AutoARTS: Taxonomy, Insights and Tools for Root Cause Labelling of Incidents in Microsoft Azure

Pradeep Dogga

UCLA

Chetan Bansal

Richie Costleigh

Gopinath Jayagopal

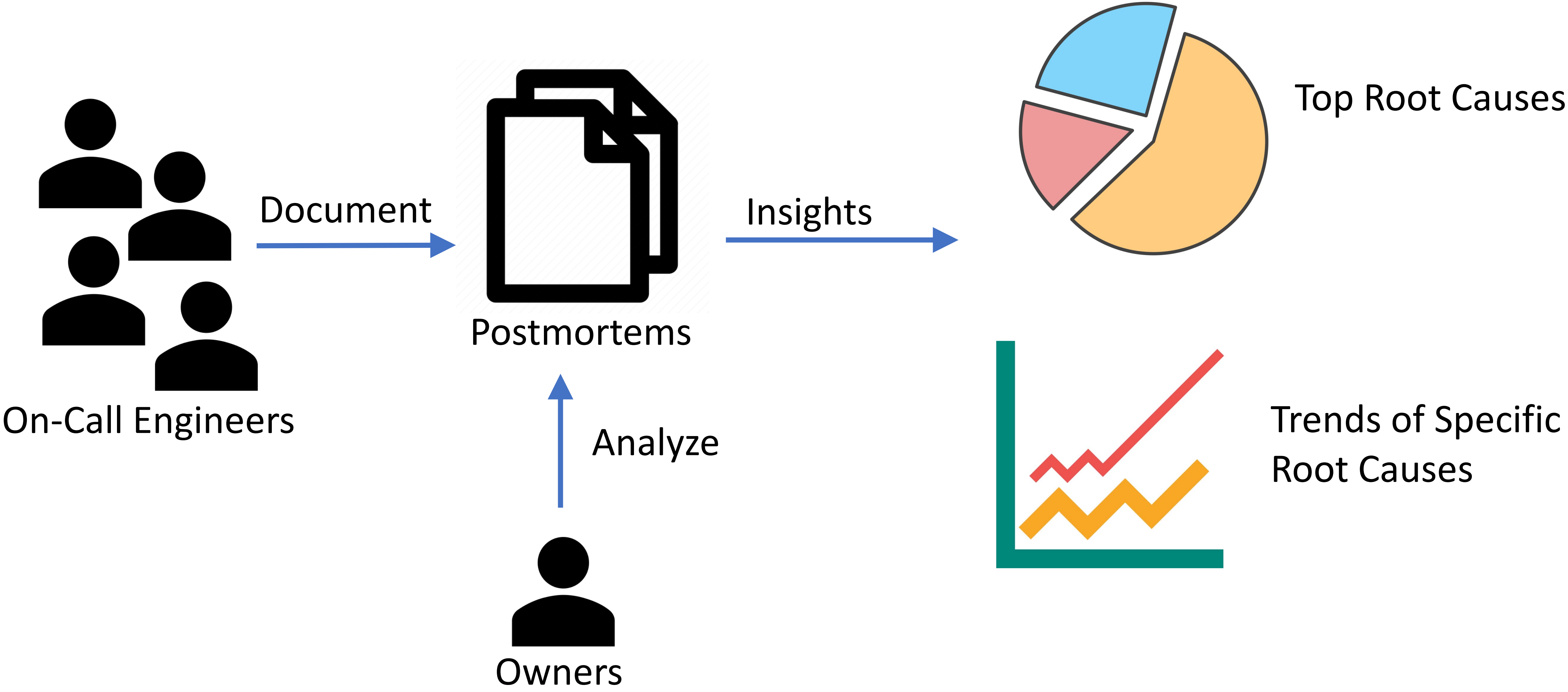
Suman Nath

Xuchao Zhang

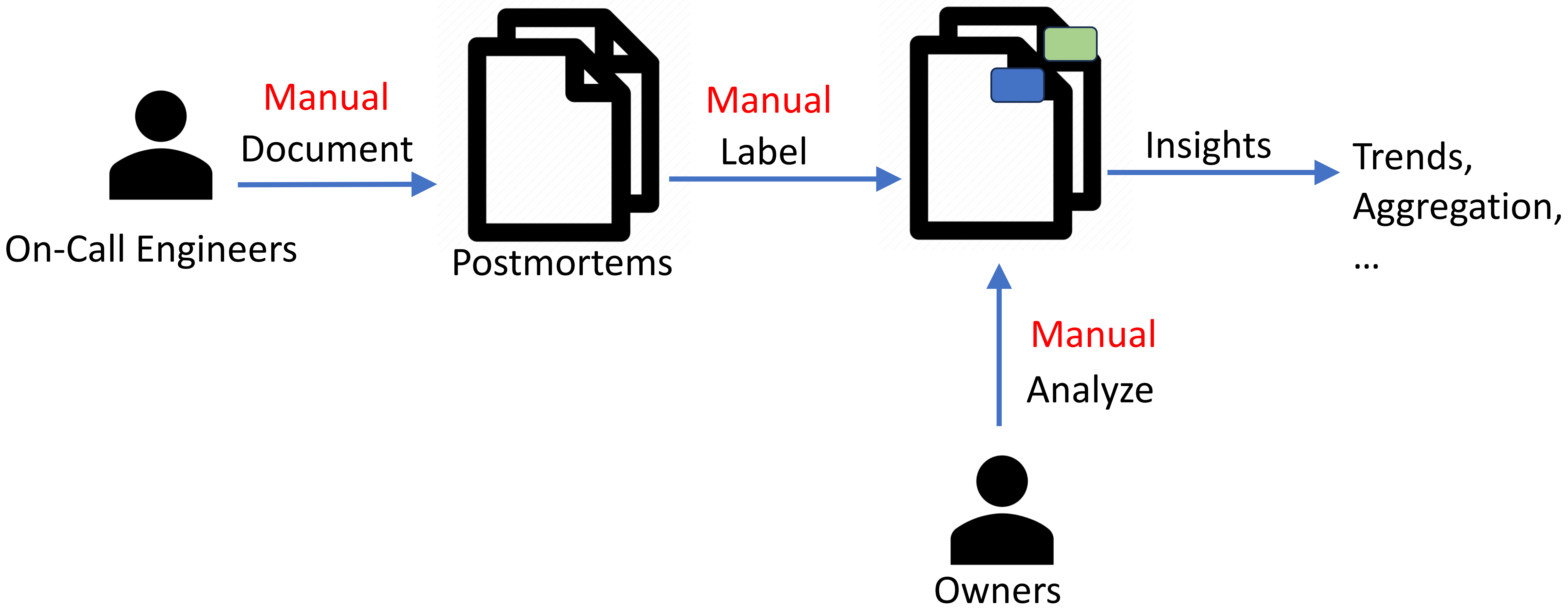
Microsoft

Incident Postmortems in Clouds

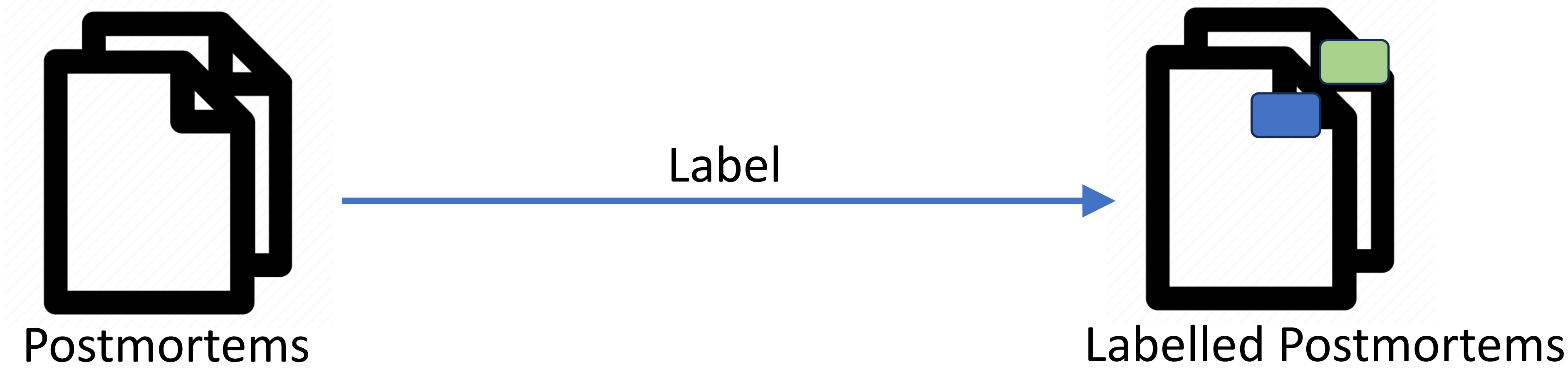
Retrospective Analysis using Postmortems



Retrospective Analysis Today



Root Cause Labelling Today – Taxonomies



Team 1
- Network
- …

Team 2
- DC Networking
- …

…

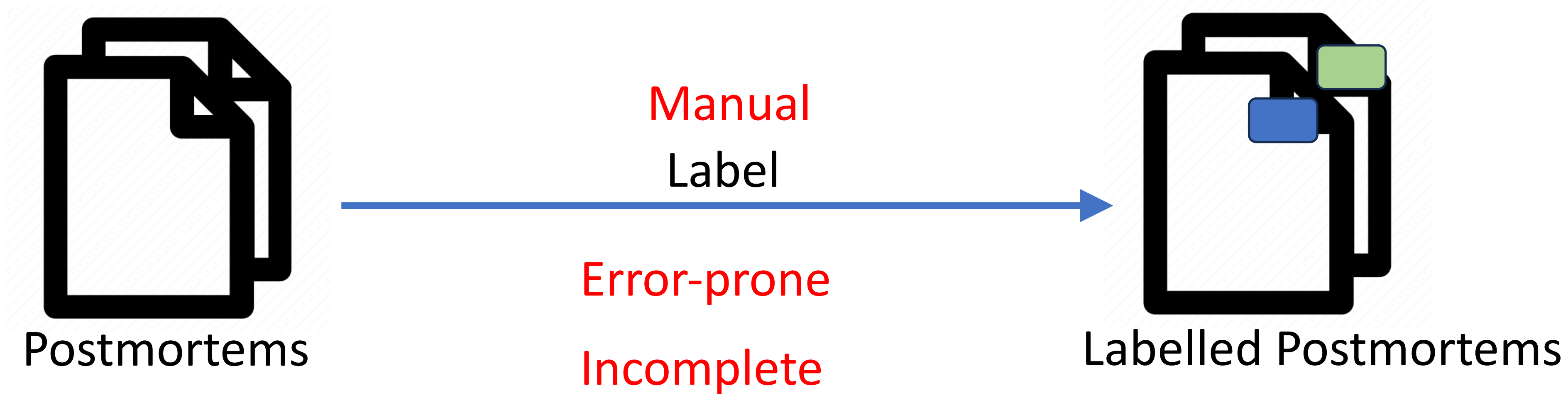
Team n
- Code bug
- …

Ambiguous

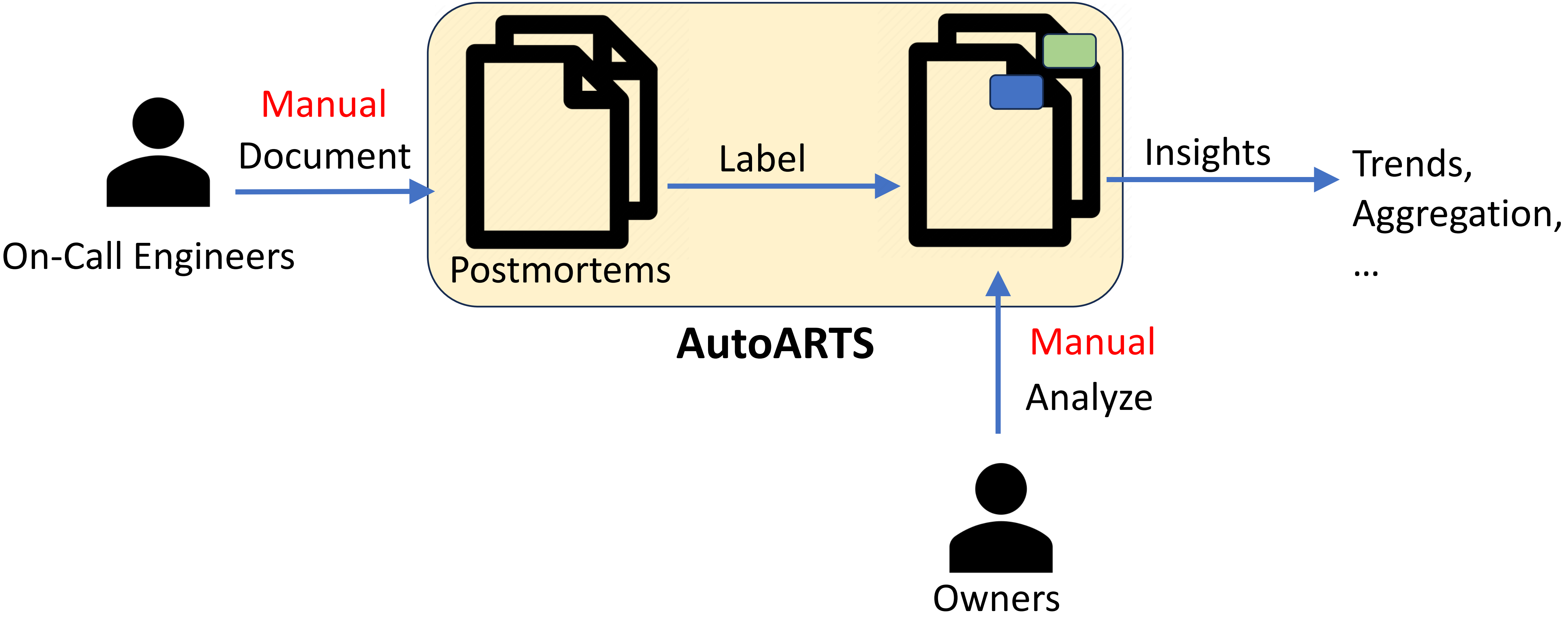
Incomplete

Flat

Root Cause Labelling Today



Retrospective Analysis Today



What AutoARTS is about

Problem: Lengthy postmortems, poor root cause taxonomies, error-prone and incomplete root cause labelling.

Solution: Develop comprehensive taxonomy, bootstrap labelling postmortems, generate succinct contexts and labels with ML.

Ideas: Leverage hierarchy in taxonomy, train text encoders w.r.to tags, finetuning gap sentence summarization.

Opensource Taxonomy: Share wide variety of contributing factors with others and develop continuously.

Postmortems – Treasure Troves of Rich Debugging Insights

- Title, symptoms, root causes, mitigation steps, 5-Whys, etc.
- Written in natural language with little to no structure.
- Valuable insights lost due to lengthy reports.

Widespread ** failures impacting multiple *** services due to overload of Azure ***** system**

Azure ******* utilizes two layers of ……. (omit)……. It must be noted that the edge caches do not cache negative responses like **** since the range of these values is infinite. A non-authoritative server like the ****** not reasonably figure out the range of values to cache. …….(omit)…….

Post-Incident Report (PIR)

Retrospective Analysis - Challenges

- Lengthy – avg. 4500 words long.
- Complex – on average, 9 engineers involved in an incident
- Written by many – 34K engineers.
  - Varying degrees of expertise and linguistic styles.

Retrospective Analysis - Challenges

- Error-prone – 20% labelled as ‘Other’.
- Incorrect – 29% labelled incorrectly.
- Incomplete – 58% incomplete labels(e.g., Networking – Other).

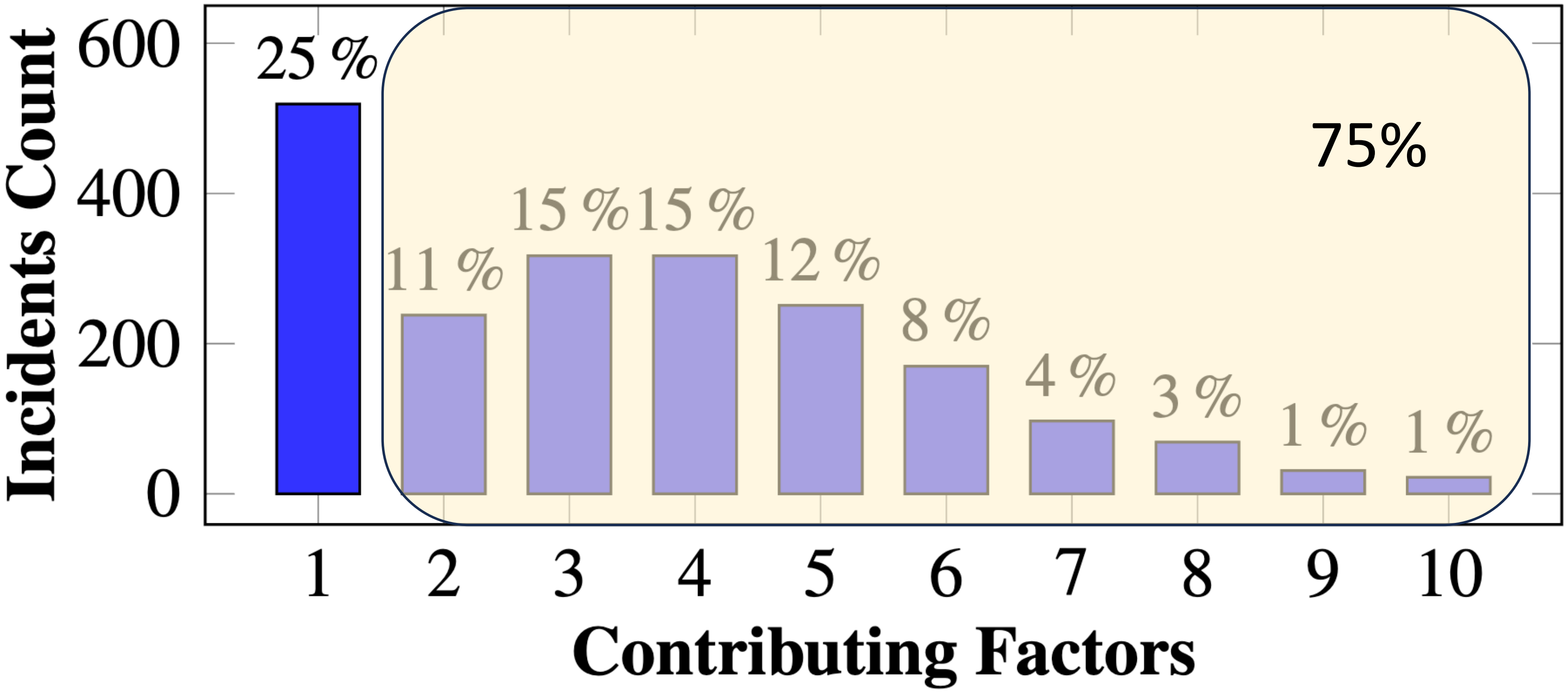
Manual Analysis at Microsoft Azure

- Extensive multiple person-year effort.
  - 2051 incidents.
  - 468 services from Microsoft Azure.
- Goals:
  - Identify all the contributing factors behind the incident.
  - Extract key context from the postmortem for each factor.
- Weekly peer review to refine analysis and develop taxonomy of contributing factors.

Manual Analysis At Microsoft Azure - Principles

- Intellectually honest
  - Involve teams and domain experts.
- Focus on depth *and* breadth
  - Extract all the contributing factors to an incident.
- Actionable findings
  - Lead to creating/updating standards to mitigate future incidents.
- Continuous evolution
  - Learn new factors and evolve the taxonomy.

Manual Analysis At Microsoft Azure – Contributing Factors



- 4 contributing factors on average – Contrary to existing work
- Addressing easiest one can reduce incidents!

Manual Analysis At Microsoft Azure - Example

- A service became unavailable after a customer pushed a load that was 60x greater than what the service can handle.

- Contributing factors:
  - Inrush of load from a single customer
  - Lack of throttling on both customer and service ends
  - High CPU, heap usage and thread count led to request failures with exceptions
  - Exception handling of failed request led to resource leaks
  - No automated watchdogs to detect early outage symptoms (or resource leaks)
  - Team cannot access metrics (collocated with service) during the outage.

- Originally chosen label: ‘Service – Load Threshold’

Manual Analysis At Microsoft Azure – Contributing Factors

- Wide Variety – 346 distinct factors!

Category	Frequency	TTM (Hrs)
Detection	**61%**	50
Authoring	**50%**	58
Dependency	37%	16
Architecture	20%	33
Deployment	20%	27
Process	18%	**123**
Load	14%	13
Auth	7%	21
Performance	6%	16
Datacenter	4%	**70**

https://autoarts-rca-taxonomy.github.io/taxonomy.html

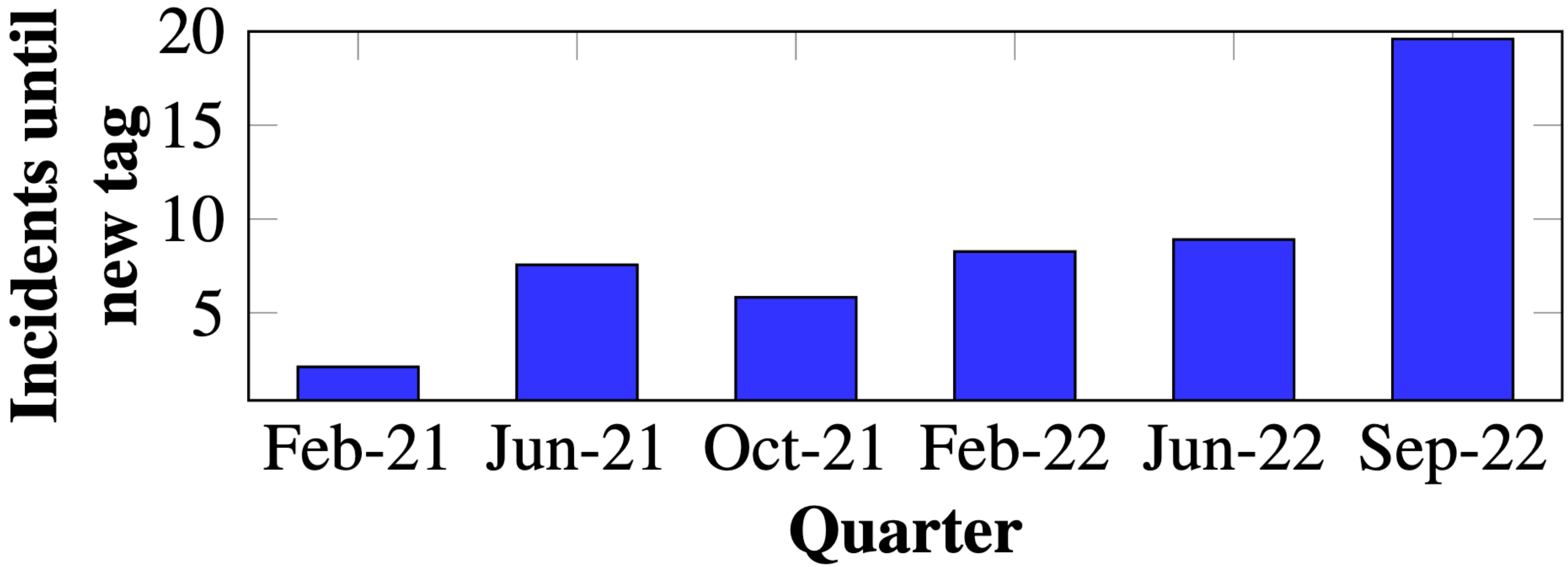
ARTS Taxonomy

- Azure Reliability Tagging System (ARTS) taxonomy to label incidents with contributing factors.

- Visualization: https://autoarts-rca-taxonomy.github.io/taxonomy.html

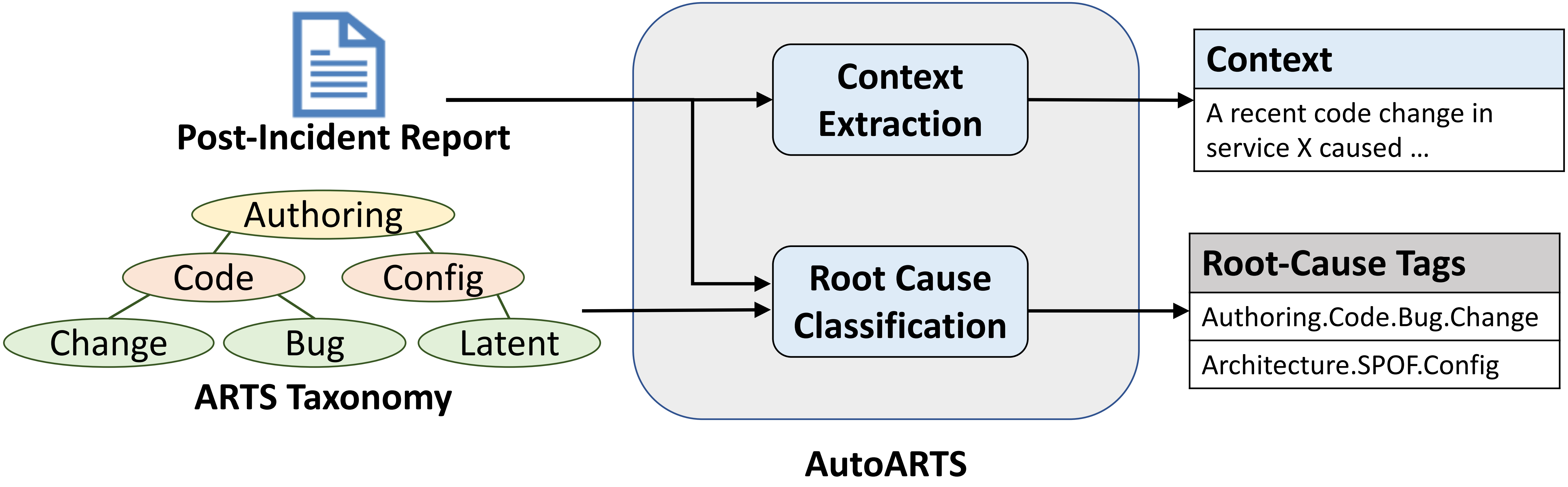
- Qualities:
  - Hierarchical (4 levels deep)
  - Comprehensive (built from analysis)
  - Unambiguous (clear separation of categories)

ARTS Taxonomy – Growing Stable



- But manual labelling is still error-prone!
- Our analysis is expensive and cannot scale to all postmortems.

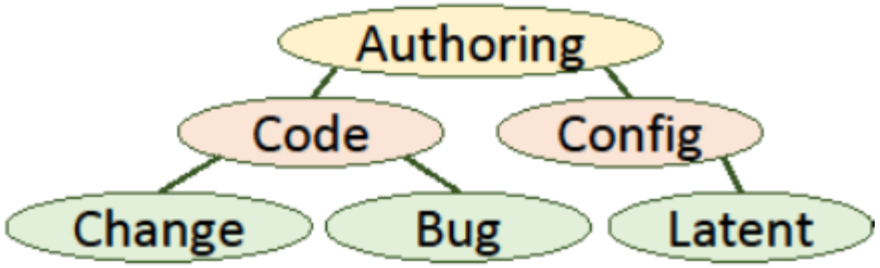
AutoARTS – Automated Root Cause Labelling



AutoARTS – Root Cause Classification

- Multi-label text classification
  - Noise: Irrelevant details in postmortems
  - Data sparsity: 68% of tags have < 10 postmortems

- Leverage hierarchy in ARTS taxonomy using GCN[1]



- LLMs need large amounts of data to encode text
  - Train custom text encoder w.r.to taxonomy

[1] Zhou, J., et al. "Hierarchy-aware global model for hierarchical text classification." ACL'20.

Can language models encode postmortems?

- 110K postmortems (20% Test split)
- Poor performance

Model	Test Perplexity
BERT-uncased	7.57
BERT-cased	6.69
XLNet-uncased	23.67

AutoARTS – Context Extraction Examples

Root-Cause Tag	Context from PIR
Authoring.Code.Bug.Change	SQL team made some recent changes to a gateway component that introduced this regression
Detection.Validation.MissingTestCoverage	NRP test infrastructure doesn't support component tests for standard public IPs.

AutoARTS – Context Extraction

- Extract key context from PIR to justify root cause tags.
- LLMs are good at summarization (abstractive/extractive)
  - But context is not a summary of PIR
- Pegasus[1] is trained for summarization by masking sentences
  - Context sentences should be extracted from PIR
  - Use labelled contexts to finetune Pegasus to extract context from PIRs

[1] PEGASUS: pre-training with extracted gap-sentences for abstractive summarization. ICML'20

AutoARTS – Evaluation

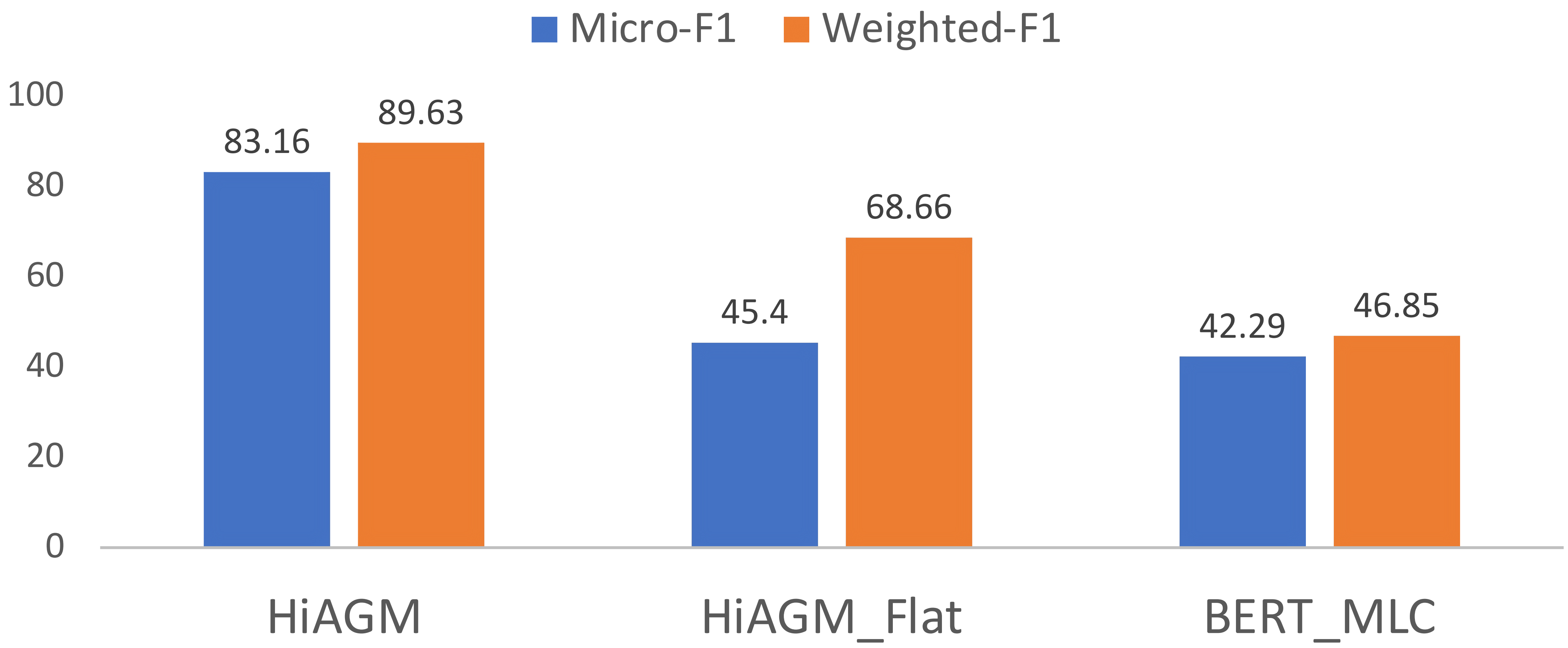
- 1120 labeled PIRs from Microsoft Azure.
- Dataset splits: Train (72%), Validation (8%), Test (20%).

Which parts of PIR to use?

Section	Micro-F1	Weighted-F1
Whole PIR	**0.55**	**0.40**
Title	0.53	0.45
Summary	0.47	0.46
RC-Details	0.52	0.45
5-Whys	0.54	0.40
Discussion	0.53	0.40
Mitigation	0.47	0.40
RC-Details + 5-Whys	**0.56**	**0.42**

Language models have limits on text sequence length!

AutoARTS – Root Cause Classification



Hierarchical structure of ARTS is beneficial for classification!

AutoARTS – Context Extraction

Model	ROUGE			BLEU			
	Rouge-1	Rouge-2	Rouge-L	BLEU	BLEU-1	BLEU-2	BLEU-3
Pegasus - Pretrained	32.55	18.72	24.30	9.61	18.03	10.31	8.93
Pegasus - Finetuned	**45.46**	**35.65**	**38.43**	**24.60**	**32.19**	**24.98**	**23.41**
T5 - Pretrained	34.38	23.31	28.03	10.06	15.68	10.83	9.43
T5 - Finetuned	41.63	33.86	35.76	23.81	29.81	24.10	22.70
BERT-cased - Pretrained	40.05	27.03	31.01	18.62	28.43	18.95	16.83
BERT-cased - Finetuned	40.08	27.35	31.20	18.80	28.32	19.03	16.95
BERT-uncased - Pretrained	39.52	26.58	30.74	17.63	27.47	17.98	15.89
BERT-uncased - Finetuned	39.92	27.44	31.57	18.64	28.08	18.91	16.90

AutoARTS – User Feedback

- 10 PIRs not previously in evaluation dataset.

- **Metric:** How useful were the AutoARTS generated contexts in identifying all contributing factors?
  - 1 – Not useful at all
  - 5 – Very useful.
- **Response:** 4.6.

- **Metric:** How many contexts were generated with unnecessary details?
- **Response:** 0.

AutoARTS – User Feedback

- **Metric:** How many new root cause labels were you able to identify using the generated contexts?
- **Response:** 2.

- **Metric:** How many crucial root cause tags were missing from the outputs?
- **Response:** 7/10.

What AutoARTS is about

Problem: Lengthy postmortems, poor root cause taxonomies, error-prone and incomplete root cause labelling.

Solution: Develop comprehensive taxonomy, bootstrap labelling postmortems, generate succinct contexts and labels with ML.

Ideas: Leverage hierarchy in taxonomy, train text encoders w.r.to tags, finetuning gap sentence summarization.

Opensource Taxonomy: Share wide variety of contributing factors with others and develop continuously.

Thank you!

Join Us: https://autoarts-rca-taxonomy.github.io/

Contact: dogga@cs.ucla.edu
http://web.cs.ucla.edu/~dogga