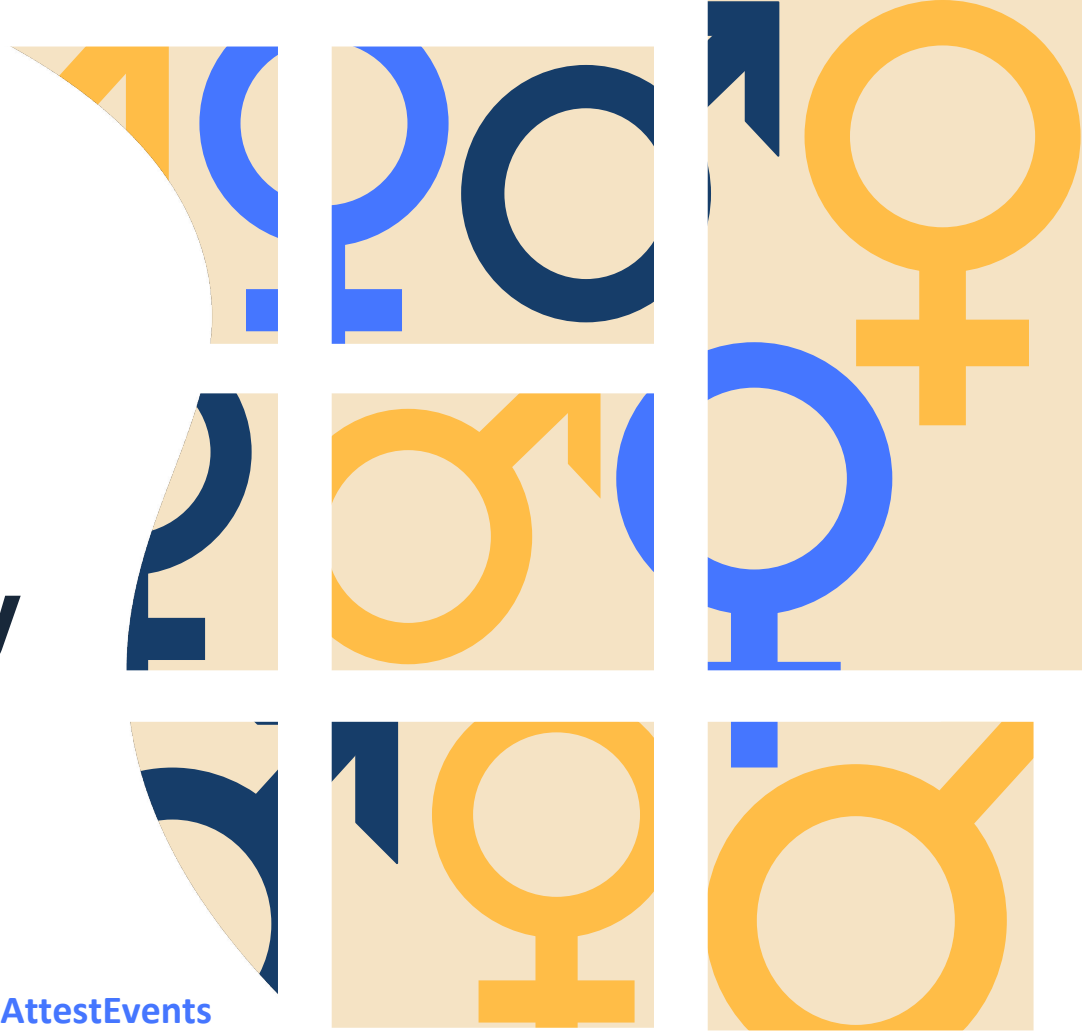
Live Webinar

How Attests Data Science Team Approach Audience Quality

23 September 2020

JOIN US NOW



ATTEST

#AttestEvents

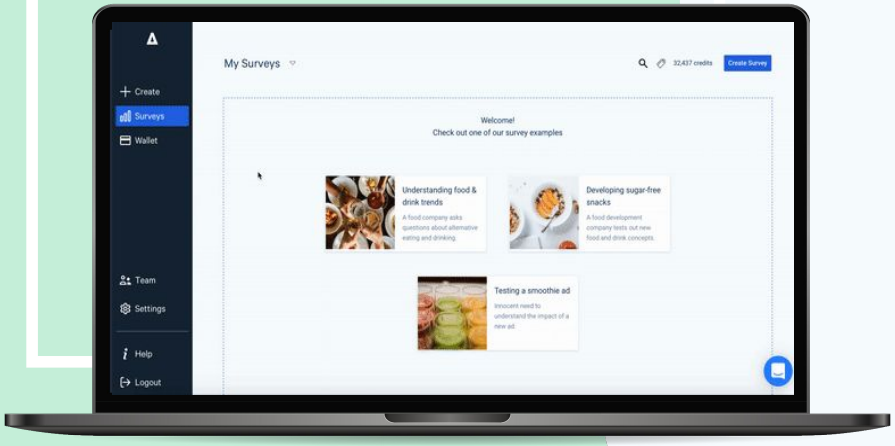
#AttestEvents

Facilitator

Lucy Chapman

Events Manager

#AttestEvents



We're Attest.

Attest was designed for ambitious marketers that aren't experts in doing research, but do need answers - fast.

Our survey platform connects you to 100 million consumers across 46 countries, making asking consumers questions and getting answers fast, easy, and scalable.

Welcome & Housekeeping

- 25 minute presentation, 10 minutes for Q&A
- Please ask questions throughout using the box on the right of your screen
- This session is being recorded - you'll receive the recording via email
- Look out for us launching some polls during the session
- You can use #AttestEvents
- Post-webinar survey

Who Am I?

Dr. Gökhan Çiflikli
Senior Data Scientist
Attest

Today's Agenda

01 Data Annotation

02 Active Learning

03 AI-Assisted Labelling

04 Fraud Detection

05 Q&A

Data Annotation: New Frontier

How obtaining labelled data at scale become the new bottleneck in the industry

AI-Assisted Active Learning

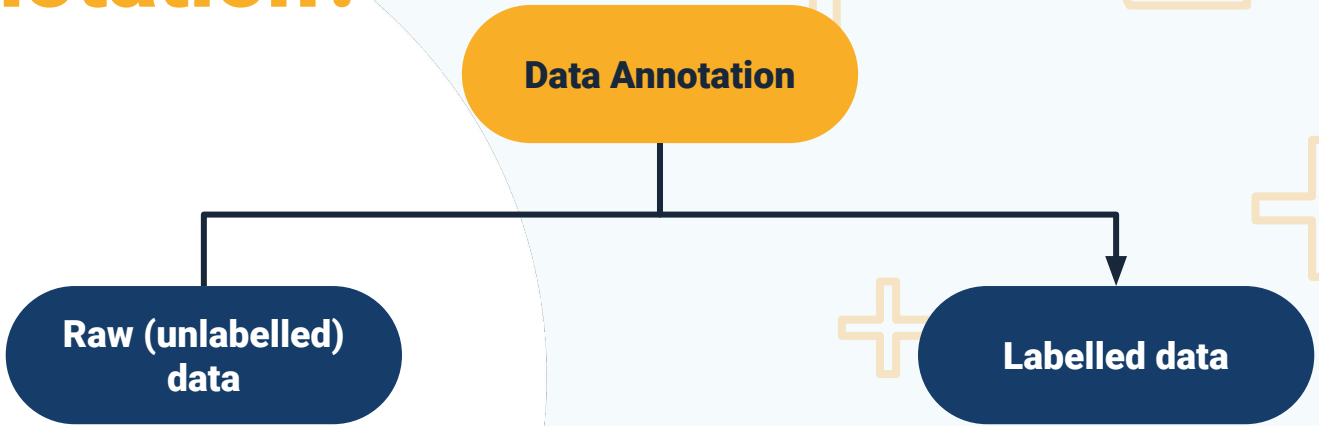
How we built our state-of-the-art machine learning infrastructure with a human-in-the-loop

Fraud Detection & Data Quality

How we leverage our algorithms to achieve unassailable level of audience quality

Data Annotation:
The New Frontier

What is annotation?



The Bottleneck

Time	Incoming Data Points	Human Annotation (120 labels/hr)	Labelled Data Percentage
1 day	138K	2 hrs/day, 2 people	0.4%
1 month	4.15M	4 hrs/day, 4 people	1.4%
1 quarter	12.5M	8 hrs/day, 8 people	5.8%
1 year	50M	8 hrs/day, 100 people	18%

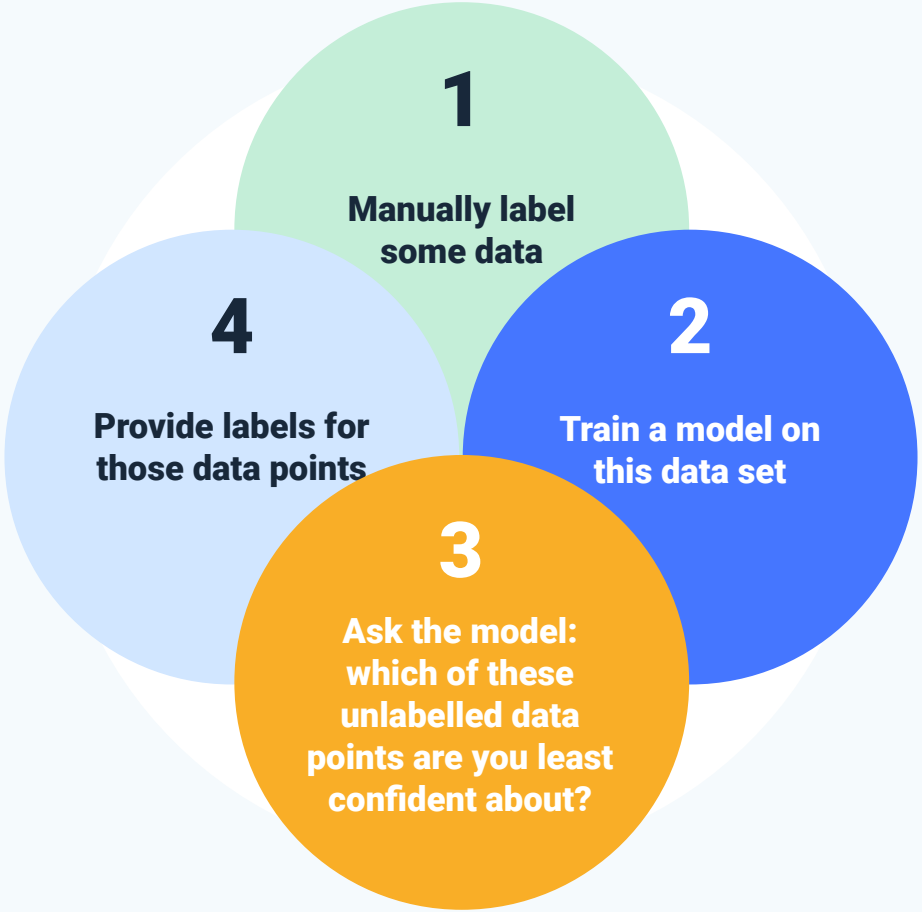
Active Learning
Human-in-the-loop

If we allow our algorithm to ask us:

Hey, I'm not feeling very confident about my prediction for this respondent answer, can you help me out?

Would it perform better?

Active Learning Cycle.



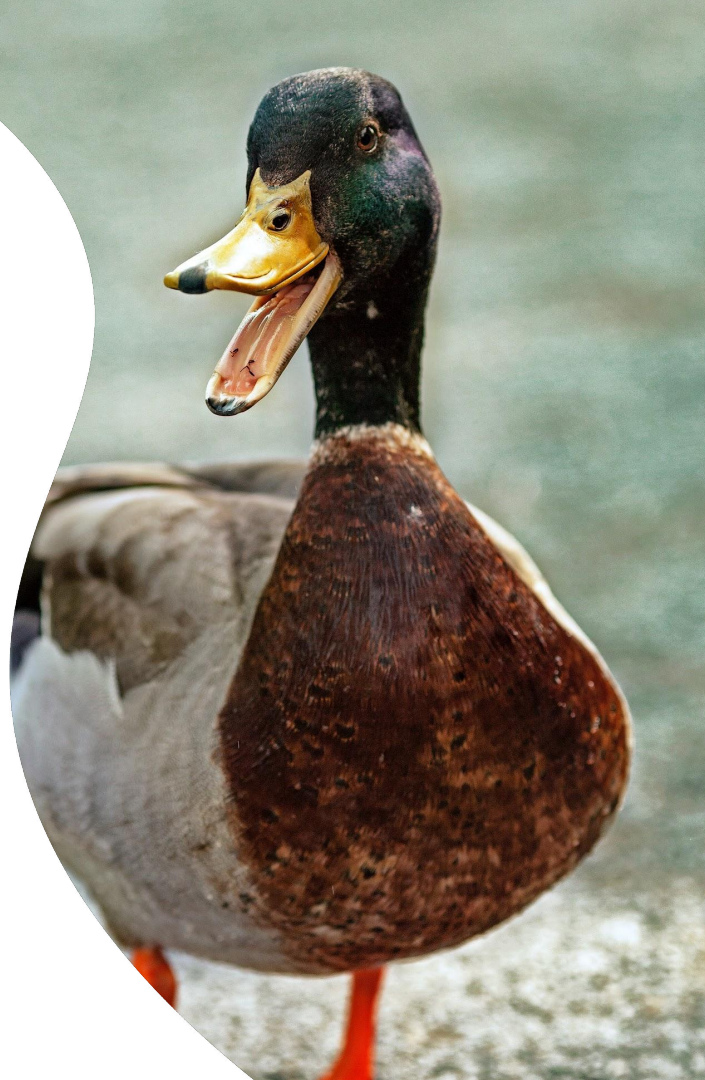
“All good.

What else can we do to expedite the labelling process?”

AI-Assisted Labelling

If it looks like a duck

1. **Predict labels** for unlabelled data points.
2. **Label the data points** that the model is highly confident about.
3. **Continue asking** for human annotation for least confident data points.



Fraud Detection

We combine active learning with AI-assisted machine learning predictions to detect fraud in real time.

Fraud Detection Algorithm
Powering Attest

Top 3
Takeaways

01. Humans provide labels for the most challenging cases to teach the model

02. Models suggest labels for similar data points, instantly labelling 1000s of data points

03. Combining both approaches, data annotation is handled at scale, internally

ATTEST

#AttestEvents

Questions

Please use the box to the right of your screen to submit

#AttestEvents

Get in touch!

Dr. Gökhan Çiflikli
gokhan.ciflikli@askattest.com

... via our website to
Book an Intro or
Try Now.

Hungry for more webinars?

ATTEST

askattest.com/webinars

/askattest @askattest #askattest