How to Read Less:

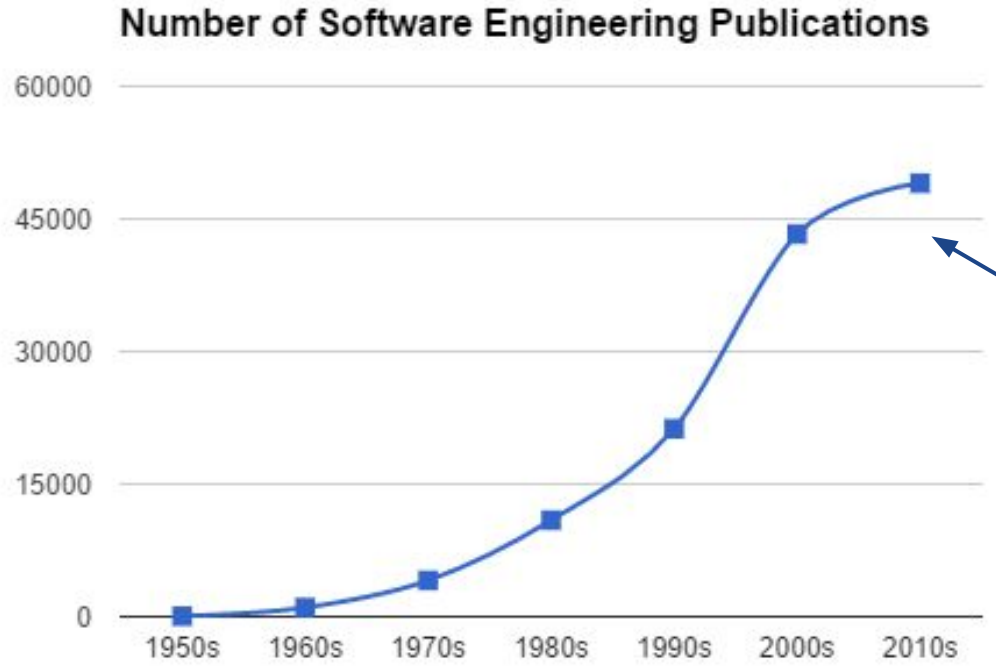
On the Benefit of Human-in-the-loop Incremental Learning for Systematic Literature Reviews

Zhe Yu
zyu9@ncsu.edu

Ph.D. Written Preliminary Exam
April 2017

Committee:
Dr. Tim Menzies (advisor)
Dr. Xipeng Shen
Dr. William Enck

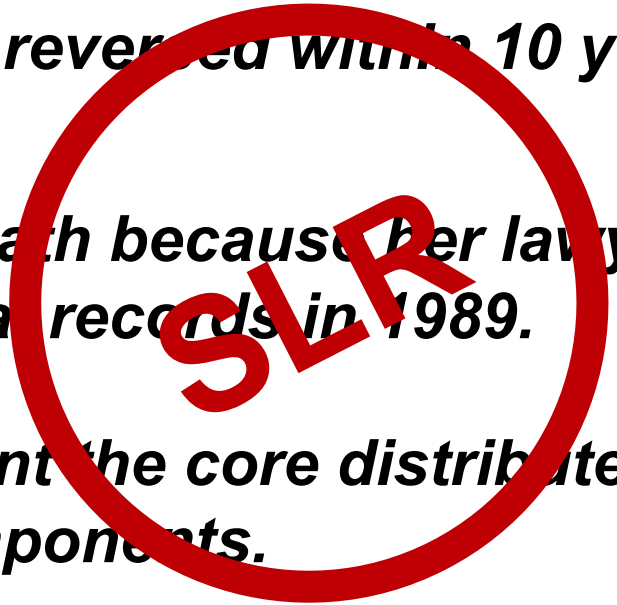
Too Much to Read



6653 new publications in 2016

Not Reading Enough

- *146 medical practises were reversed within 10 years.*
- *A lady was sentenced to death because her lawyer failed to find related hospital records in 1989.*
- *Software developers reinvent the core distributed software concepts and components.*



Prasad, Vinay, et al. "A decade of reversal: an analysis of 146 contradicted medical practices." Mayo Clinic Proceedings. Vol. 88. No. 8. Elsevier, 2013.
Bright, Stephen B. "Counsel for the poor: The death sentence not for the worst crime but for the worst lawyer." The Yale Law Journal 103.7 (1994): 1835-1883.
Fayad, Mohamed E., et al. "Lessons learned building reusable OO frameworks for distributed software." Communications of the ACM 40.10 (1997): 85-87.

Done

- *Apply human-in-the-loop method to facilitate SLRs*
- *FASTREAD, a better method than state-of-the-art*
- *Save 90% cost to retrieve 90% relevant studies*
- *A tool to implement FASTREAD*

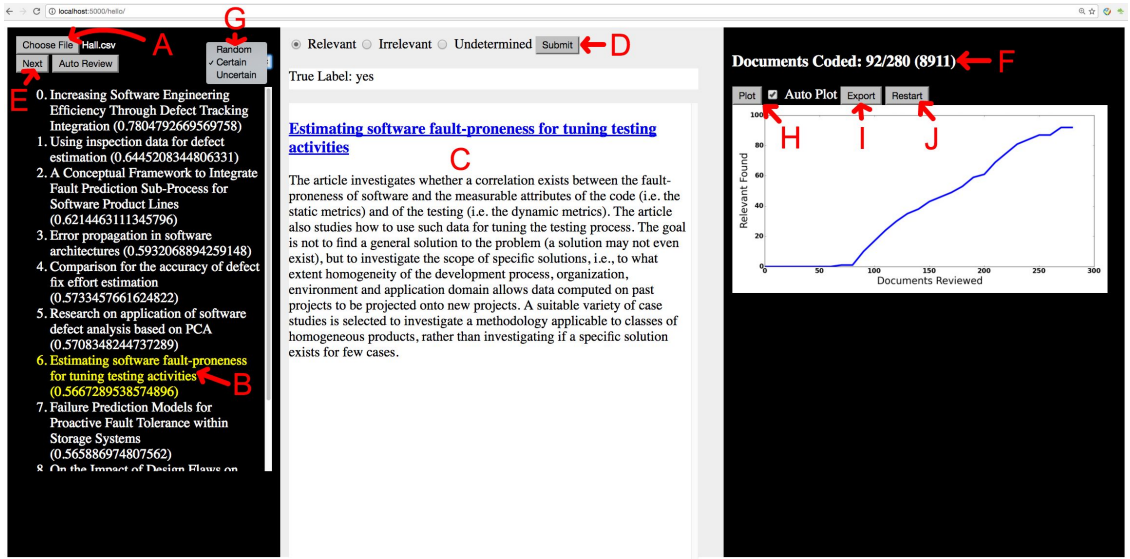
Opensource

Paper:

- Submitted to EMSE
- arxiv

Tool:

- SeaCraft Zenodo
- Github



Todo

- ***More data***
- ***More detailed, real data***
- ***Massive data***

Outline

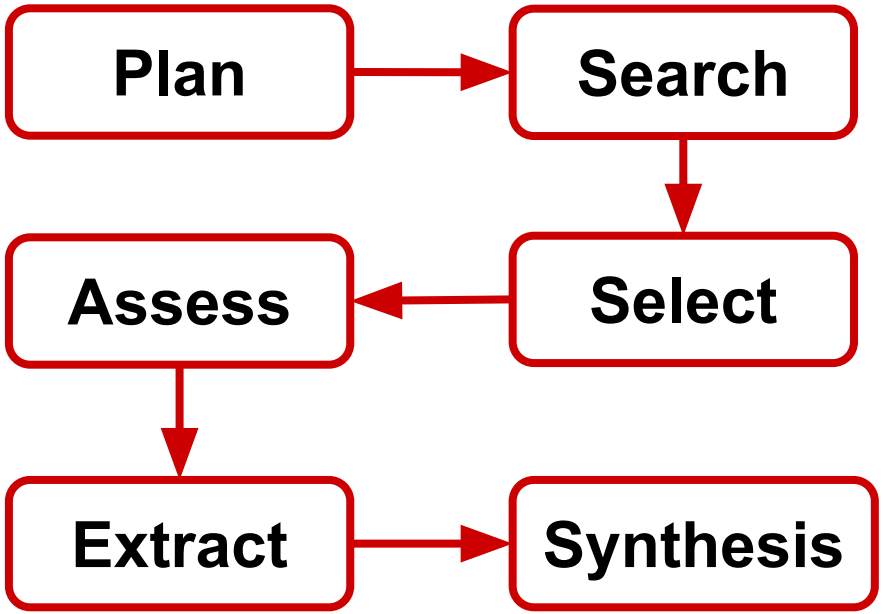
- ***Background***
- ***Method***
- ***Experiment***
- ***Result***
- ***Conclusion and Future works***

Outline

- ***Background***
  - Systematic Literature Review
  - Primary Study Selection (Select)
- ***Method***
- ***Experiment***
- ***Result***
- ***Conclusion and Future works***

Systematic Literature Review (SLR)

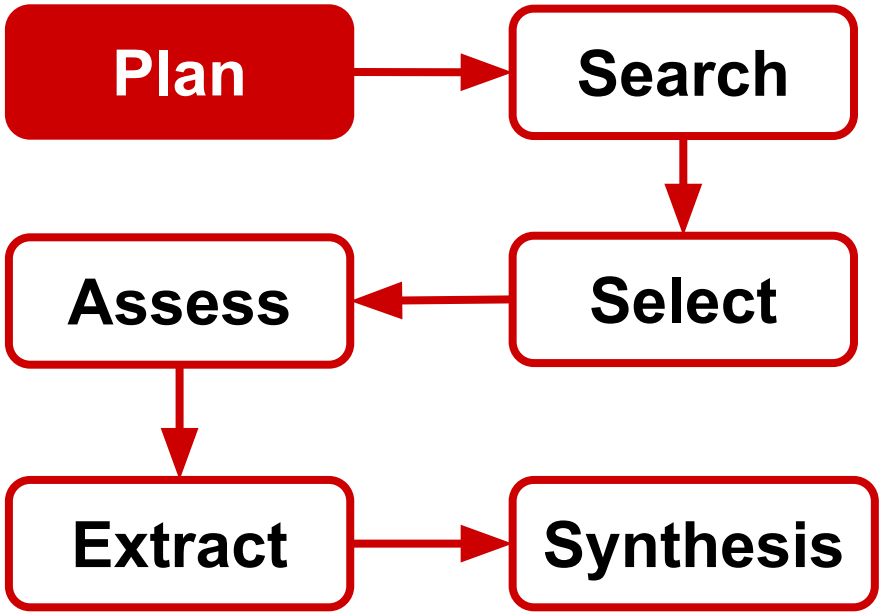
SLR: A systematic guide to review literature



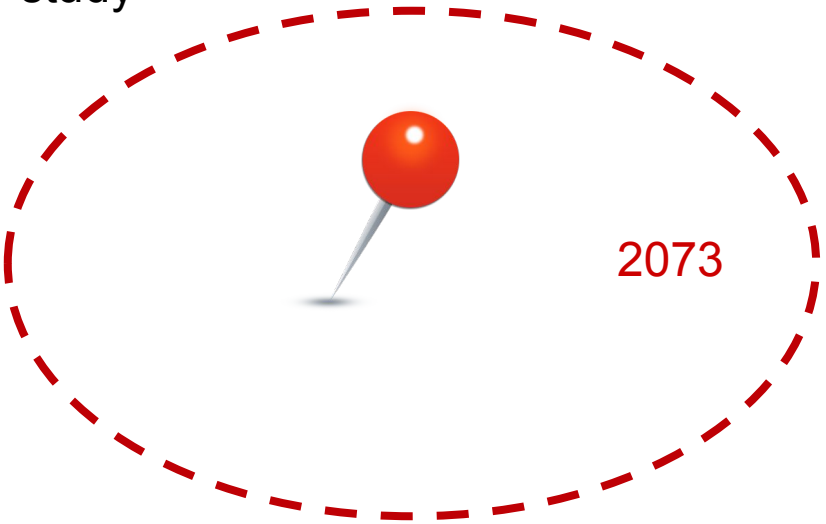
Kitchenham, Barbara A., Tore Dyba, and Magne Jorgensen. "Evidence-based software engineering." ICSE'04.

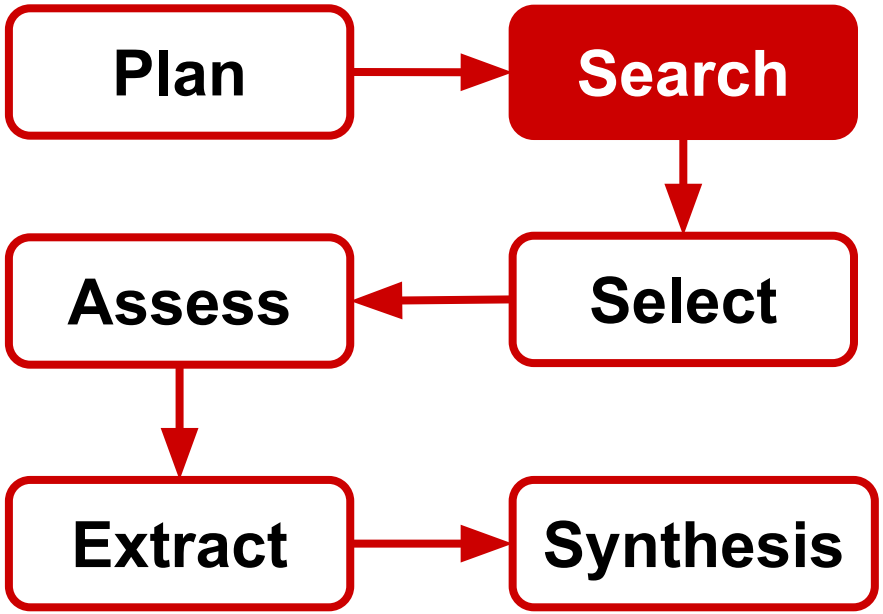
Defect Prediction Papers, from 2000 to 2010 **[Hall'12]**

- Context
- Feature
- Modeling



Search:
(Fault* OR bug* OR defect* OR errors OR corrections OR corrective OR fix*) in title only AND (Software) anywhere in study

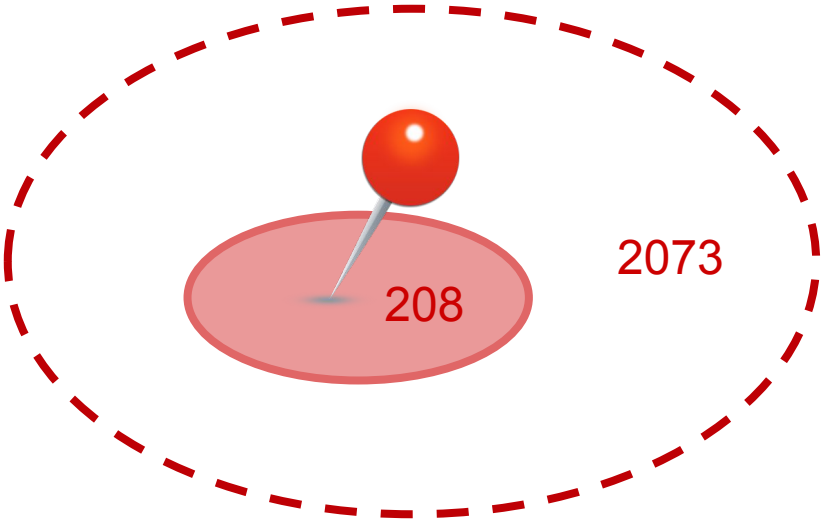


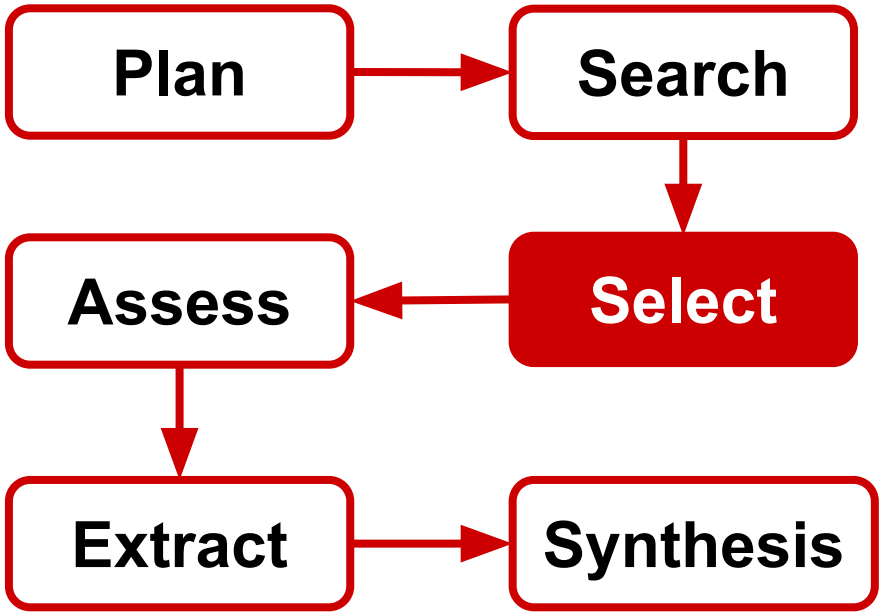


Hall, Tracy, et al. "A systematic literature review on fault prediction performance in software engineering." TSE'12.

Exclude:

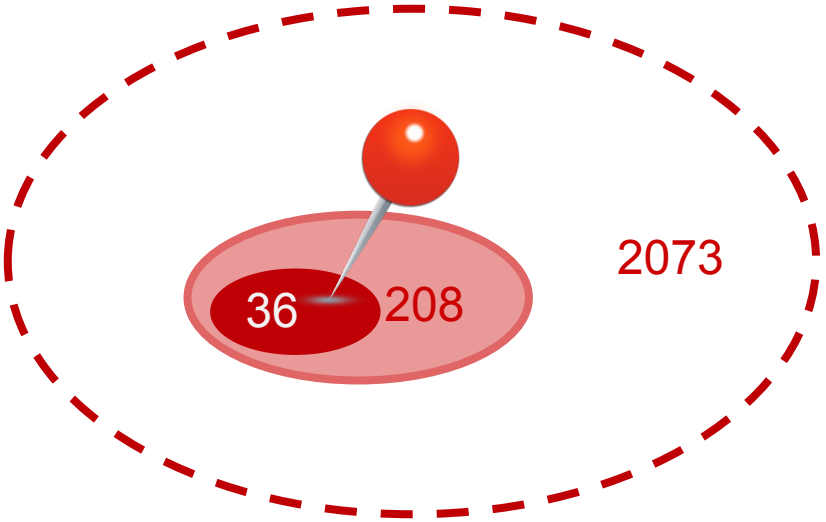
- not an empirical study
- not in software systems
- main output is not faults in code

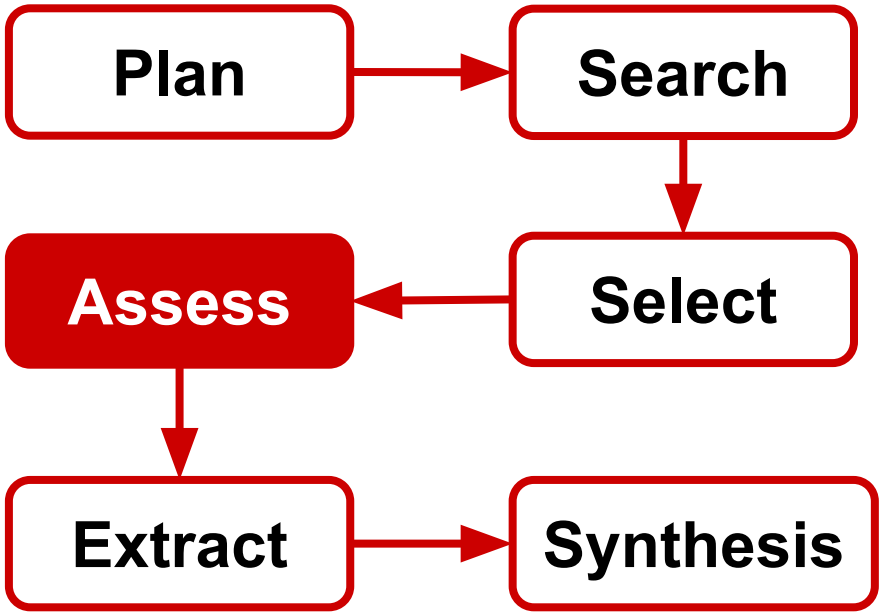




Assess:

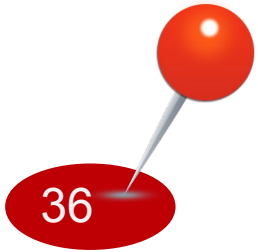
- is a prediction study
- sufficient contextual information
- sufficient model building information
- has data

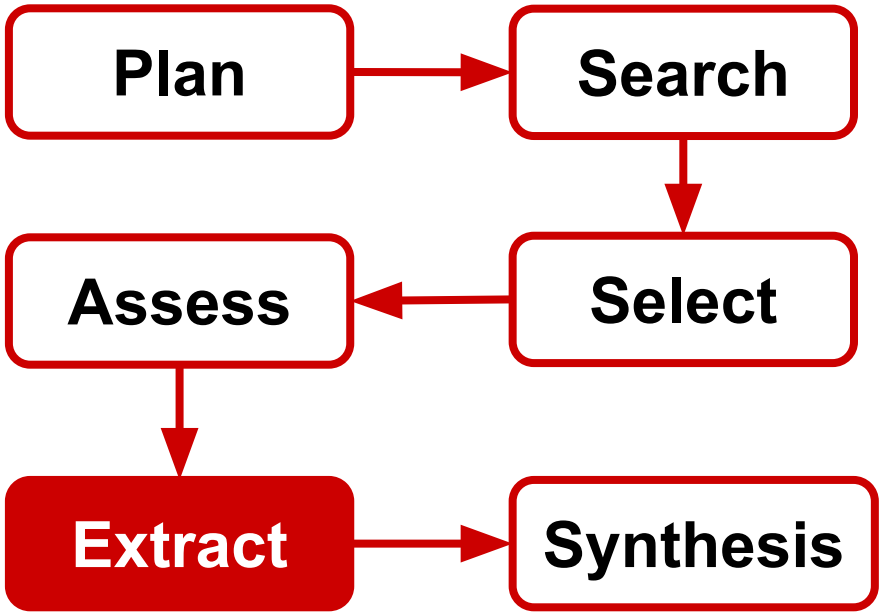




Extract:

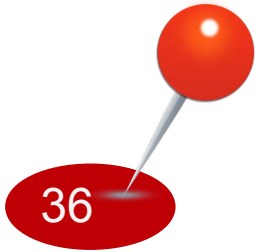
- Context data
- Qualitative data
- Quantitative data

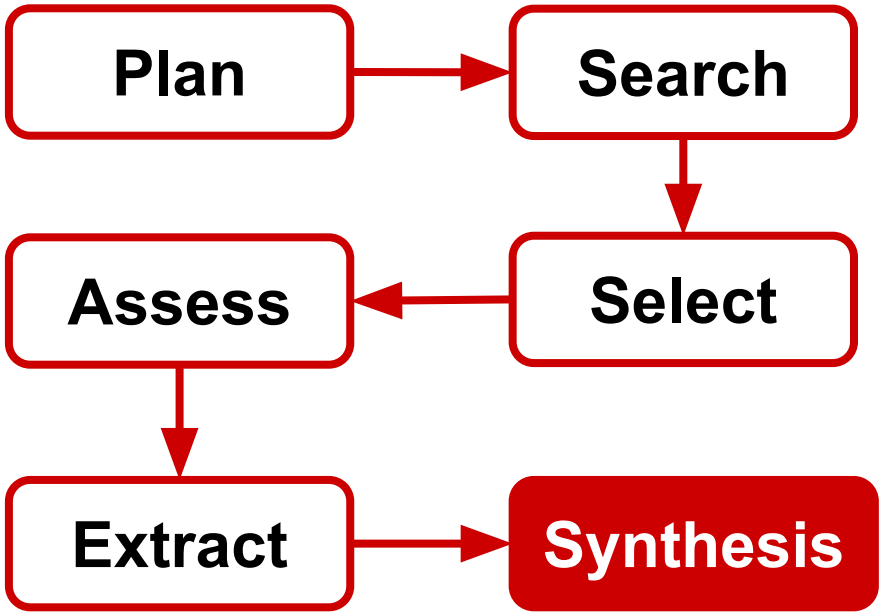




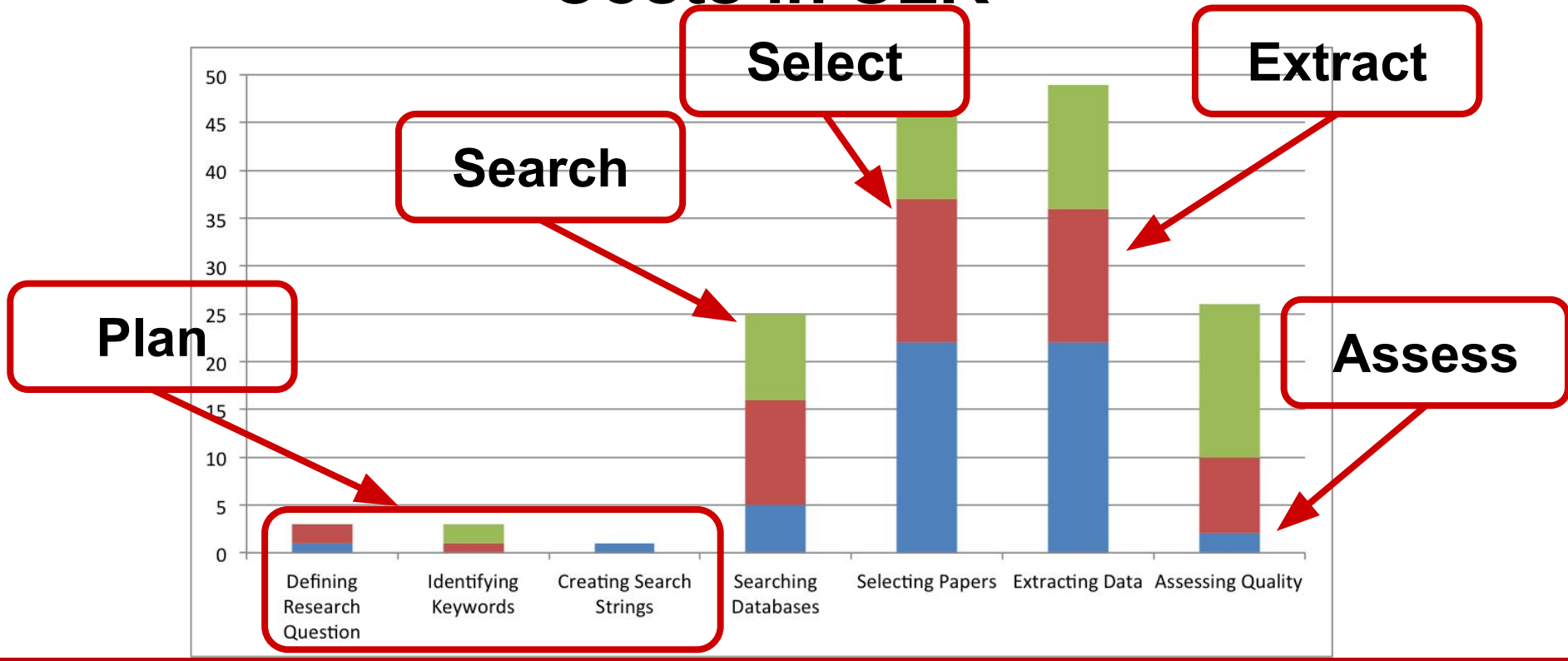
Synthesis:

- **What features are used most frequently?** LOC, OO, etc.
- **Which model works best?** Naive Bayes, Decision Tree, Logistic Regression, etc.



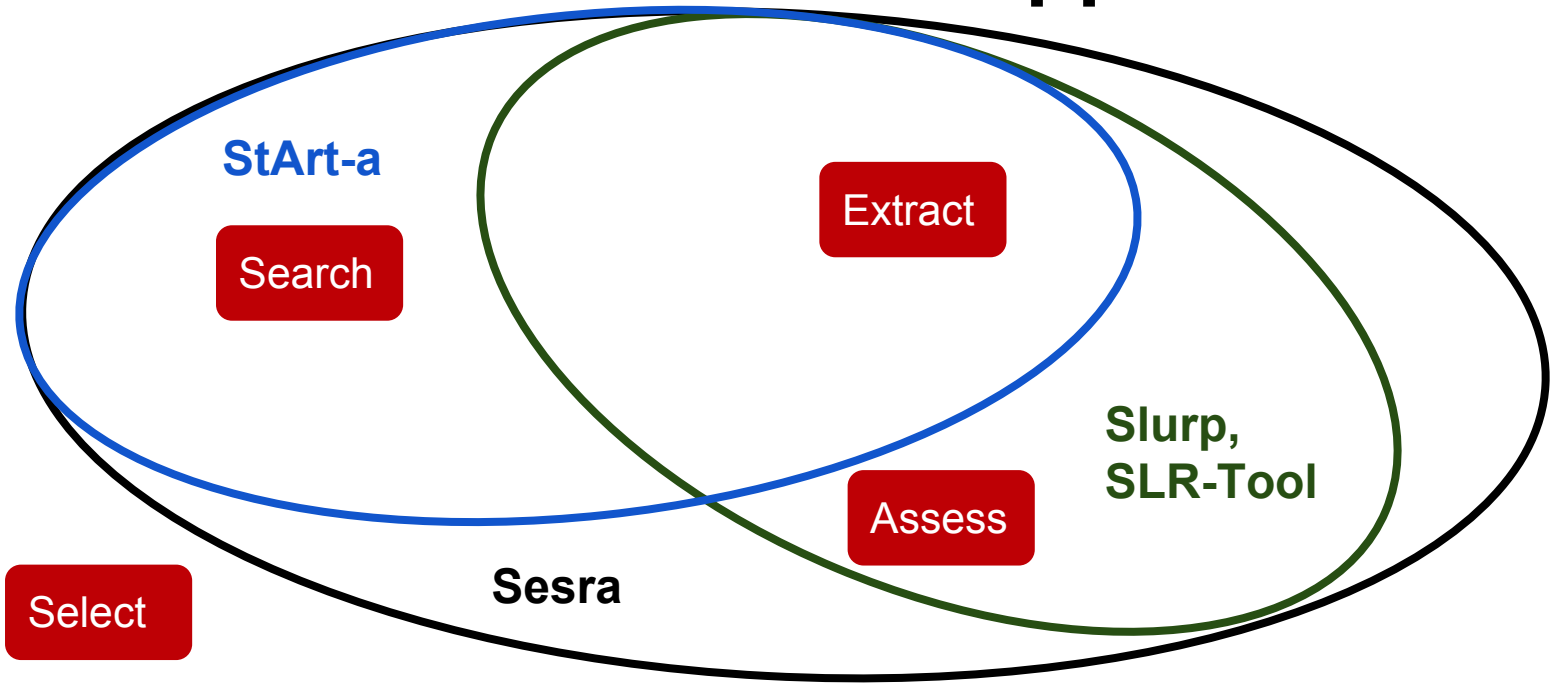


Costs in SLR



Carver, Jeffrey C., et al. "Identifying barriers to the systematic literature review process." ESEM'13.

Automated Tool Support



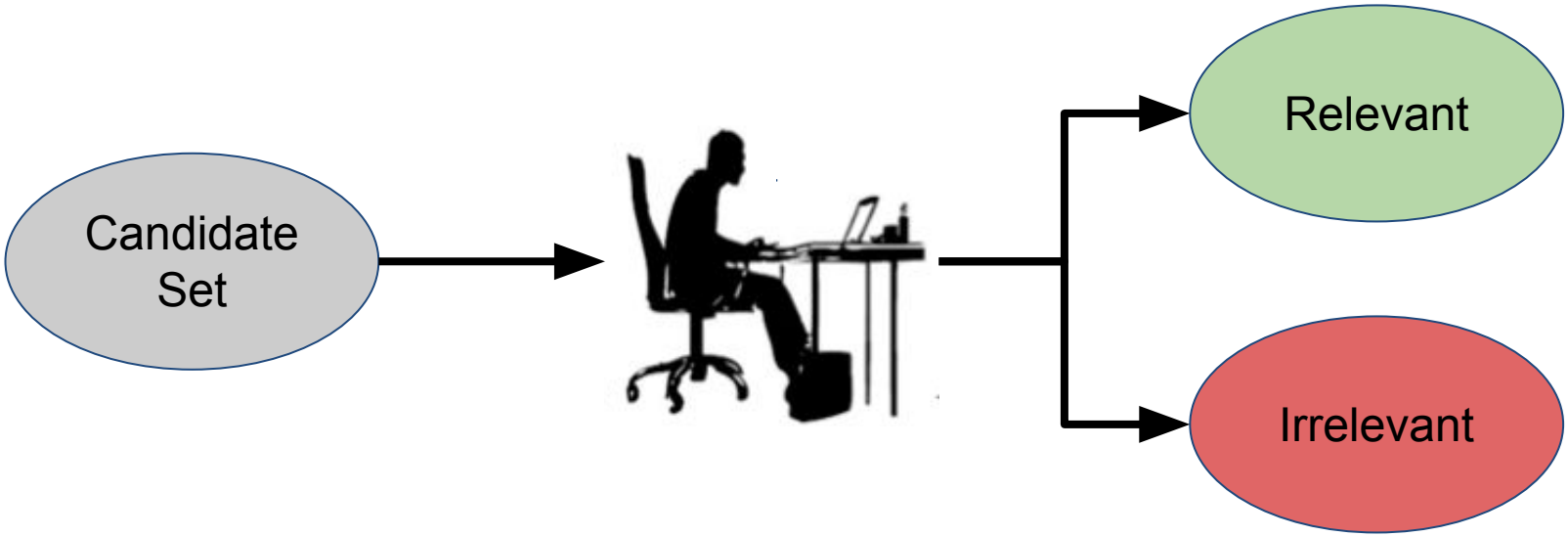
Human-in-the-loop, VTM, Snowballing

Marshall, Christopher, et al. "Tools to support systematic reviews in software engineering: a cross-domain survey using semi-structured interviews." EASE'15.

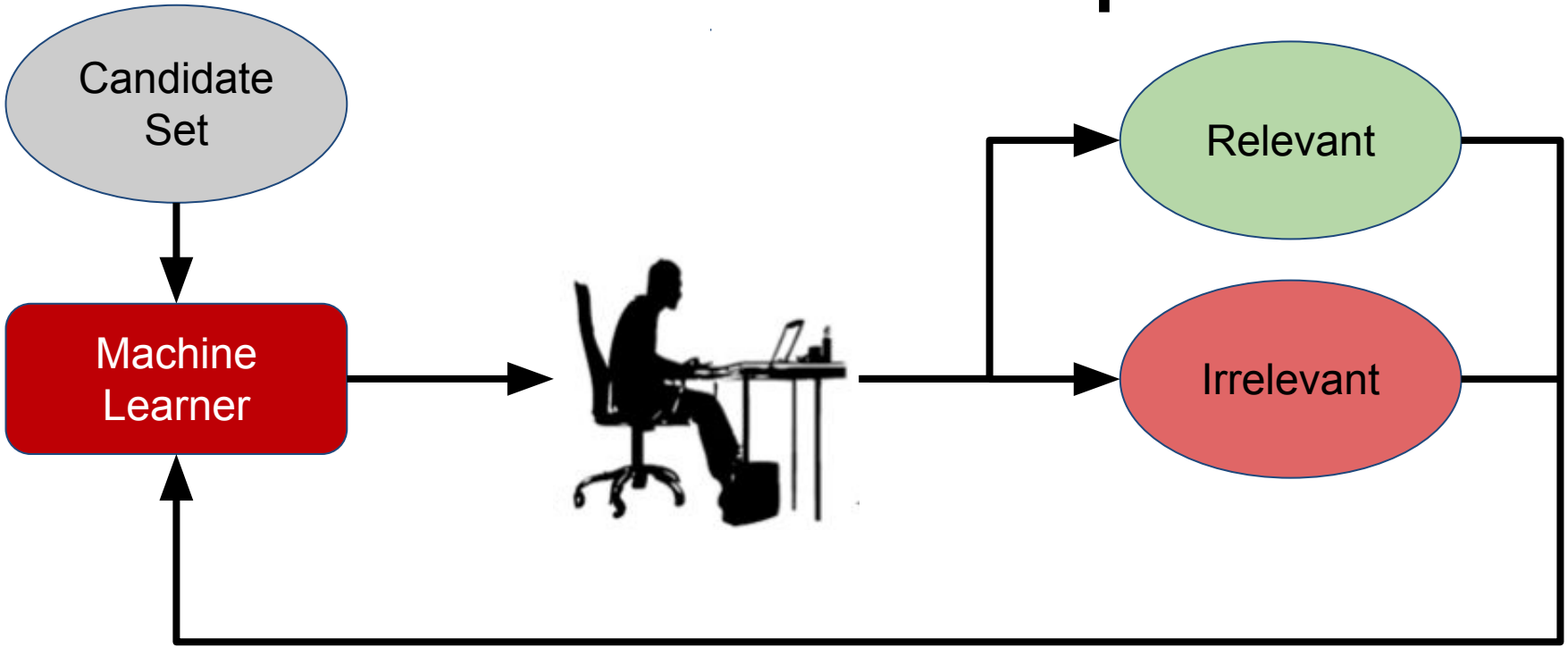
Outline

- ***Background***
  - Systematic Literature Review
  - Primary Study Selection (Select)
- ***Method***
- ***Experiment***
- ***Result***
- ***Conclusion and Future works***

Select: Linear Review

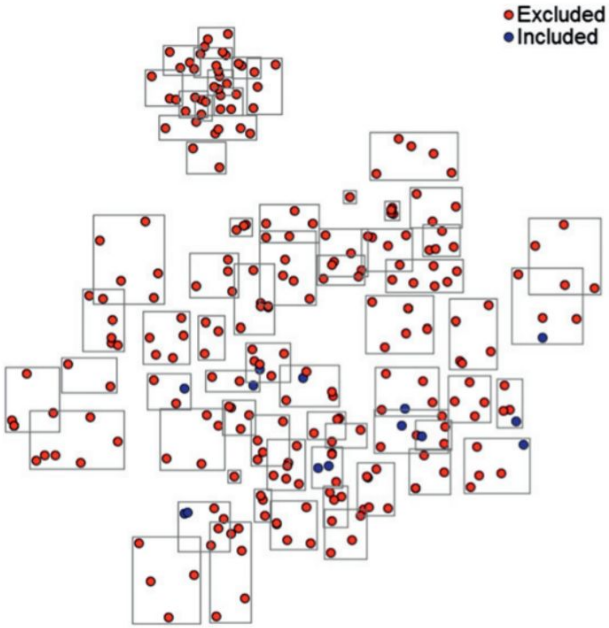


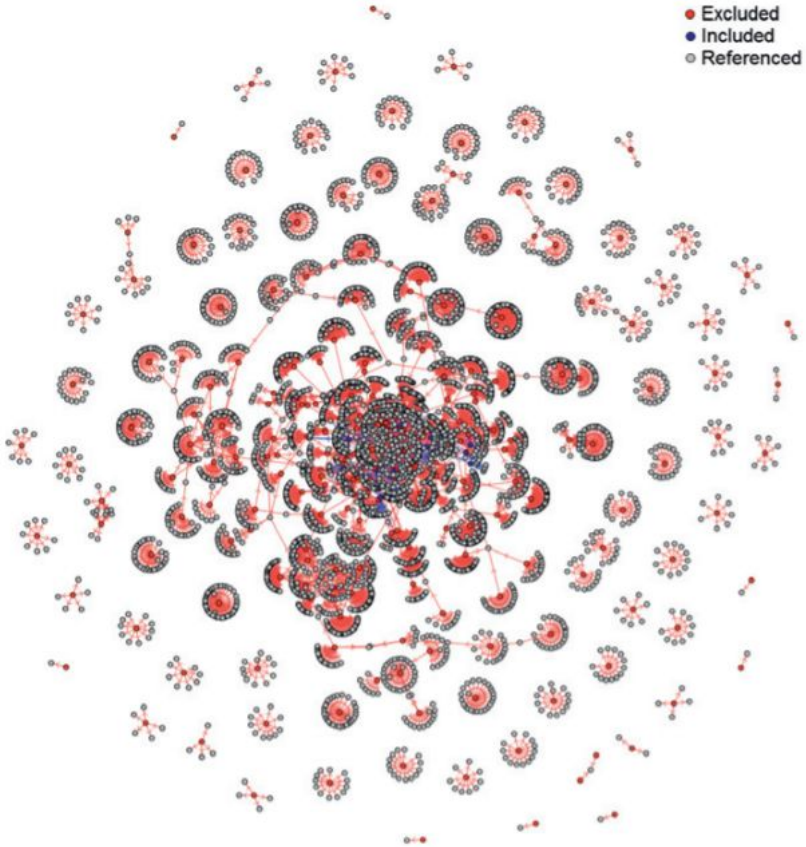
Human-in-the-loop



Sheng, Provost, Ipeirotis. 2008. “Get another label? improving data quality and data mining using multiple, noisy labelers”, KDD ‘08
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR’14.

Visual Text Mining





Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.

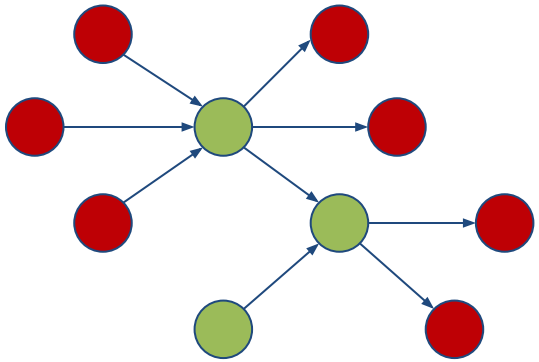
Snowballing

Forward:

Papers cite the known one.

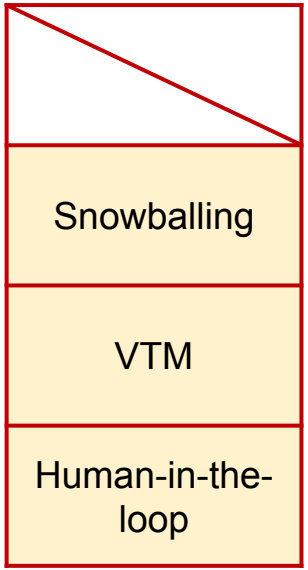
Backward:

The known one's references



Kitchenham, Barbara A., Tore Dyba, and Magne Jorgensen. "Evidence-based software engineering." ICSE'04.
Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.

Techniques to Select



Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.
Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Techniques to Select

	Machine Learning?
Snowballing	no
VTM	Unsupervised Learning
Human-in-the-loop	Active Learning

Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.
Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Techniques to Select

	Machine Learning?	Require initial papers?
Snowballing	no	yes
VTM	Unsupervised Learning	no
Human-in-the-loop	Active Learning	no

Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.
Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Techniques to Select

	Machine Learning?	Require initial papers?	Easy to Validate?
Snowballing	no	yes	yes
VTM	Unsupervised Learning	no	no
Human-in-the-loop	Active Learning	no	yes

Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.
Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Techniques to Select

	Machine Learning?	Require initial papers?	Easy to Validate?	Scale up?
Snowballing	no	yes	yes	?
VTM	Unsupervised Learning	no	no	?
Human-in-the-loop	Active Learning	no	yes	yes

Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.
Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Techniques to Select

	Machine Learning?	Require initial papers?	Easy to Validate?	Scale up?	Tested on SE data?
Snowballing	no	yes	yes	?	yes
VTM	Unsupervised Learning	no	no	?	yes
Human-in-the-loop	Active Learning	no	yes	yes	no

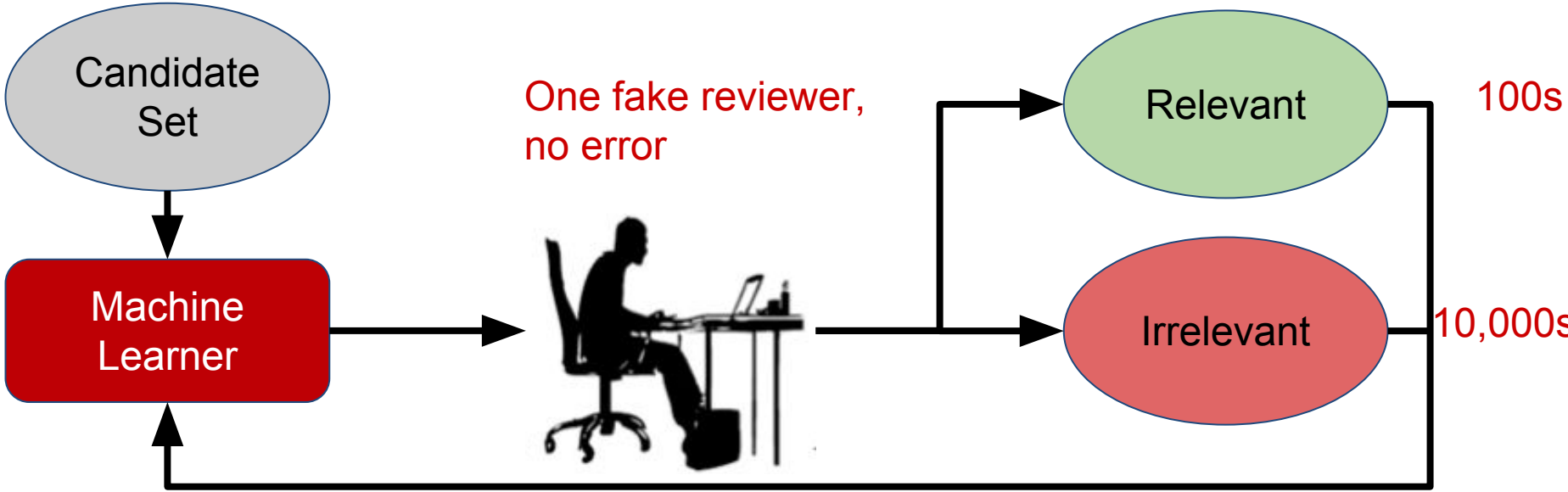
Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." EASE'14.
Felizardo, Katia R., et al. " visual analysis approach to validate the selection review of primary studies" Information and Software Technology 54.10 (2012): 1079-1091.
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Outline

- *Background*
- ***Method***
  - Problem Statement
  - State-of-the-art Methods
  - Algorithm Code
- *Experiment*
- *Result*
- *Conclusion and Future works*

Problem Statement

Title+Abstract, unlabeled, 10,000s



Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.
Hall, Tracy, et al. "A systematic literature review on fault prediction performance in software engineering." TSE'12.

Outline

- *Background*
- ***Method***
  - Problem Statement
  - State-of-the-art Methods
  - Algorithm Code
- *Experiment*
- *Result*
- *Conclusion and Future works*

State-of-the-art

Medicine:

[Wallace'10] *Core algorithm stays unchanged in subsequent works.*

Legal:

[Cormack'14] *Still state-of-the-art in legal domain.*

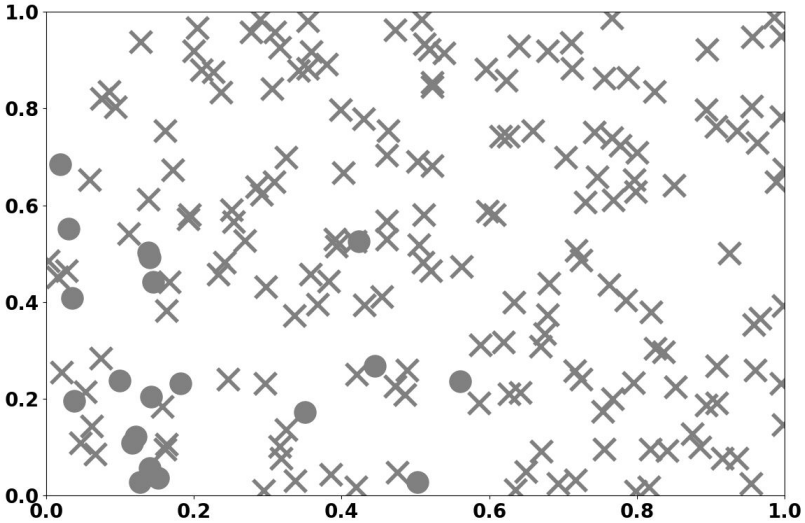
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Outline

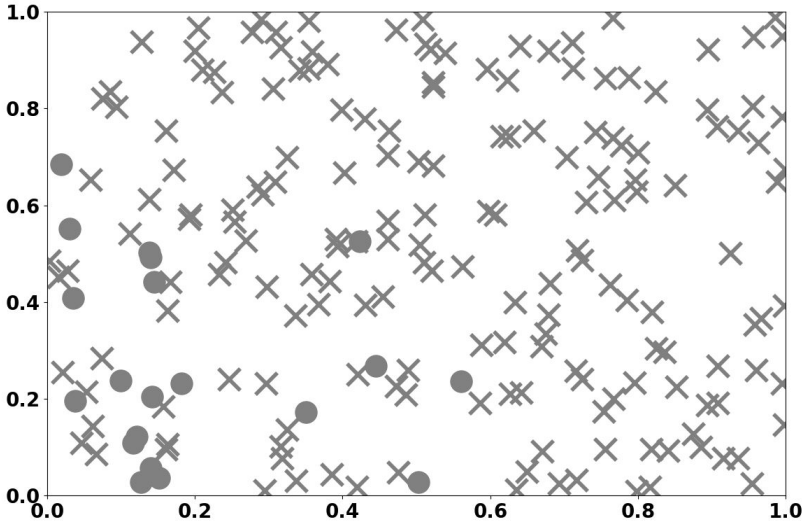
- *Motivation and Background*
- ***Method***
  - Problem Statement
  - State-of-the-art
  - Algorithm Code
- *Experiment*
- *Result*
- *Conclusion and Future works*

Initial Candidates

Wallace'10 (0/0)



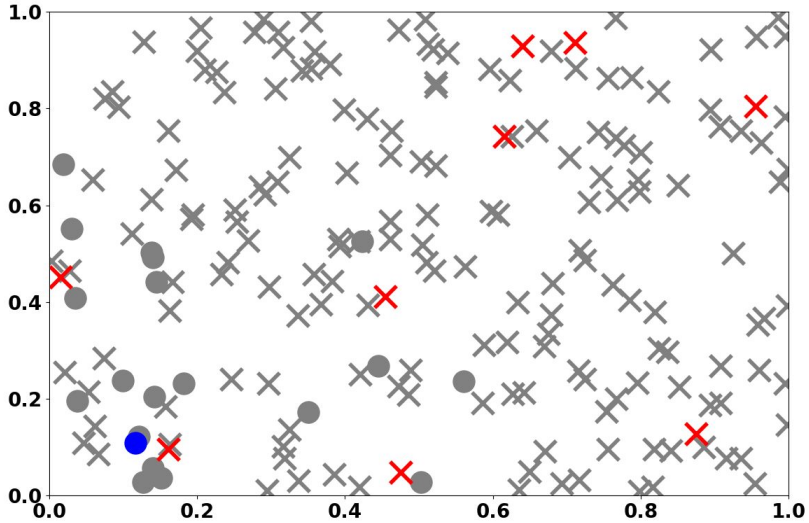
Cormack'14 (0/0)



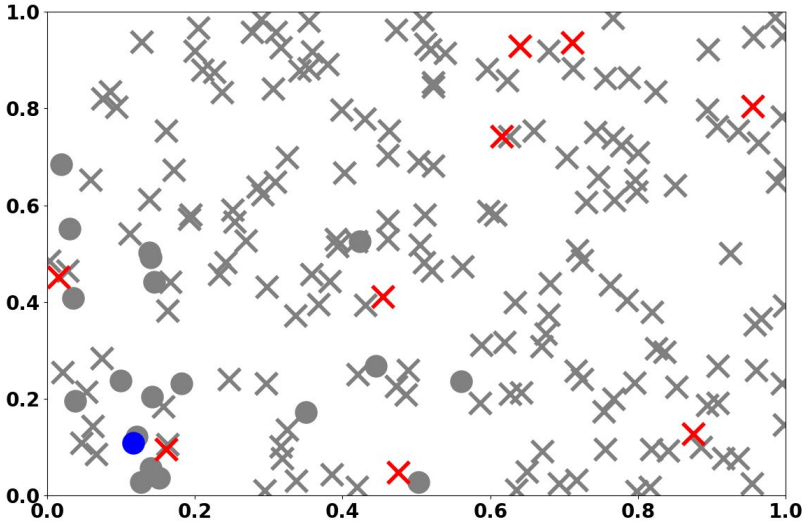
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Random Sampling

Wallace'10 (1/10)



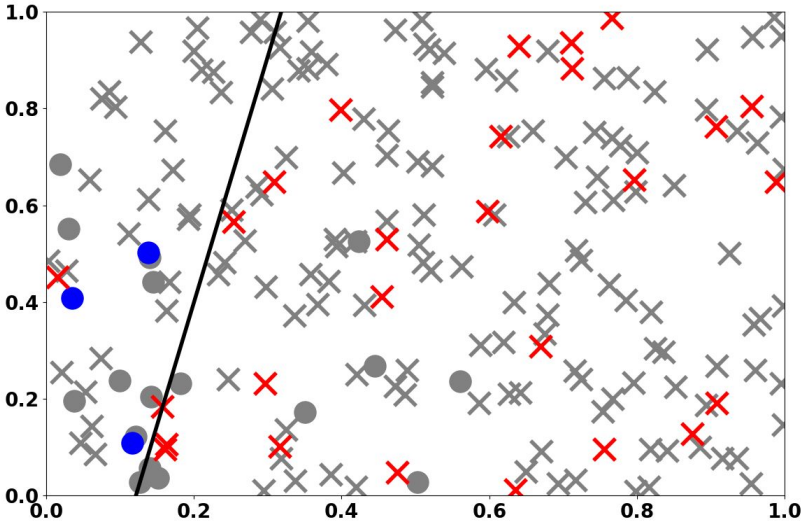
Cormack'14 (1/10)



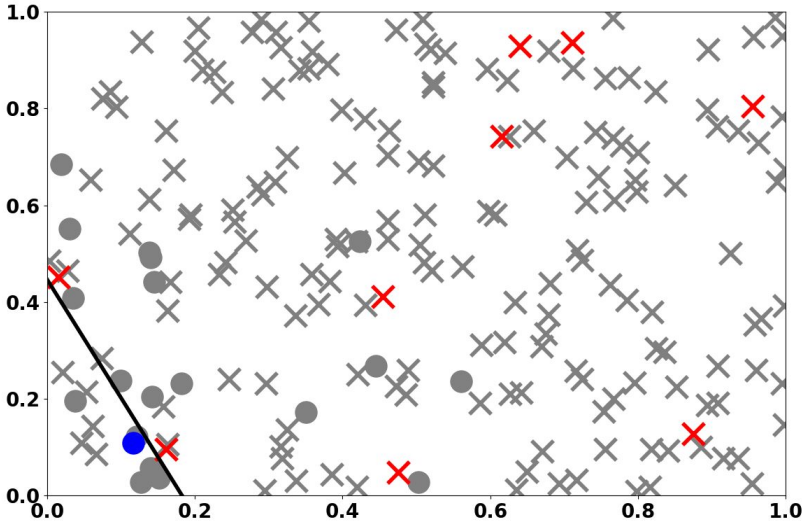
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Start Point

Wallace'10 (3/30)



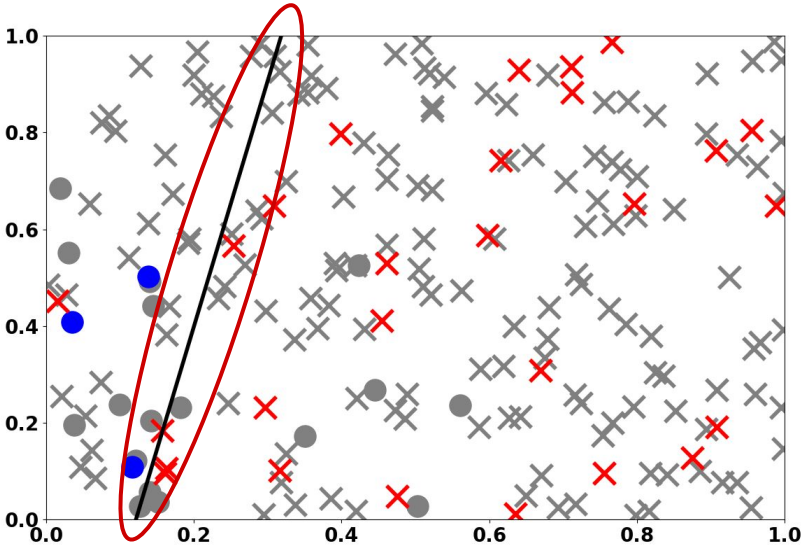
Cormack'14 (1/10)



Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

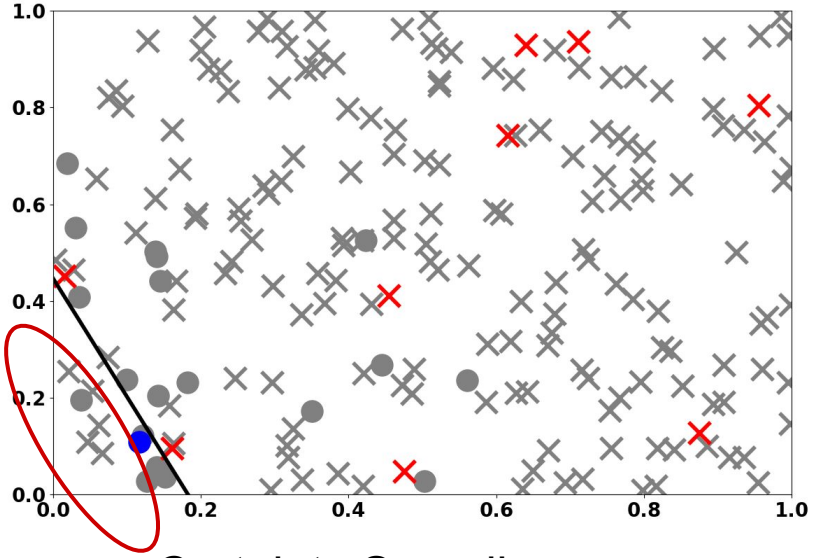
Query Strategy

Wallace'10 (3/30)



Uncertainty Sampling

Cormack'14 (1/10)

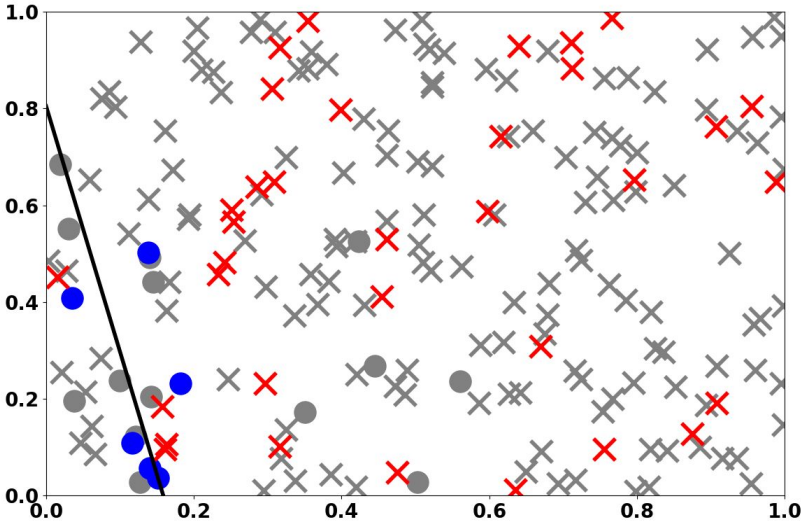


Certainty Sampling

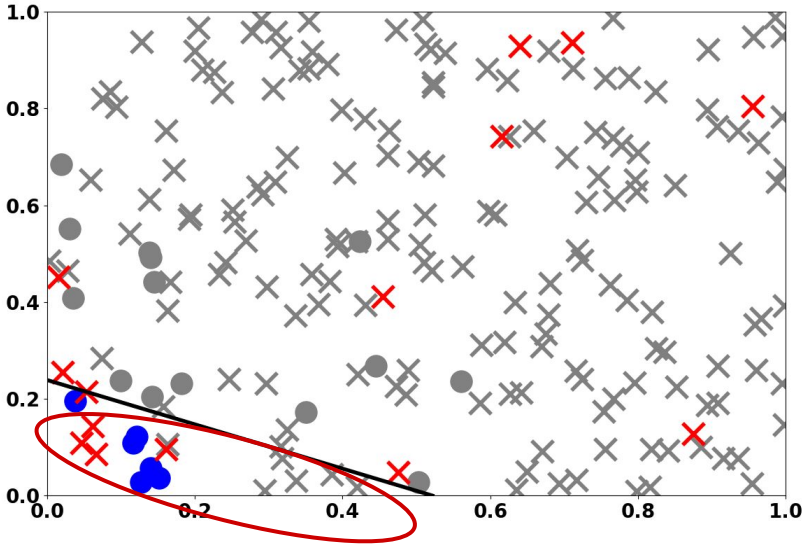
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Data Balancing

Wallace'10 (6/40)



Cormack'14 (6/20)

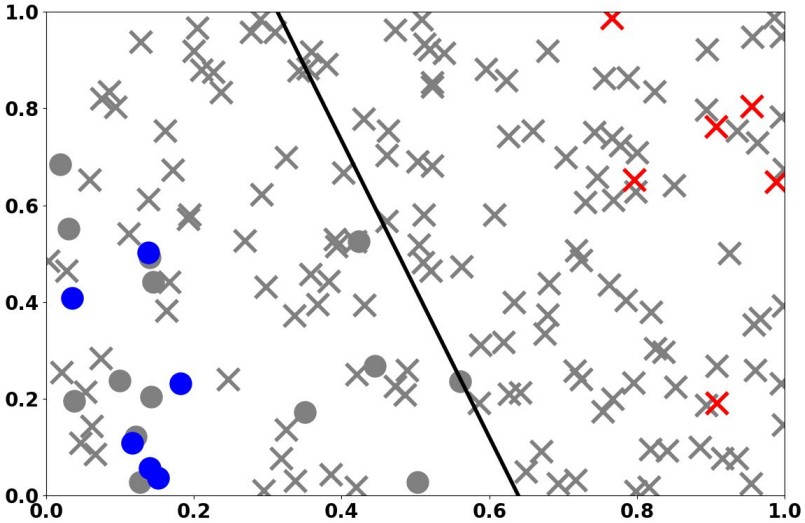


Certainty Sampling

Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

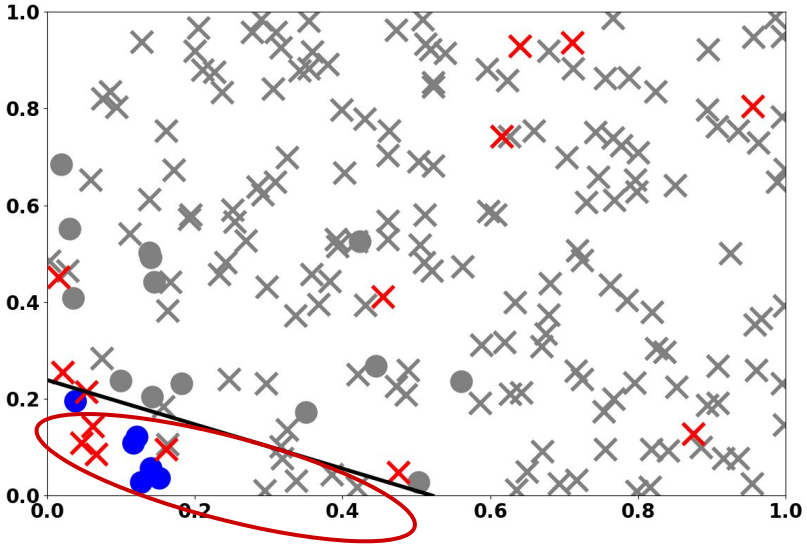
Data Balancing

Wallace'10 (6/40)



Aggressive Undersampling

Cormack'14 (6/20)

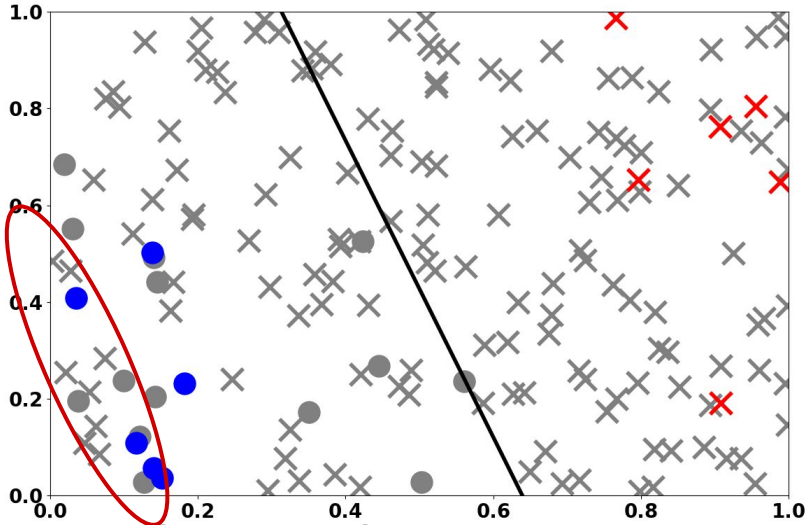


Certainty Sampling

Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

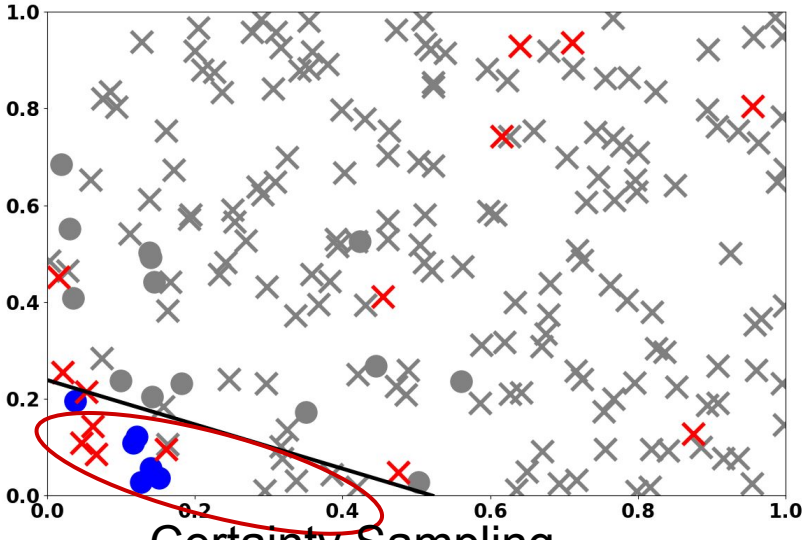
Continuity

Wallace'10 (6/40)



Certainty Sampling,
Stop updating model

Cormack'14 (6/20)



Certainty Sampling,
Keeps updating model

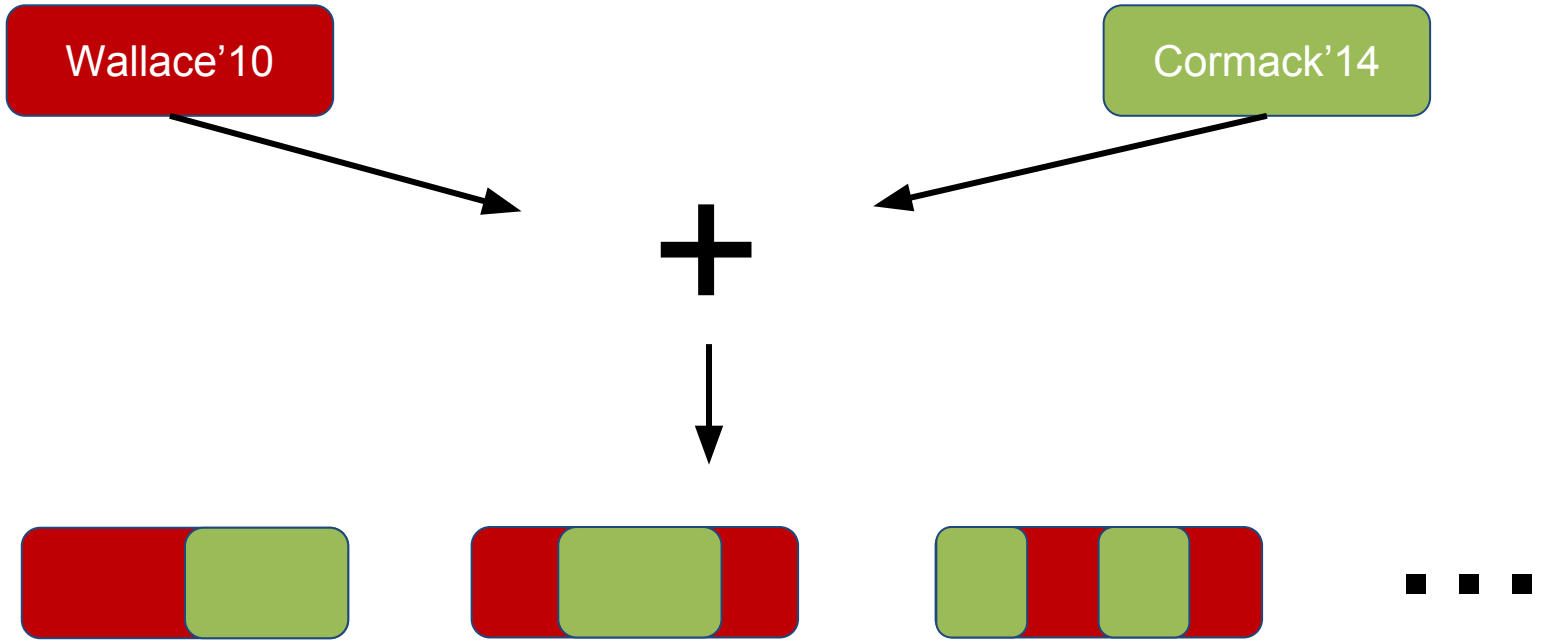
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Algorithm Code

	Start point	Query Strategy	Continuity	Data Balancing
Wallace'10	late (P)	uncertainty sampling (U)	stop training (S)	aggressive undersampling (A)
Cormack'14	early ($\bar{P}$)	certainty sampling ($\bar{U}$)	non-stop ($\bar{S}$)	none ($\bar{A}$)

Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Refactor

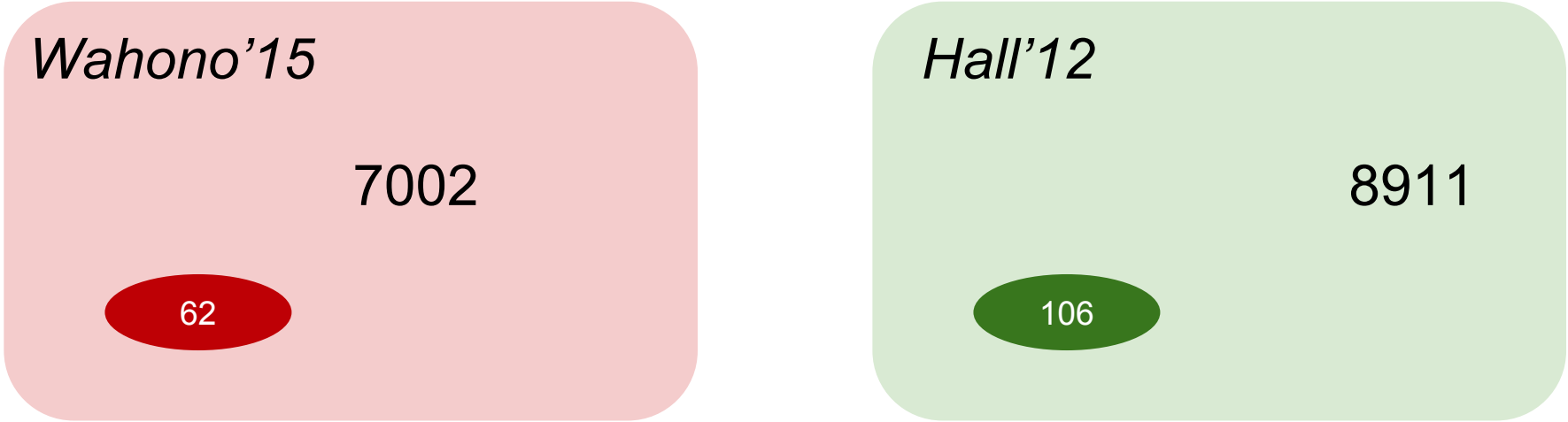


Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Outline

- *Background*
- *Method*
- ***Experiment***
  - Data
  - Framework
  - Evaluation
- *Result*
- *Conclusion and Future works*

Data Sets

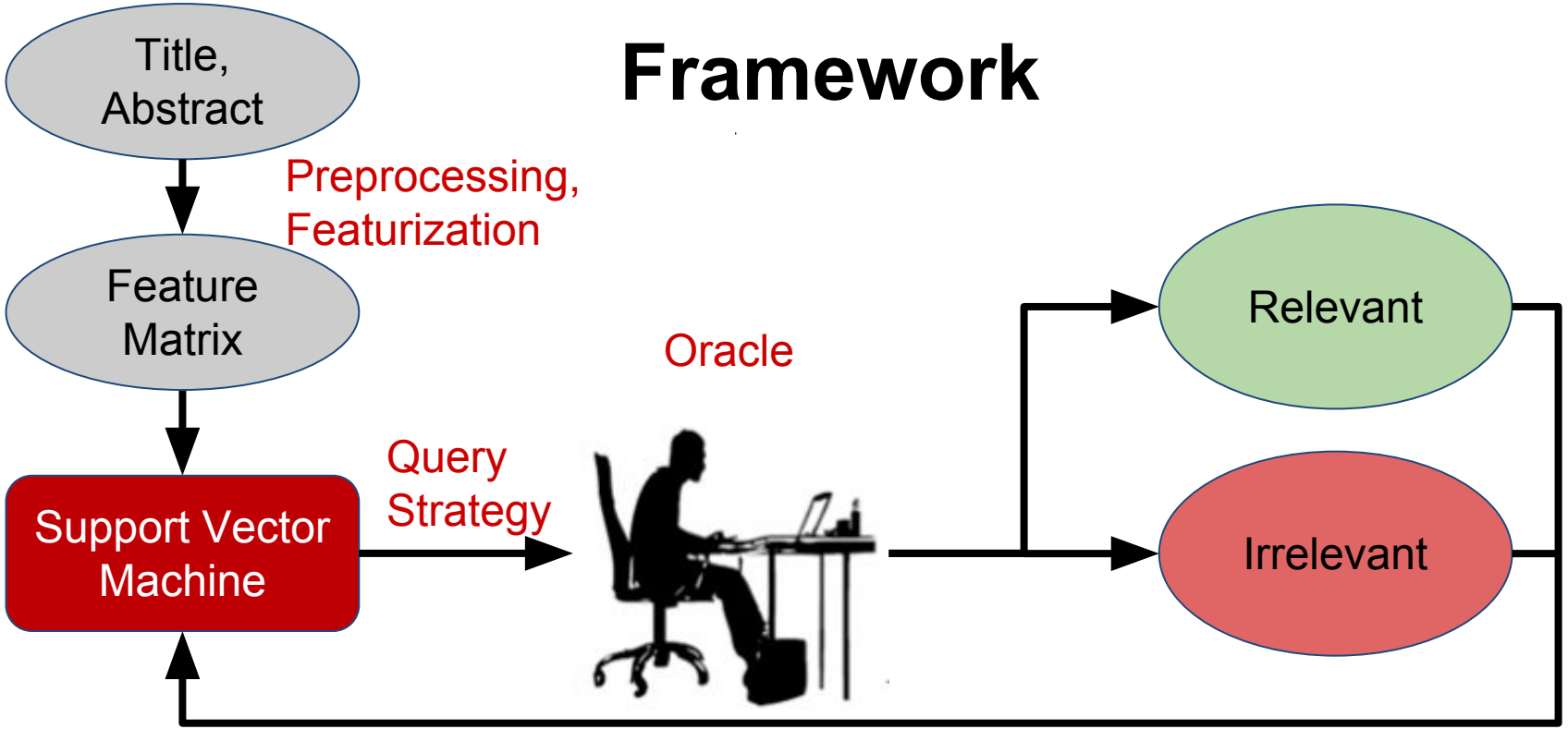


Wahono, Romi Satria. "A systematic literature review of software defect prediction: Research trends, datasets, methods and frameworks." JSE'15.
Hall, Tracy, et al. "A systematic literature review on fault prediction performance in software engineering." TSE'12.

Outline

- *Background*
- *Method*
- ***Experiment***
  - Data
  - Framework
  - Evaluation
- *Result*
- *Conclusion and Future works*

Framework

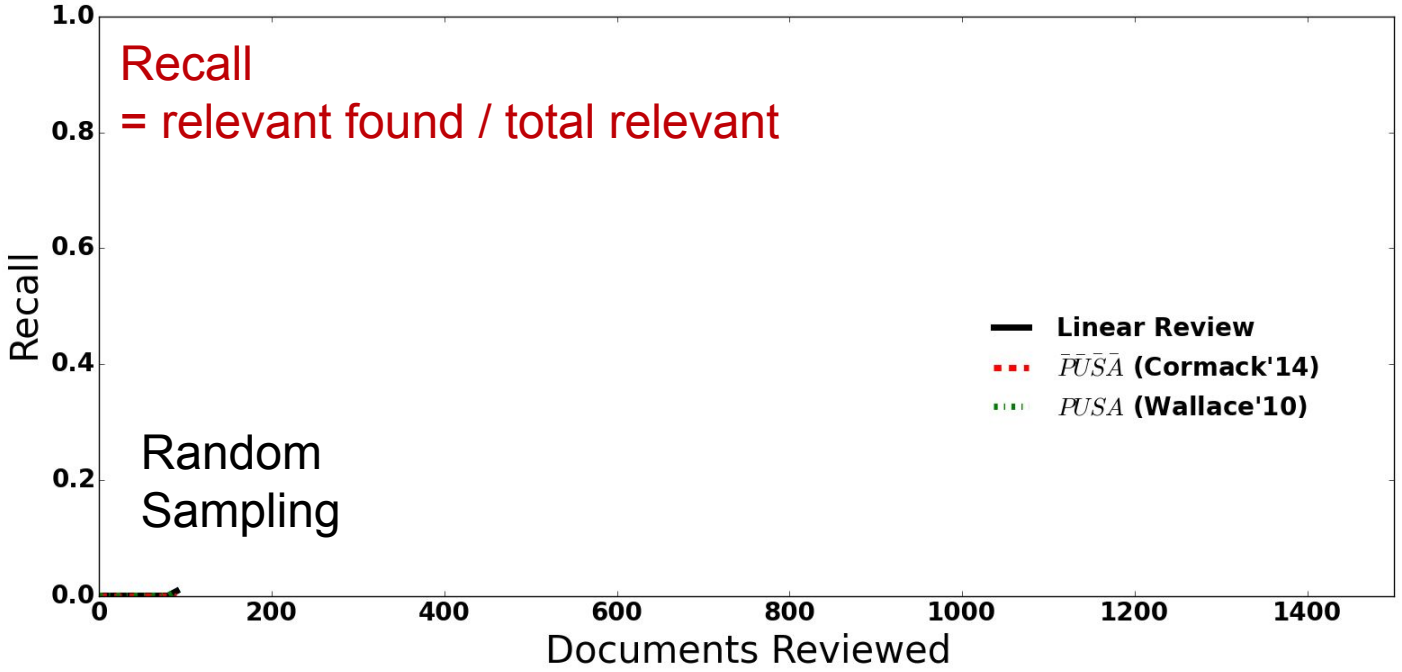


Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.
Krishna, Rahul, et al. "The BigSE project: lessons learned from validating industrial text mining." BIGDSE'16.

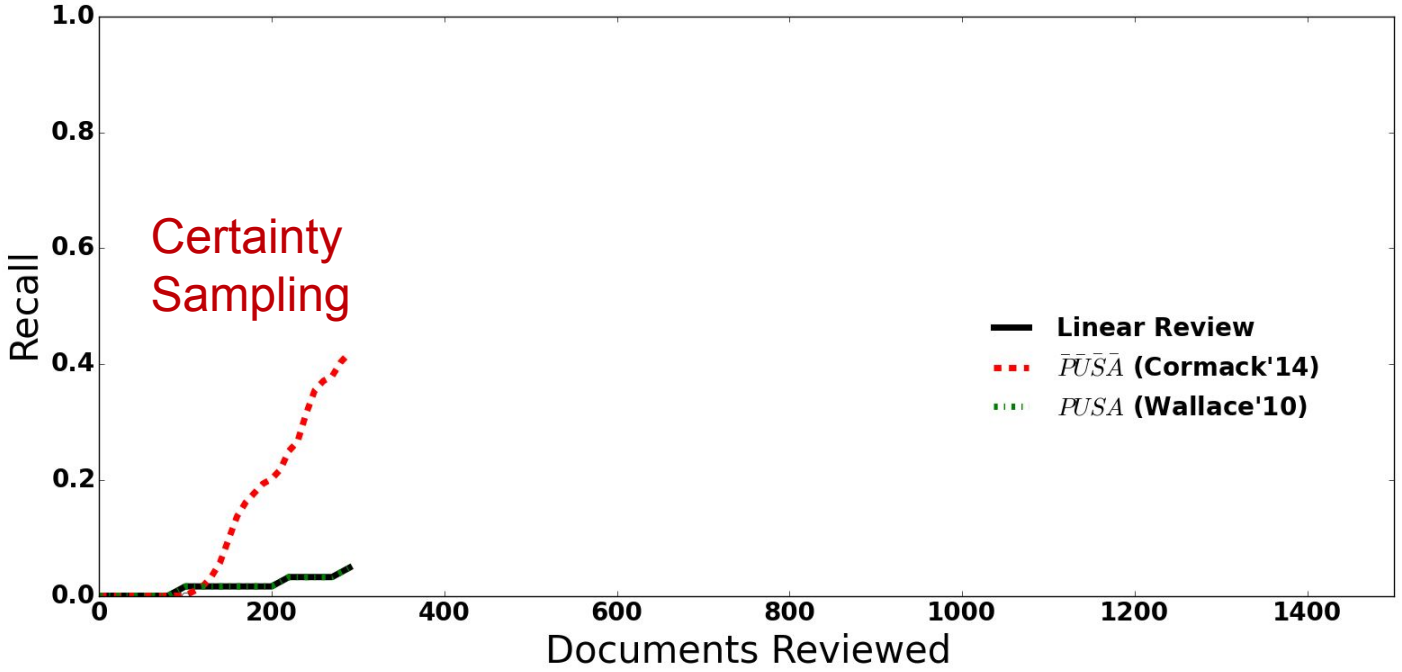
Outline

- *Background*
- *Method*
- ***Experiment***
  - Data
  - Framework
  - Evaluation
- *Result*
- *Conclusion and Future works*

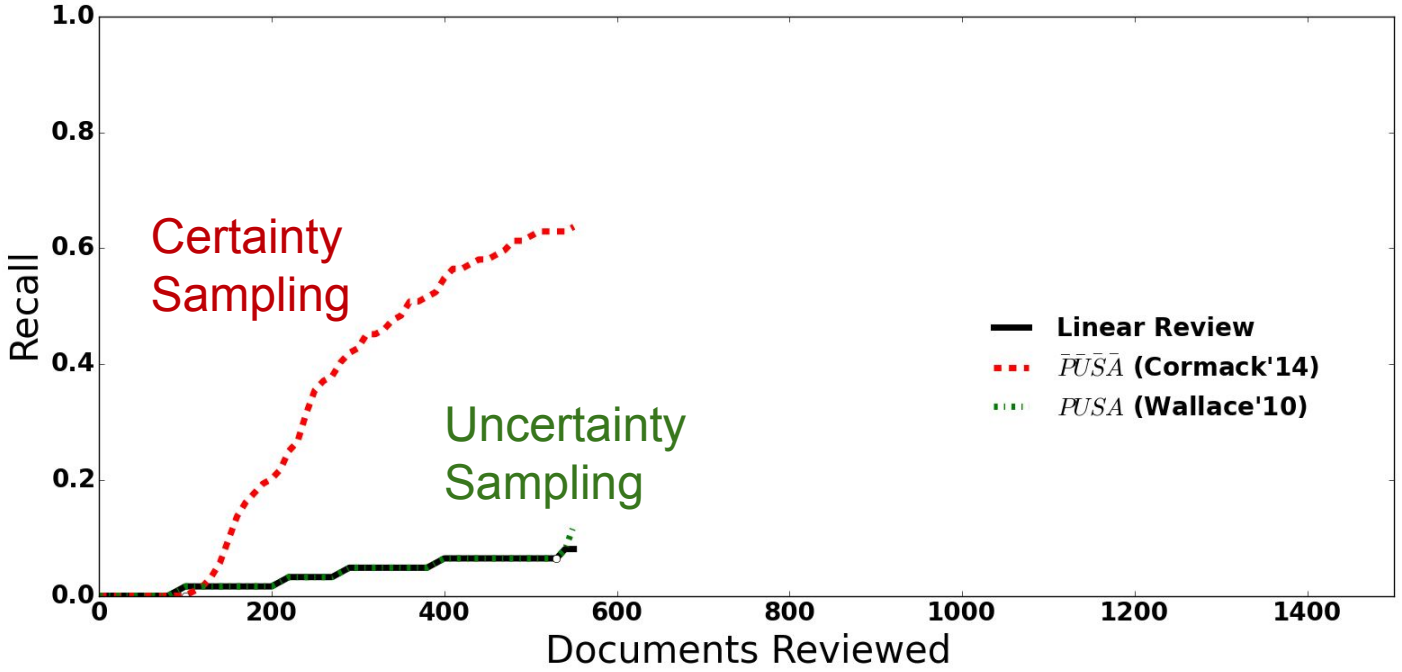
Evaluation 1



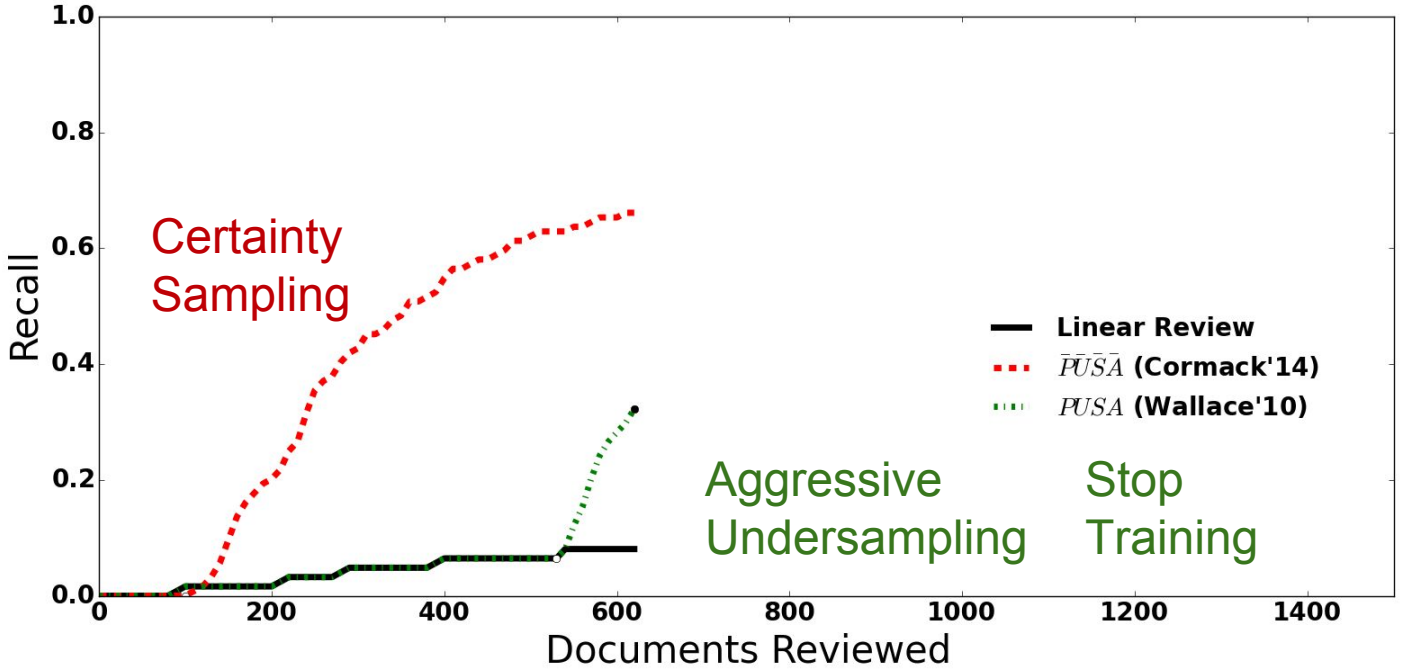
Evaluation 1



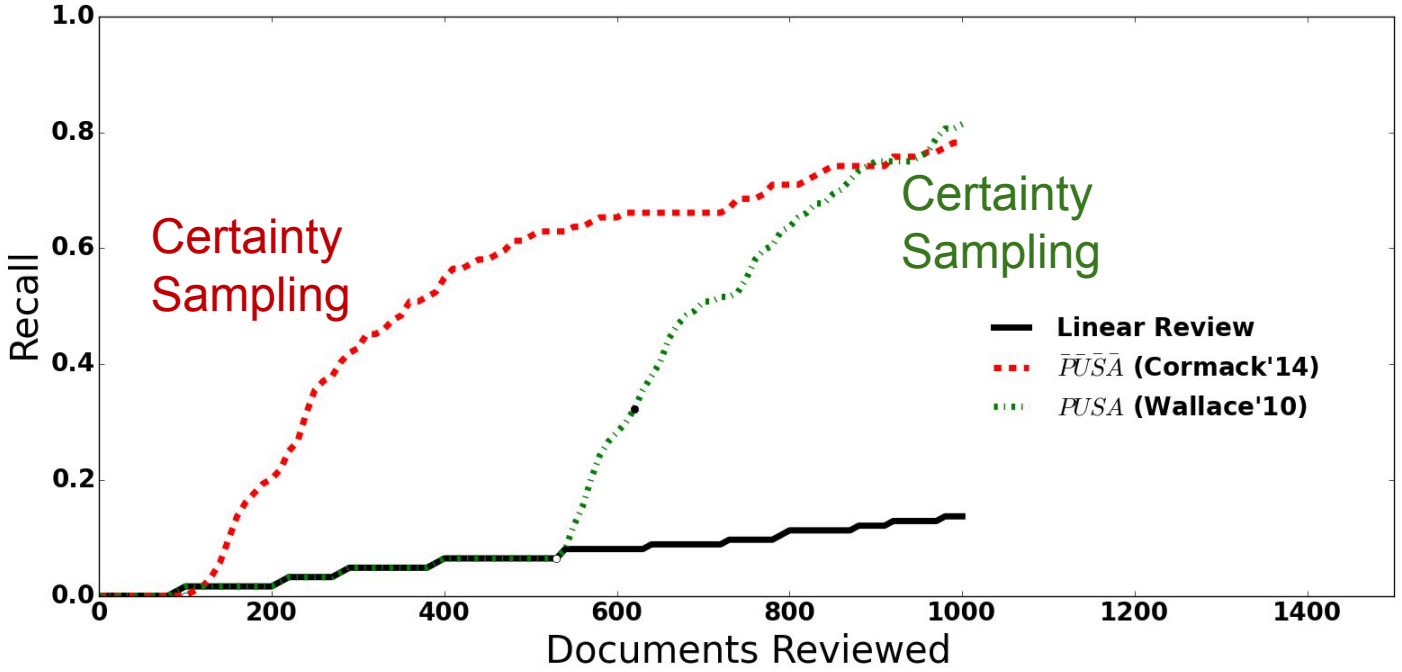
Evaluation 1



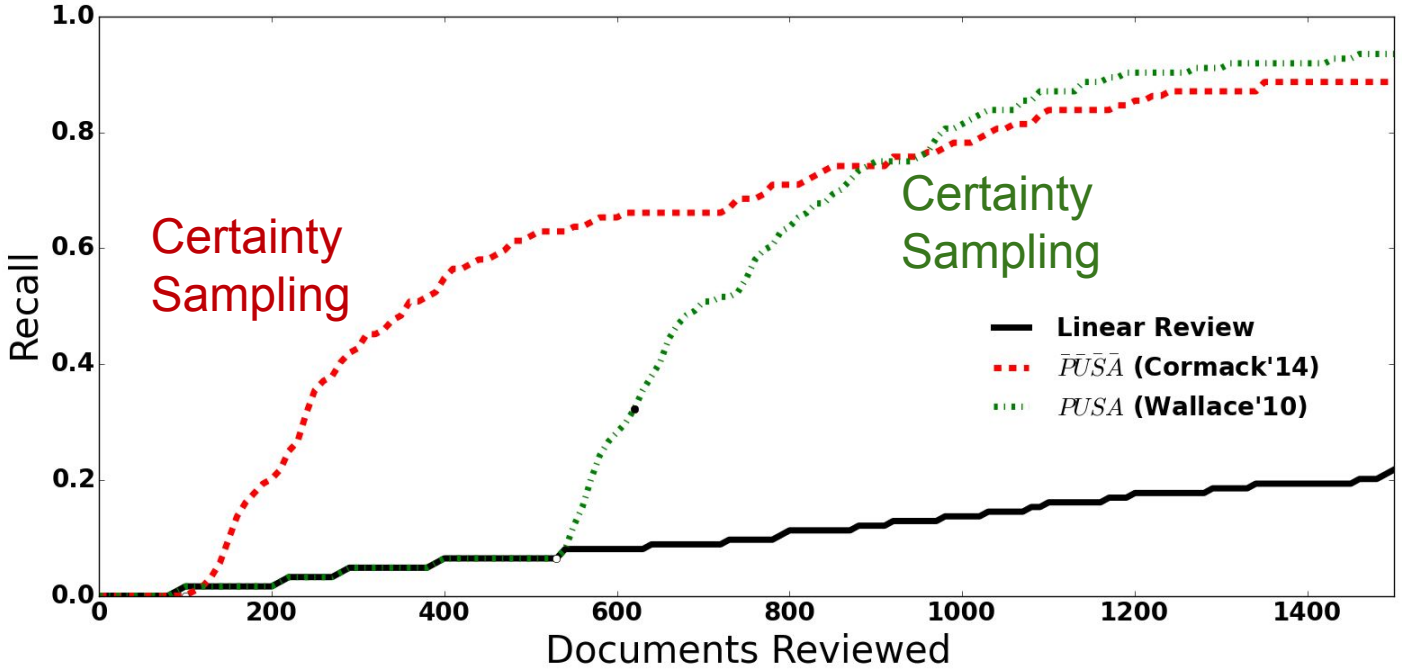
Evaluation 1



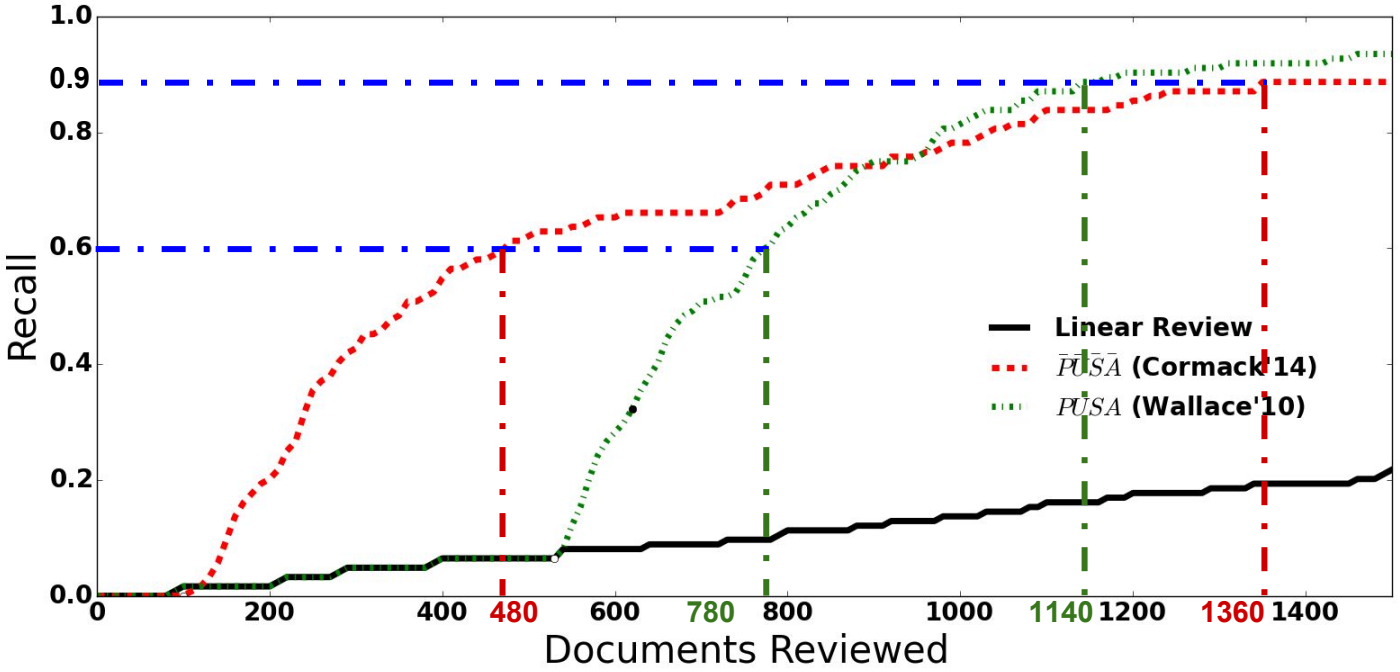
Evaluation 1



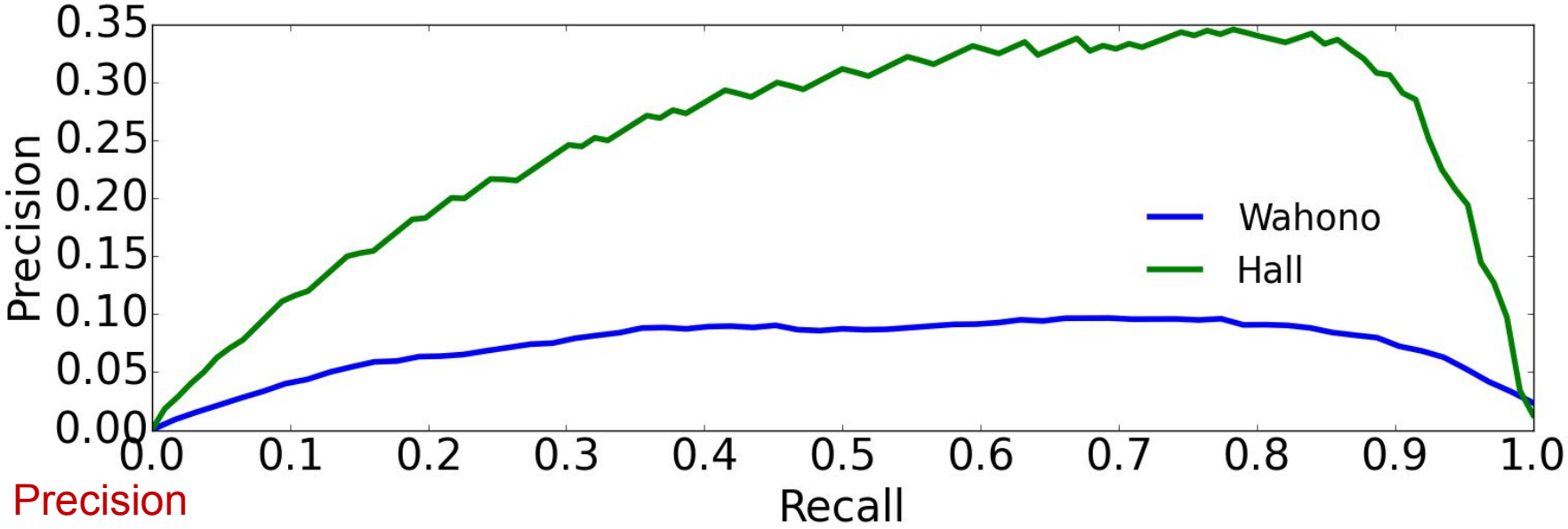
Evaluation 1



Evaluation 1



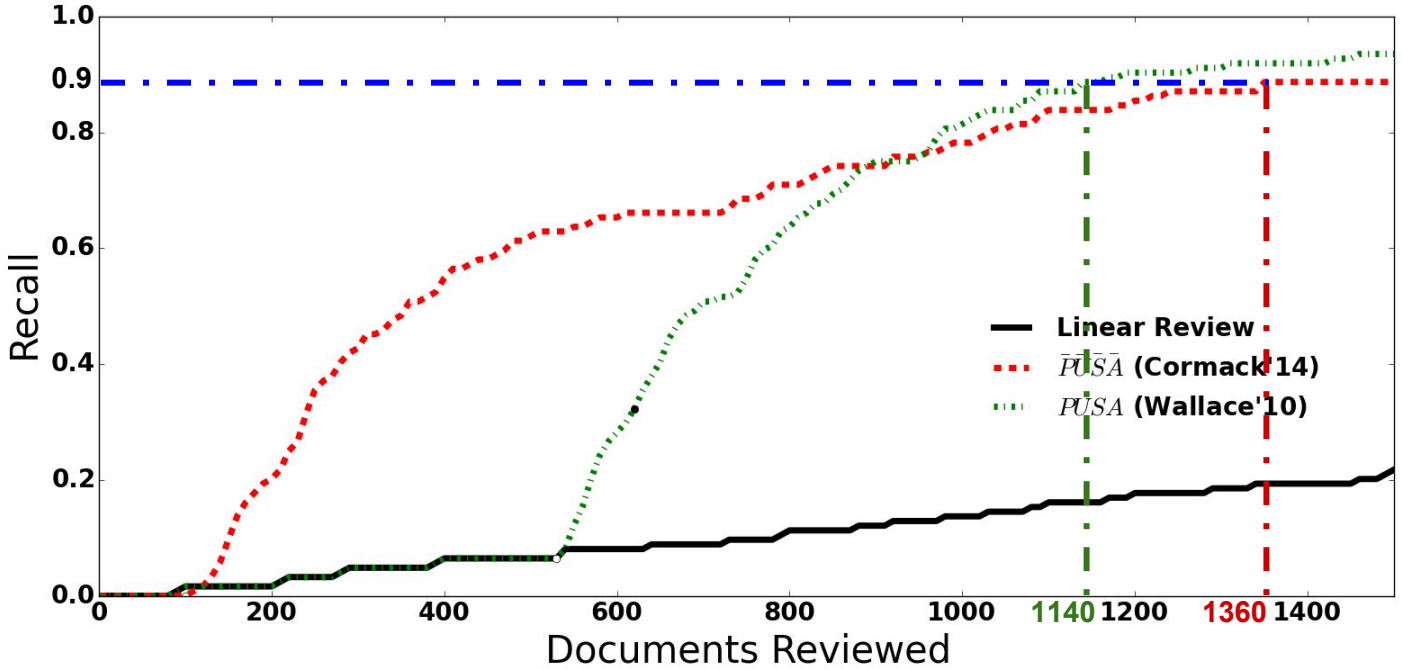
When to Stop



Precision
= relevant found / studies reviewed

Evaluation 2

X90 = studies reviewed to reach 90% recall



Research Questions

RQ1:

Can human-in-the-loop methods improve SE SLR?

RQ2:

Should we adopt state-of-the-art?

RQ3:

How much can we save?

Research Questions

RQ1:

Can human-in-the-loop methods improve SE SLR?

RQ2:

Should we adopt state-of-the-art?

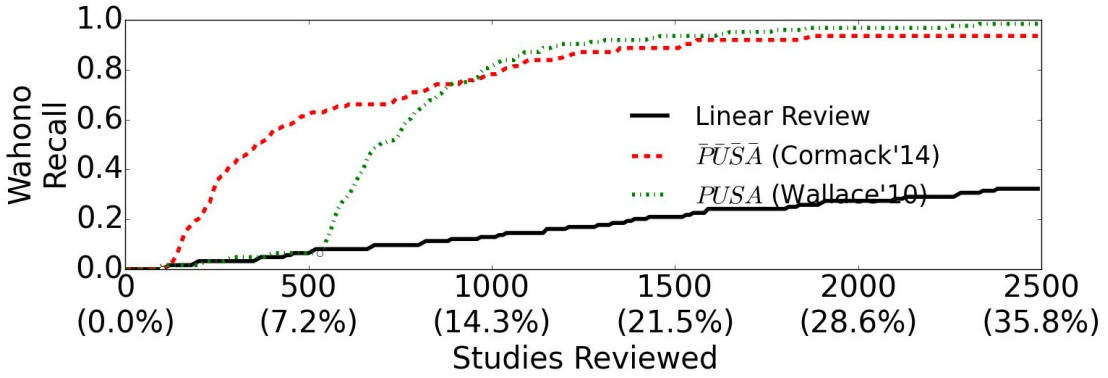
RQ3:

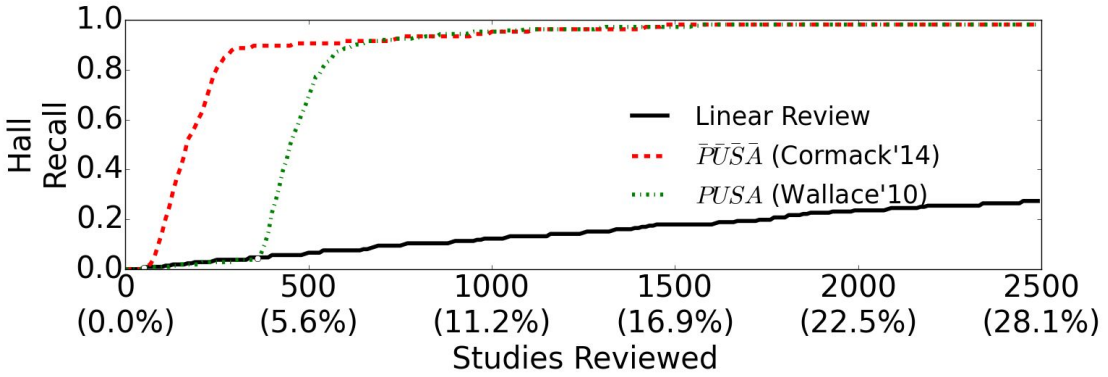
How much can we save?

Outline

- *Background*
- *Method*
- *Experiment*
- ***Result***
  - Human-in-the-loop vs. linear review
  - FASTREAD vs. state-of-the-art
  - Cost Reduction
- *Conclusion and Future works*

Human-in-the-loop vs. Linear Review





Research Questions

RQ1:

Can human-in-the-loop methods improve SE SLR? *Yes*

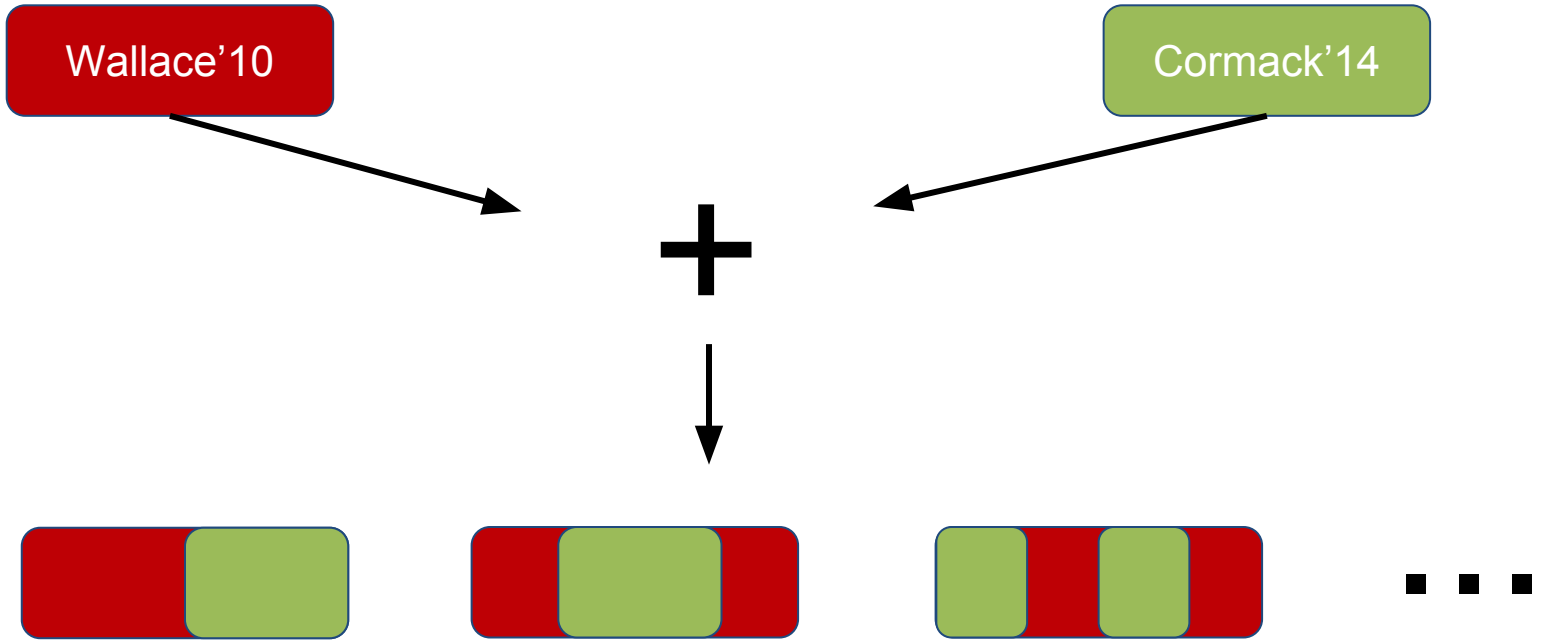
RQ2:

Should we adopt state-of-the-art?

RQ3:

How much can we save?

Refactor



Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Refactor

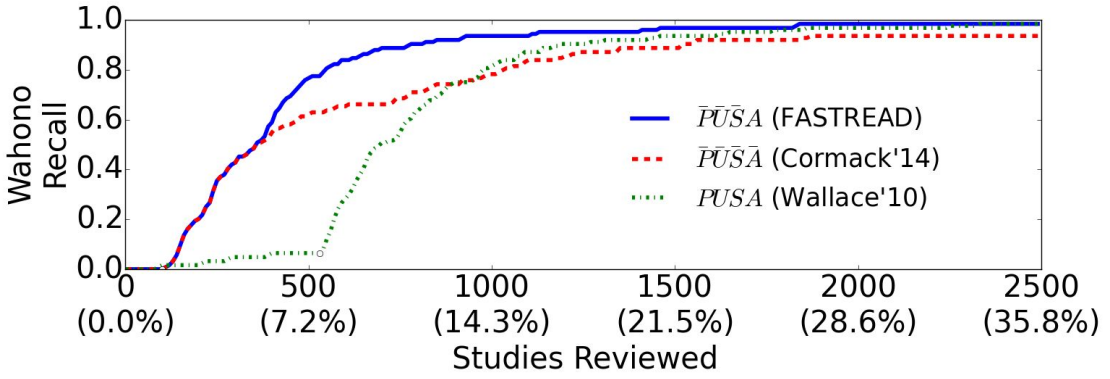
	Start point	Query Strategy	Continuity	Data Balancing
Wallace'10	late (P)	uncertainty sampling (U)	stop training (S)	aggressive undersampling (A)
Cormack'14	early ($\bar{P}$)	certainty sampling ($\bar{U}$)	non-stop ($\bar{S}$)	none ($\bar{A}$)
FASTREAD	early ($\bar{P}$)	certainty sampling ($\bar{U}$)	non-stop ($\bar{S}$)	aggressive undersampling (A)

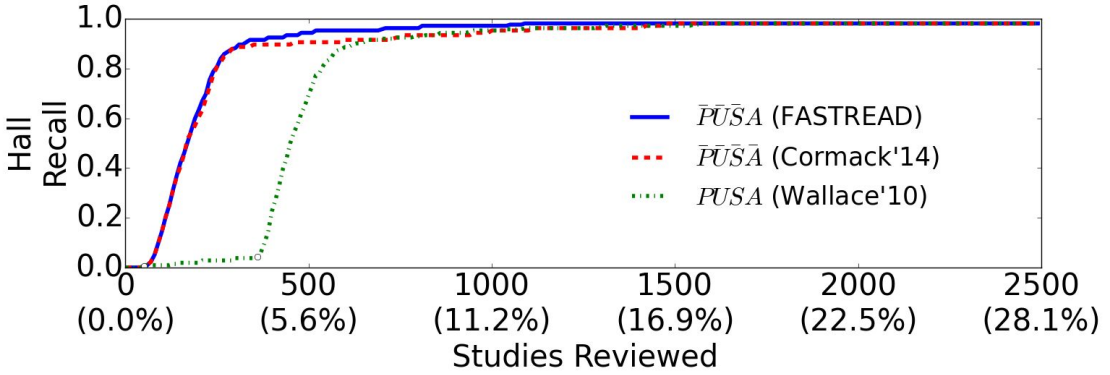
Wallace, Byron C., et al. "Semi-automated screening of biomedical citations for systematic reviews." BMC bioinformatics 11.1 (2010): 55.
Cormack, Gordon V., and Maura R. Grossman. "Evaluation of machine-learning protocols for technology-assisted review in electronic discovery." SIGIR'14.

Outline

- *Background*
- *Method*
- *Experiment*
- ***Result***
  - Human-in-the-loop vs. linear review
  - FASTREAD vs. state-of-the-art
  - Cost Reduction
- *Conclusion and Future works*

FASTREAD
vs.
State-of-the-art





Research Questions

RQ1:

Can human-in-the-loop methods improve SE SLR? *Yes*

RQ2:

Should we adopt state-of-the-art? *No*

RQ3:

How much can we save?

Outline

- *Background*
- *Method*
- *Experiment*
- ***Result***
  - Human-in-the-loop vs. linear review
  - FASTREAD vs. state-of-the-art
  - Cost Reduction
- *Conclusion and Future works*

X90 = studies reviewed to reach 90% recall

Wahono:

Review 680 (9.7%) to retrieve 90% relevant

Hall:

Review 310 (3.5%) to retrieve 90% relevant

X90 from 30 experiments

Rank	Treatment	Median	IQR	Wahono
1	$\bar{P}\bar{U}\bar{S}A$ (FASTREAD)	680	130	
1	$\bar{P}U\bar{S}A$	700	150	
2	$\bar{P}USA$	840	350	
2	$\bar{P}\bar{U}\bar{S}A$	1060	320	
2	$PU\bar{S}A$	1090	330	
2	$PUSA$ (Wallace'10)	1140	280	
2	$\bar{P}U\bar{S}\bar{A}$	1340	140	
2	$\bar{P}\bar{U}\bar{S}A$ (Cormack'14)	1350	130	
3	$\bar{P}\bar{U}\bar{S}A$	1640	370	
3	$PU\bar{S}\bar{A}$	1640	380	
4	$\bar{P}USA$	2240	1440	
4	$PUS\bar{A}$	2490	770	
5	Linear Review	6240	310	

Rank	Treatment	Median	IQR	Hall
1	$\bar{P}U\bar{S}\bar{A}$	290	70	
1	$\bar{P}\bar{U}\bar{S}A$	300	70	
1	$\bar{P}\bar{U}\bar{S}A$ (FASTREAD)	310	30	
2	$\bar{P}USA$	320	80	
2	$\bar{P}\bar{U}\bar{S}\bar{A}$ (Cormack'14)	350	70	
2	$\bar{P}USA$	350	100	
3	$PU\bar{S}A$	620	260	
3	$PUSA$ (Wallace'10)	630	270	
3	$\bar{P}\bar{U}\bar{S}A$	660	340	
3	$\bar{P}\bar{U}\bar{S}\bar{A}$	680	240	
3	$PU\bar{S}\bar{A}$	680	230	
3	$PUS\bar{A}$	750	230	
4	Linear Review	7880	250	

Scott, ANDREW JHON, and M. Knott. "A cluster analysis method for grouping means in the analysis of variance." Biometrics (1974): 507-512.

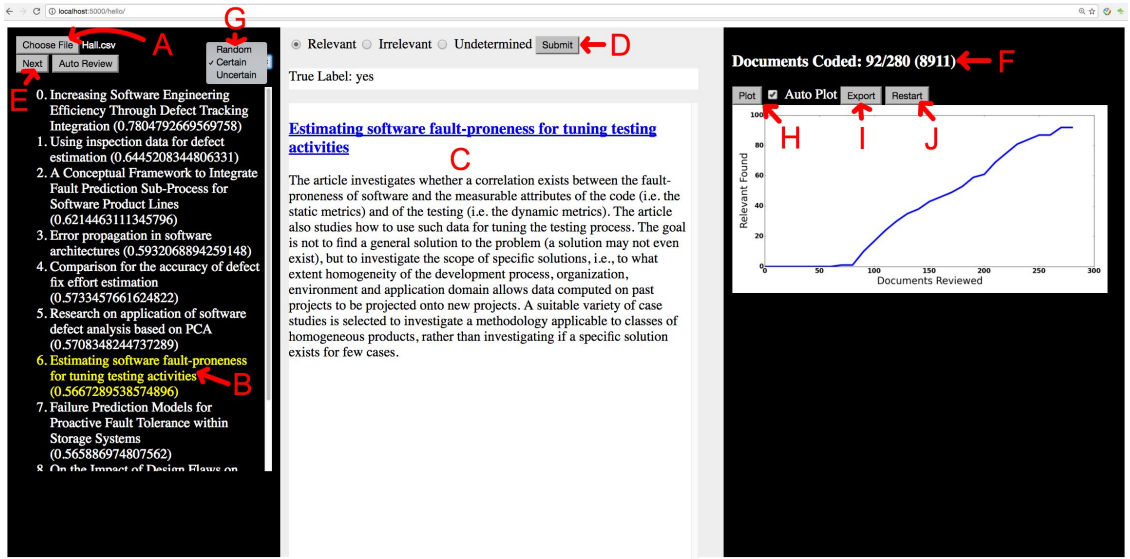
Resources

Data:

- SeaCraft Zenodo

Tool:

- SeaCraft Zenodo
- Github



Outline

- *Motivation and Background*
- *Method*
- *Experiment*
- *Result*
- ***Conclusion and Future works***

Conclusion

- *Apply human-in-the-loop method to facilitate SLRs*
- *FASTREAD, a better method than state-of-the-art*
- *Save 90% cost to retrieve 90% relevant studies*
- *A tool to implement FASTREAD*

Future Roadmap

Assumptions:

- no external domain knowledge
- binary classification
- one reviewer, no error

Future Roadmap

Assumptions:

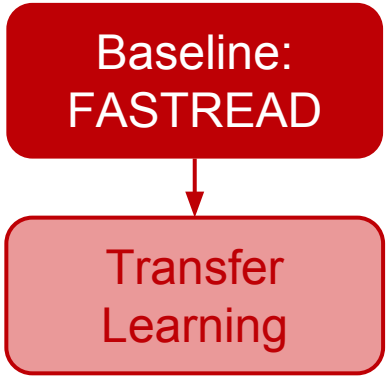
- no external domain knowledge
- binary classification
- one reviewer, no error

Baseline: FASTREAD

Future Roadmap

Assumptions:

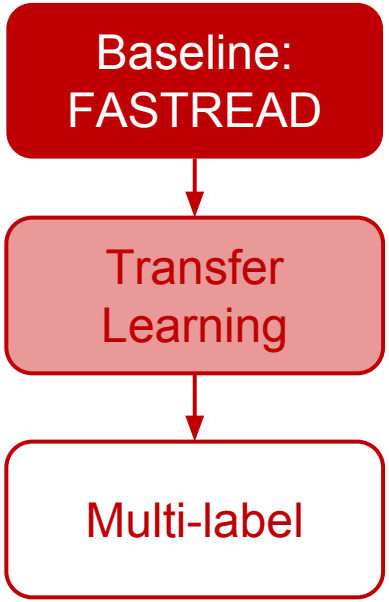
- no external domain knowledge
- binary classification
- one reviewer, no error



Future Roadmap

Assumptions:

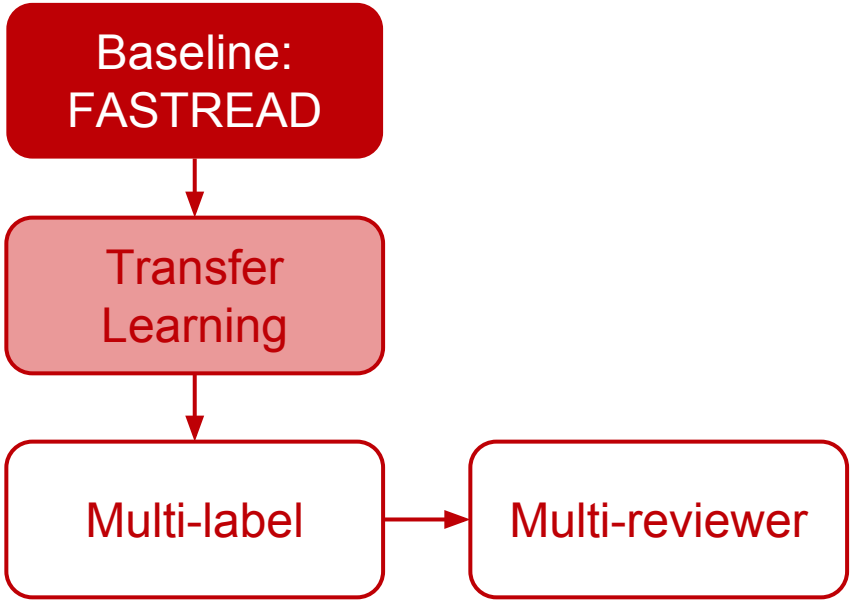
- no external domain knowledge
- binary classification
- one reviewer, no error



Future Roadmap

Assumptions:

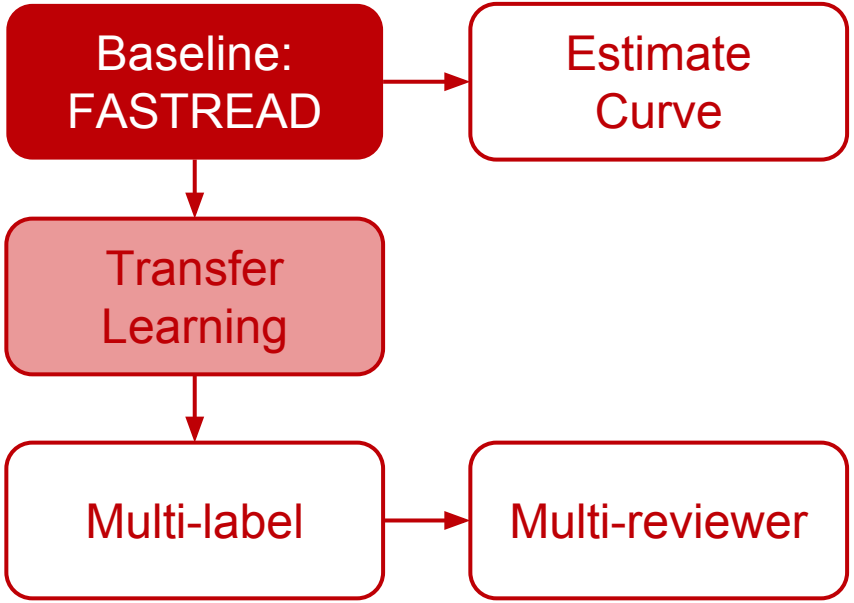
- no external domain knowledge
- binary classification
- one reviewer, no error



Future Roadmap

Assumptions:

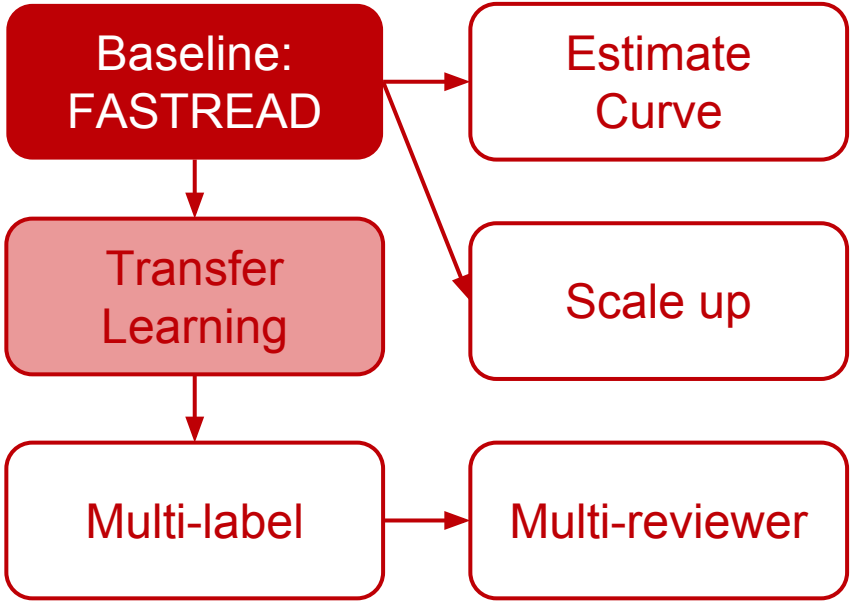
- no external domain knowledge
- binary classification
- one reviewer, no error



Future Roadmap

Assumptions:

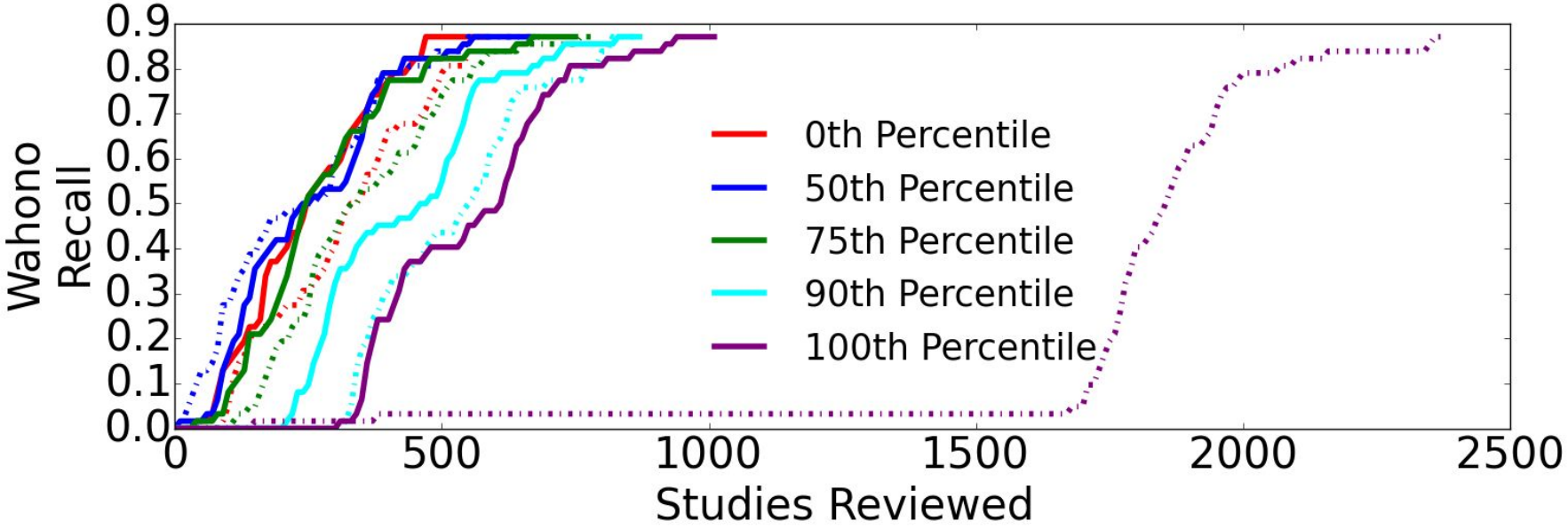
- no external domain knowledge
- binary classification
- one reviewer, no error



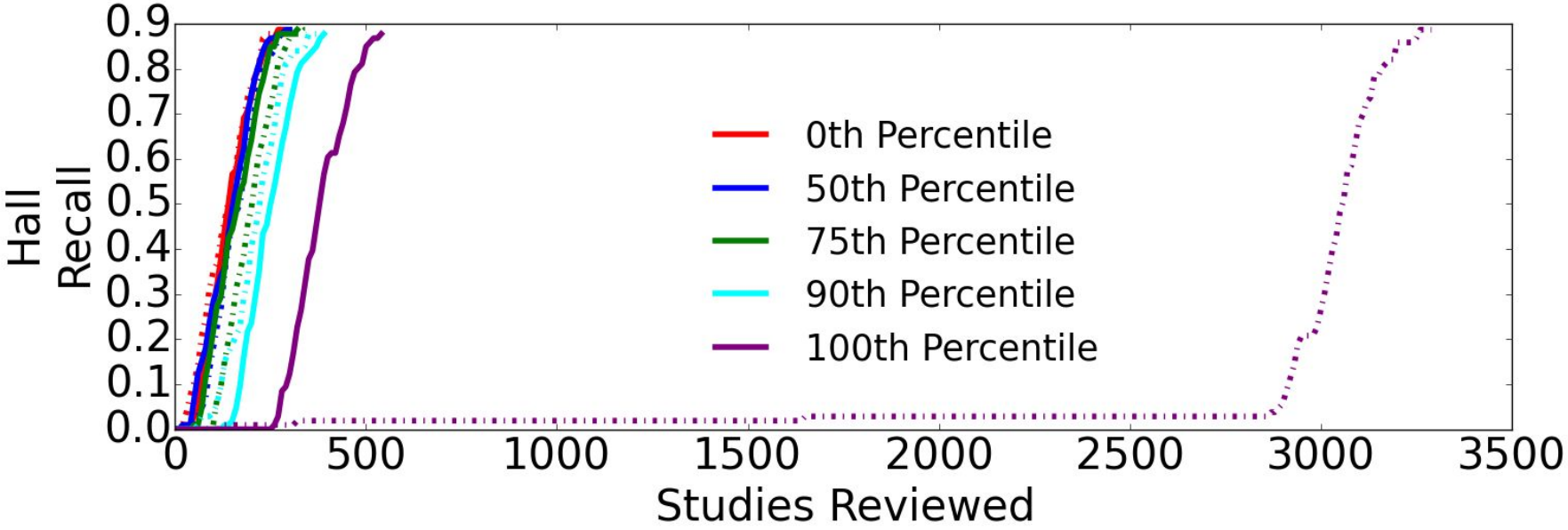

THANK
YOU


Questions
???

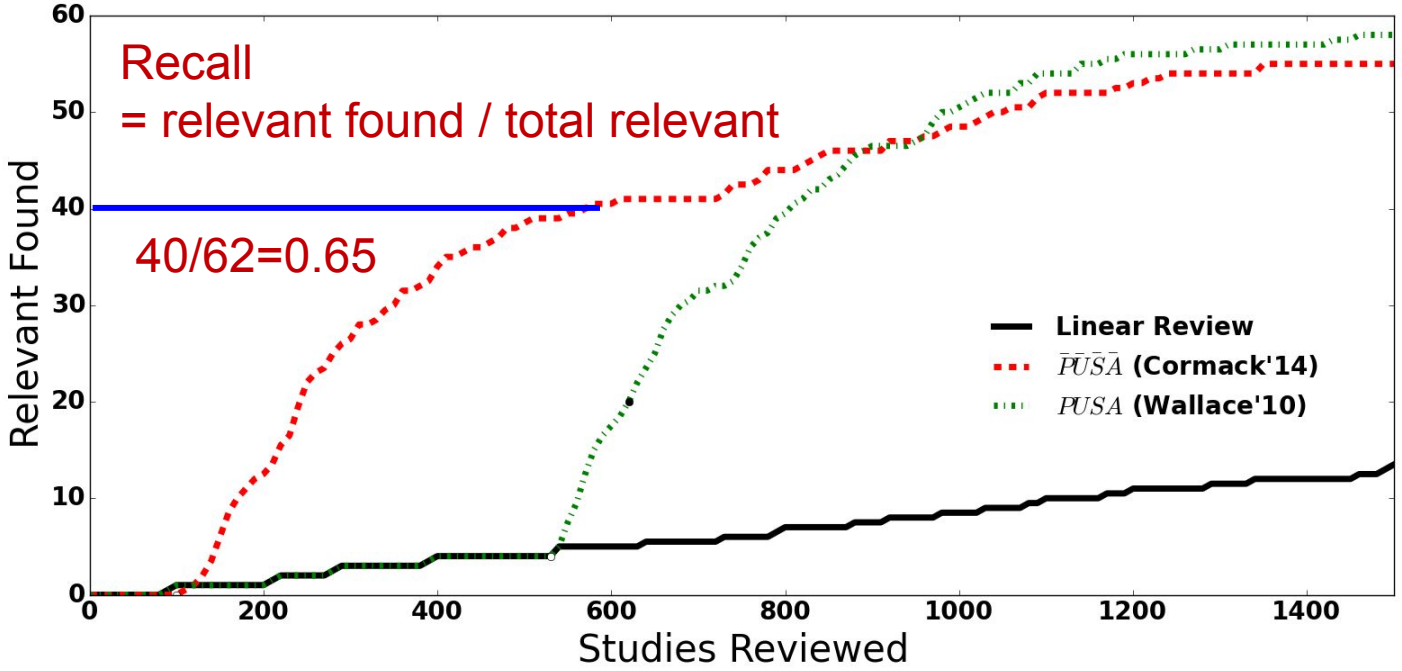
Back-up Slides



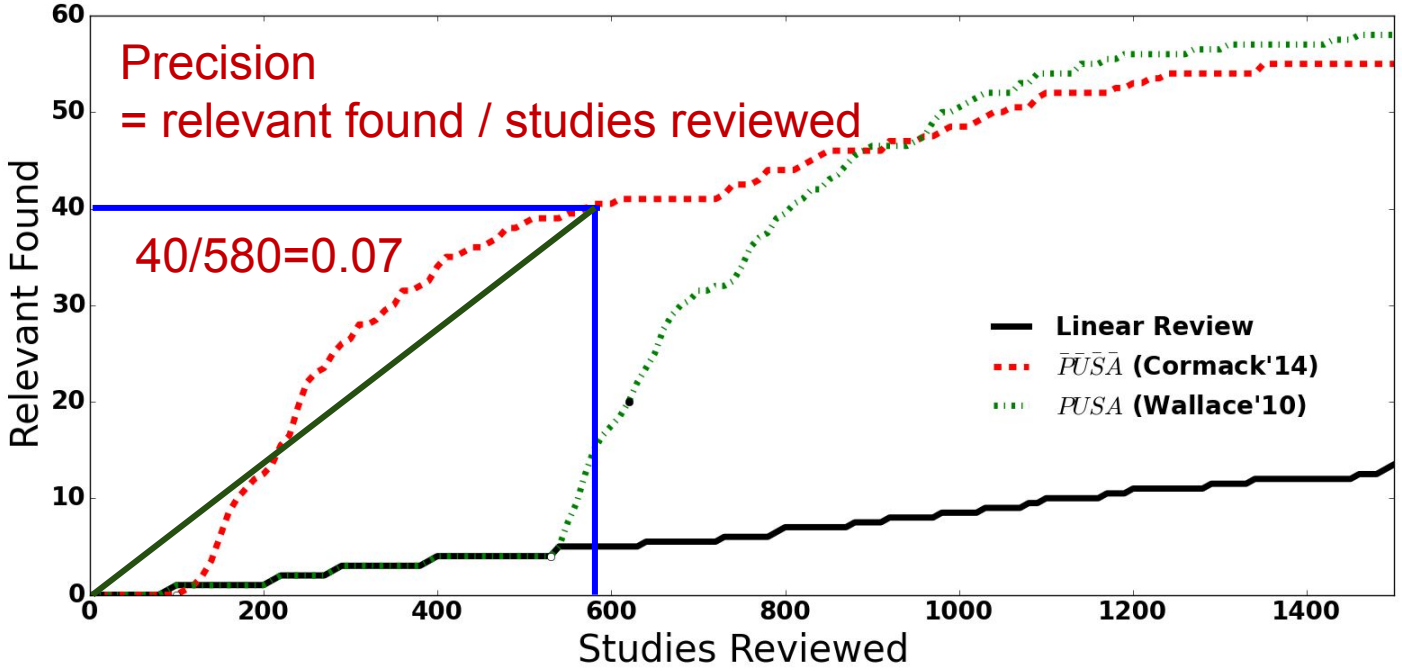
Back-up Slides



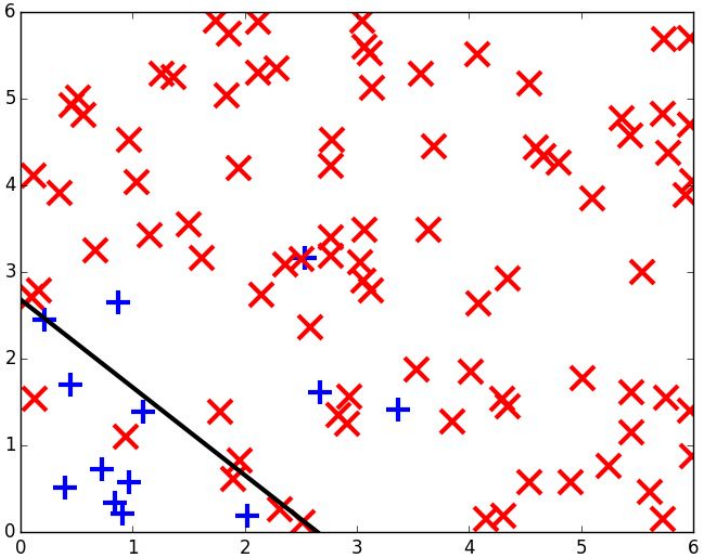
Evaluation

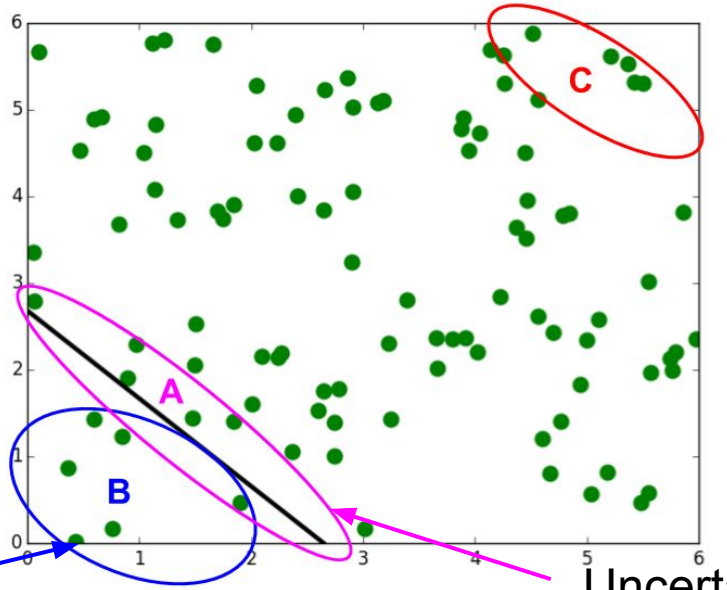

Recall
= relevant found / total relevant
40/62=0.65
Relevant Found
Studies Reviewed
0
10
20
30
40
50
60
0
200
400
600
800
1000
1200
1400
Linear Review
$\bar{P}\bar{U}\bar{S}\bar{A}$ (Cormack'14)
$PUSA$ (Wallace'10)

Evaluation


Precision
= relevant found / studies reviewed
40/580=0.07
Relevant Found
Studies Reviewed
0
10
20
30
40
50
60
0
200
400
600
800
1000
1200
1400
Linear Review
$\bar{P}\bar{U}\bar{S}\bar{A}$ (Cormack'14)
$PUSA$ (Wallace'10)

Query Strategy

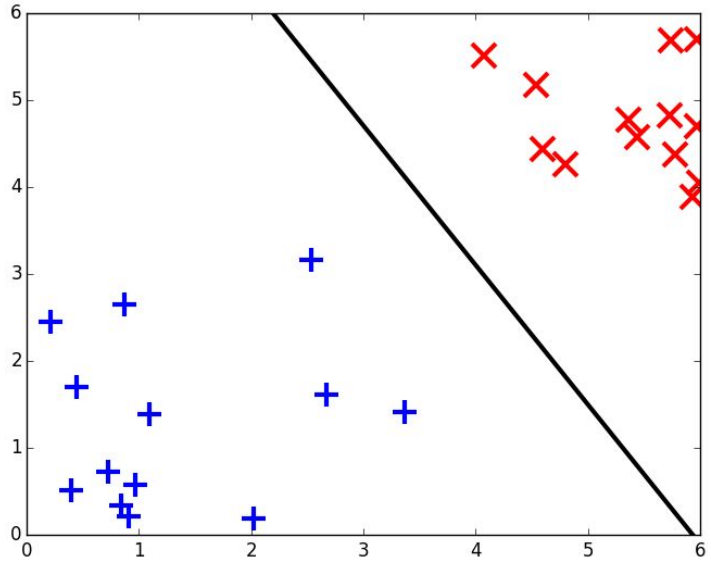


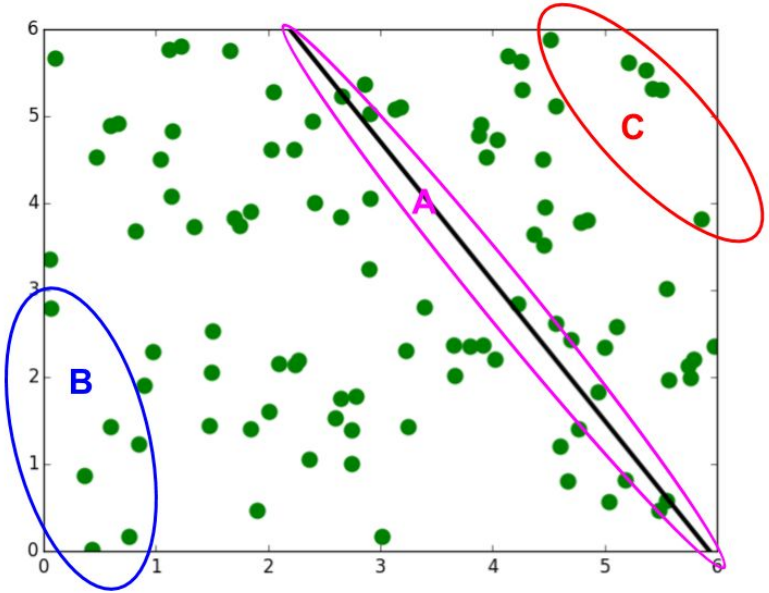


Certainty Sampling

Uncertainty Sampling

Data Balancing





Data Sets

	Wahono		Hall	
	Stated	Retrieved	Stated	Retrieved
Initial List	2117	7002	2073	8911
Final List	72	62	136	106

Hall, Tracy, et al. "A systematic literature review on fault prediction performance in software engineering." TSE'12.
Wahono, Romi Satria. "A systematic literature review of software defect prediction: Research trends, datasets, methods and frameworks." JSE'15.

Roadmap

Systematic literature reviews:

- Useful. Important.
  - In medicine, law, SE and elsewhere
- Expensive.

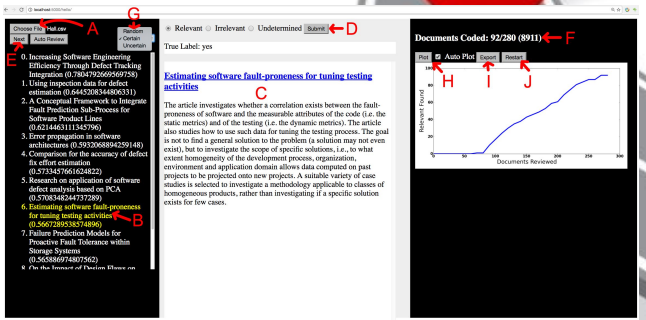
Cost Reduction

- Primary study selection= part of each reviews. Hard.
- Tools to make it easier?
  - Problems with state of the art

Human-in-the-loop Incremental Learning

- Used in other domains.
- Q: Useful in SE?
- A1: Yes.
- A2: Can be used to tailor a even better SE method.


SCALING UP


A
B
C
D
E
F
G
H
I
J
Relevant
Irrelevant
Undetermined
Submit
True Label: yes
Estimating software fault-proneness for tuning testing activities
Documents Coded: 92/280 (8911)
Auto Plot
Relevant Found
Documents Reviewed

LexisNexis®