LESSONS & PITFALLS

DATA AT SWEDEN'S TELEVISION

Ismail Elouafiq
The great moose walk

For thousands of years, the moose have walked the same paths to reach the summer's rich pastures. For several thousand years the moose have walked the same path to get to the rich pastures of summer. Follow the walk live from Kullberg in the north of Sweden.
svt
NYHETER

FASTLANDET
DAG 15
A wide spectrum of Apps

svt
NYHETER

svt
BARN

svt

DUO
svt
SPORT
A wide spectrum of Apps

Running on different platforms
A wide spectrum of Users

ANALYSTS

STRATEGY

PRODUCT OWNERS
A wide spectrum of Users

ANALYSTS  STRATEGY  PRODUCT OWNERS  DEVELOPERS  AUTHORS/EDITORS
tl;dr:

Defining what to prioritise
tl;dr:

Defining what to prioritise

“Data:
I could be chasing an untamed ornithoid without cause.”

• — Star Trek The Next Generation
tl;dr:

Defining what to prioritise

“Data:
I could be chasing an untamed ornithoid without cause.”

• — Star Trek The Next Generation
**tl;dr:**

Defining what to prioritise

Experimenting and iterating in small increments

**Spoilers**: how and why we now use protobuf, functional data engineering and ETL practices
tl;dr:

Defining what to prioritise

Experimenting and iterating in small increments

Spoilers: how and why we now use protobuf, functional data engineering and ETL practices

BLOCKCHAIN
AI
Deep reinforcement learning
tl;dr:

Defining what to prioritise

Experimenting and iterating in small increments
tl;dr:

Defining what to prioritise

Experimenting and iterating in small increments

ismail.land/velocity
tl;dr:

Defining what to prioritise

Experimenting and iterating in small increments

ismail.land/velocity
What **events** should you collect?

**Collect**

**ALL THE THINGS!!**
What **events** should you collect?
what we want to know

How many people read the article per day
what we want to know

How many people read the article per day

what we can observe

click scroll share
what we want to know

How many people read the article per day

what we can observe

click scroll share events
what we want to know

How many people read the article per day

explicit model

click scroll share

events

what we can observe
what we want to know

How many people read the article per day

explicit model

click
scroll
share

events

what we can observe
let's start with views
• If you could do **anything** with data...
• What would you **actually** use for decision making
• If you could do *anything* with data...
• What would you actually use for decision making

A/B tests... Hell yeah!
tl;dr:

Defining what to prioritise

Experimenting and iterating in small increments

ismail.land/velocity
First we need to collect data

1 COLLECT

2 INGEST

SDK
COLLECT

SDK

events

INGEST

Event API
1 COLLECT

SDK

events

2 INGEST

Event API

publish
COLLECT

SDK

events

Event API

publish

INGEST

pub/sub
INGEST

pub/sub
INGEST

pub/sub

STORE
**INGEST**

- pub/sub

judge-judi

subscribe

write

**STORE**

Events table
pub/sub

judge-judi

Events table

subscribe

write
{event_type: click}

{ eventType: click}

{eventType: klick}
COLLECT

{event_type: click}
{eventType: click}
{eventType: klick}

INGEST

STORE
More Issues

- Multiple teams/platforms => takes time to update the clients
- The schema is sent with every event
- Unclear types (arbitrary memory allocation)
More Issues

- Multiple teams/platforms => takes time to update the clients
- The schema is sent with every event
- Unclear types (arbitrary memory allocation)

We know the schema on all levels
we have a common model for the data..
how can we make use of that...
ENTER PROTOBUF

Keeping a centralized Event Schema
ENTER PROTOBUF

Keeping a centralized Event Schema

```protobuf
message Person {
  required string name = 1;
  required int32 id = 2;
  optional string email = 3;
}
```
ENTER PROTOBUF
Keepign a centralized Event Schema
ENTER PROTOBUF

Keeping a centralized Event Schema
ENTER PROTOBUF

Keepign a centralized Event Schema

1 - Define the Schema
As a .proto file
ENTER PROTOBUF
Keepign a centralized Event Schema

1 - Define the Schema
As a .proto file

2 - Publish libraries
Publish using CI pipeline

go, js, java, swift
1 - Define the Schema
As a .proto file

2 - Publish libraries
Publish using CI pipeline

2 - Fetch
- Fetch in SDKs (serialization)
- Fetch in Judy (deserialization)
- Use to generate table
My work here is done!
Not really...

- Backward and forward compatibility
- Table changes
- Language agnostic but nor really
- Lack of support
The Data Pyramid

Collection and ingestion

Storage, transformation, monitoring

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collection and ingestion

Collecti...
The Data Pyramid

Collection and ingestion

Storage, transformation, monitoring
The Data Pyramid

- Collection and ingestion
- Storage, transformation, monitoring
- Metrics, aggregations, KPIs
- Learn, Optimise, Experiment
Nirvana AI, machine learning

Metrics, aggregations, KPIs

Collection and ingestion

Storage, transformation, monitoring

Learn, Optimise, Experiment

The Data Pyramid
"The **pyramids of Egypt** could be explained as symbolic stairways to the stars, **according to a British scientist**" _The Guardian_
The Data Pyramid

"The pyramids of Egypt could be explained as symbolic stairways to the stars, according to a British scientist" _The Guardian_

"The data pyramid could be explained as a symbolic stairway to the A.I., according to myself" _Me_
Endorse me on Linkedin

Ismail Elouafiq
I tell stories, some of which are true, using data

Skills & Endorsements

Take skill quiz

Data Pyramid Worshipper 7
Endorsed by 3 people who are highly skilled at this
We have the data

Now what?
We have the data

Now what?

1. COLLECT
2. INGEST
3. STORE

Define

Collect

Ingest

Store

Analyze
Batch jobs etl
Streaming

Present
Service/API
Dashboard
Reports
We have the data

Now what?

1. COLLECT
2. INGEST
3. STORE

Define

4. Analyze
   - Batch jobs
   - ETL
   - Streaming

5. Present
   - Service/API
   - Dashboard
   - Reports

4 5
Everybody
ETLs
Everybody
ETLs
**Inputs**
- Some data to be aggregated

**Output**
- Aggregated Table (article reads) Per DAY

---

<table>
<thead>
<tr>
<th>click events</th>
<th>article titles</th>
</tr>
</thead>
</table>

| article reads per day | NEWS |
today- partition

magic job

Append
today - partition

Failed

magic job
Principle: Ensuring reproducibility

- Immutable data partitions
- Versioned logic
On ETL design

- Ensure reproducibility
- Practice failure in small increments
- Defining conventions in one place
keeping a tidy pipeline
summary...
summary...
Cyberconflict: A new era of war, sabotage, and fear

We're living in a new era of constant sabotage, misinformation, and fear, in which everyone is a target, and you're often the collateral damage in a growing conflict among states. From crippling infrastructure to sowing discord and doubt, cyber is now the weapon of choice for democracies, dictators, and terrorists.

David Sanger explains how the rise of cyberweapons has transformed geopolitics like nothing since the invention of the atom bomb. Moving from the White House Situation Room to the dens of Chinese, Russian, North Korean, and Iranian hackers to the boardrooms of Silicon Valley, David reveals a world coming face-to-face with the perils of technological revolution—a conflict that the United States helped start when it began using cyberweapons against Iranian nuclear plants and North Korean missile launches. But now we find ourselves in a conflict we're uncertain how to control, as our adversaries exploit vulnerabilities in our interconnected nation and we struggle to figure out how to deter these complex, covert war attacks.
summary... (what worked for us)
summary...  (what worked for us)
Thank You

ismail.land/velocity