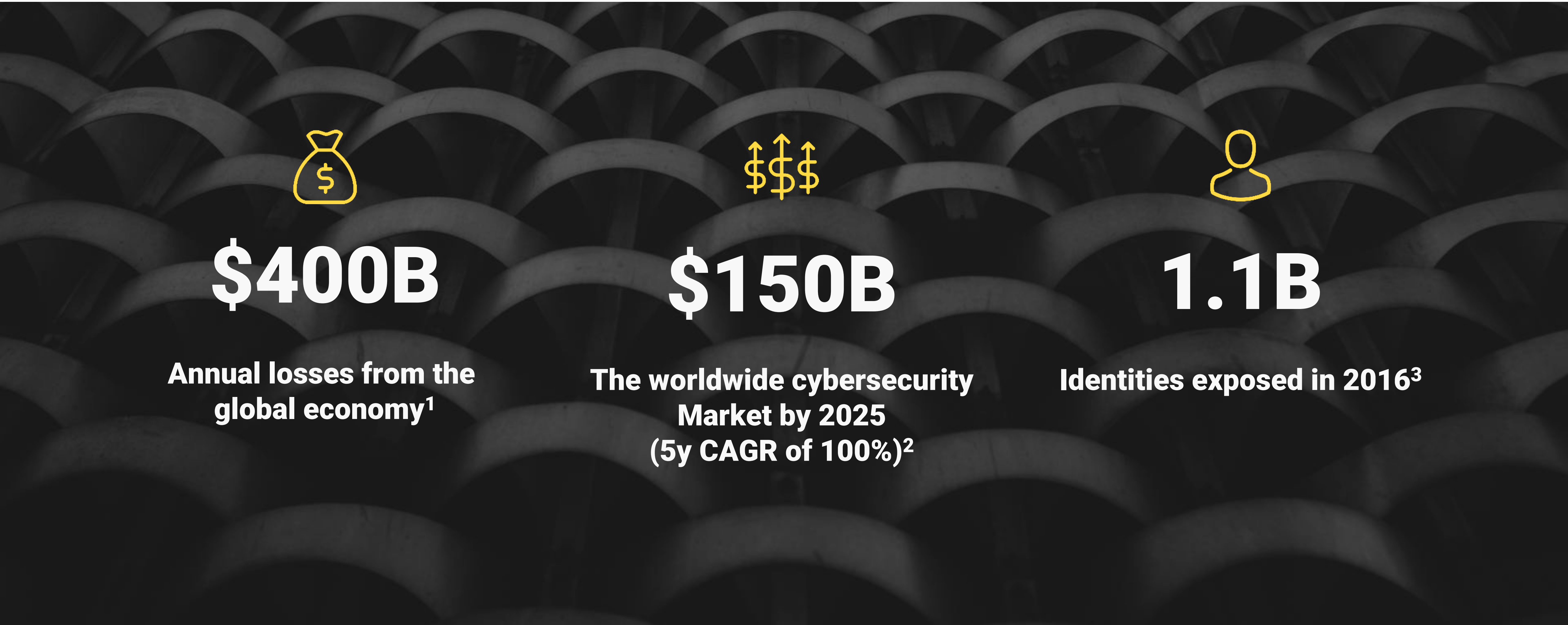
BIOKEY

Brandon Kucera Tony Wu Connor Giles
Sankalp Hariharan Josh Weinstein

Advisor: Dr. Jagath Samarabandu

Cybersecurity is a growing problem



$400B

Annual losses from the global economy[1]

$150B

The worldwide cybersecurity Market by 2025
(5y CAGR of 100%)[2]

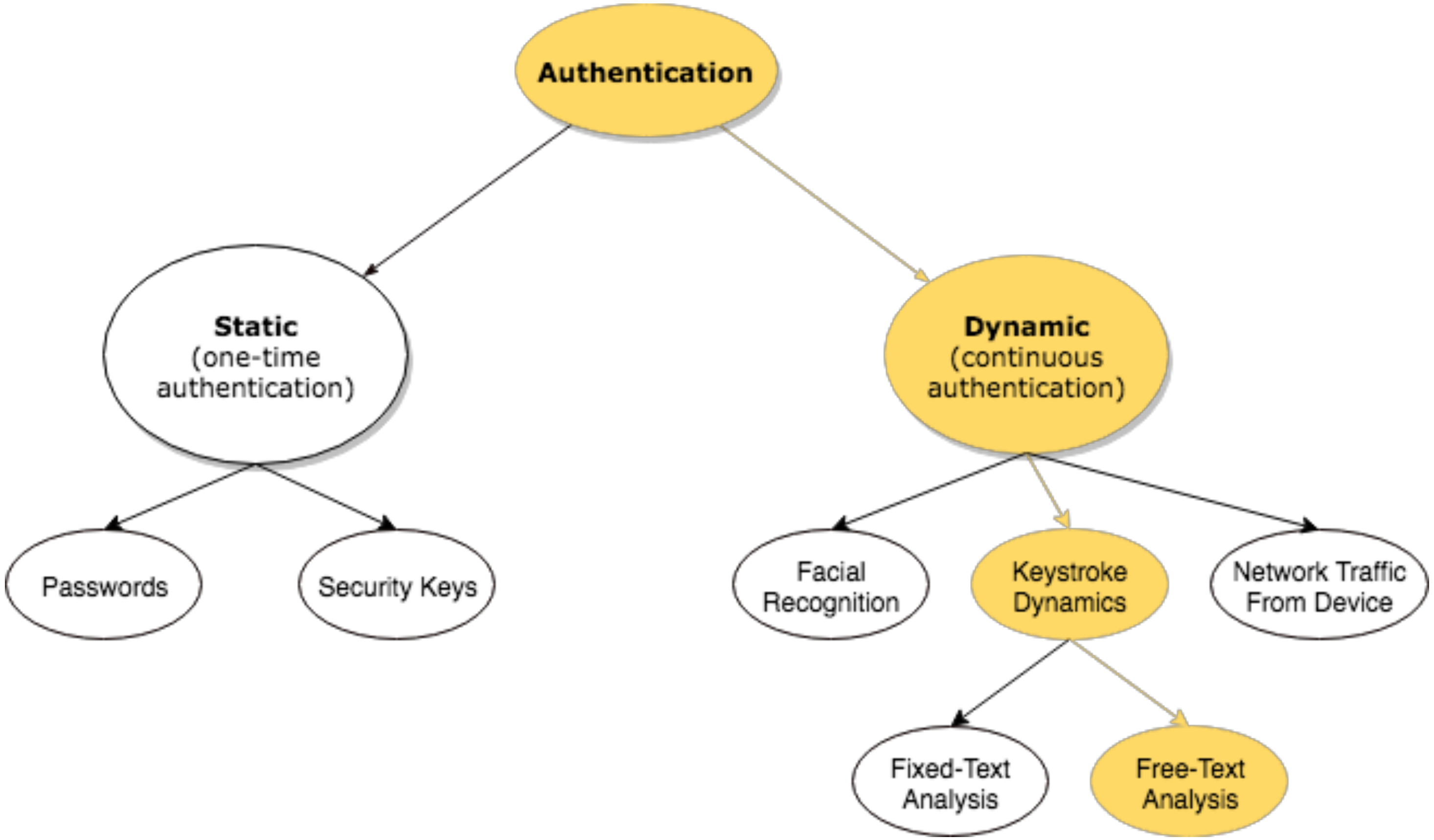
1.1B

Identities exposed in 2016[3]

Authentication plays a key role in security

General trend toward dynamic authentication in the industry because it is harder to imitate behaviour than it is to get a password.

Keystroke dynamics have academic support, but have not been implemented commercially because it has historically been computationally expensive.



Implemented Functionality

BioKey will constantly monitor the behaviour of the user for suspicious activity

Run Locally

Application runs locally
in the background of machines

Detect Suspicious Behaviour

Checks keystroke input against the model to detect suspicious behaviour

Lock Out Imposters

If suspicious behaviour is detected, BioKey locks the machine

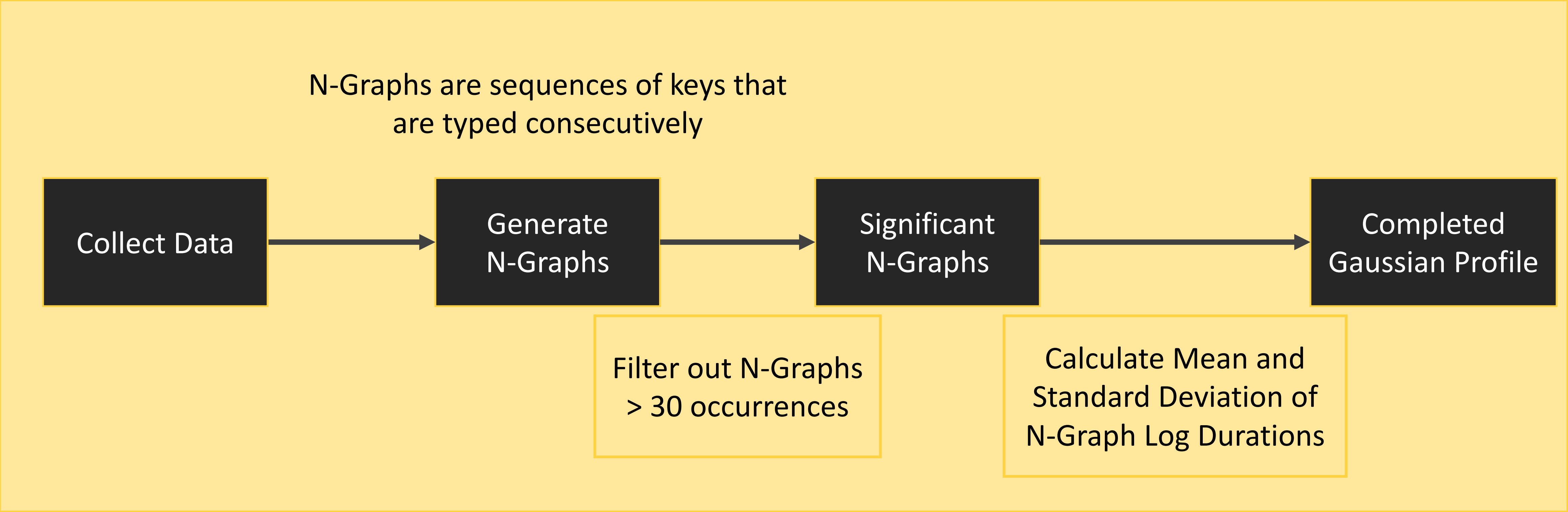
Online and Offline Reauthentication

Users can reauthenticate using either SMS or Google Authenticator

Remote Locking

Admins are able lock and unlock machines remotely through the BioKey companion web application

Generating Gaussian Profiles



N-Graphs are sequences of keys that are typed consecutively

Collect Data

Generate N-Graphs

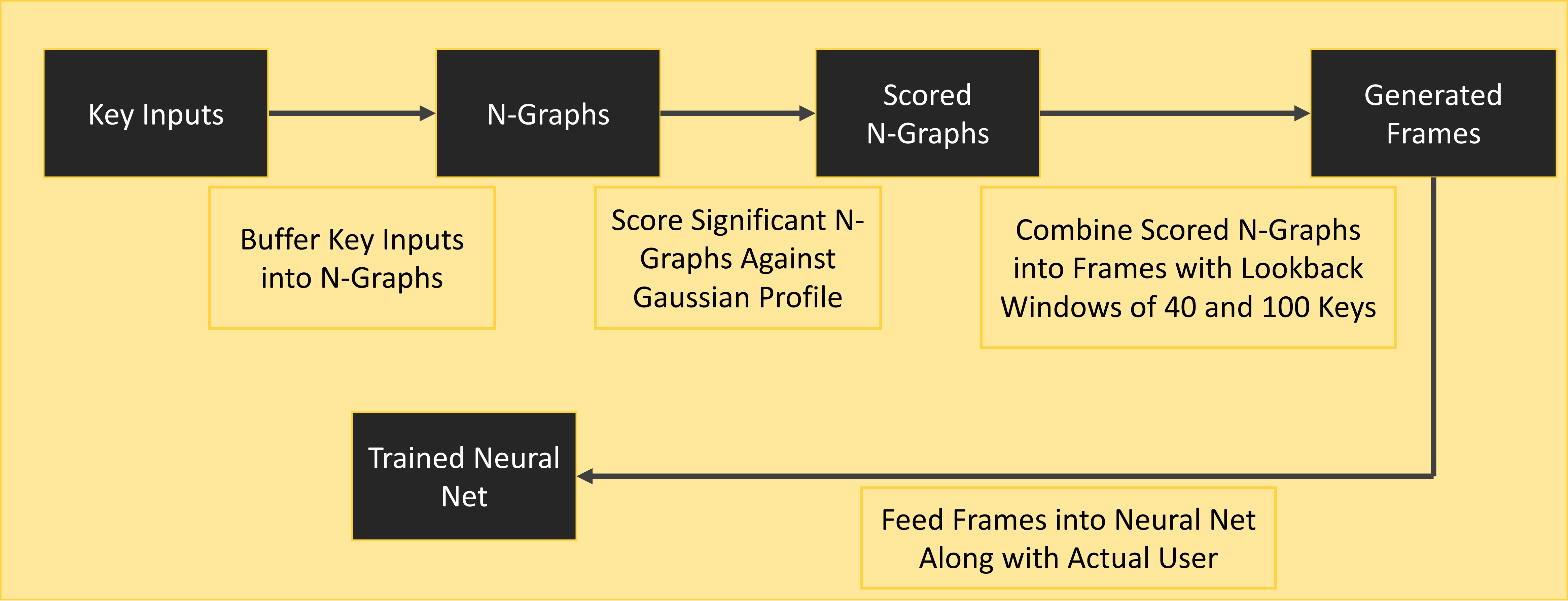
Filter out N-Graphs > 30 occurrences

Significant N-Graphs

Calculate Mean and Standard Deviation of N-Graph Log Durations

Completed Gaussian Profile

Training Neural Nets



Key Inputs

N-Graphs

Scored N-Graphs

Generated Frames

Buffer Key Inputs into N-Graphs

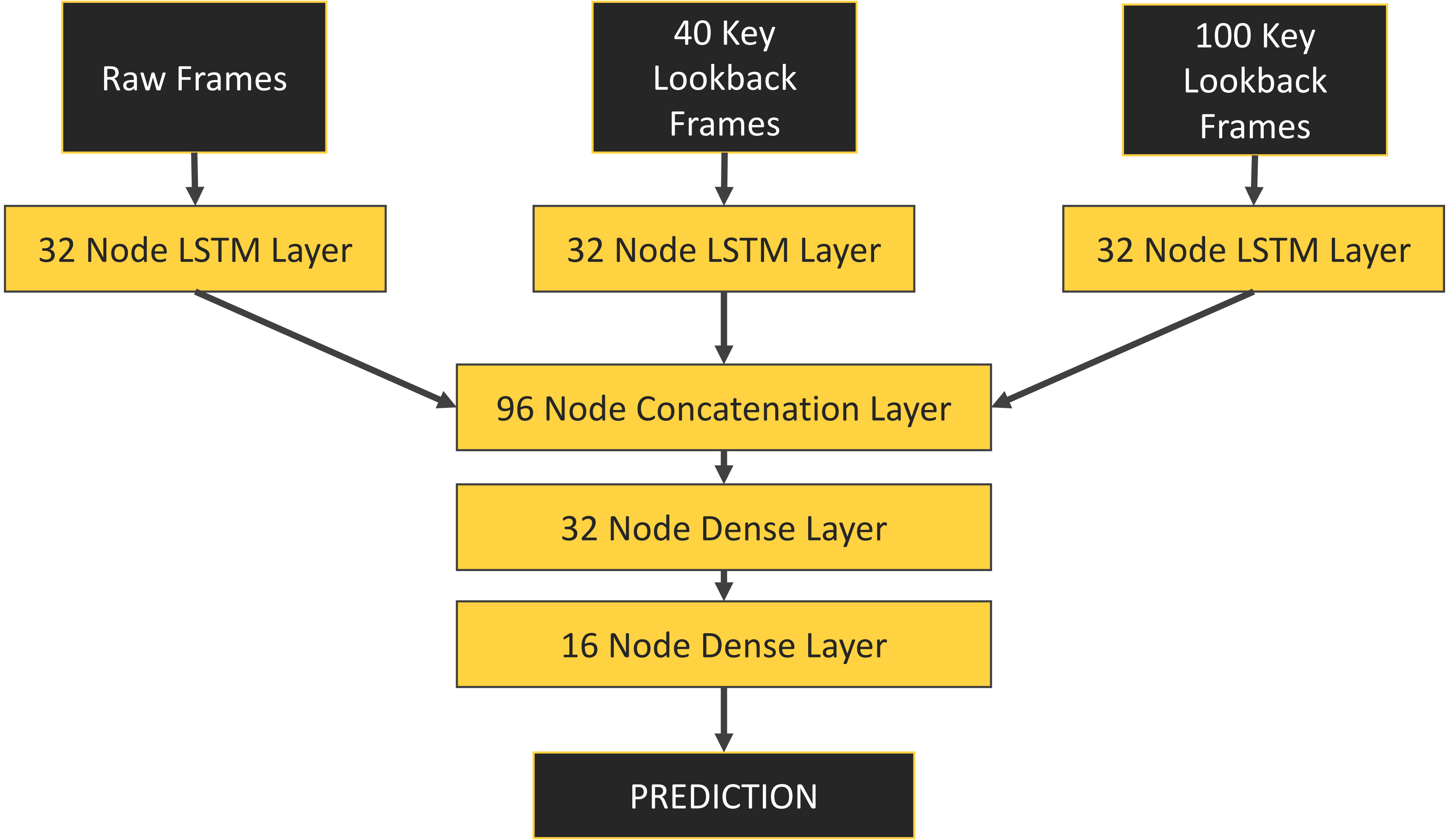
Score Significant N-Graphs Against Gaussian Profile

Combine Scored N-Graphs into Frames with Lookback Windows of 40 and 100 Keys

Trained Neural Net

Feed Frames into Neural Net Along with Actual User

Neural Net Architecture



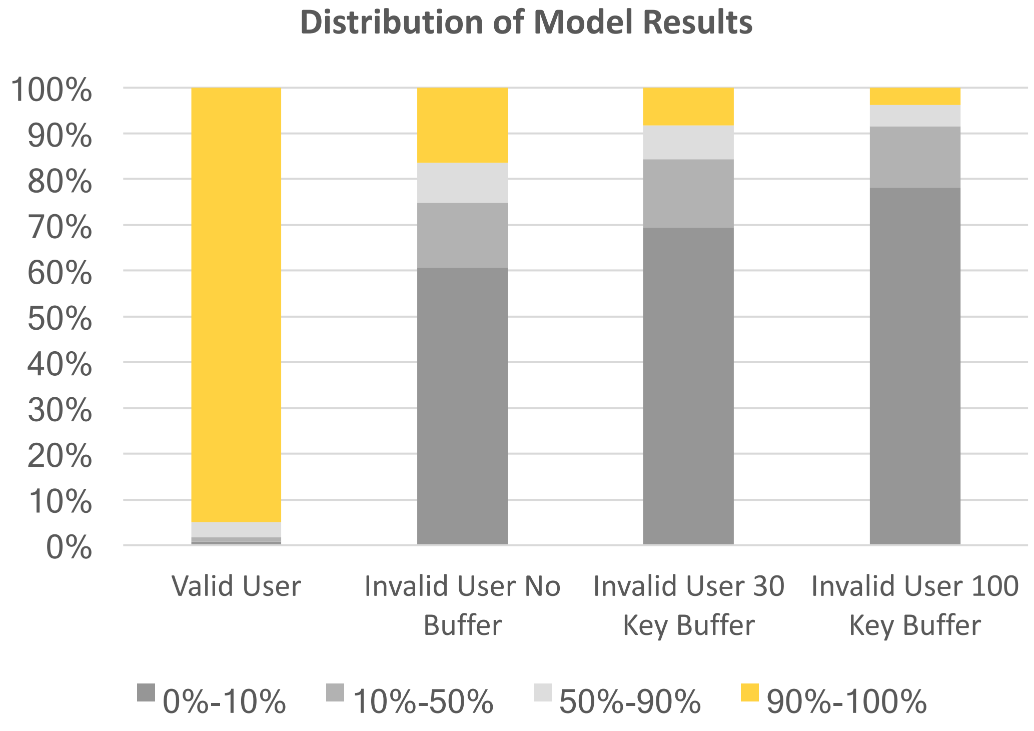
Model Prediction Results

We evaluated our model by comparing model prediction results to actual users

On average the model predicted a value of **96.65% for actual users** and **26.54% for imposters**

Looking at all imposter results is a tough bar. The model needs some time to determine when an imposter has started typing. When a grace period was added, the model performed significantly better.

With a **buffer period of 30 and 100 keys**, average imposter scores dropped to **17.95% and 10.80%**

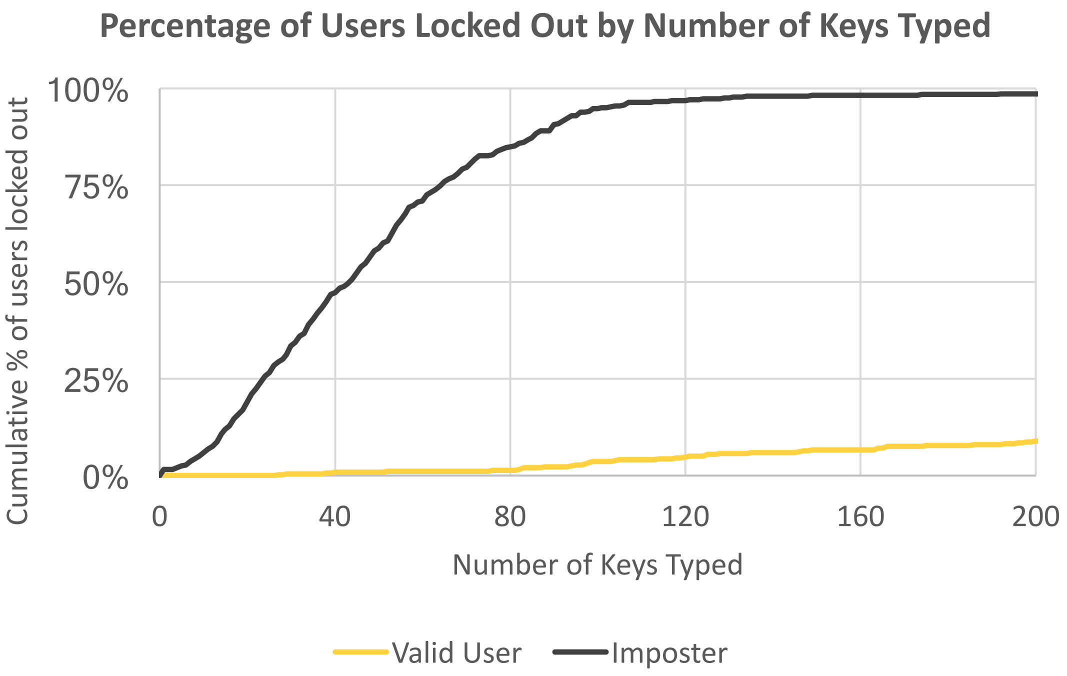


Simulated Real World Results

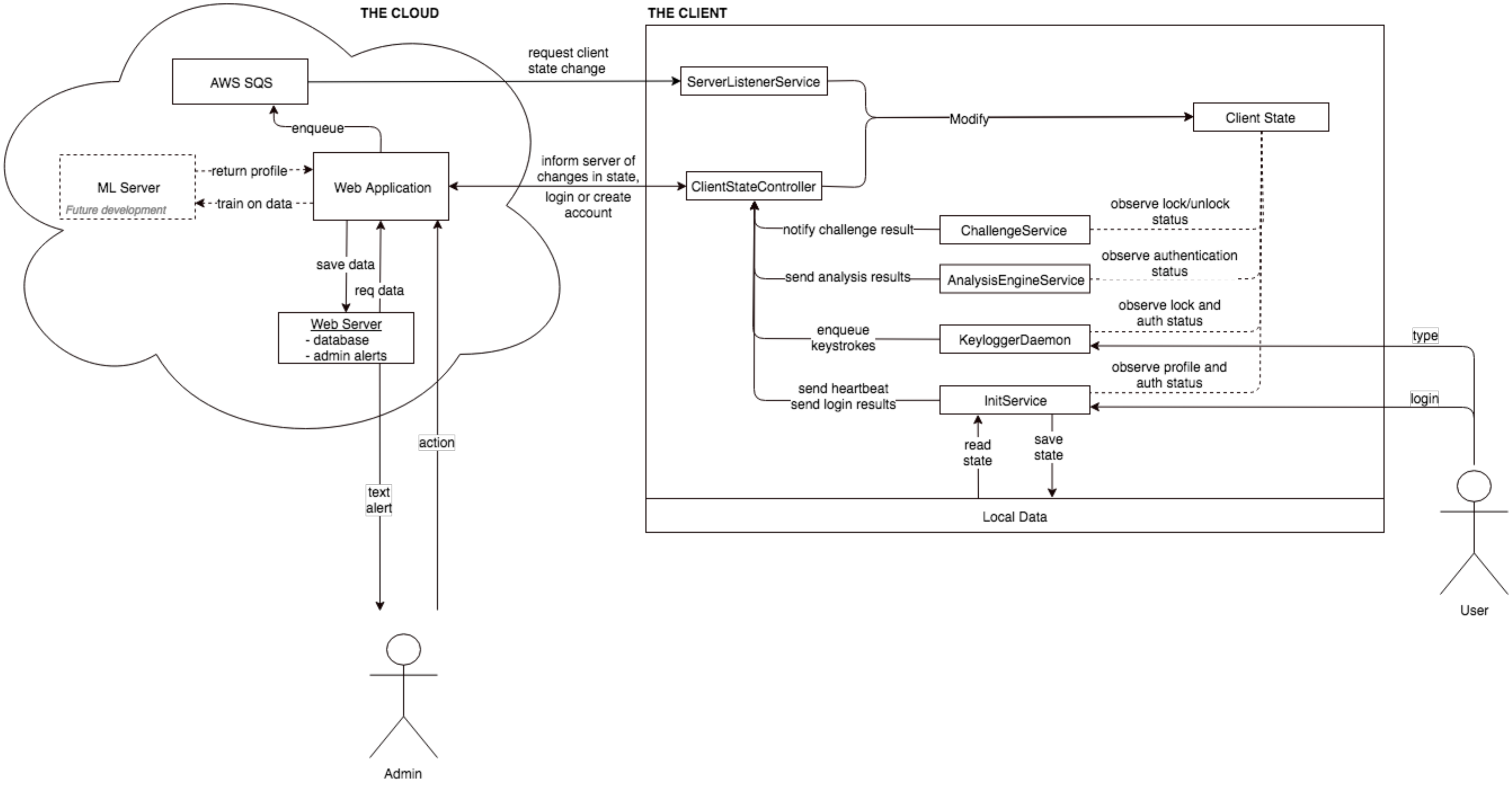
We simulated a series of 200 user keystrokes followed by 200 imposter keystrokes and tracked results by session.

98.8% of imposters were locked out in an average of **47 keystrokes.**

9% of users are locked out in an average of 120 keystrokes. This translates to approximately **2200** keystrokes on average **between improper lockouts.**



High-Level Architecture



We leveraged many technologies

ON THE WEB

JS

node JS

ember

HEROKU

mongoDB

amazon web services

ML TECH

python™

Keras

TensorFlow

CLIENT

Java

spring

DATA VIZ

ALTERYX

We leveraged many technologies

PROCEDURAL

Travis CI

git

slack

APIDOC

Trello

Challenges Surmounted

BioKey's ambition posed many obstacles

Obscurity of the Field

No standard or best practice for keystroke CA so we had to get creative to improve

Sparse Data

Recurrent neural networks struggle to converge with highly sparse data

Improving Upon Existing Results

Developed a novel algorithm as existing methods did not show reasonable results

Large Amounts of Data

Processing efficiency was a constant consideration

High Security Expectations

We had to consider many possible ways in which BioKey could be circumvented

Maintaining Client-Server Agreement

Especially considering network disconnects

LIVE DEMO

Resources

1) https://www.reuters.com/article/us-cybersecurity-mcafee-csis/cyber-crime-costs-global-economy-445-billion-a-year-report-idUSKBN0EK0SV20140609
2) http://blogs.wsj.com/venturecapital/2016/02/17/the-daily-startup-increased-spending-in-cybersecurity-drives-funding-surge/
3) 2016 Symantec Internet Security Threat Report