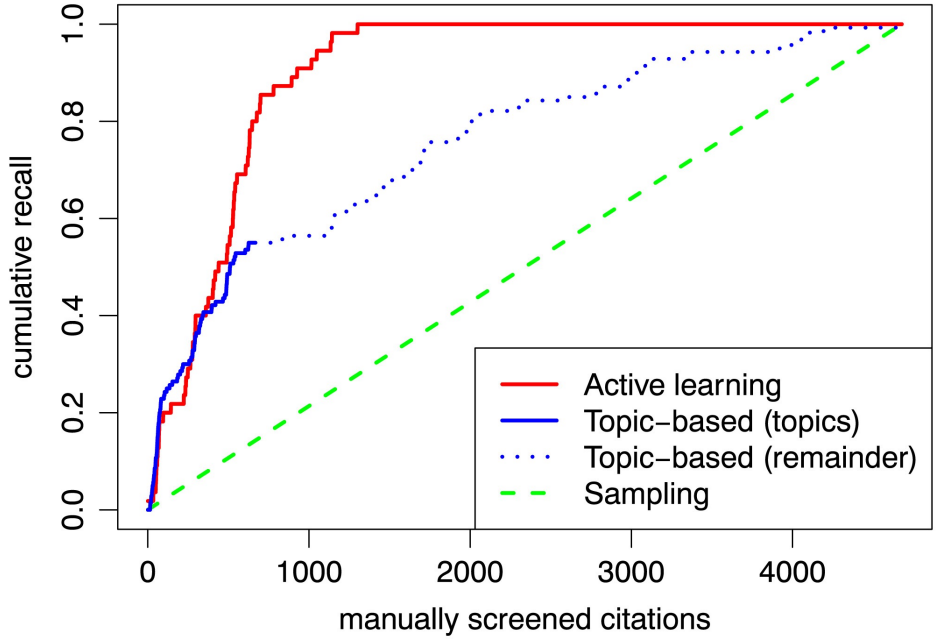
Active Learning



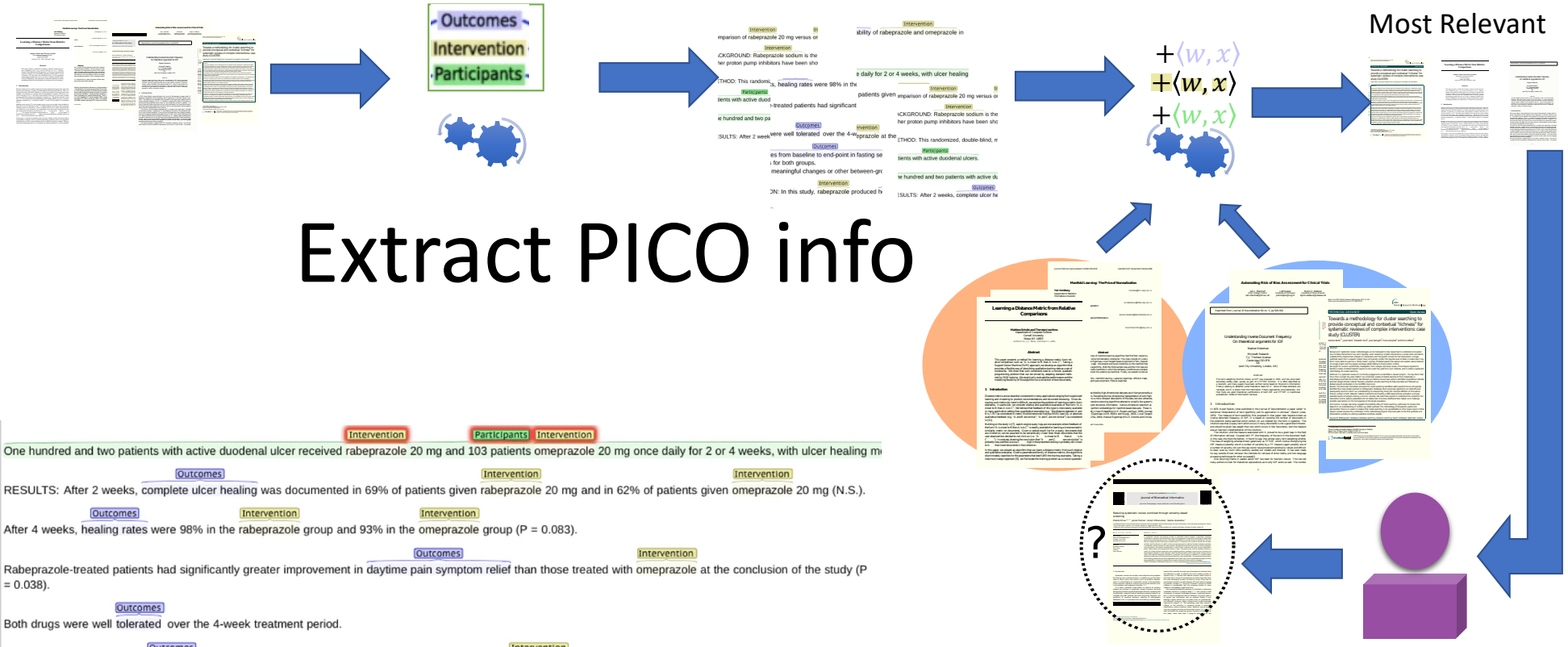
Prioritising references for systematic reviews with RobotAnalyst: A user study

Identify the majority of relevant references while only manually screening a portion





Research Synthesis Methods, Volume: 9, Issue: 3, Pages: 470-488, First published: 28 June 2018, DOI: (10.1002/jrsm.1311)



Extract PICO info

Intervention | Participants | Intervention

5 One hundred and two patients with active duodenal ulcer received rabeprazole 20 mg and 103 patients omeprazole 20 mg once daily for 2 or 4 weeks, with ulcer healing m

Outcomes | Intervention | Intervention

6 RESULTS: After 2 weeks, complete ulcer healing was documented in 69% of patients given rabeprazole 20 mg and in 62% of patients given omeprazole 20 mg (N.S.).

Outcomes | Intervention | Intervention

7 After 4 weeks, healing rates were 98% in the rabeprazole group and 93% in the omeprazole group (P = 0.083).

Outcomes | Intervention

8 Rabeprazole-treated patients had significantly greater improvement in daytime pain symptom relief than those treated with omeprazole at the conclusion of the study (P = 0.038).

Outcomes

9 Both drugs were well tolerated over the 4-week treatment period.

Outcomes | Intervention

10 Mean changes from baseline to end-point in fasting serum gastrin were significantly greater in the rabeprazole group, but at end-point mean values were well within normal limits for both groups.

11 No clinically meaningful changes or other between-group differences were observed in laboratory parameters.

Intervention | Outcomes | Intervention

12 CONCLUSION: In this study, rabeprazole produced healing rates equivalent to omeprazole at weeks 2 and 4, and provided significantly greater improvement in

Outcomes

daytime pain.

Outcomes

13 Both treatments were well tolerated.

Brockmeier *et al.* "Improving reference prioritisation with PICO recognition". *BMC Med Inform Decis Mak* **19,** 256 (2019). https://doi.org/10.1186/s12911-019-0992-8

When human experts fail

- Biased
  - Becomes rare with enough **training**
- Ambiguous cases
  - Need second opinion or more data
- Out of distribution
  - Outside of expertise

When statistical models fail

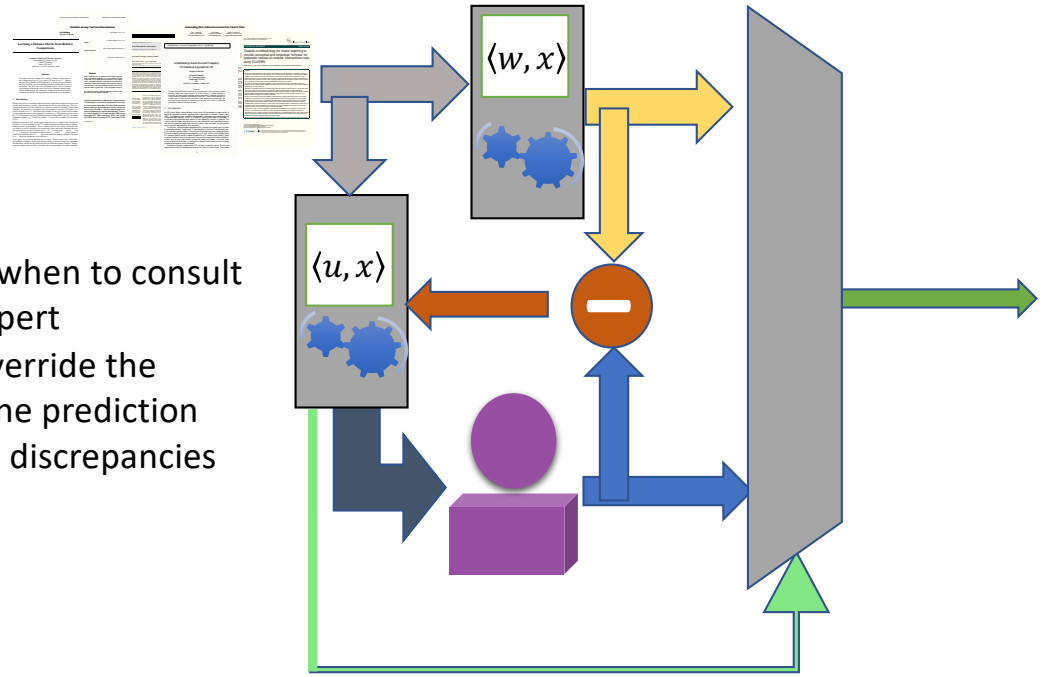
- Systematic error
  - Becomes rare with enough **correct** training data
- Ambiguous cases
  - Expert also needs to be careful
- Out of distribution
  - Corrupted or unseen case
  - Expert can easily recognize

Human-machine systems fail when

- they propagate/exacerbate biases
  - machine as a productivity multiplier
- excessive trust is given to machine
  - loss of vigilance
- insufficient data is collected to improve future versions
  - acquire more unbiased labeled data to validate

Need intelligent infrastructure for Artificial Intelligence

- Improve synergistic performance of the machine and human experts
- maintain vigilance and even heighten the performance of groups of experts
- collect training data for the future model instances



Learn when to consult the expert and override the machine prediction due to discrepancies