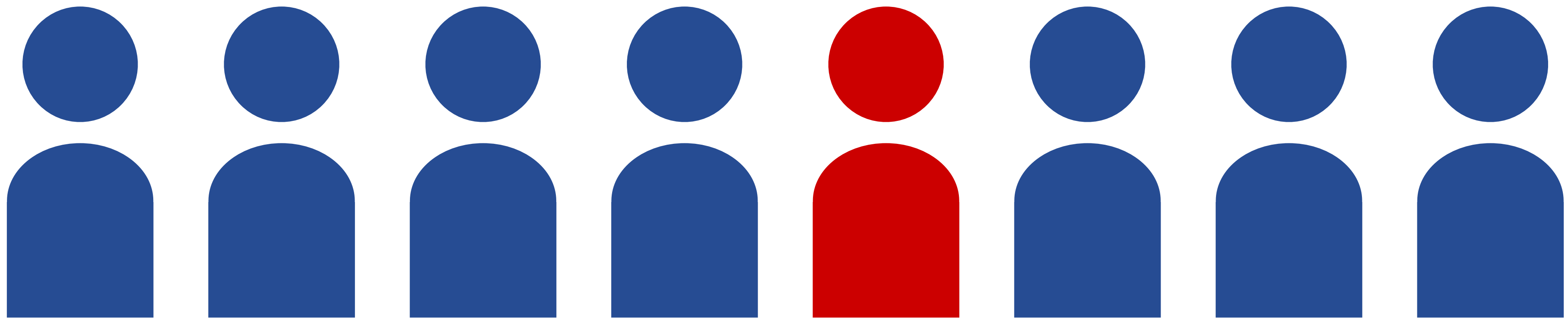
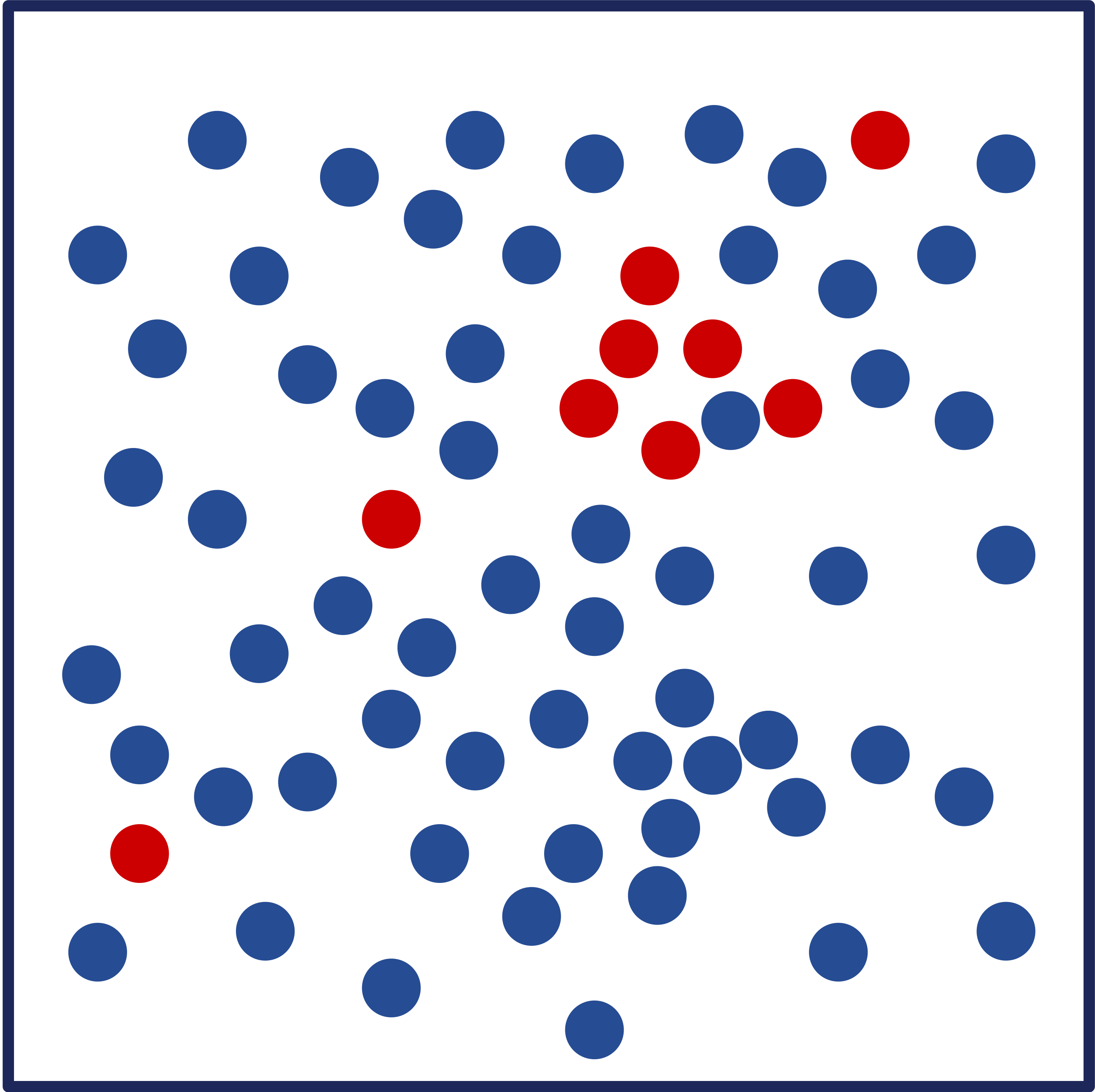


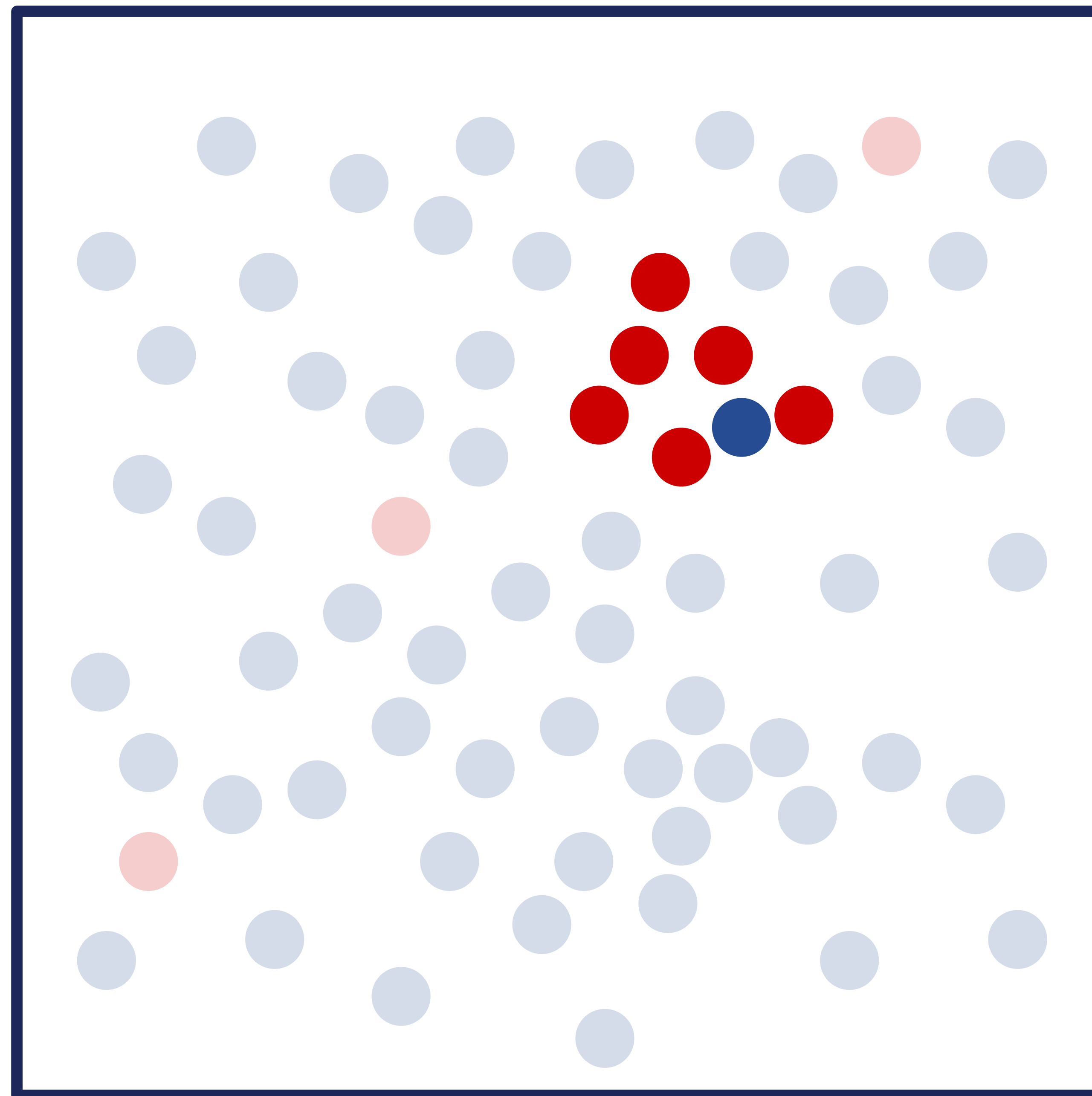
Machine learning and discovery with Kubernetes

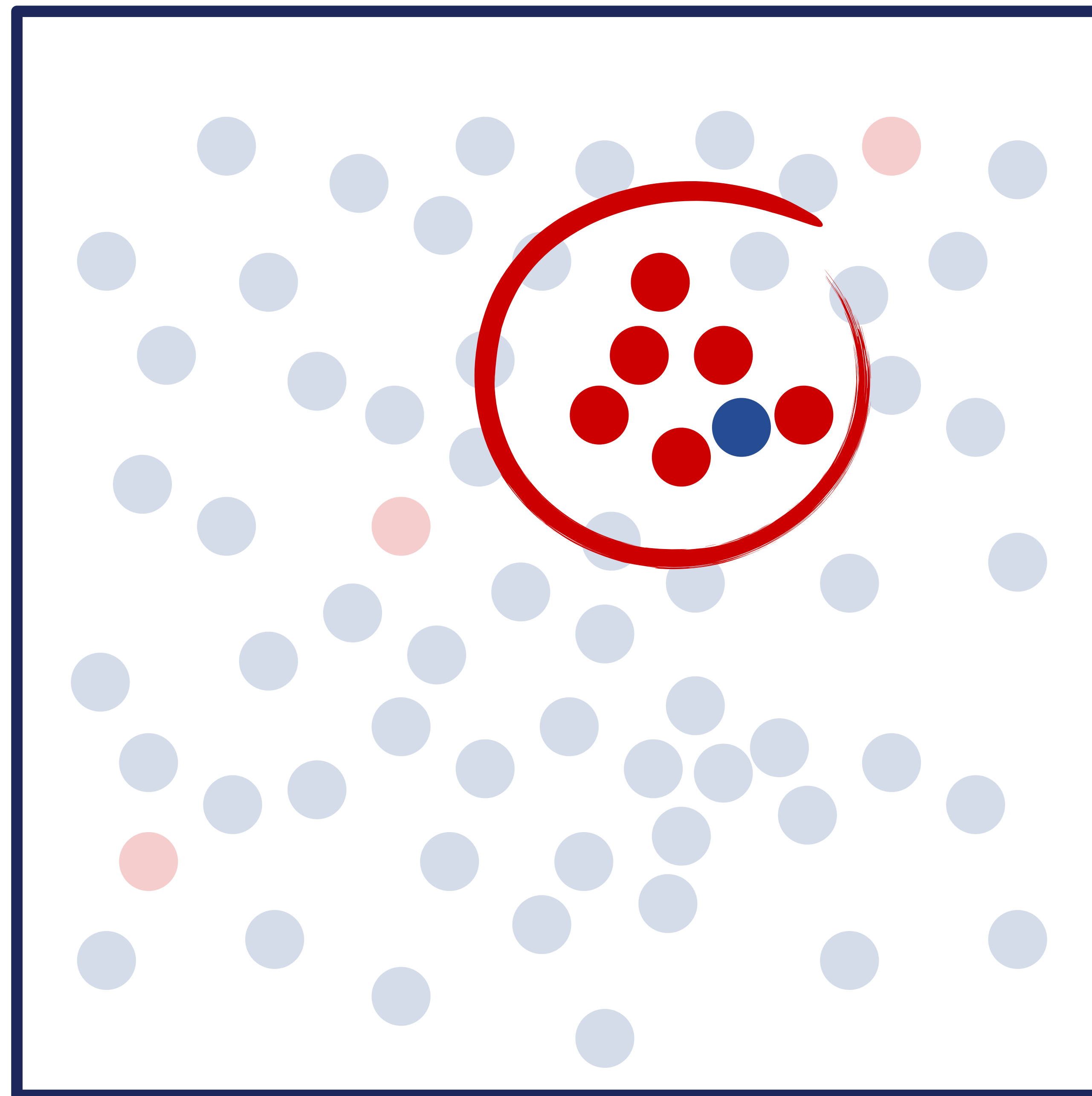
William Benton • @willb • willb@redhat.com

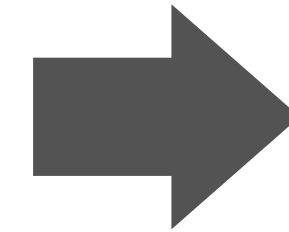
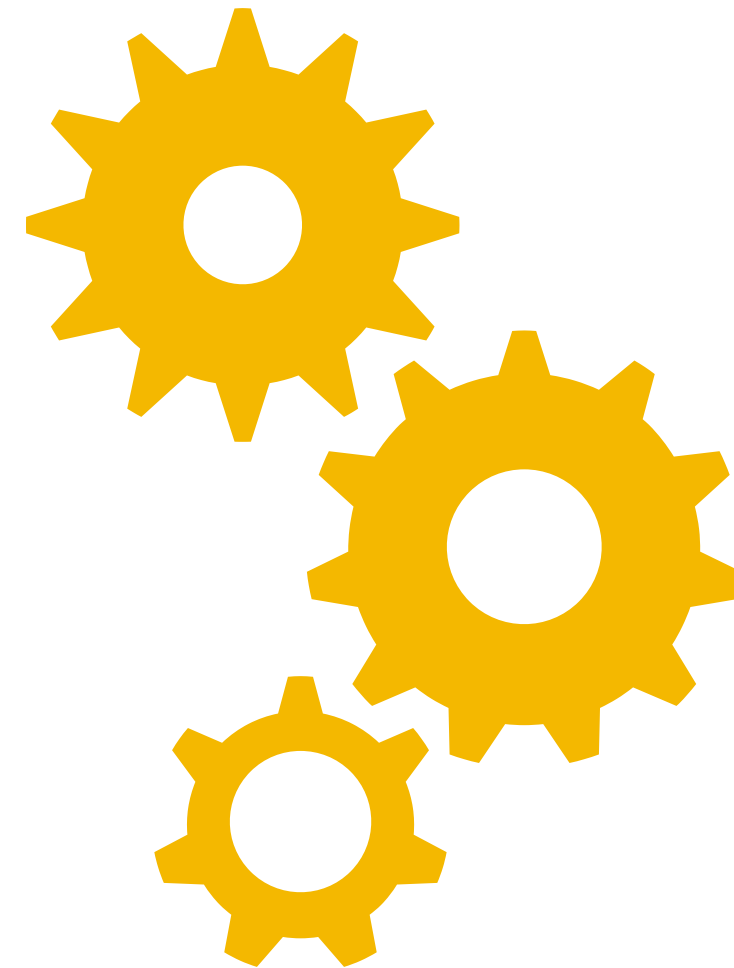
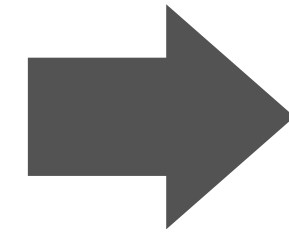
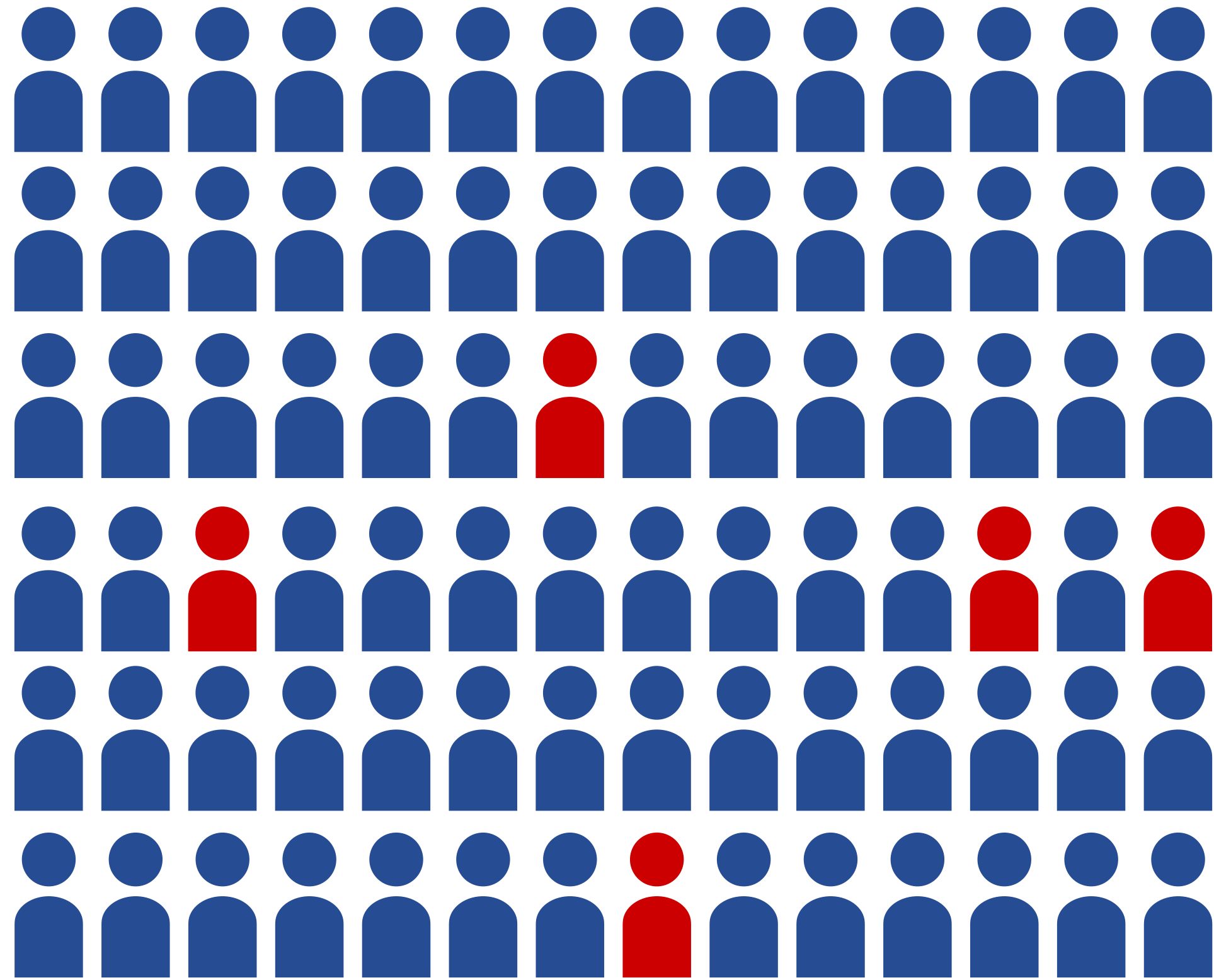
**What do machine learning
workflows look like?**

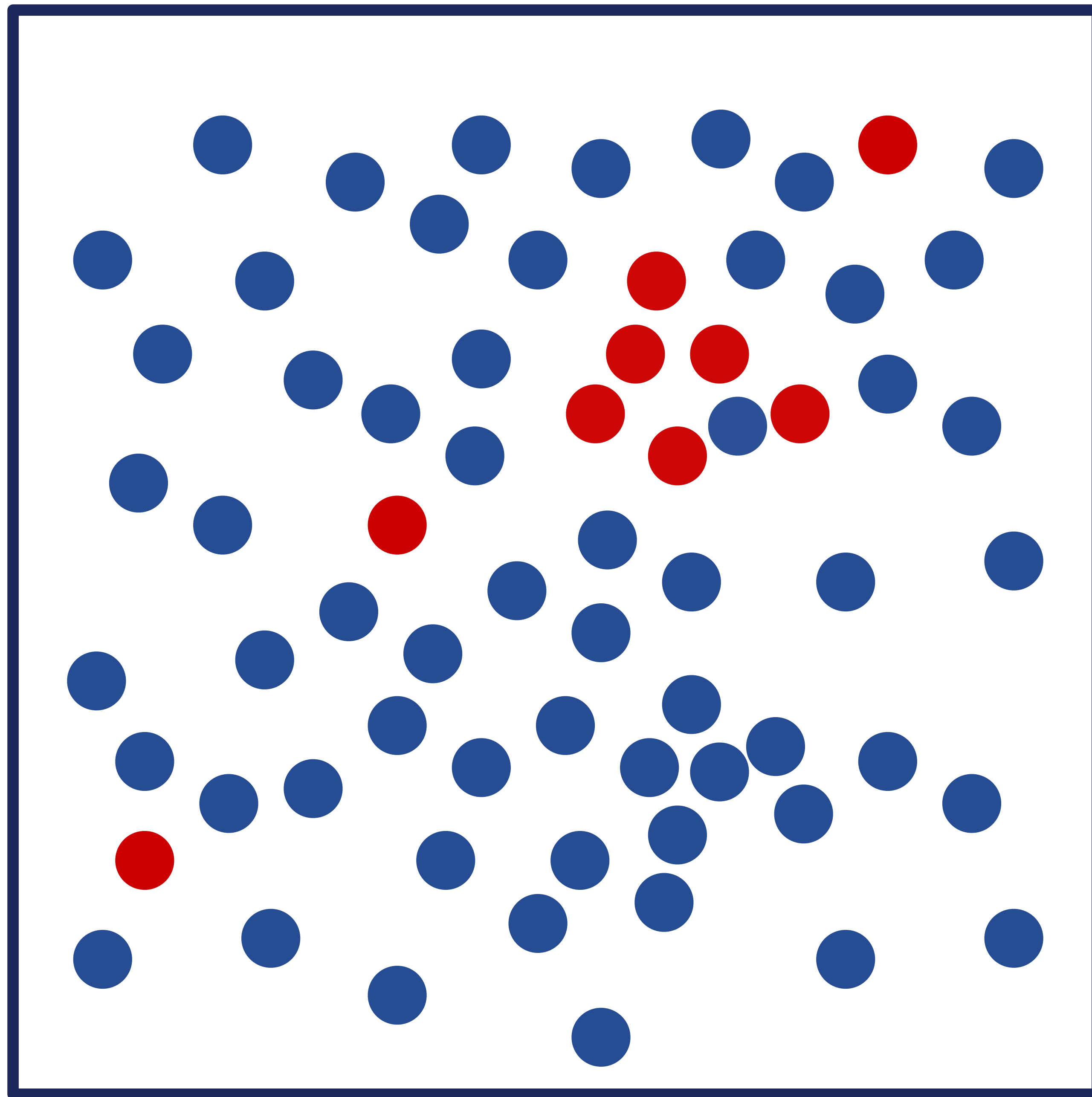


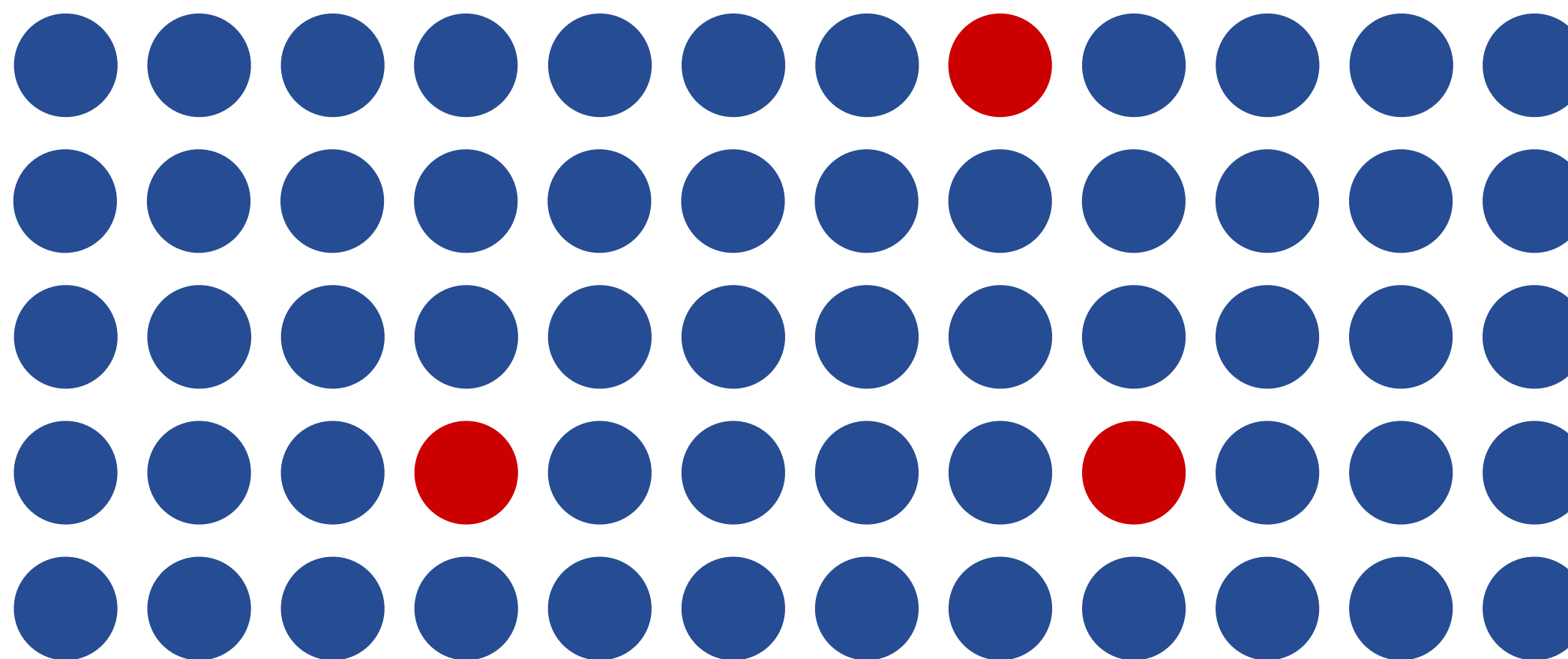
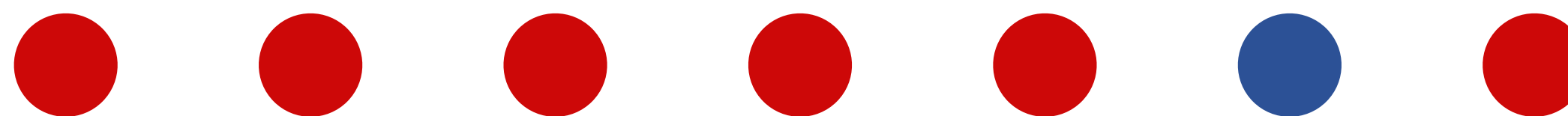




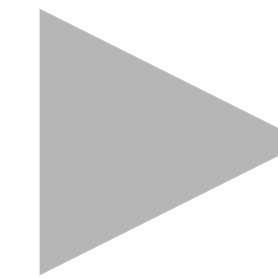








**codifying problem
and metrics**



**data collection
and cleaning**



**data collection
and cleaning**

**feature
engineering**

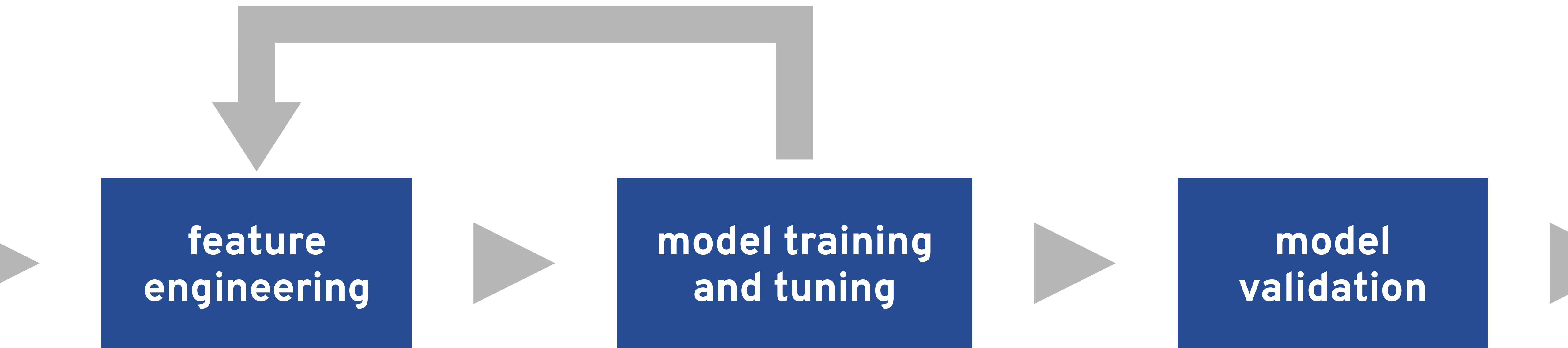
**model training
and tuning**

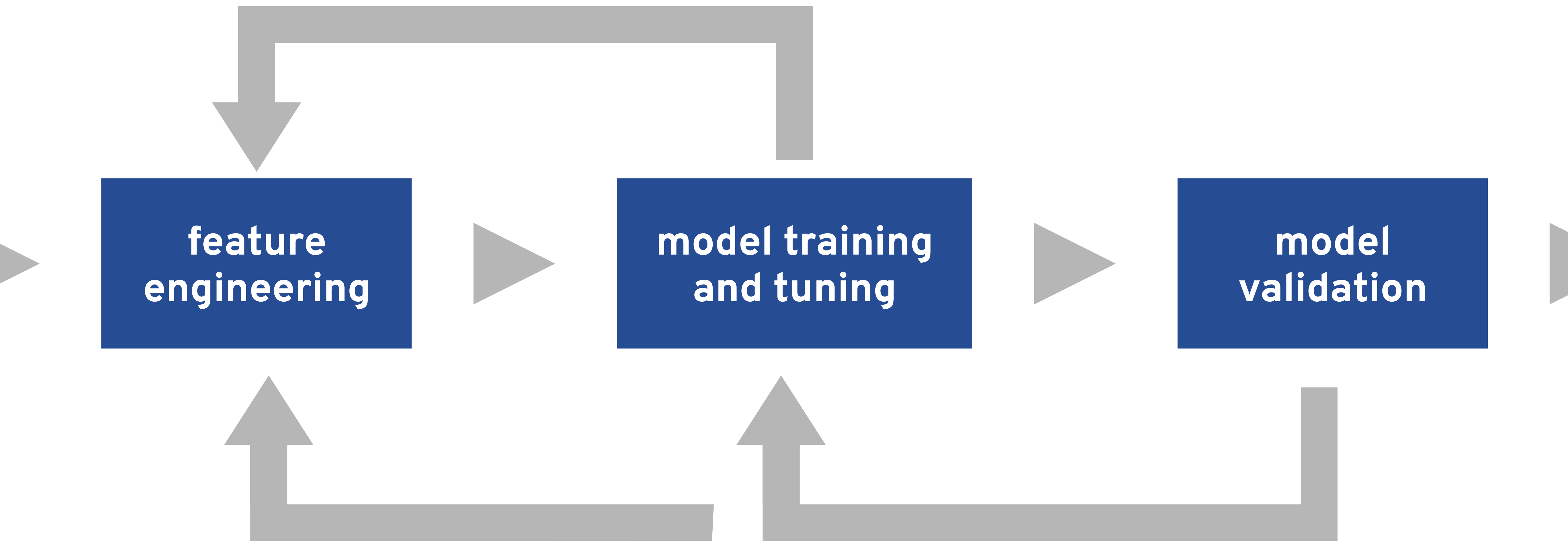
```
graph LR; A[data collection and cleaning] --> B[feature engineering]; B --> C[model training and tuning]; C --> B;
```

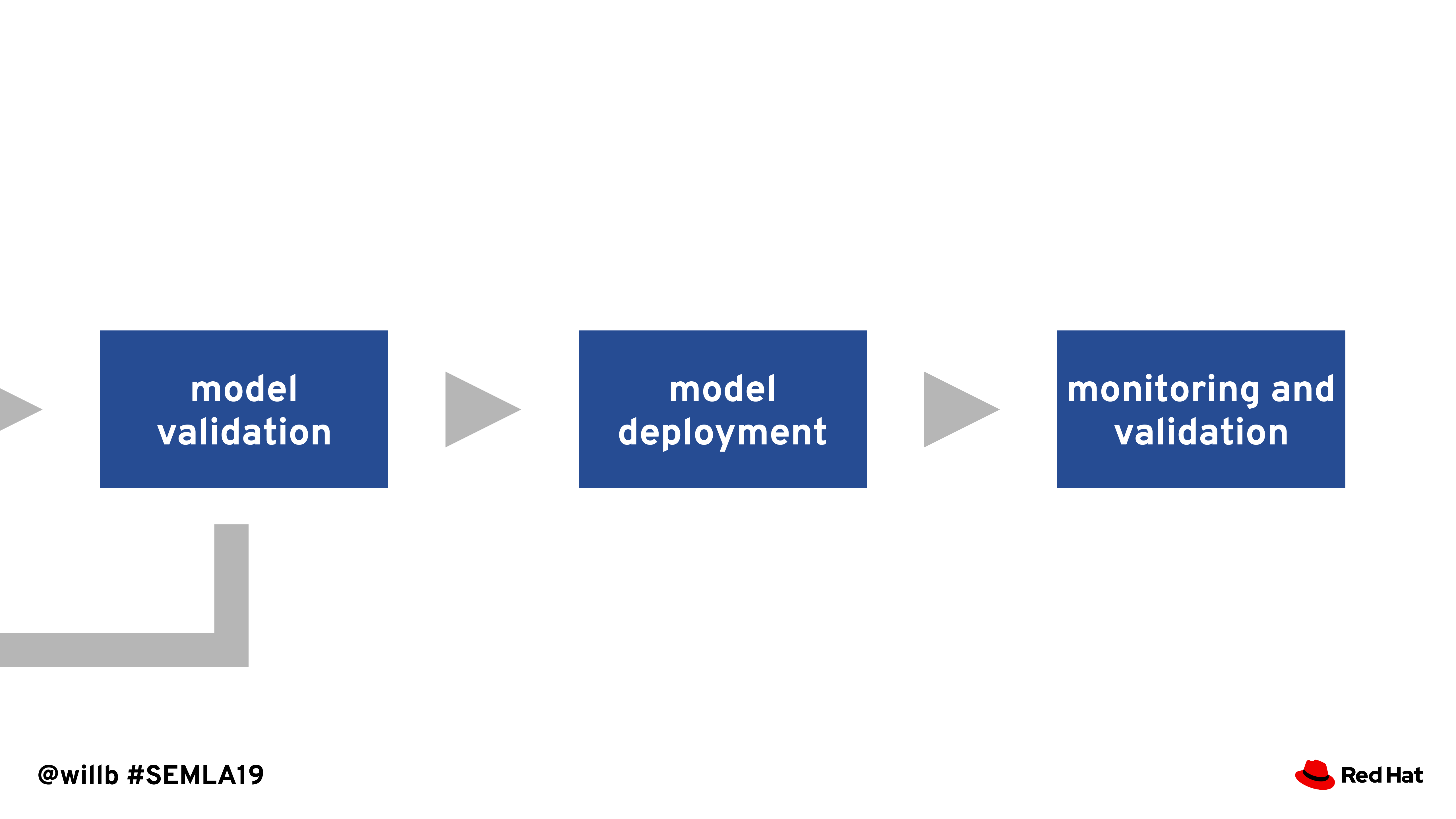
**data collection
and cleaning**

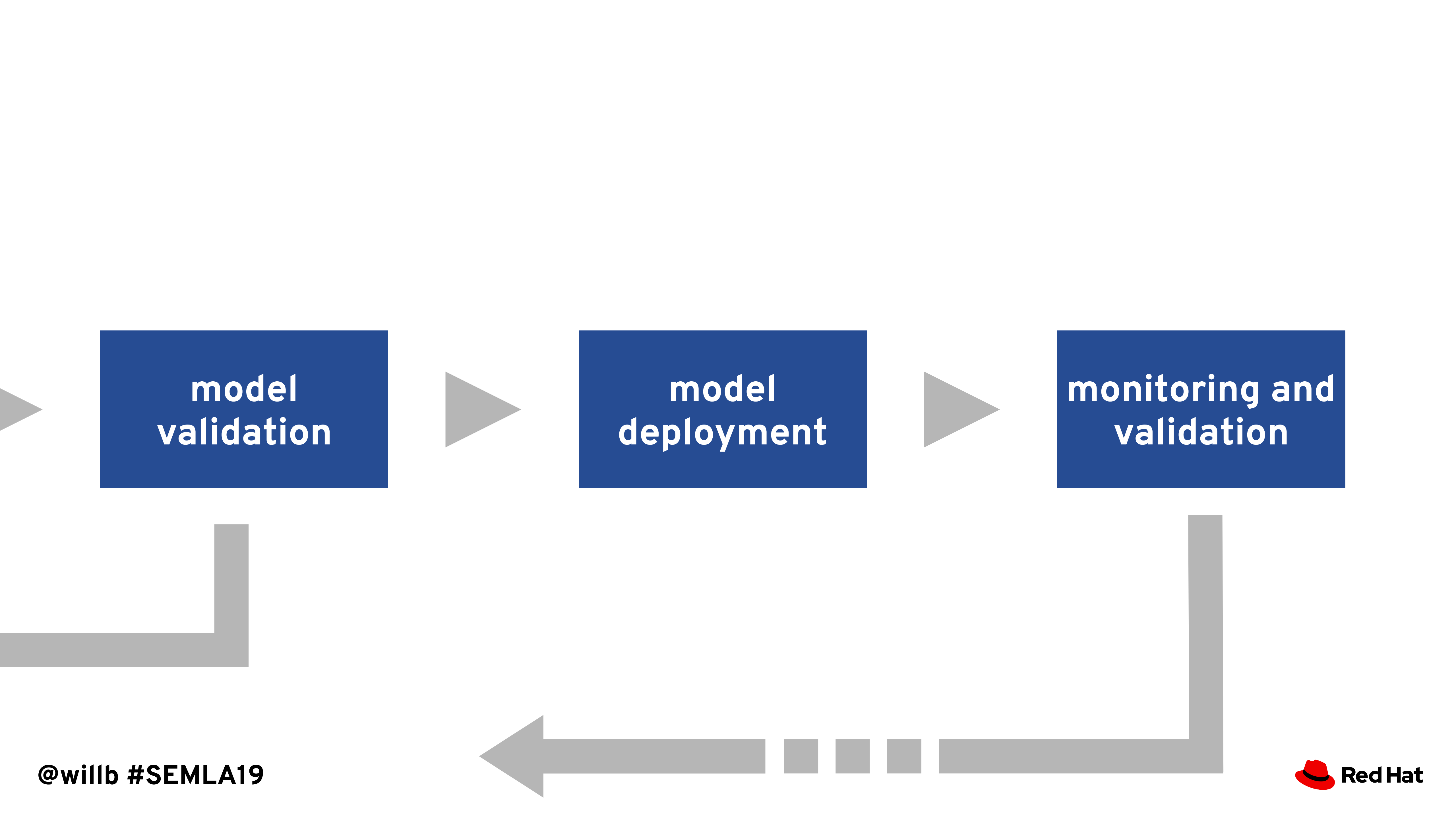
**feature
engineering**

**model training
and tuning**





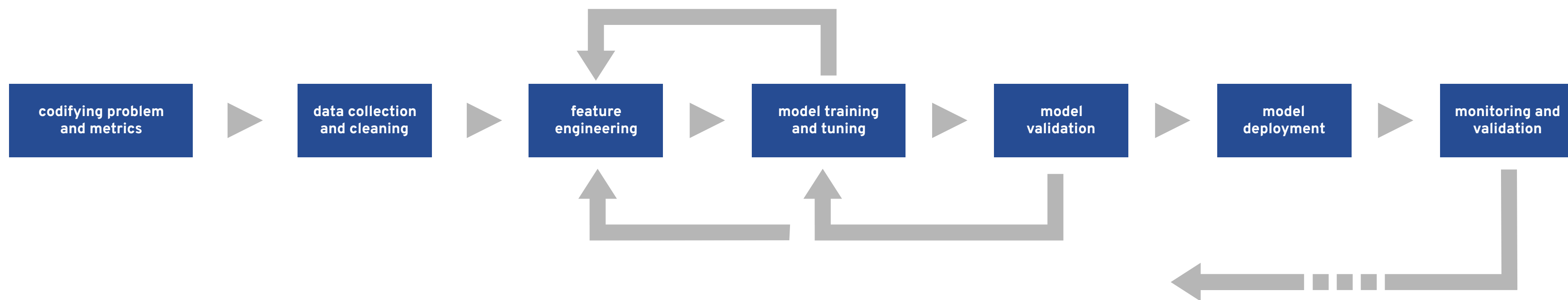


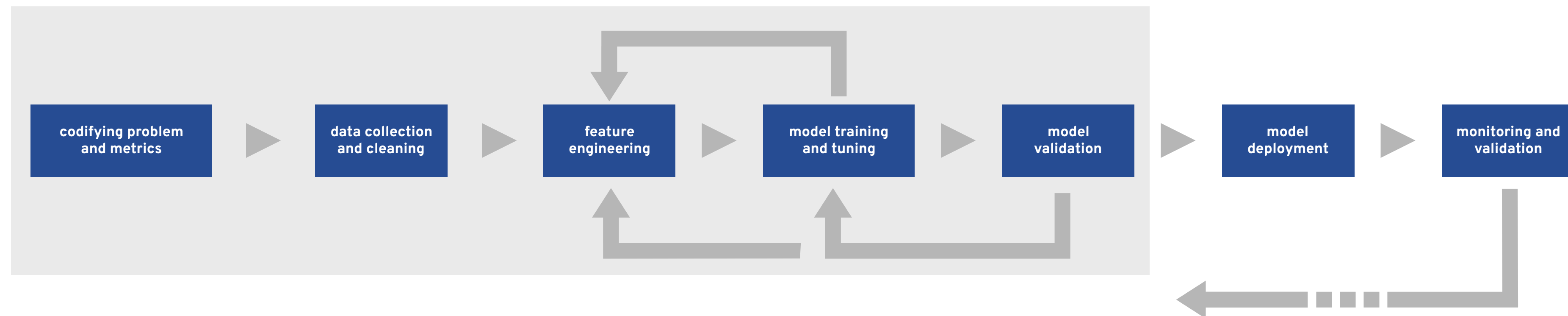


**model
validation**

**model
deployment**

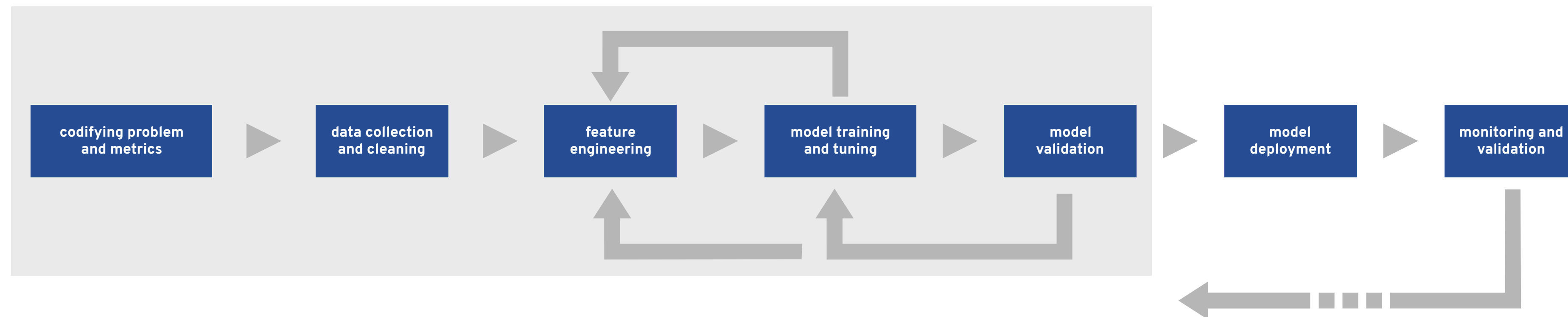
**monitoring and
validation**





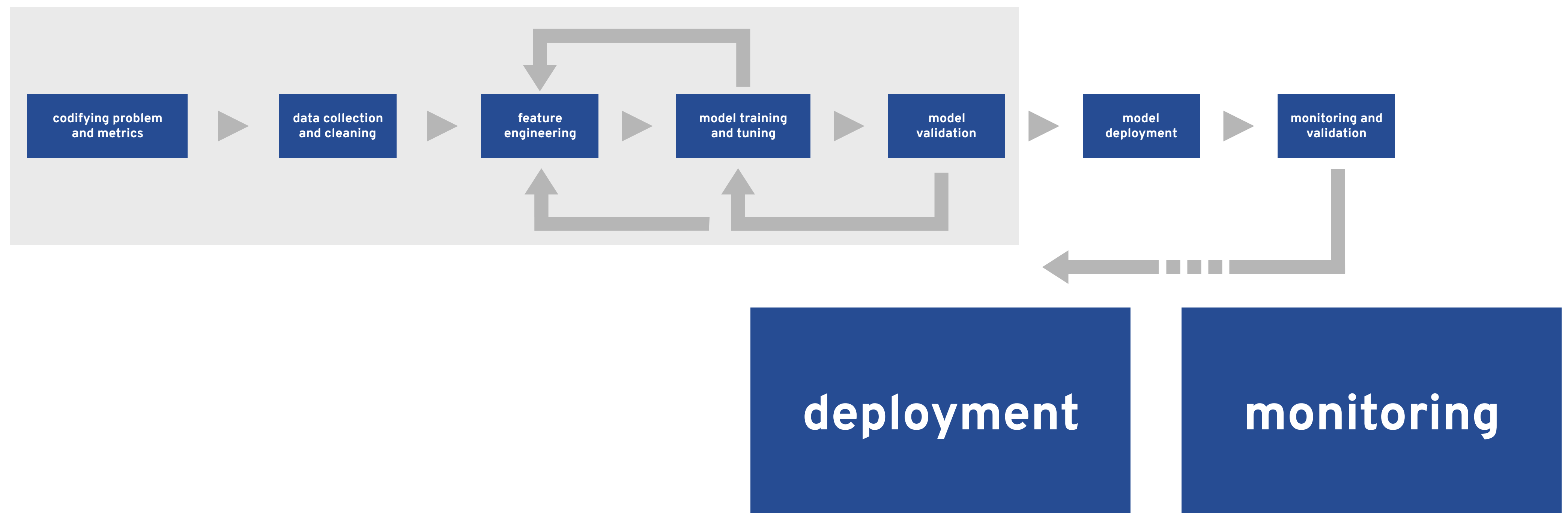
**defining types
and interfaces**

prototyping

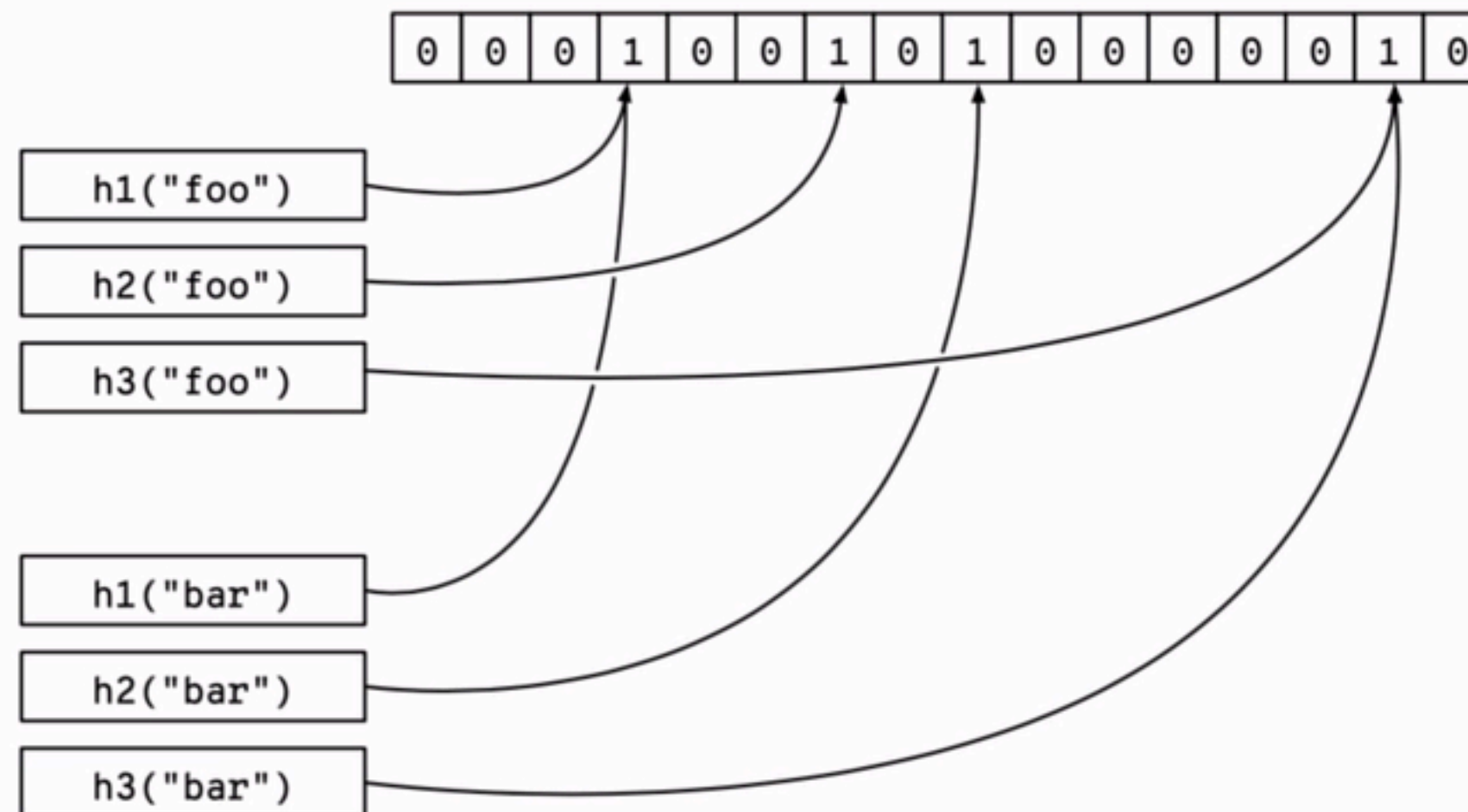


**unit, behavioral, and
integration testing**

**formal
verification**



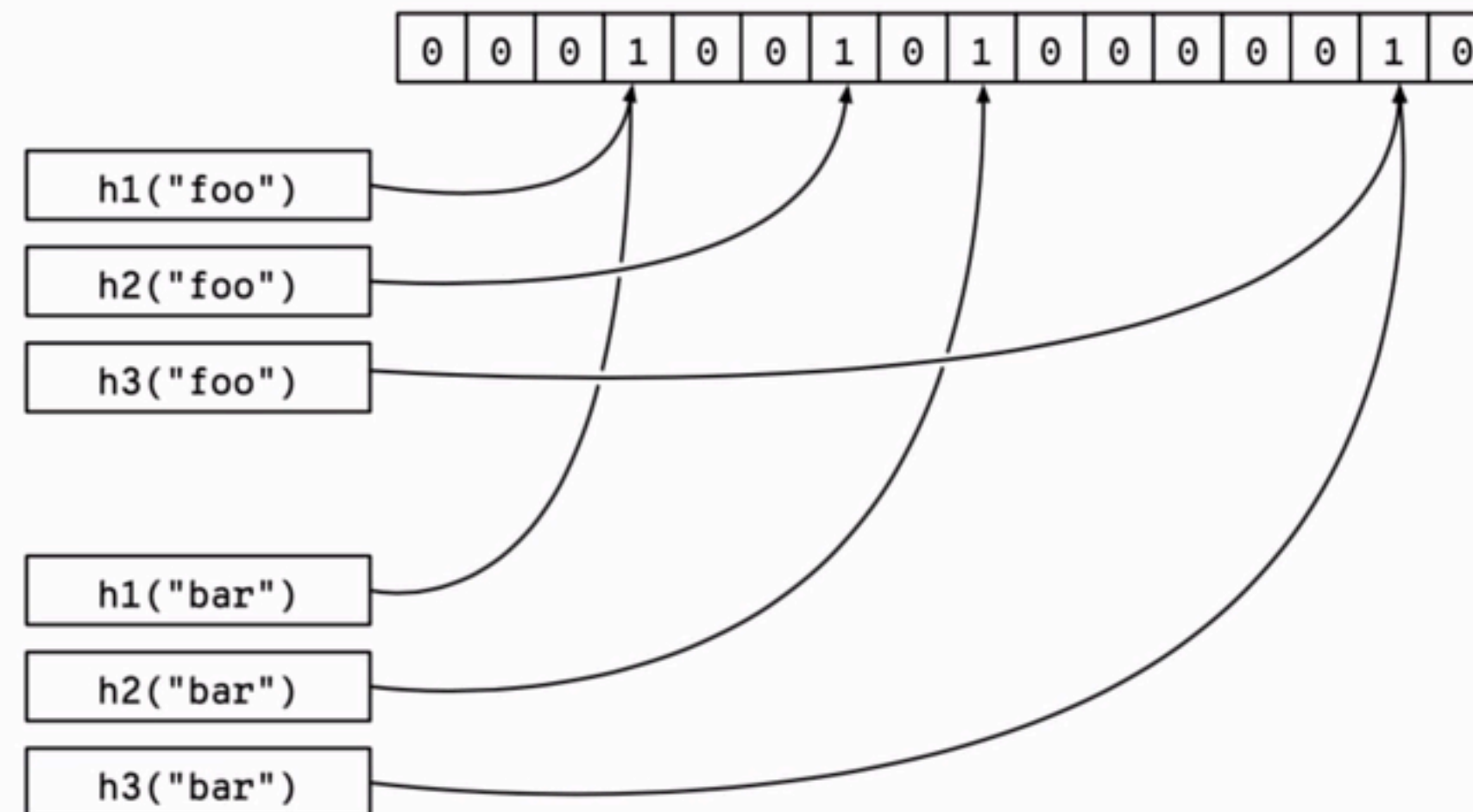
Bloom filter ¶



A conventional hash table (or hash table-backed set structure) consists of a series of *buckets*. Hash table insert looks like this:

1. First, use the hash value of the key to identify the index of the bucket that should contain it.
2. If the bucket is empty, update the bucket to contain the key and value (with a trivial value in the case of a hashed set).
3. If the bucket is not empty and the key stored in it is not the one you've hashed, handle this *hash collision*. There are several strategies to handle hash collisions precisely; most involve extra lookups (e.g., having a second hash function or going to the next available bucket) or

Bloom filter ¶



A conventional hash table (or hash table-backed set structure) consists of a series of *buckets*. Hash table insert looks like this:

1. First, use the hash value of the key to identify the index of the bucket that should contain it.
2. If the bucket is empty, update the bucket to contain the key and value (with a trivial value in the case of a hashed set).
3. If the bucket is not empty and the key stored in it is not the one you've hashed, handle this *hash collision*. There are several strategies to handle hash collisions precisely; most involve extra lookups (e.g., having a second hash function or going to the next available bucket) or

What's a container?

%


```
%pip install numpy
```

executable

`/usr/bin/pip`

arguments

`pip install numpy`

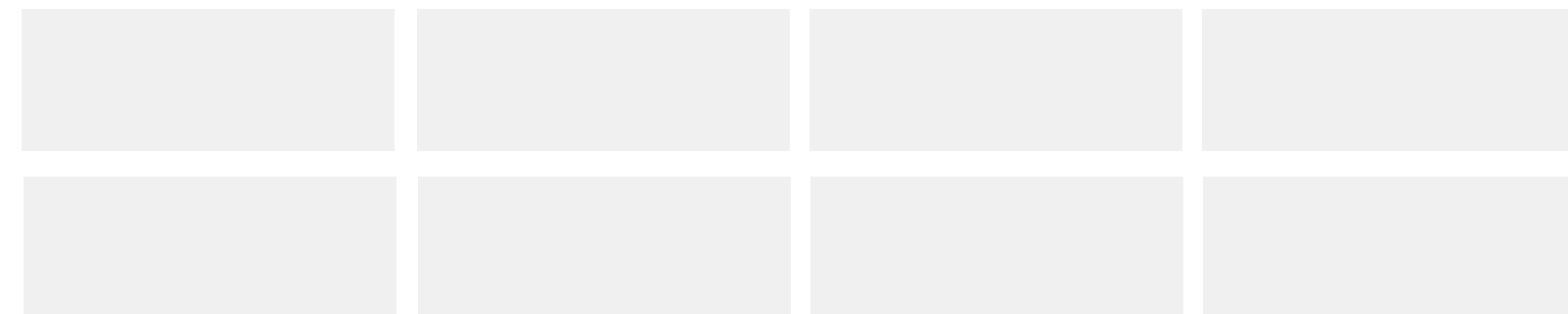
environment

`LANG=en_US USER=willb ...`

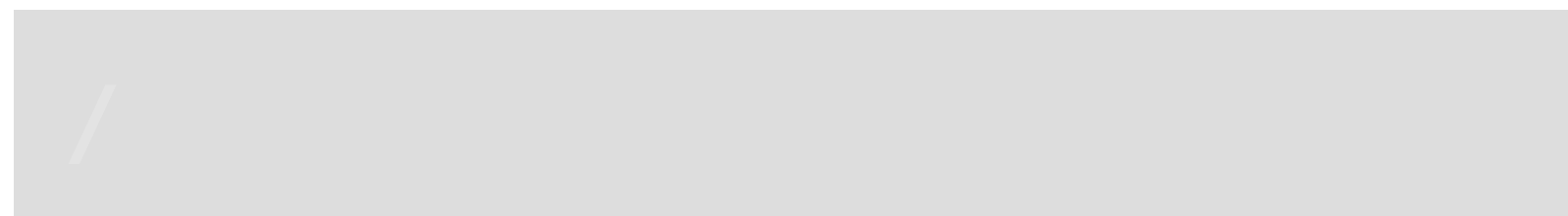
virtual memory



file handles



root filesystem



process table



network routes



executable

/usr/bin/pip

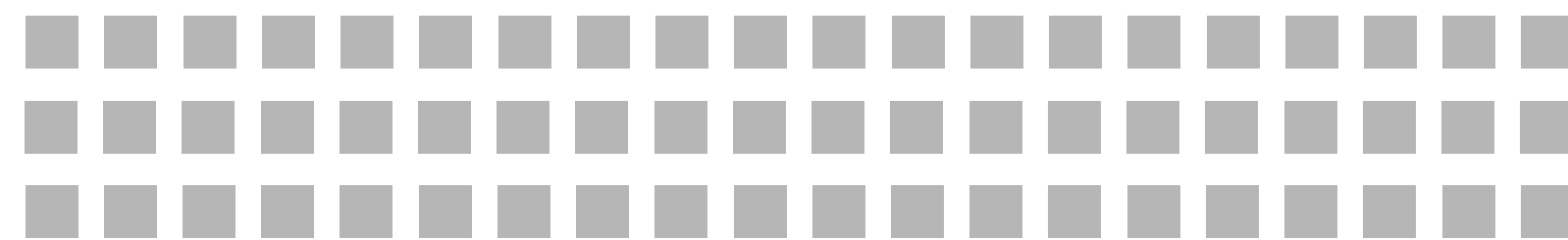
arguments

pip install numpy

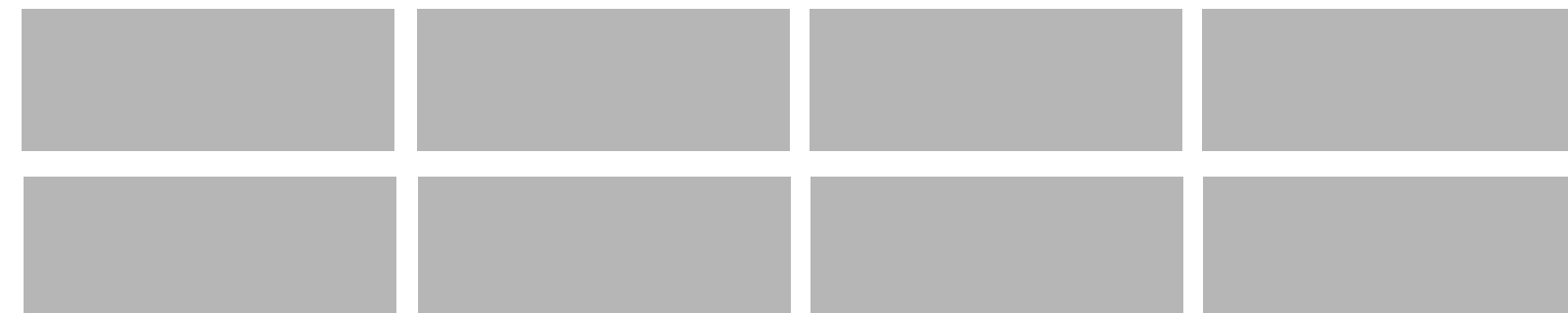
environment

LANG=en_US USER=willb ...

virtual memory



file handles



root filesystem

/

process table



network routes



executable

/usr/bin/pip

arguments

pip install numpy

Software Failure. Press left mouse button to continue.

Guru Meditation #00000004.0000AAC0

root filesystem

/

process table

network routes

executable

/usr/bin/pip

arguments

pip install numpy

Software Failure. Press left mouse button to continue.

Guru Meditation #00000004.0000AAC0

root filesystem

/

process table

network routes

executable

/usr/bin/pip

arguments

pip install numpy

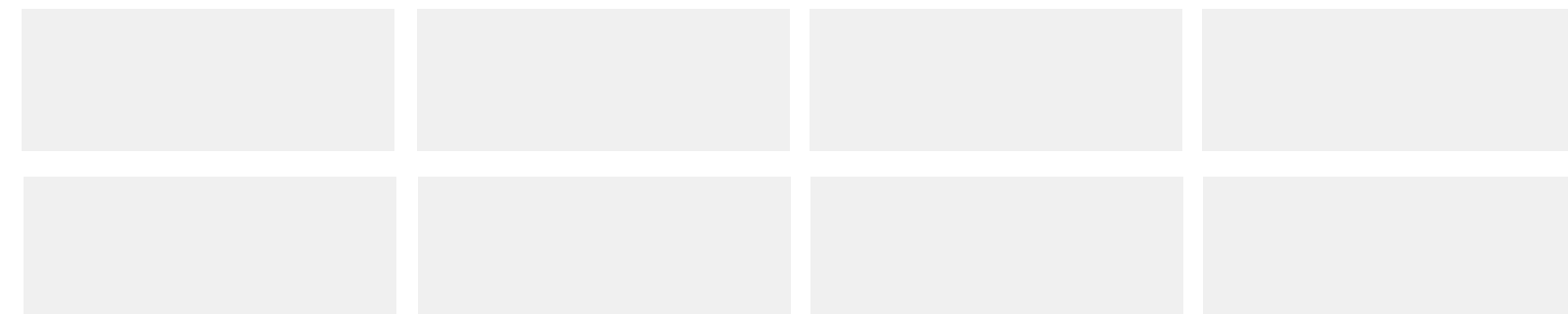
environment

LANG=en_US USER=willb ...

virtual memory



file handles



root filesystem

/

process table



network routes



executable

/usr/bin/pip

arguments

pip install numpy

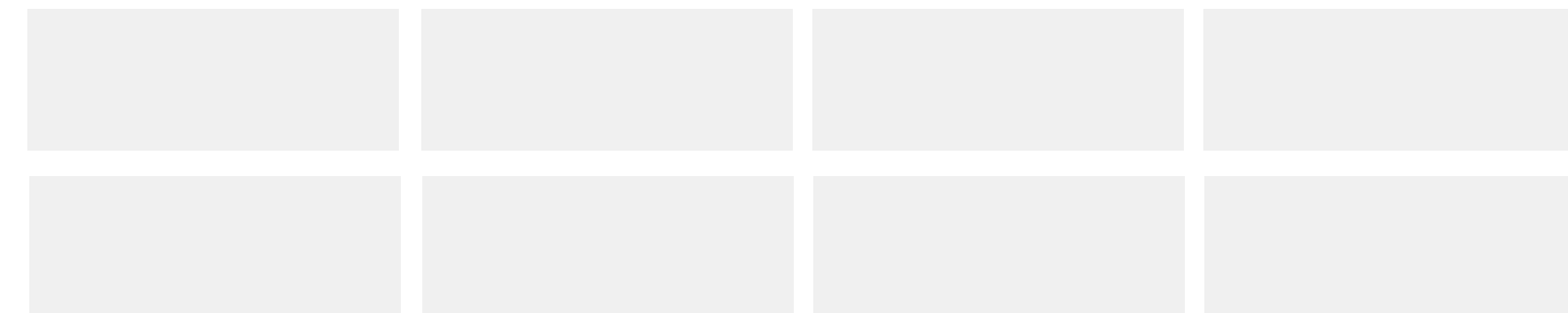
environment

LANG=en_US USER=willb ...

virtual memory



file handles



root filesystem

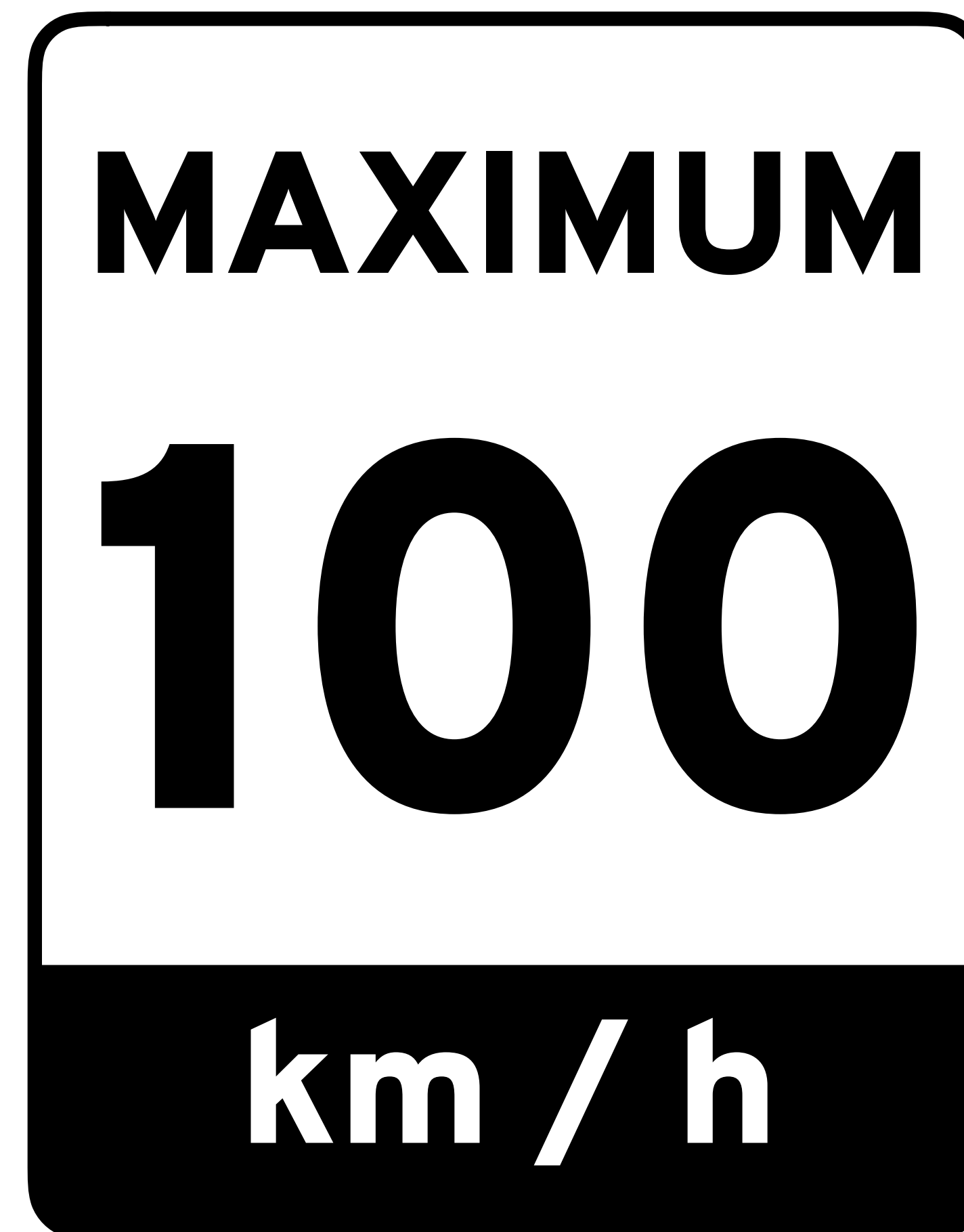
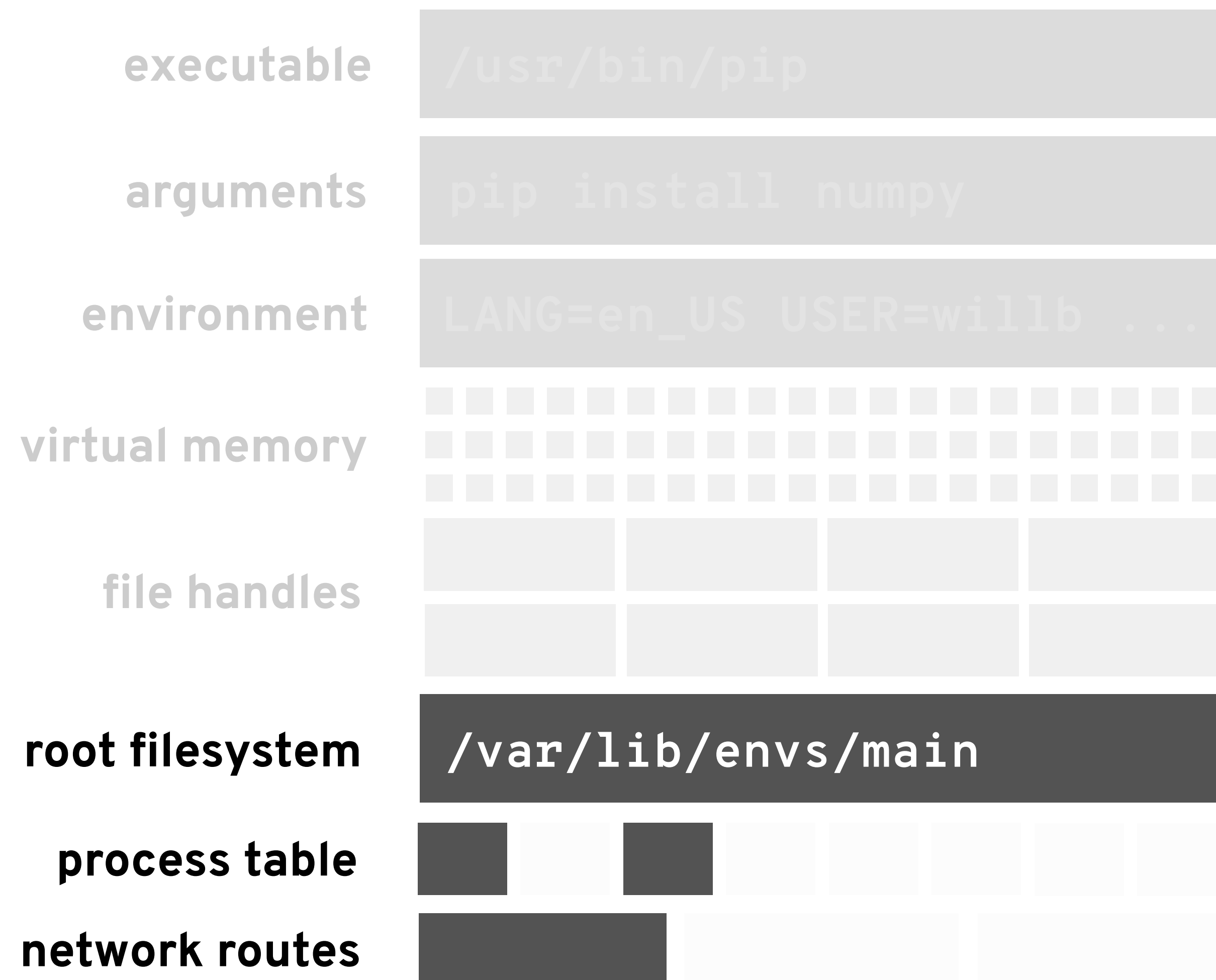
/var/lib/envs/main

process table



network routes





Immutable images

user application code

a6afd91e
6b8cad3e

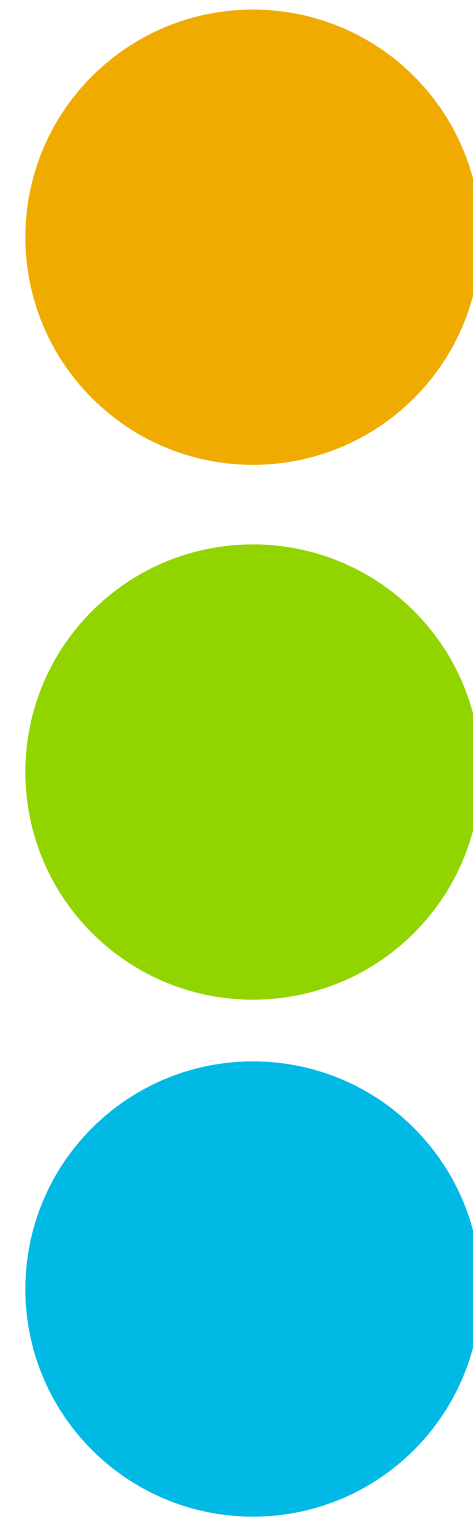
**configuration and
installation recipes**

33721112
e8cae4f6
2bb6ab16
a8296f7e

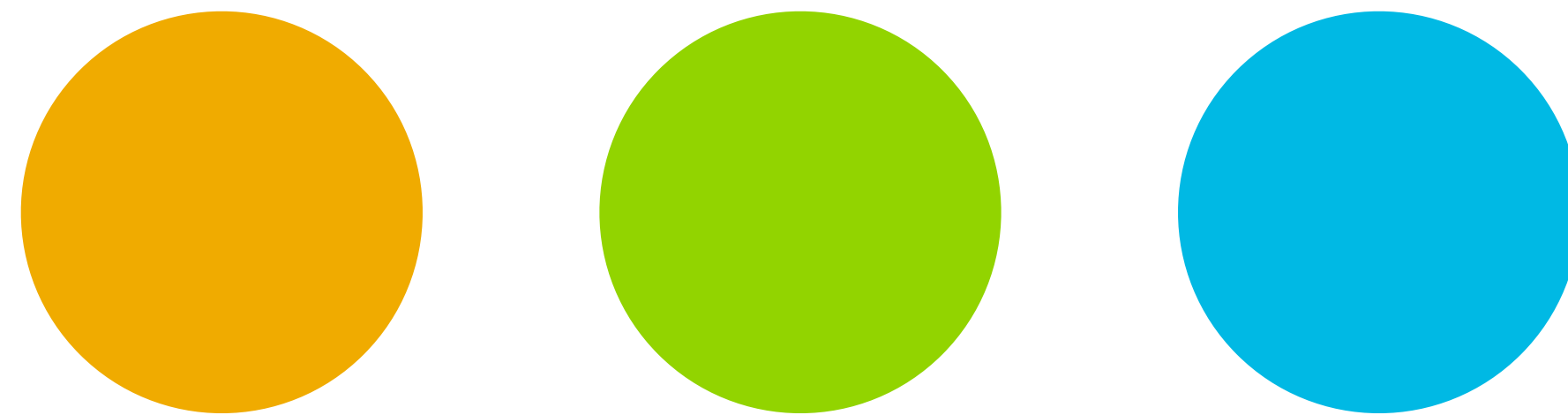
base image

979229b9

Stateless microservices



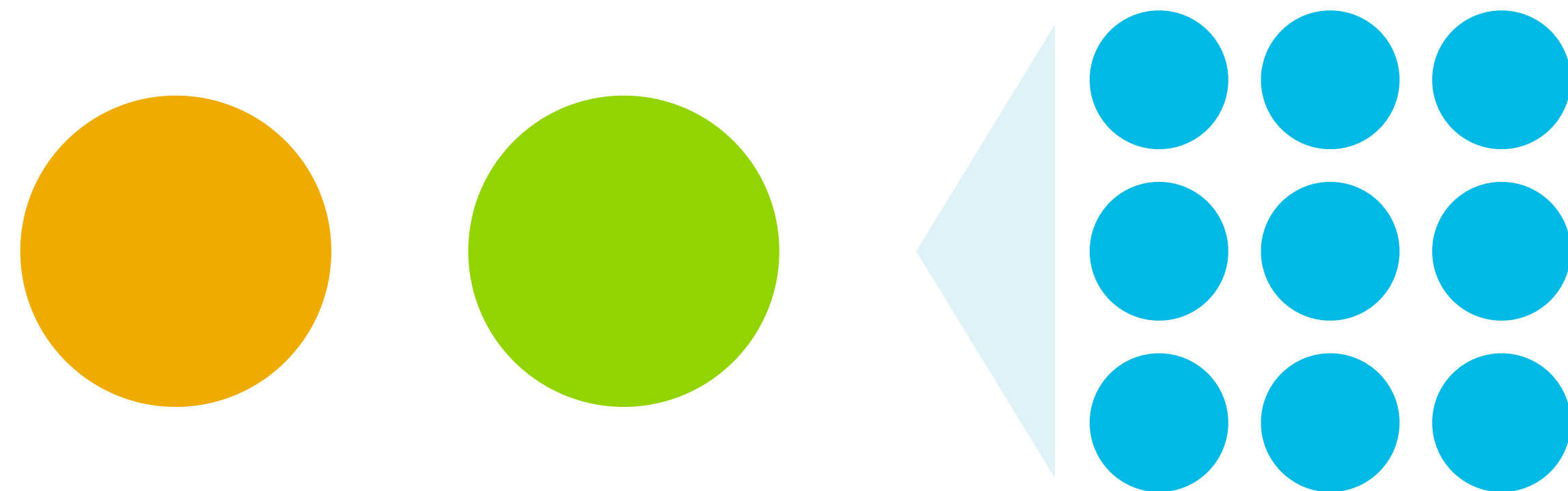
Stateless microservices



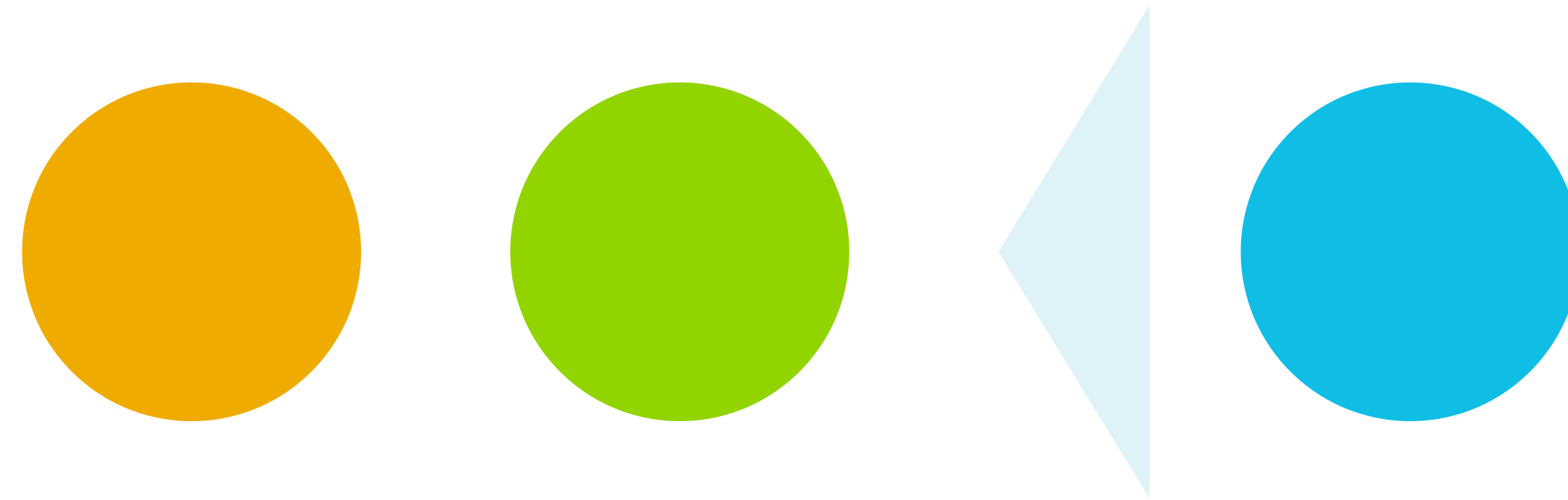
Stateless microservices



Stateless microservices



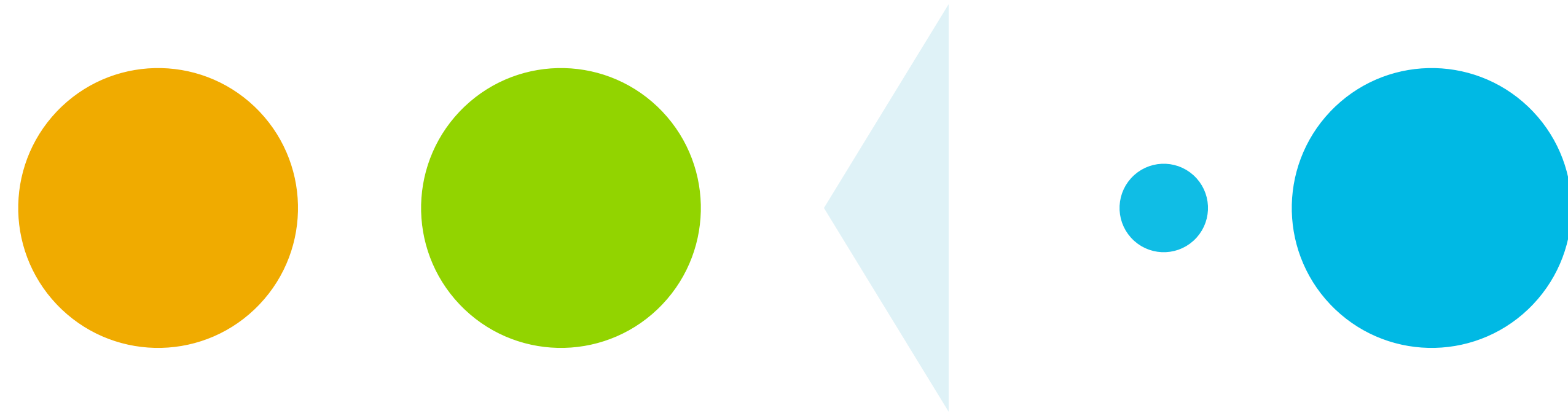
Stateless microservices



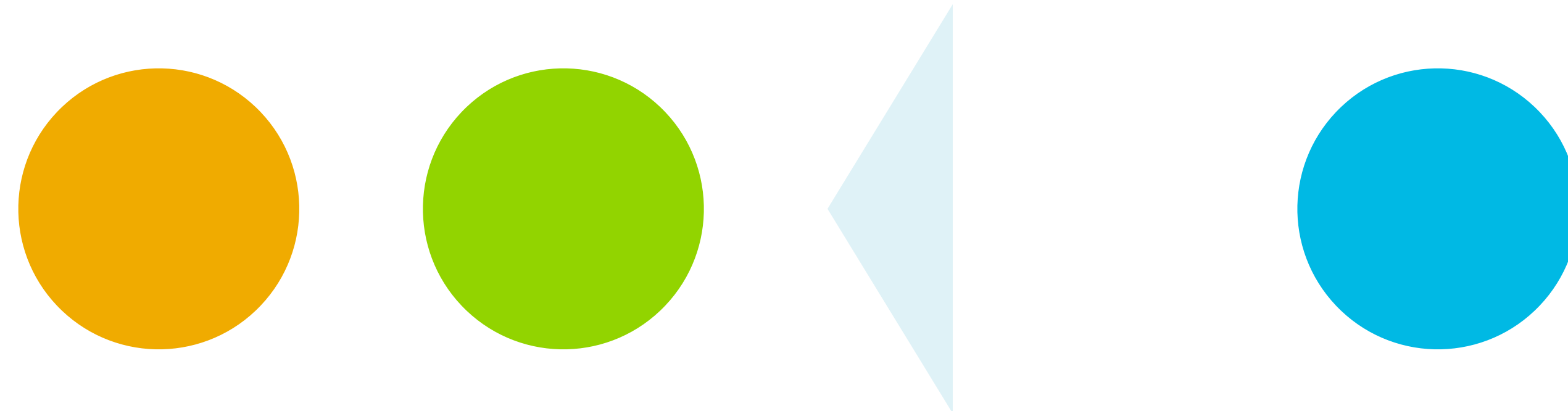
Stateless microservices



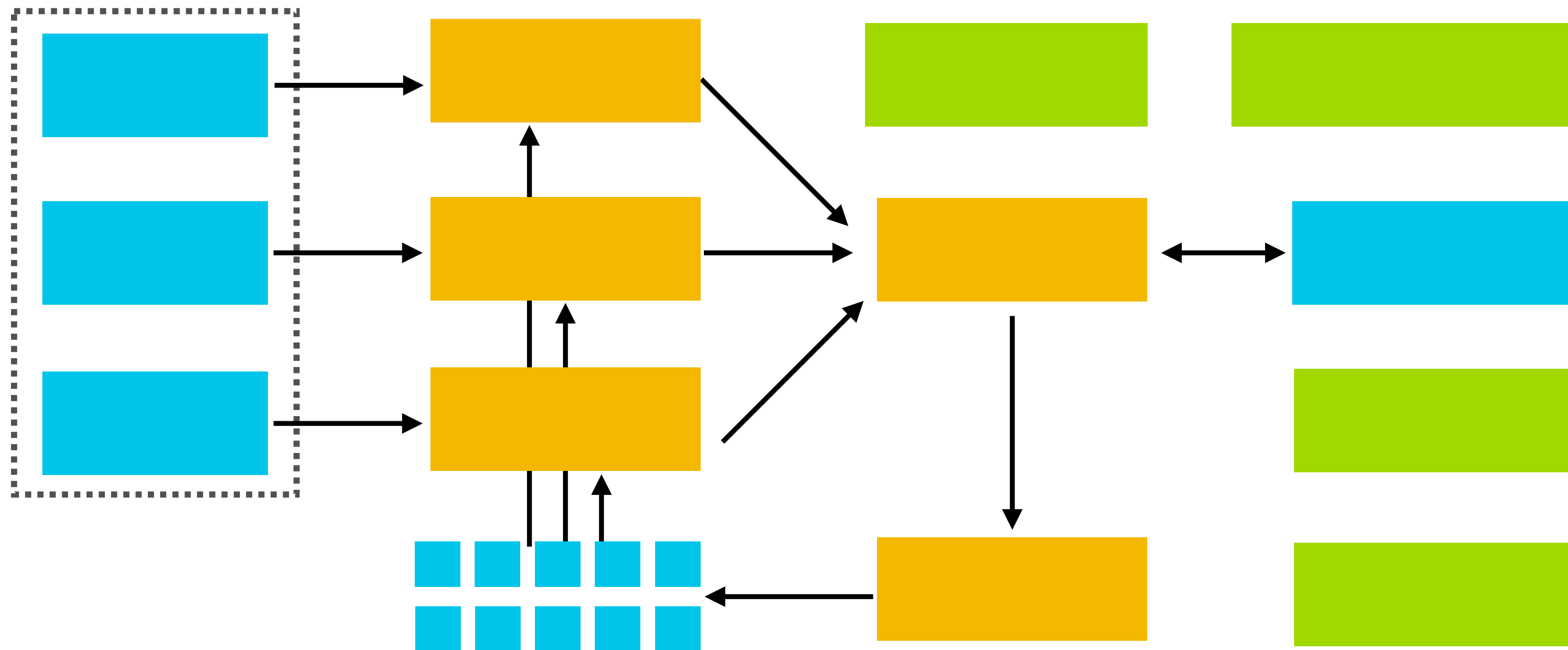
Stateless microservices



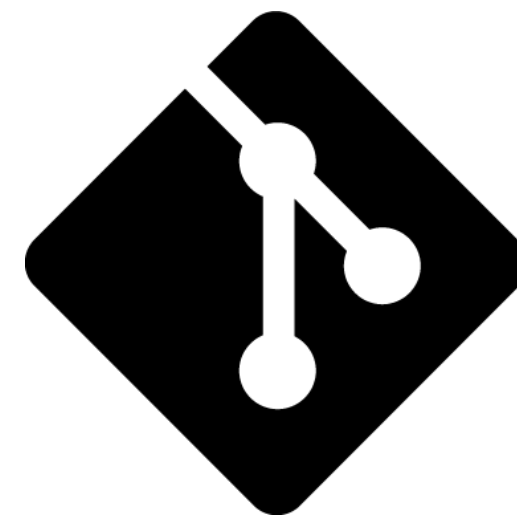
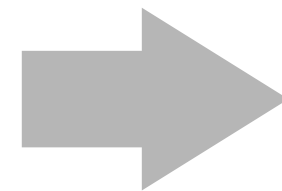
Stateless microservices



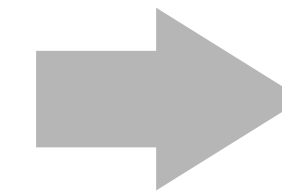
Declarative app configuration



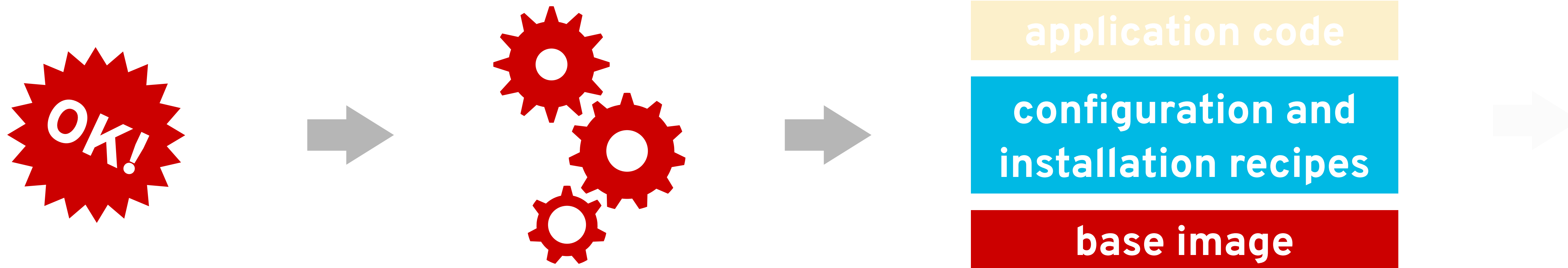
Integration and deployment



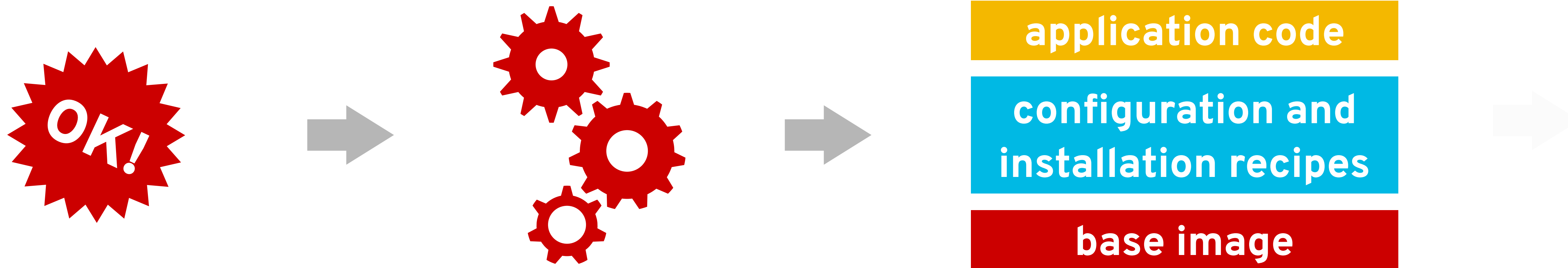
git



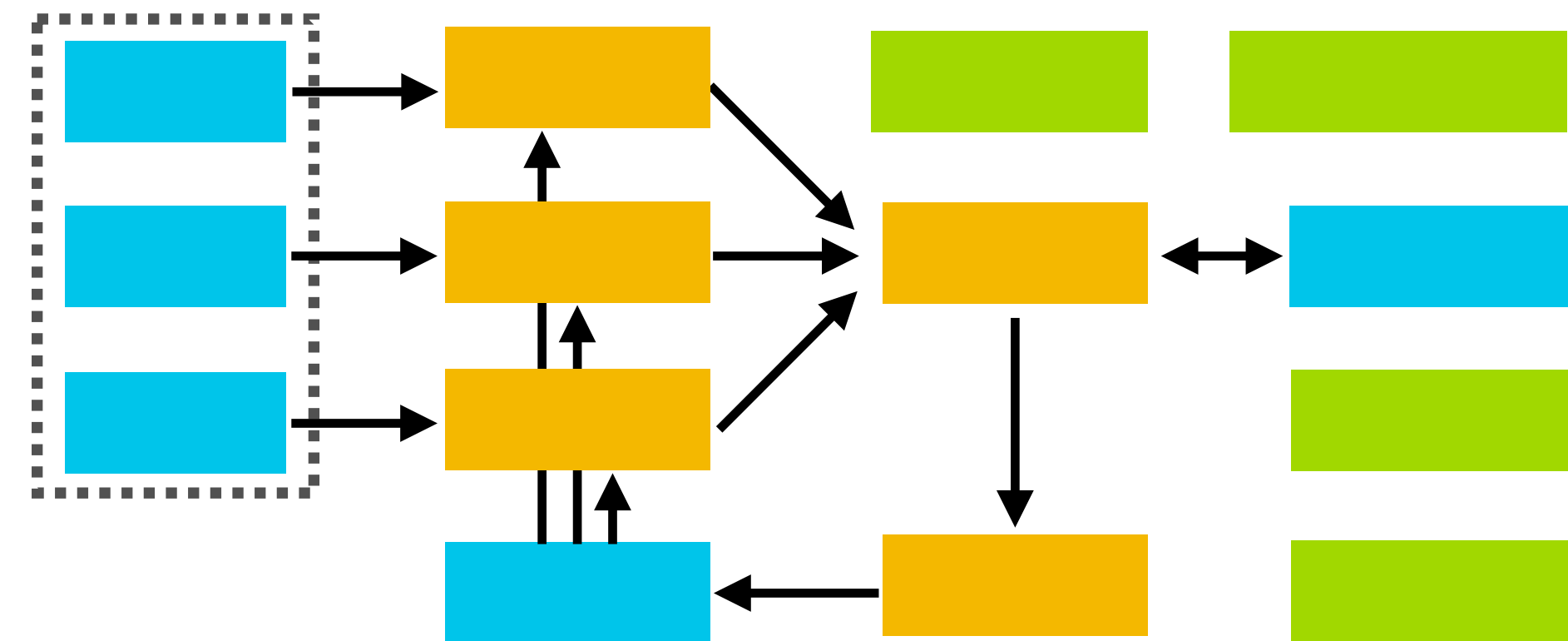
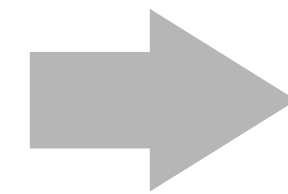
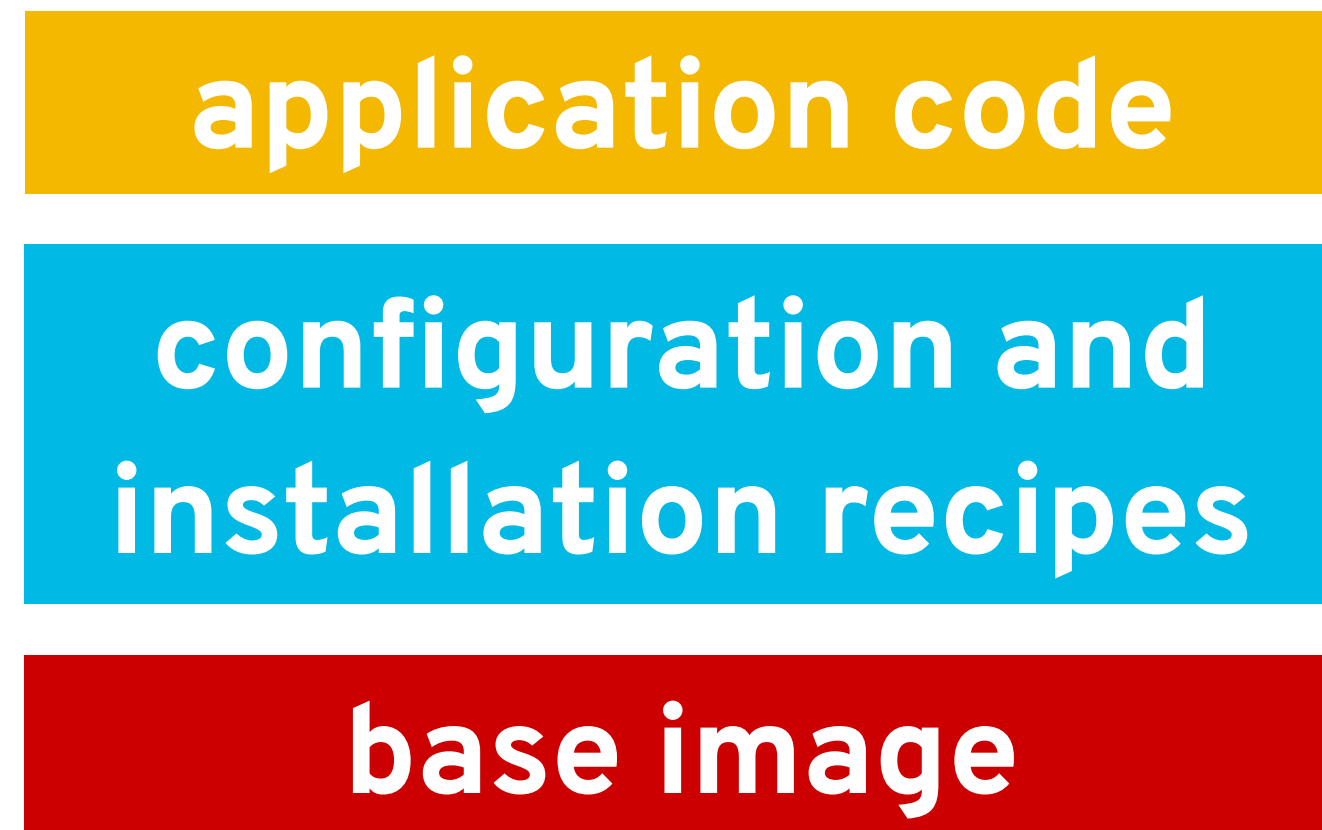
Integration and deployment



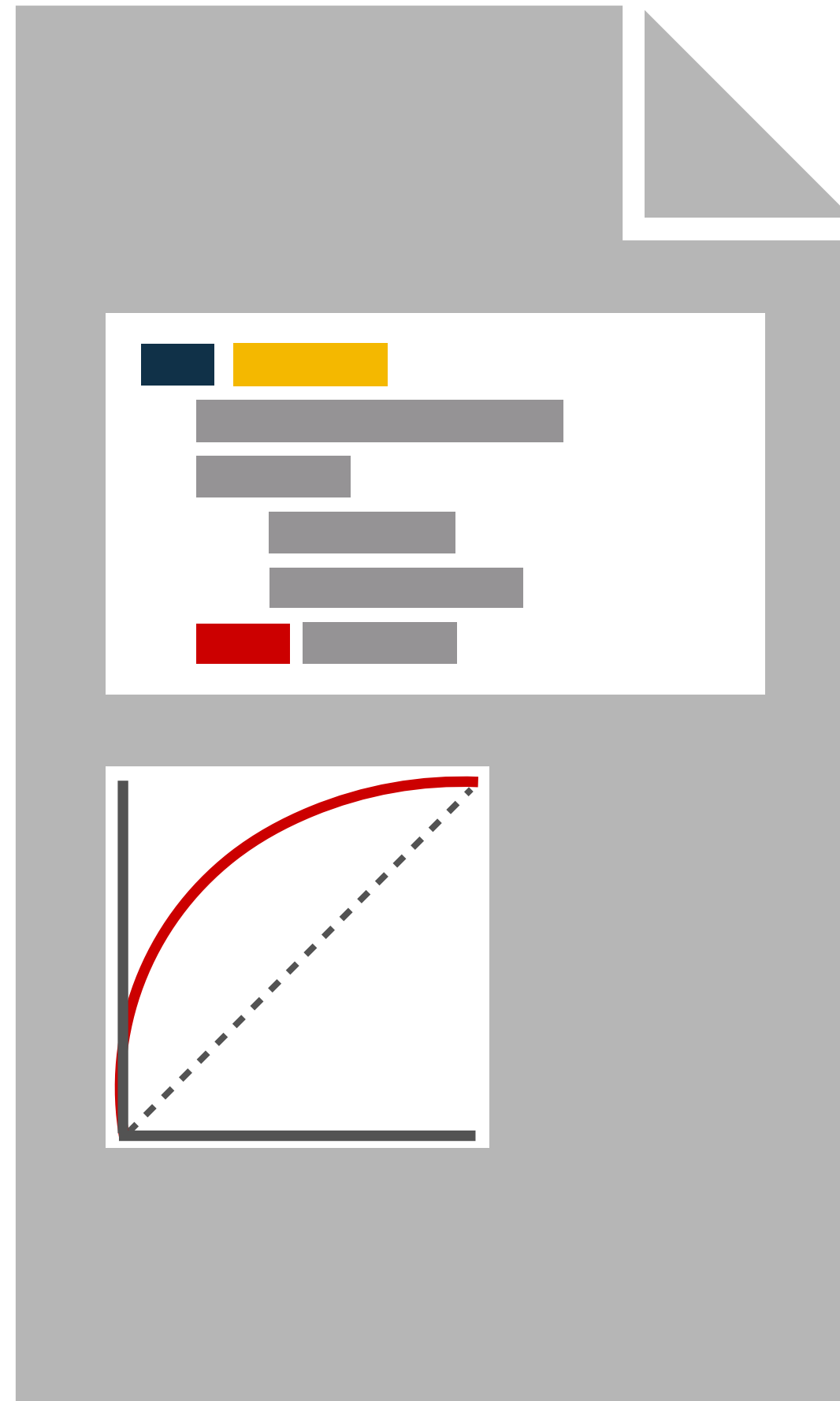
Integration and deployment

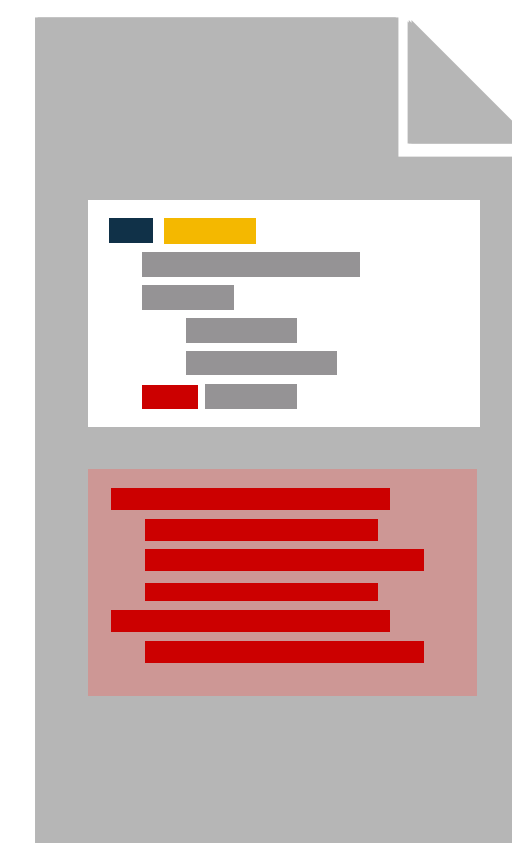
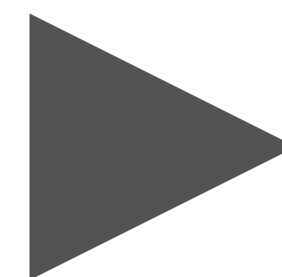
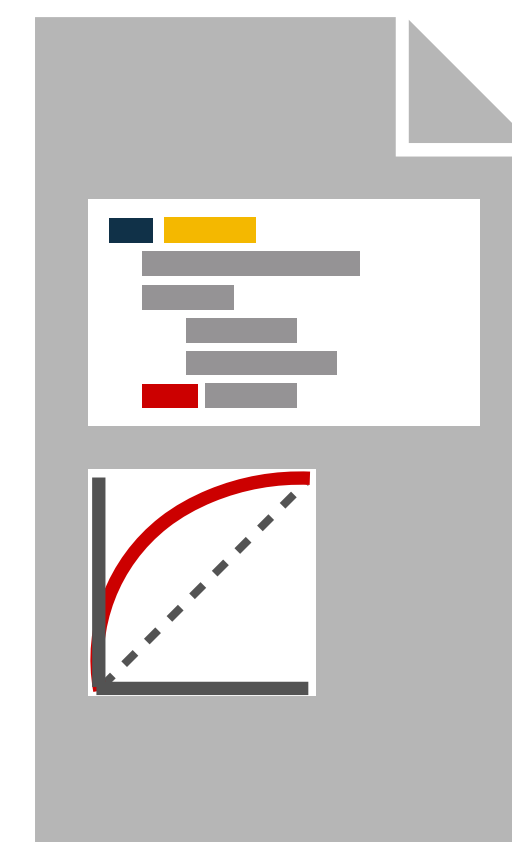
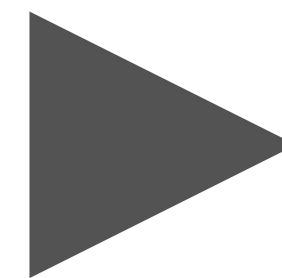
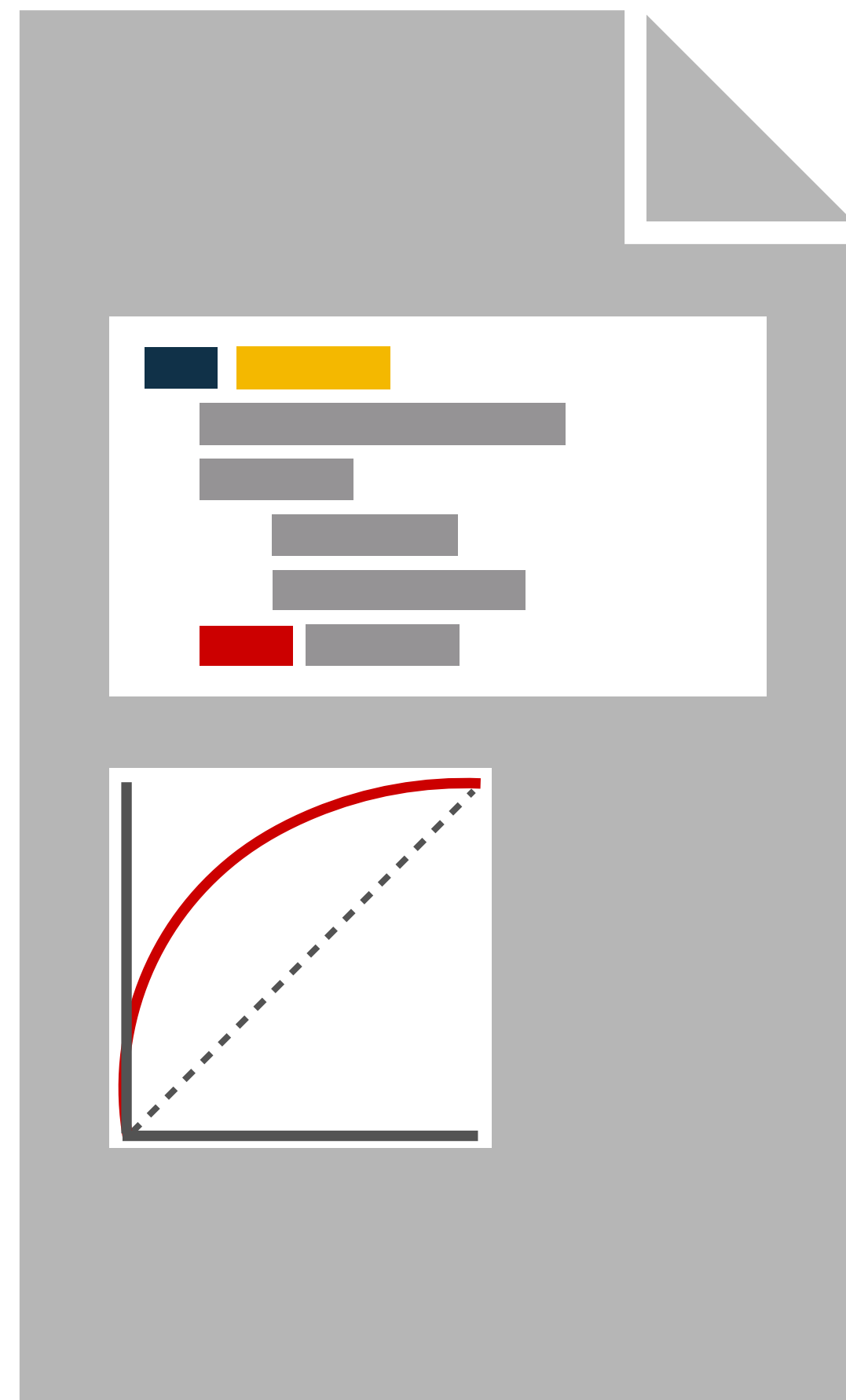


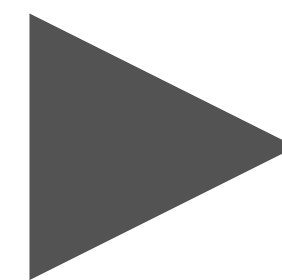
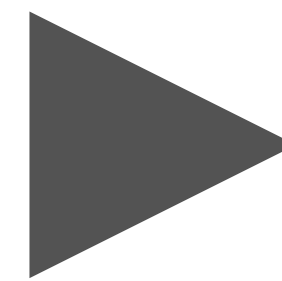
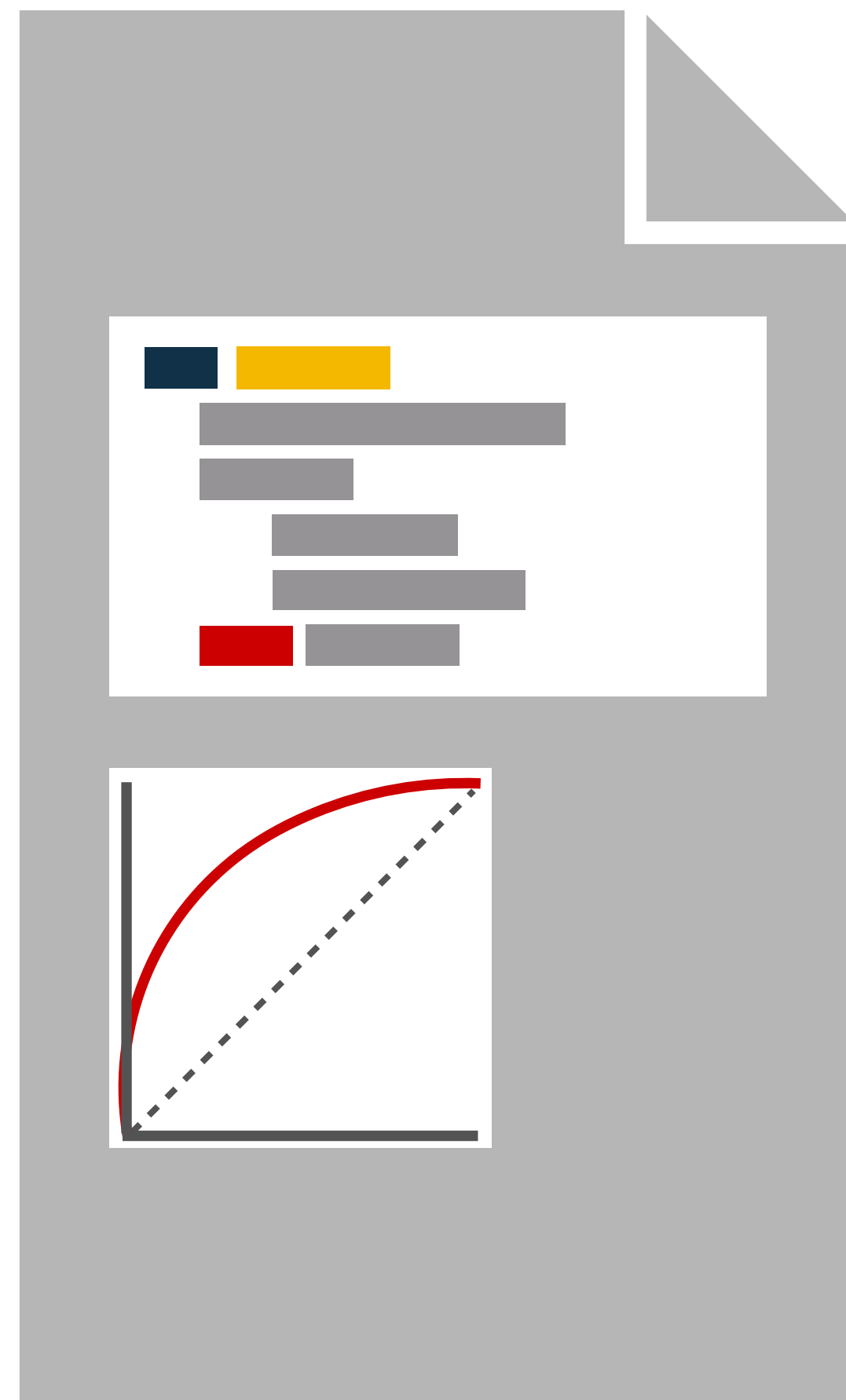
Integration and deployment



**What containers offer
data scientists**



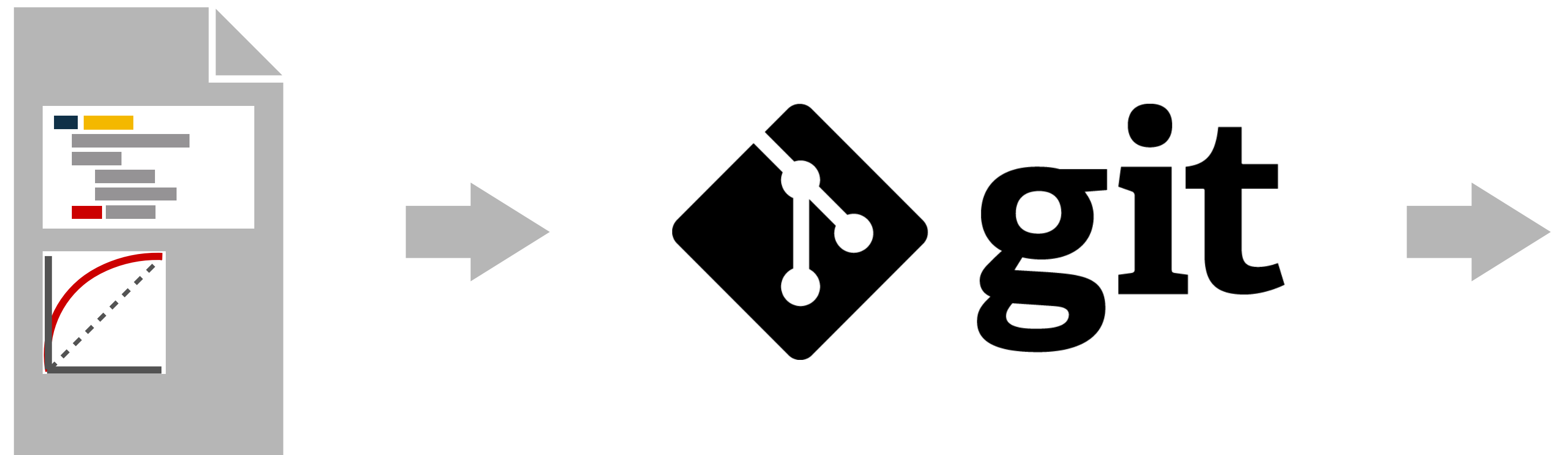




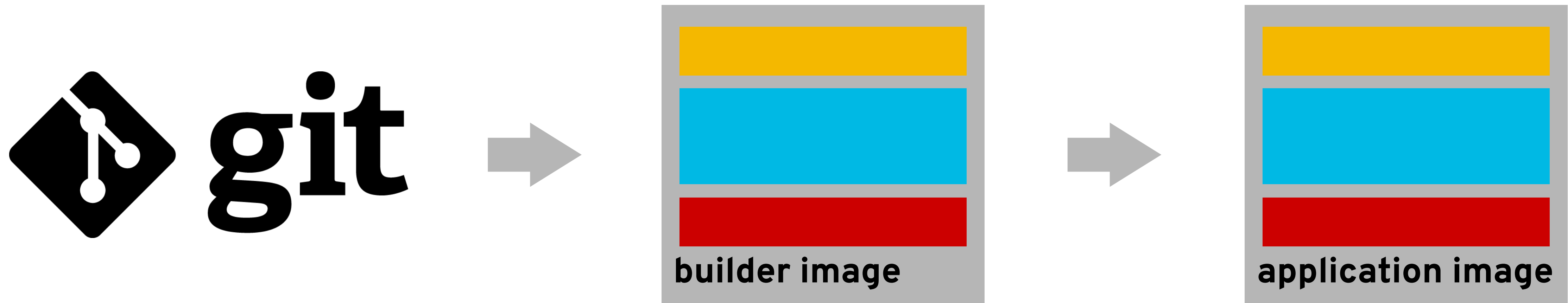
No friction: mybinder.org



More flexible: source-to-image



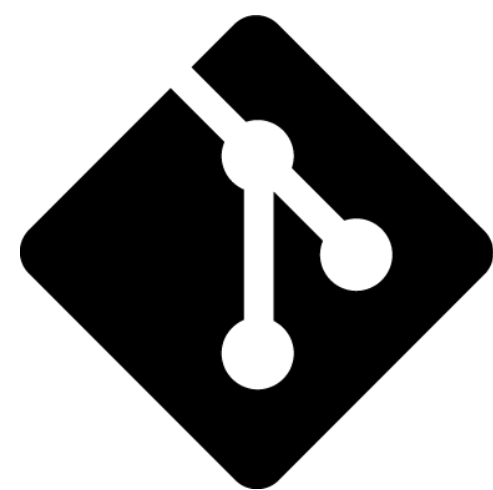
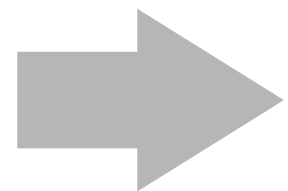
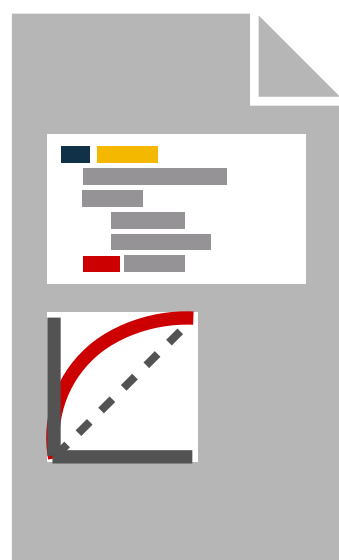
More flexible: source-to-image



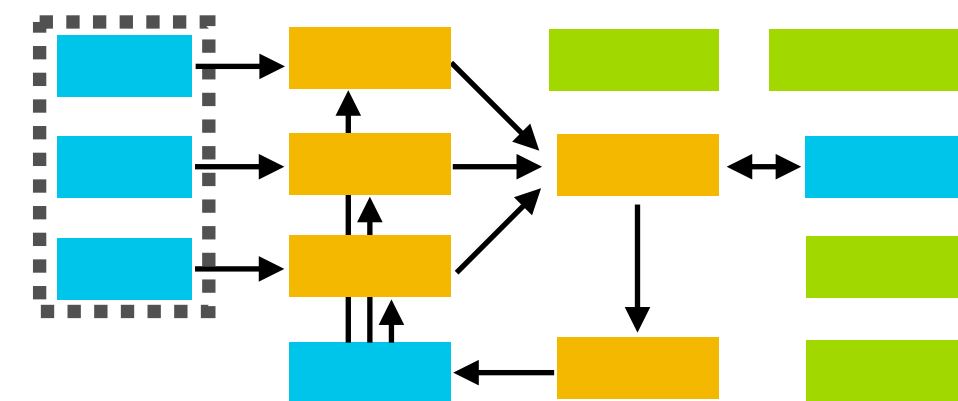
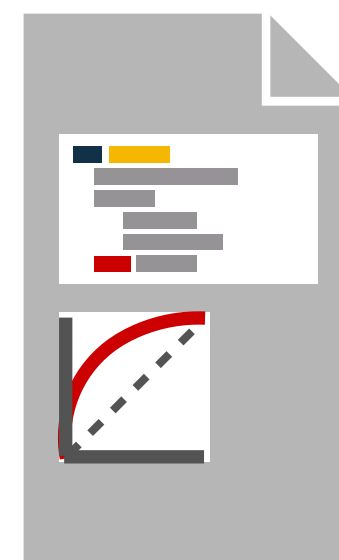
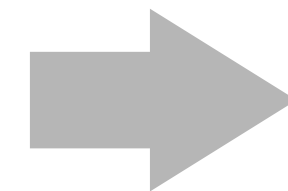
<https://github.com/openshift/source-to-image>

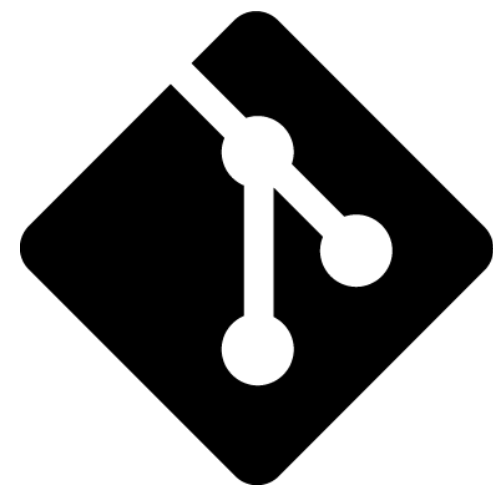
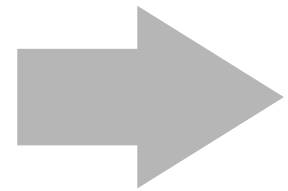
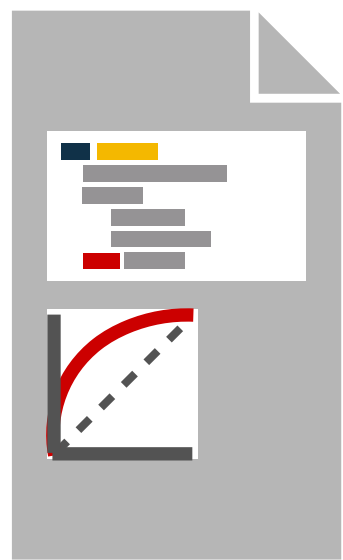
```
willb@echo % oc new-app \  
getwarped/s2i-minimal-notebook:latest~https://github.com/willb/probabilistic-structu  
res \  
-e JUPYTER_NOTEBOOK_PASSWORD=developer
```

```
willb@echo % oc new-app \  
getwarped/s2i-minimal-notebook:latest~https://github.com/willb/probabilistic-structu  
res \  
-e JUPYTER_NOTEBOOK_PASSWORD=developer
```

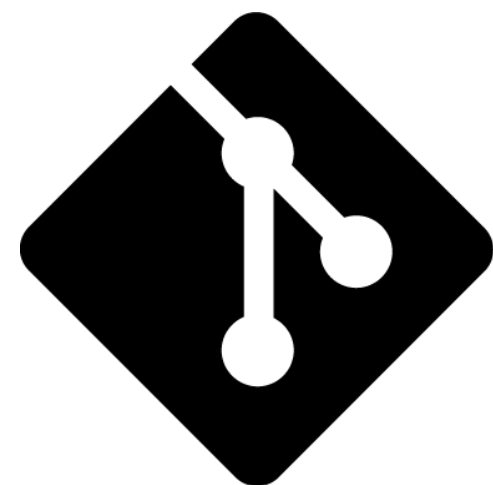
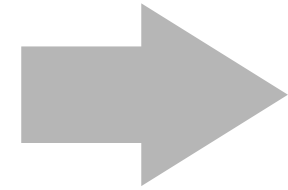
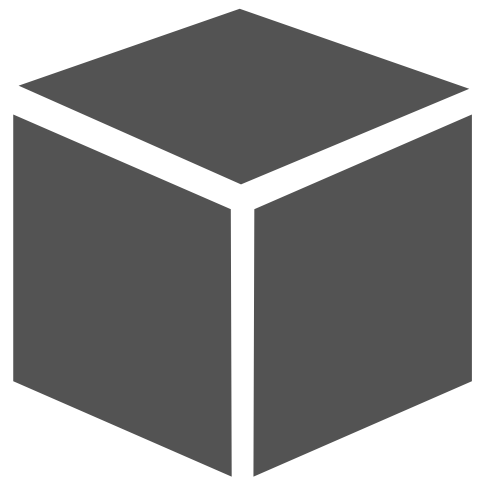
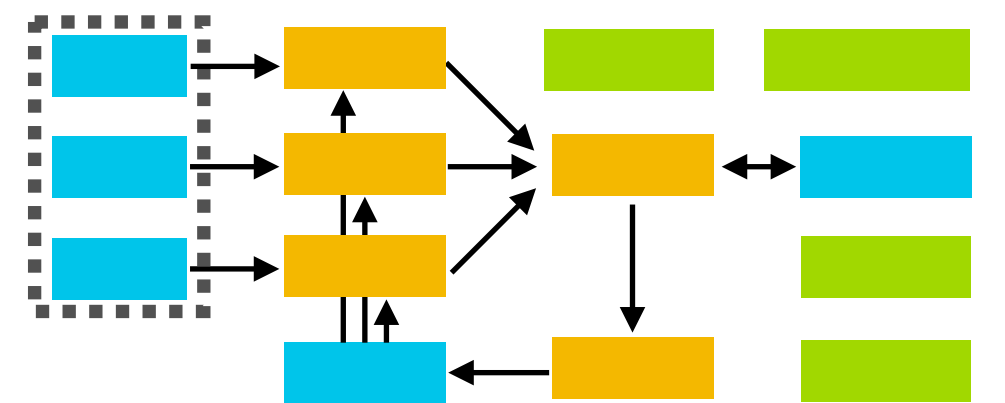
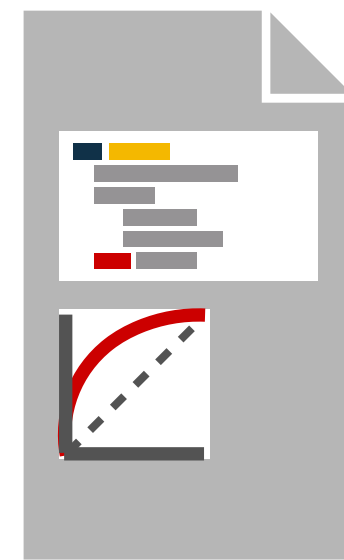
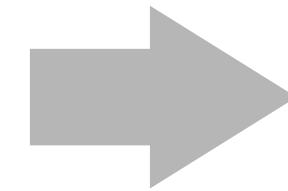


git

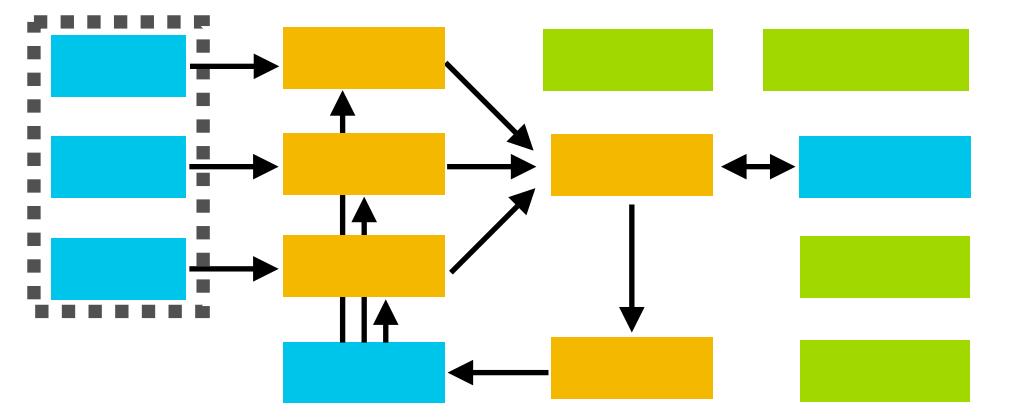
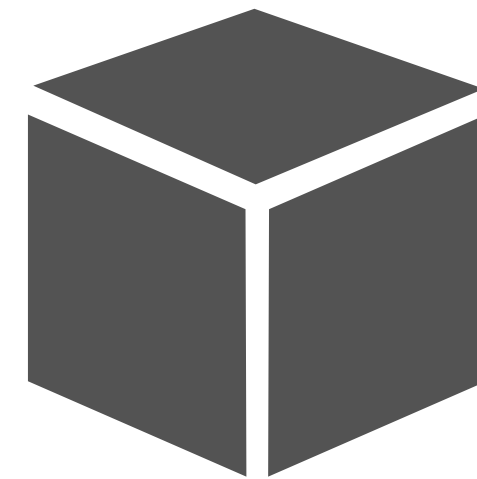
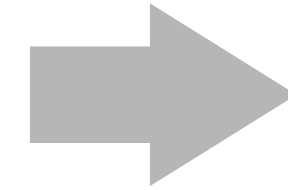




git

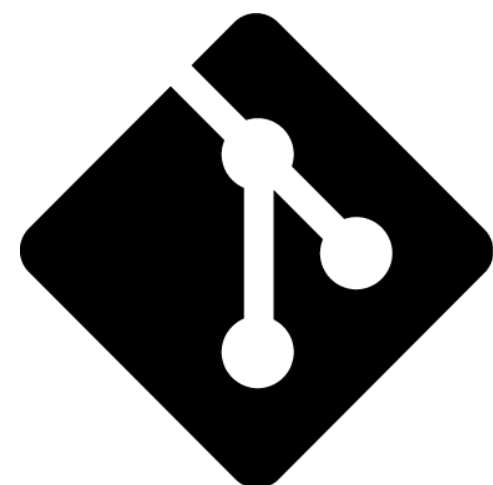
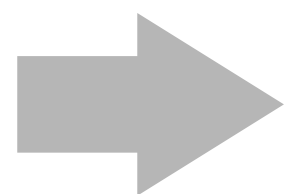
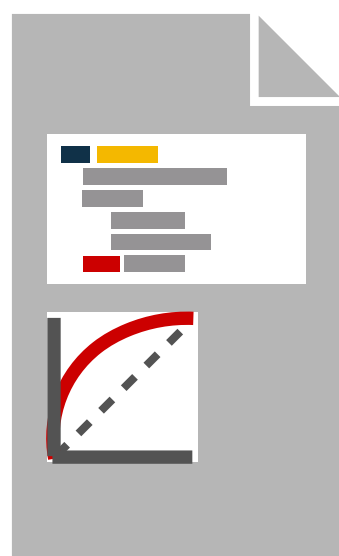


git

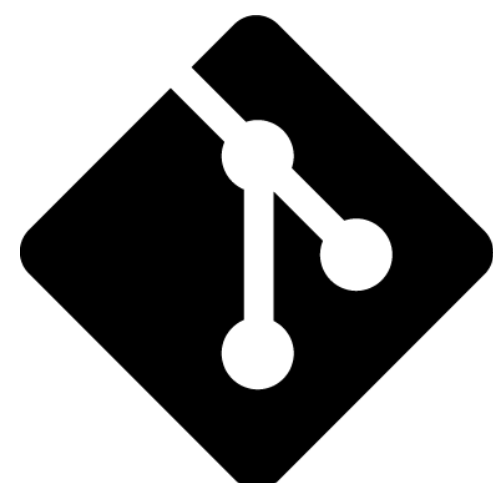
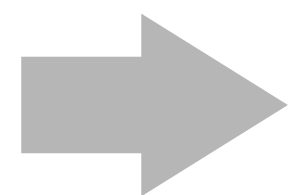
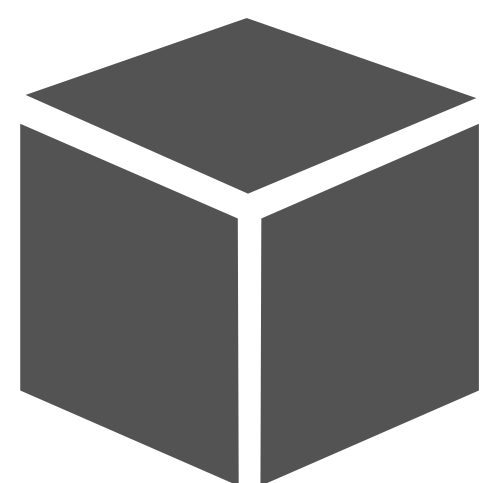
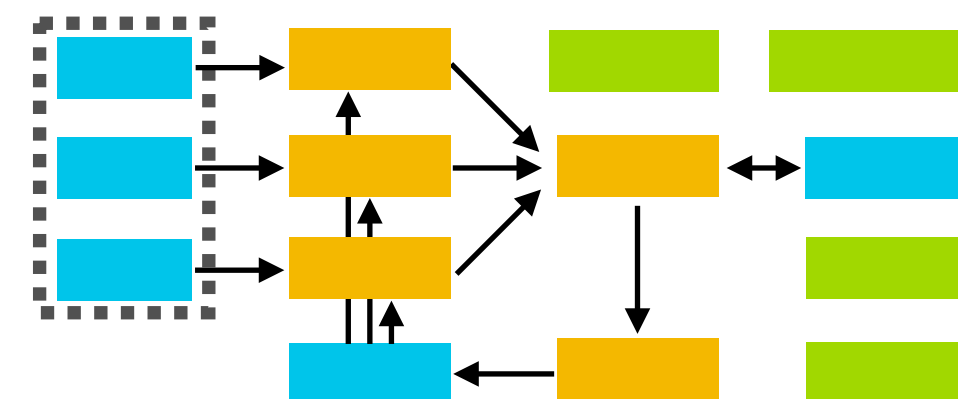
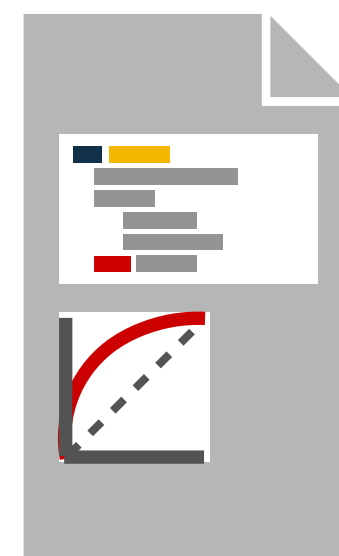
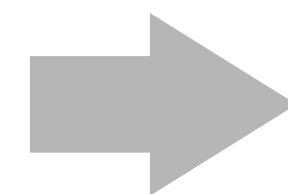


```
willb@echo % oc new-app --name model \  
quay.io/willbenton/simple-model-s2i:demo\  
~https://github.com/willb/example-model-s2i-notebook
```

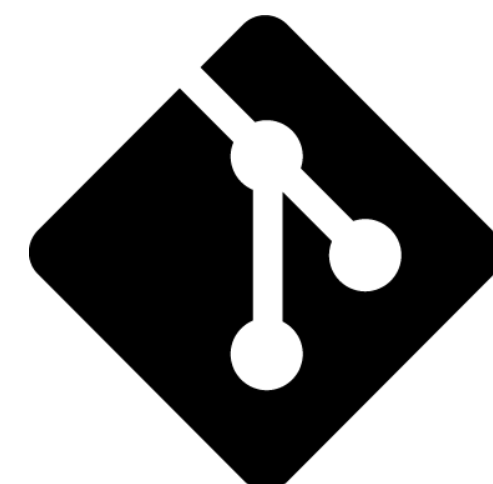
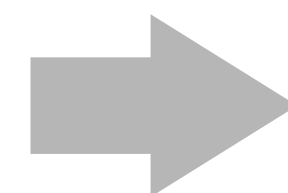
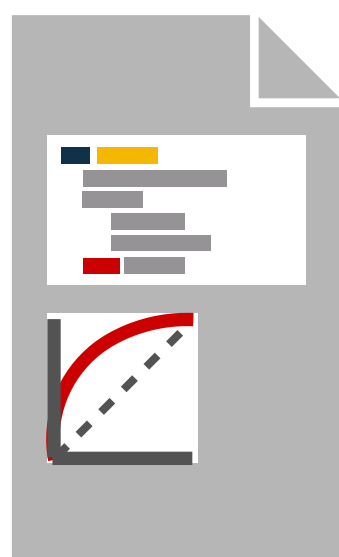
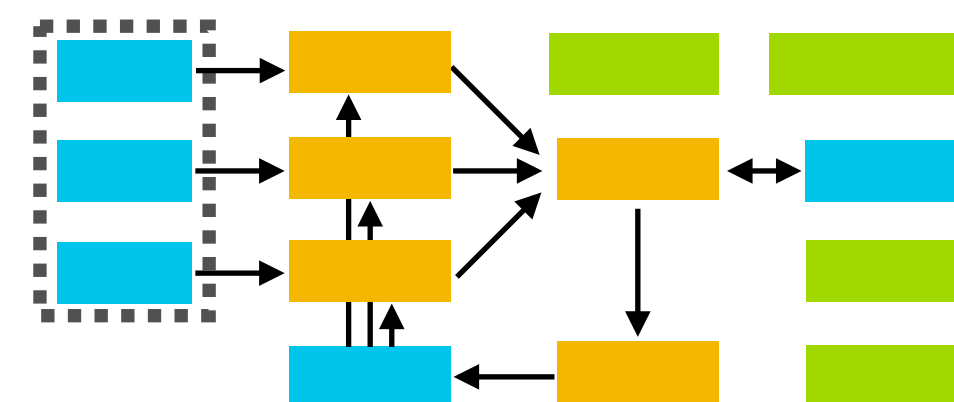
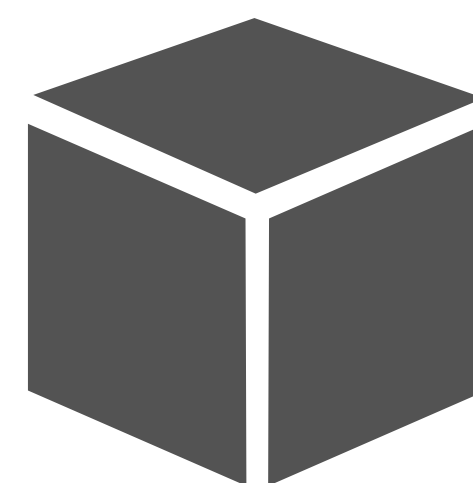
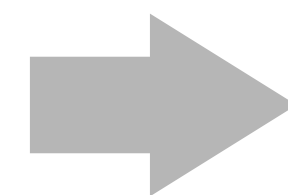
```
willb@echo % oc new-app --name model \  
quay.io/willbenton/simple-model-s2i:demo\  
~https://github.com/willb/example-model-s2i-notebook
```



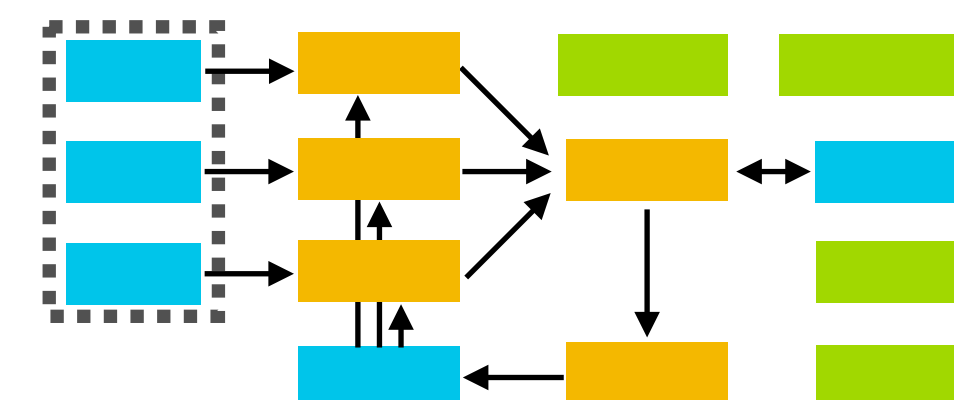
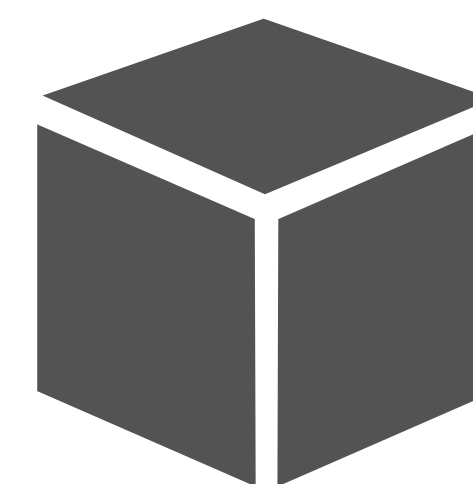
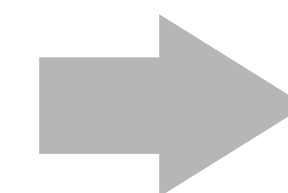
git

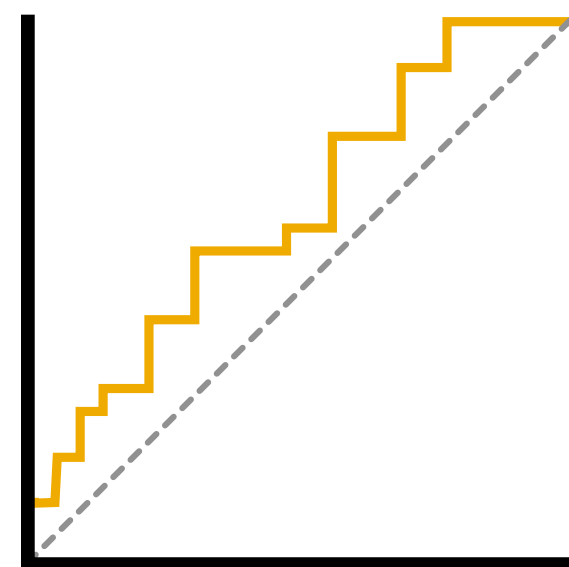


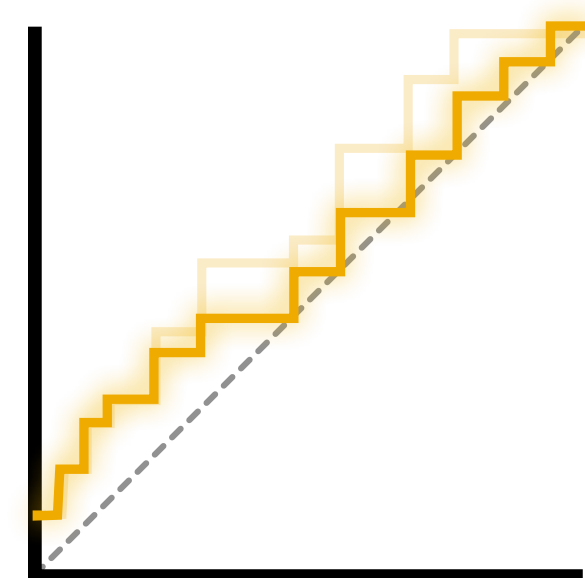
git



git

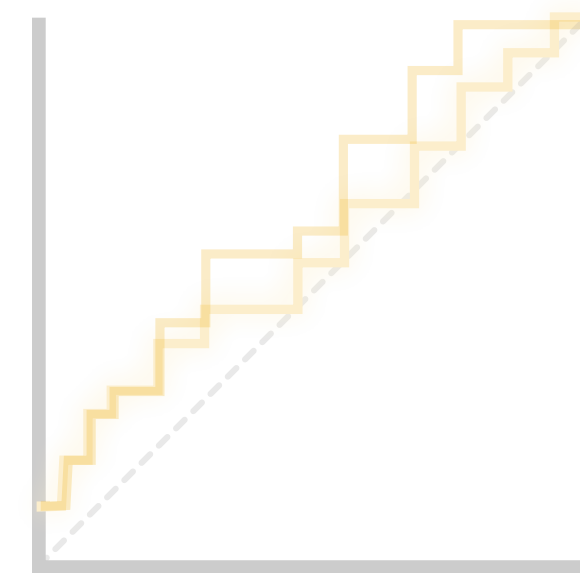




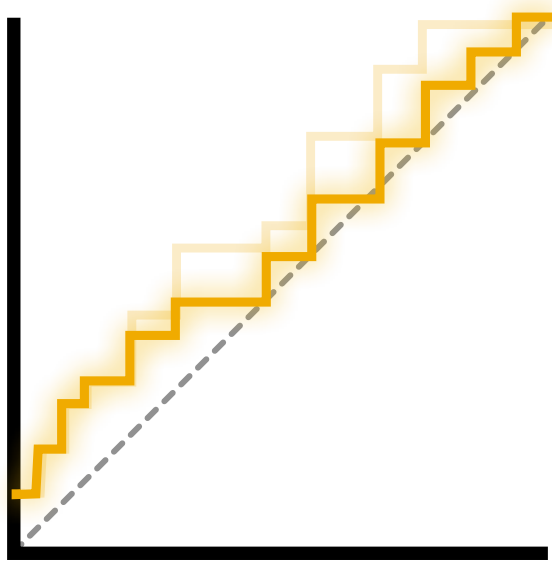
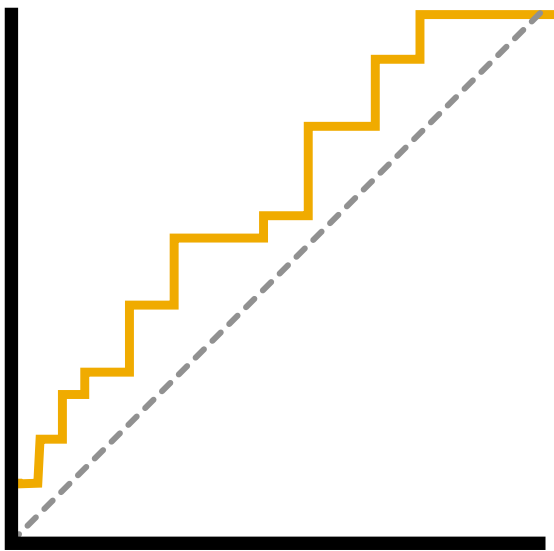
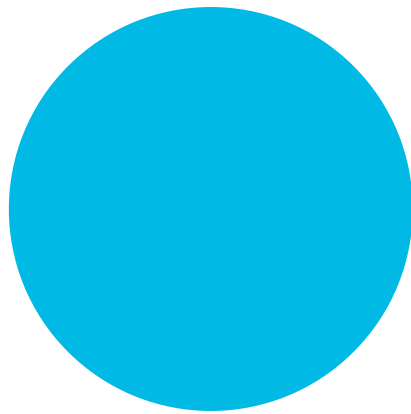
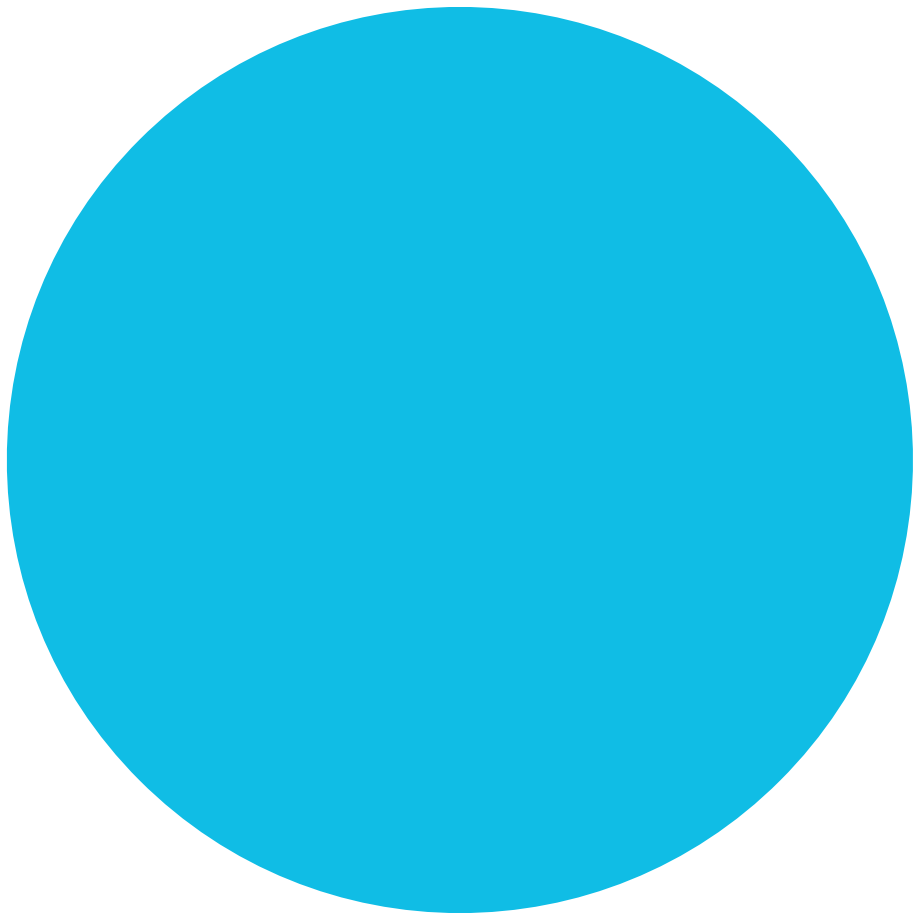


(joint) distribution of input data?

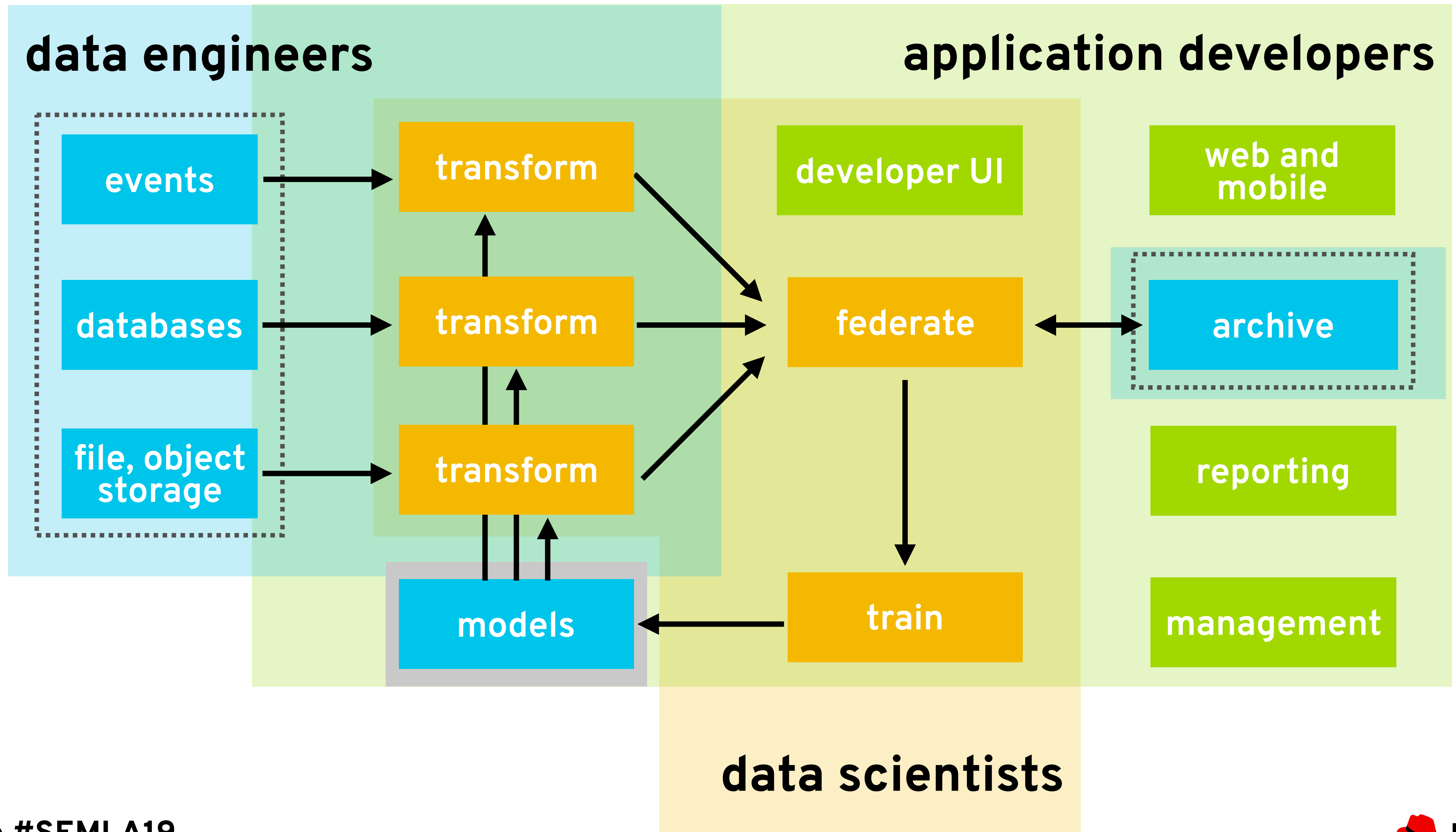
distribution of predictions?



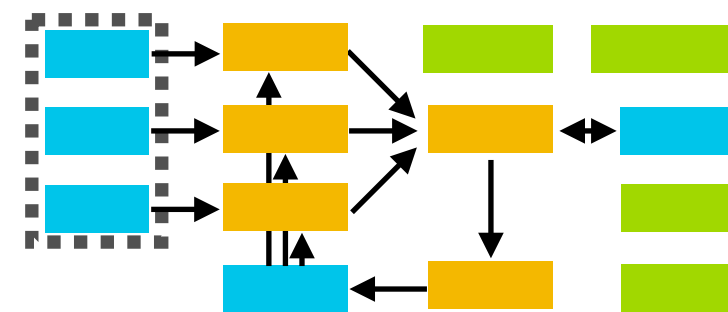
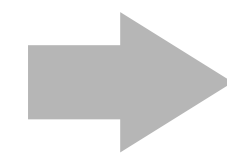
distribution of acyclic paths
taken through scoring code?



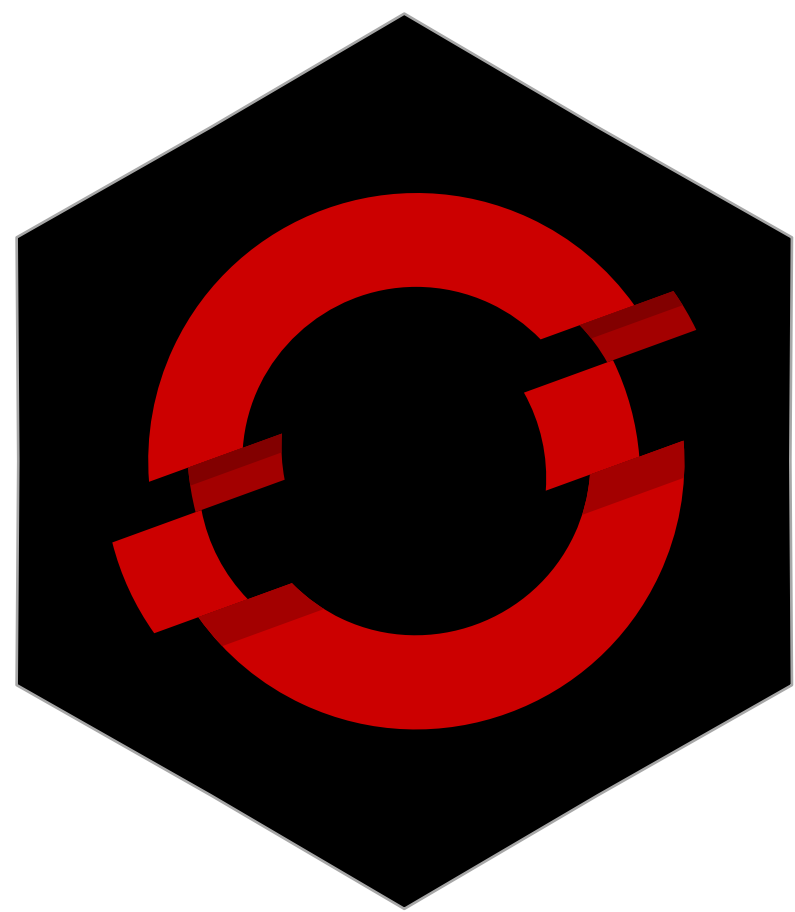
Where from here?



radanalytics.io



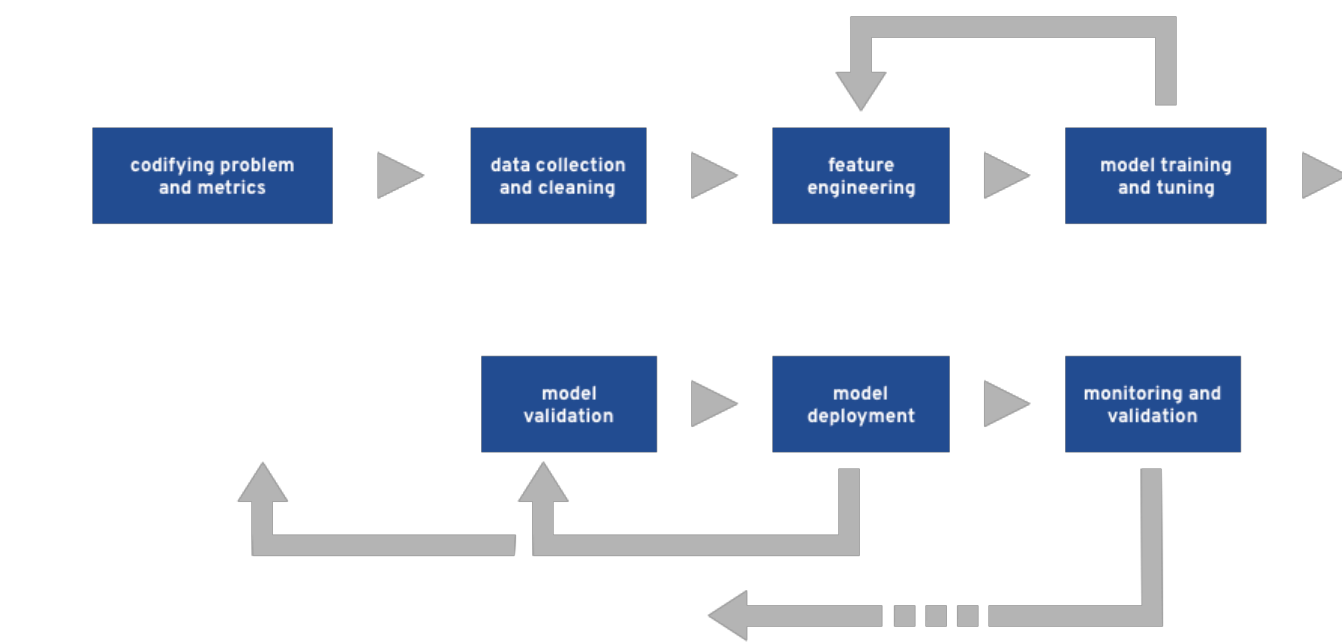
opendatahub.io



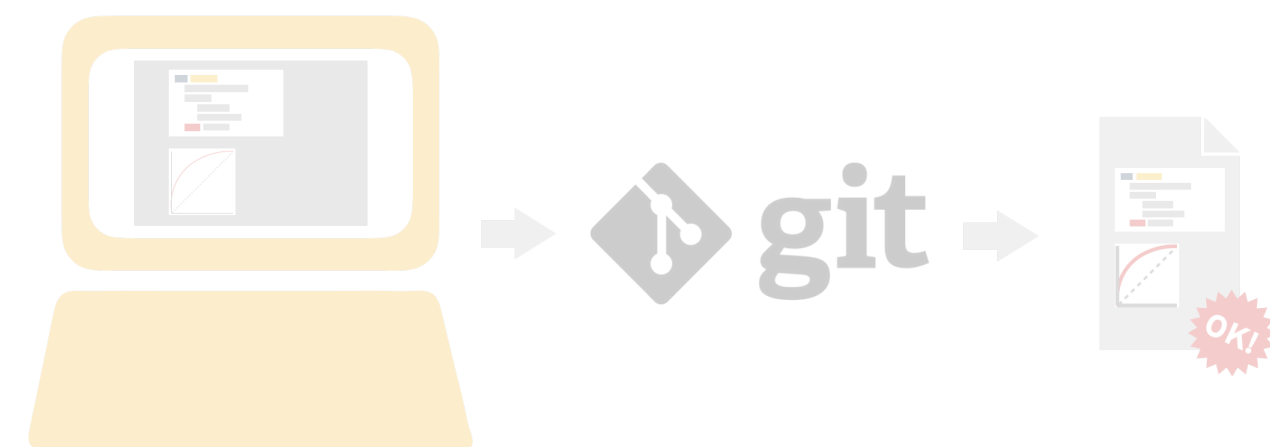
Kubeflow

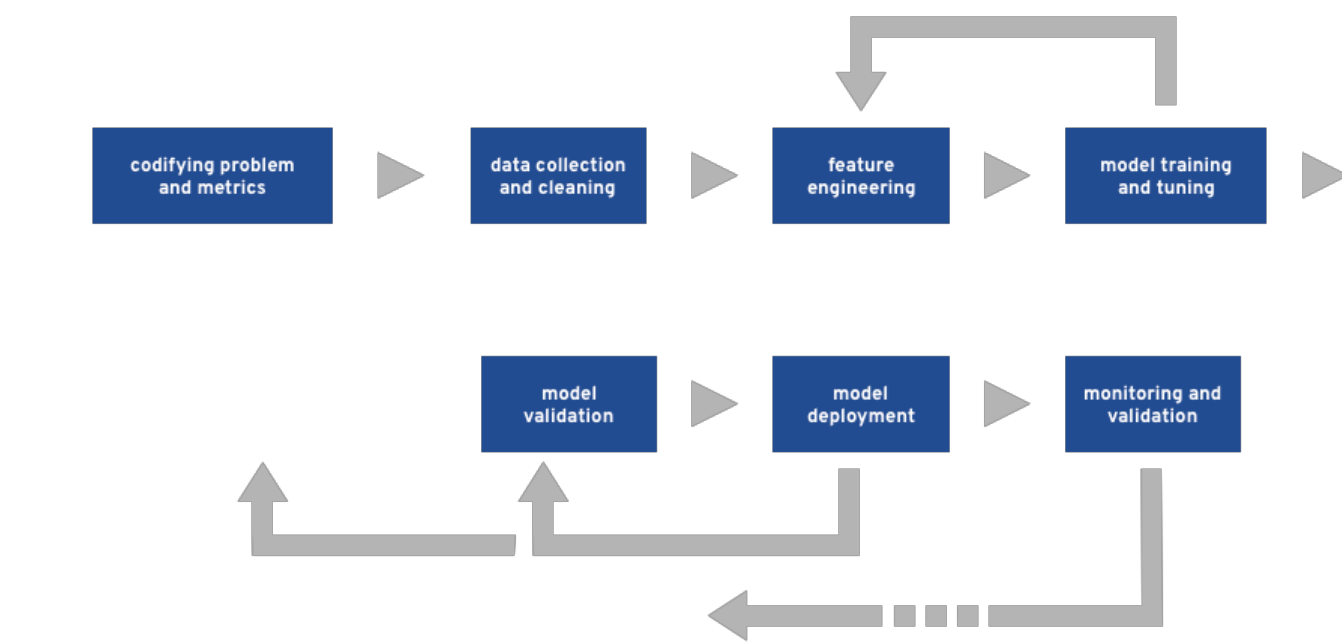


What did we talk about today?

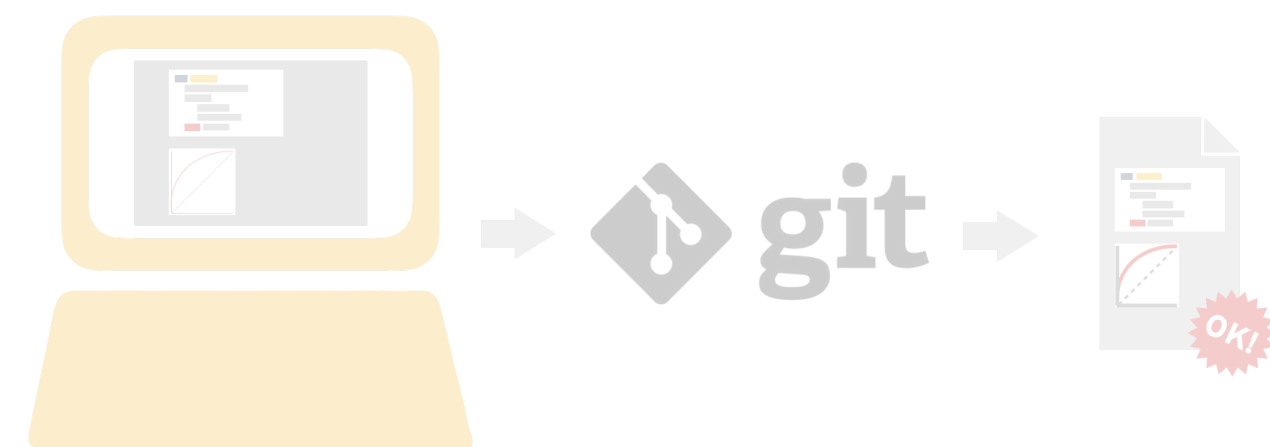


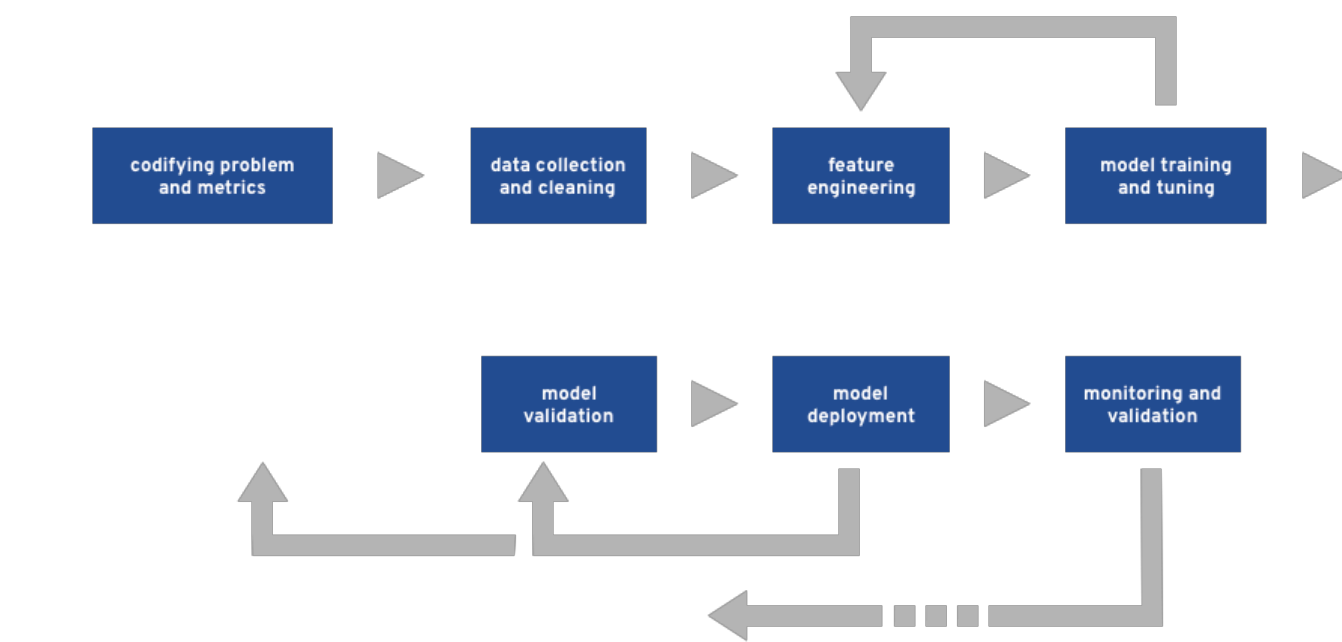
executable		
arguments	pip install numpy	pip install riskylib
environment		
virtual memory		
file handles		
root filesystem	/var/lib/envs/main	/var/lib/envs/risky
process table		
network routes		





executable	/usr/bin/pip	/usr/bin/pip
arguments	pip install numpy	pip install riskylib
environment	LANG=en_US USER=willb	LANG=en_US USER=willb
virtual memory		
file handles		
root filesystem	/var/lib/envs/main	/var/lib/envs/risky
process table		
network routes		

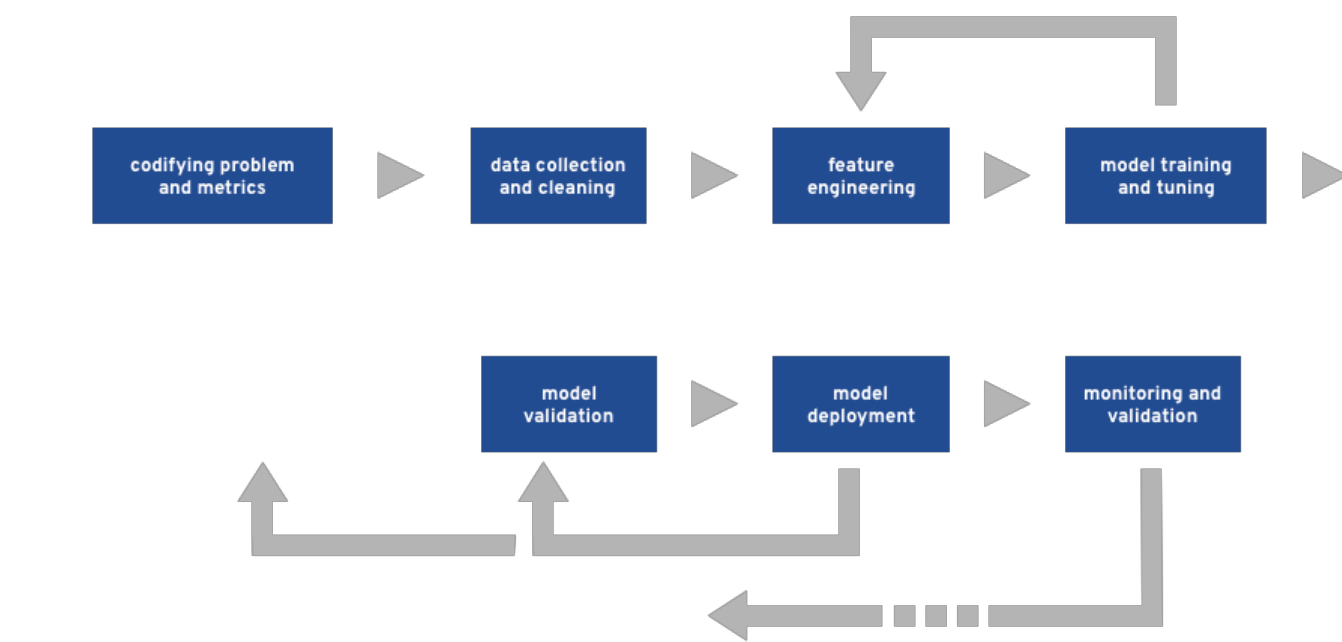




executable	/usr/bin/pip	/usr/bin/pip
arguments	pip install numpy	pip install riskylib
environment	LANG=en_US USER=willb	LANG=en_US USER=willb
virtual memory		
file handles		
root filesystem	/var/lib/envs/main	/var/lib/envs/risky
process table		
network routes		



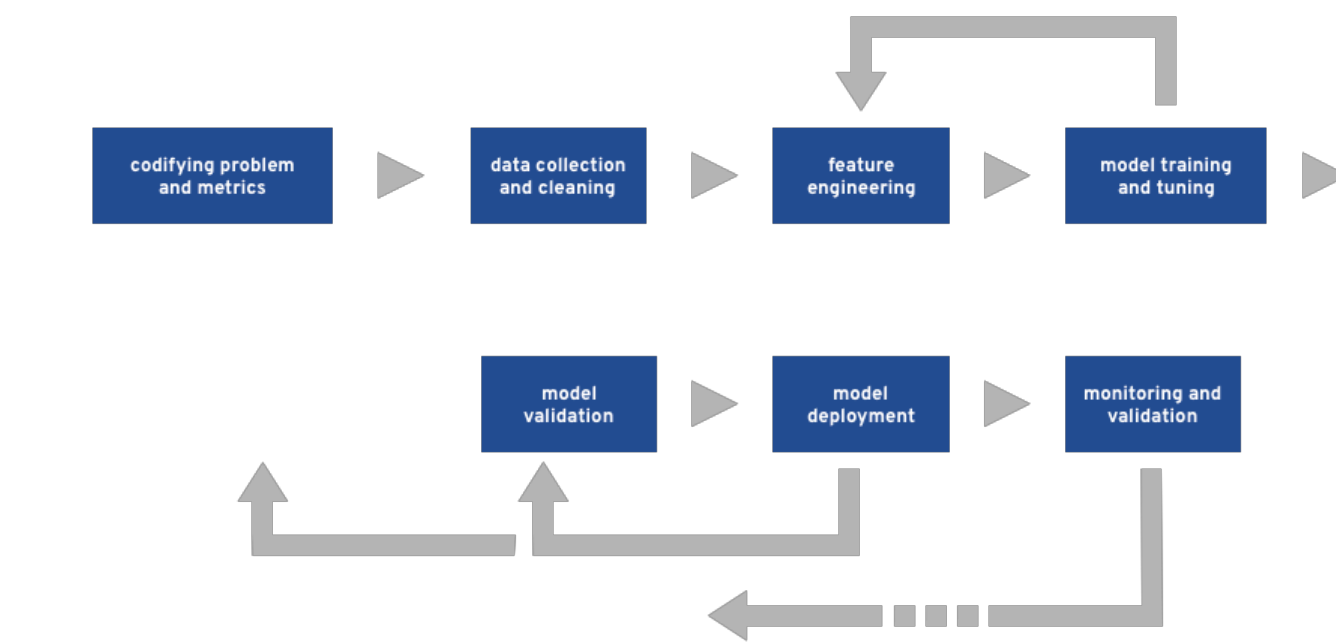
@willb #SEMLA19



executable	/usr/bin/pip	/usr/bin/pip
arguments	pip install numpy	pip install riskylib
environment	LANG=en_US USER=willb	LANG=en_US USER=willb
virtual memory		
file handles		
root filesystem	/var/lib/envs/main	/var/lib/envs/risky
process table		
network routes		



@willb #SEMLA19



executable	/usr/bin/pip	/usr/bin/pip
arguments	pip install numpy	pip install riskylib
environment	LANG=en_US USER=willb	LANG=en_US USER=willb
virtual memory		
file handles		
root filesystem	/var/lib/envs/main	/var/lib/envs/risky
process table		
network routes		



THANKS

willb@redhat.com • @willb
<https://chapeau.freevariable.com>