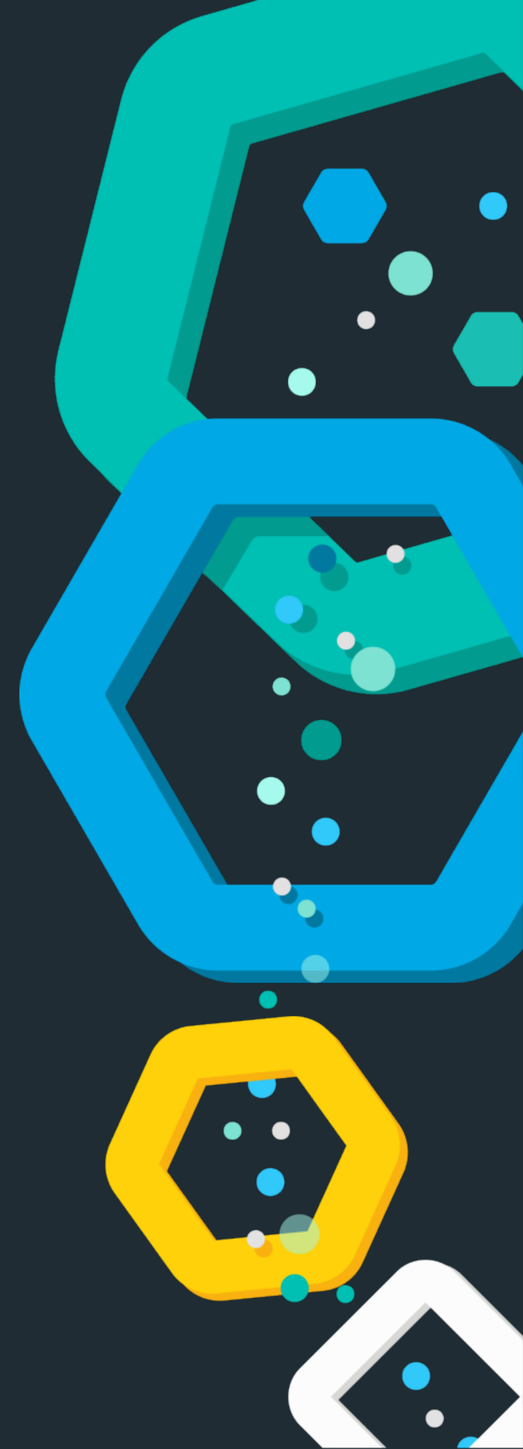
Elasticsearch Ingest Processors

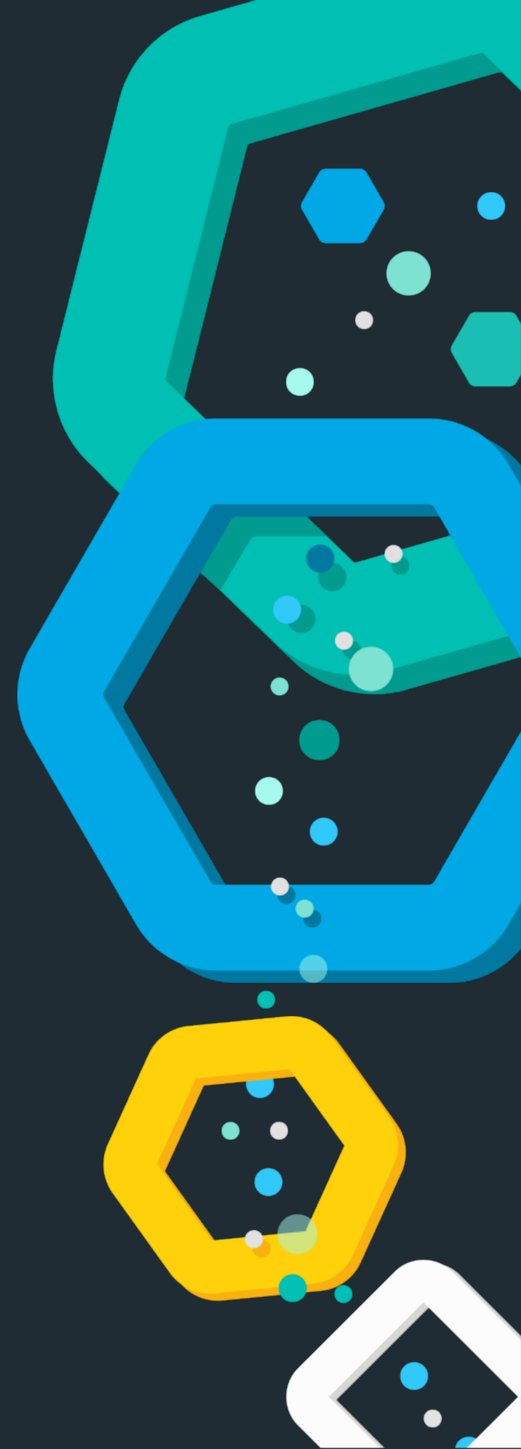
Alexander Reelsen
alex@elastic.co
@spinscale

Luca Wintergerst
luca.wintergerst@elastic.co
@LucaWintergerst

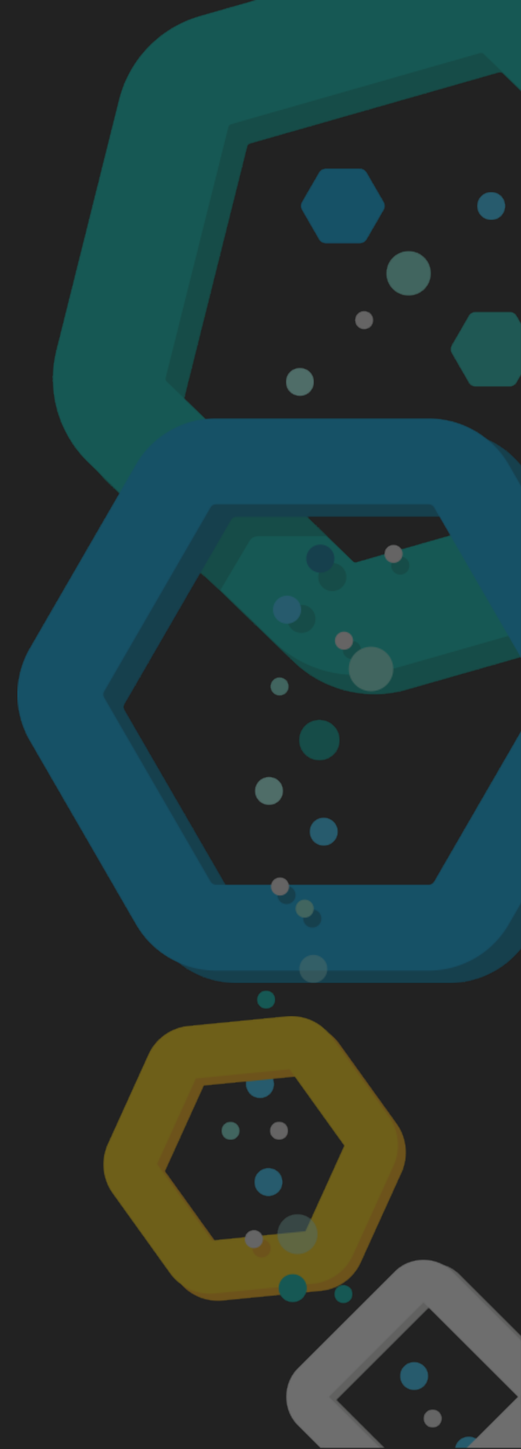


Agenda

- Update
- Writing your own processors
- Use-Cases
- Discussion



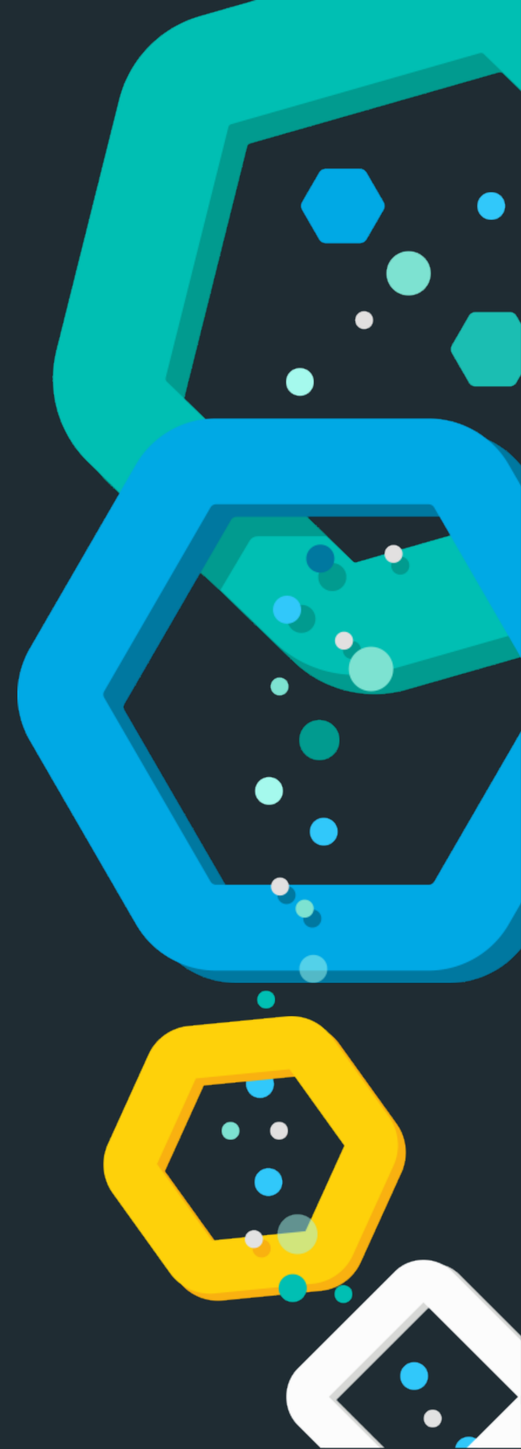
Update



elastic

New processors

- bytes (convert to human readable bytes)
- dissect (grok without regexes, much faster)
- pipeline processor, referring to other pipelines

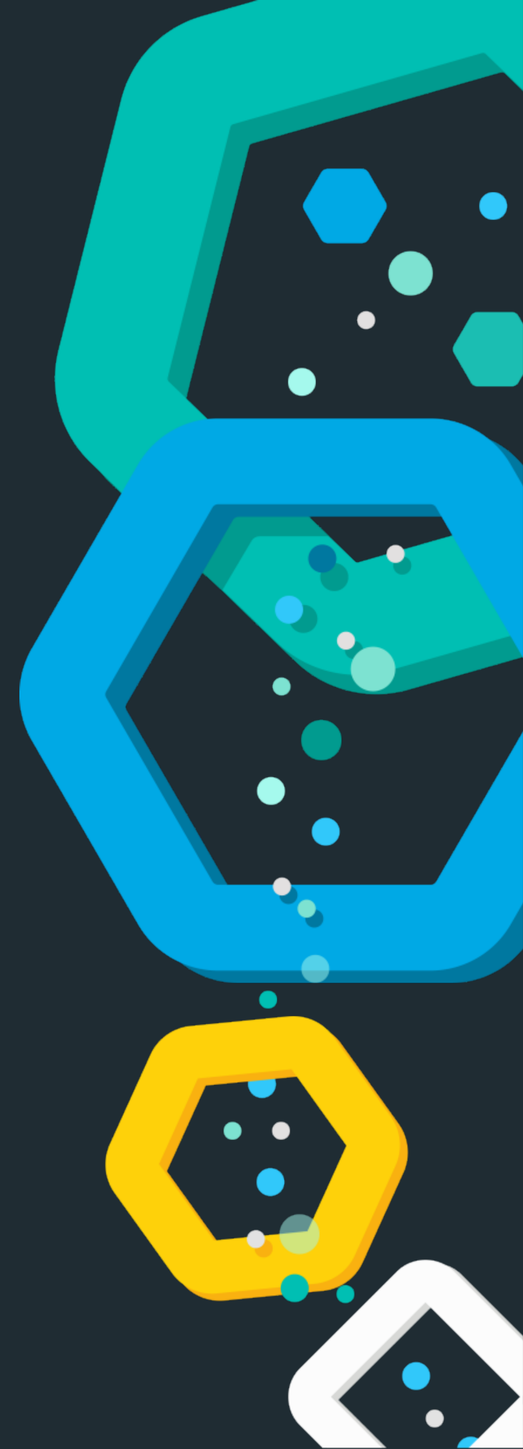


New processors

- - drop processor to fully drop an event
  - `"drop" : { "if": "ctx.foo == 'bar'" }`
- - scripting can invoke other processors
  - `"ctx.target_field = Processors.bytes(ctx.source_field)"`
- `if` in every processor using scripting

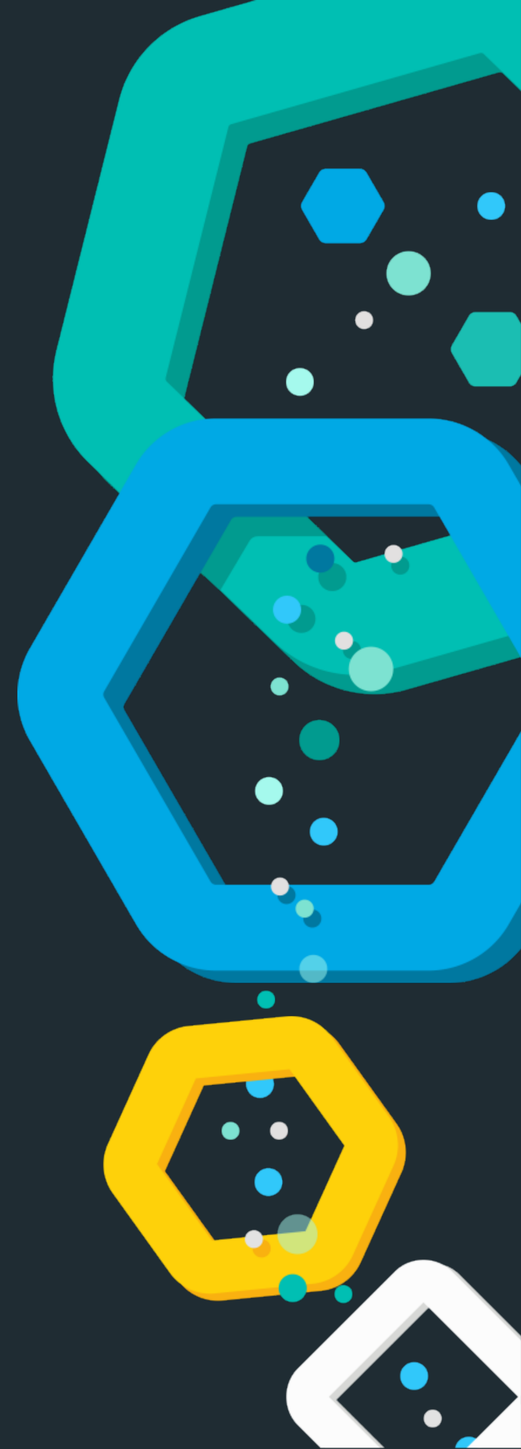
Others

- performance bump in **`geoip`** processor
- per processor metrics
- index default pipeline:
  - **`settings.index.default_pipeline: "my_pipeline"`**

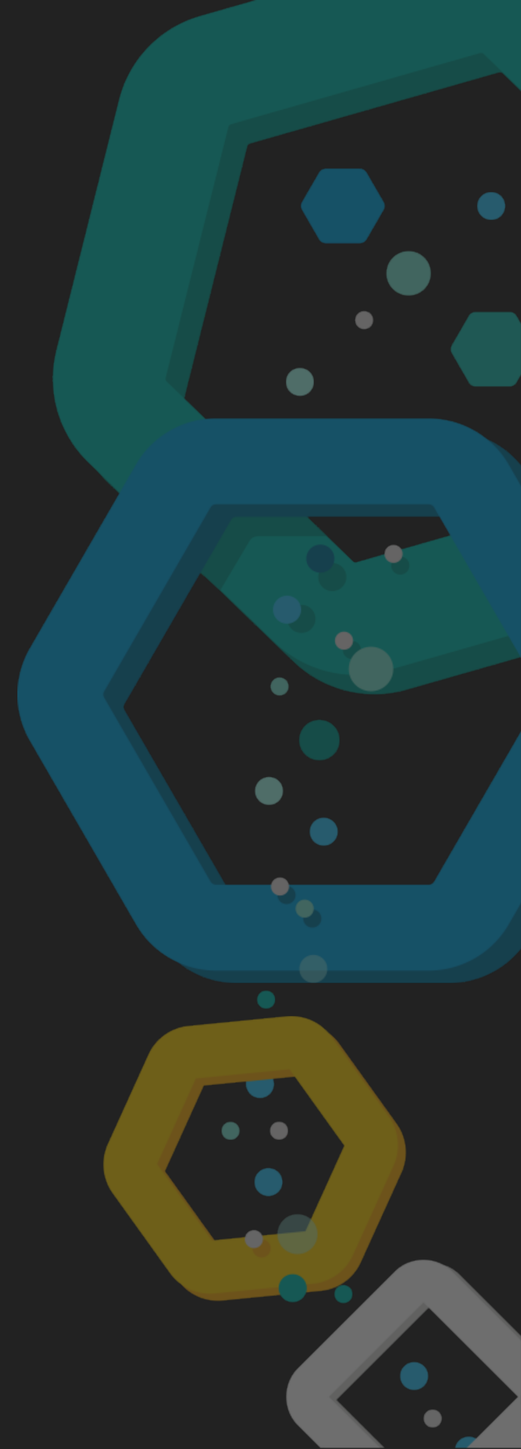


Future

- Aligning dissect filters in logstash/beats/ES
  - https://github.com/elastic/dissect-specification
- UI



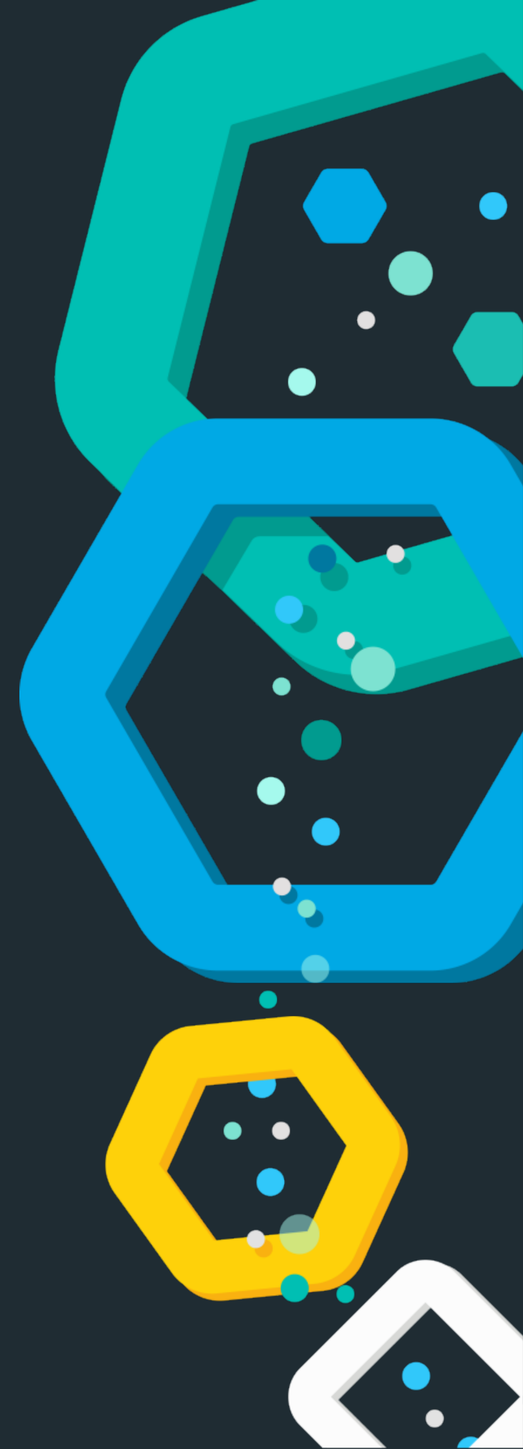
Writing your own



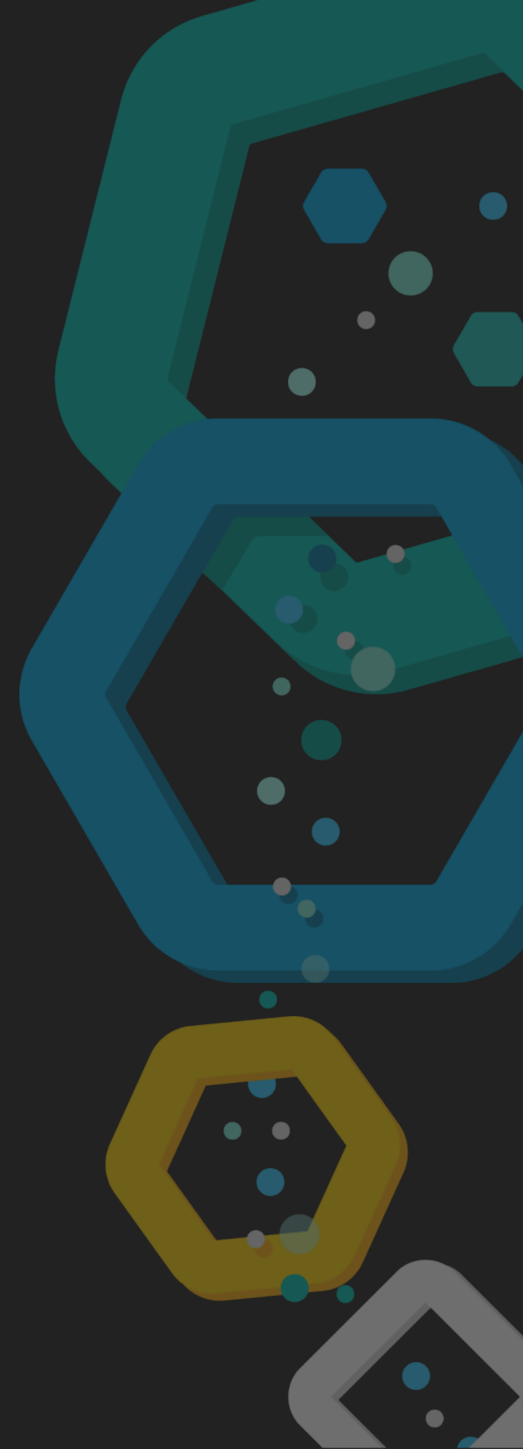
elastic

Write your own ingest plugin

- https://github.com/spinscale/cookiecutter-elasticsearch-ingest-processor
- https://github.com/spinscale/elasticsearch-ingest-langdetect
- https://github.com/spinscale/elasticsearch-ingest-opennlp



Use-Cases



elastic

Discussion

… ask all the things!