

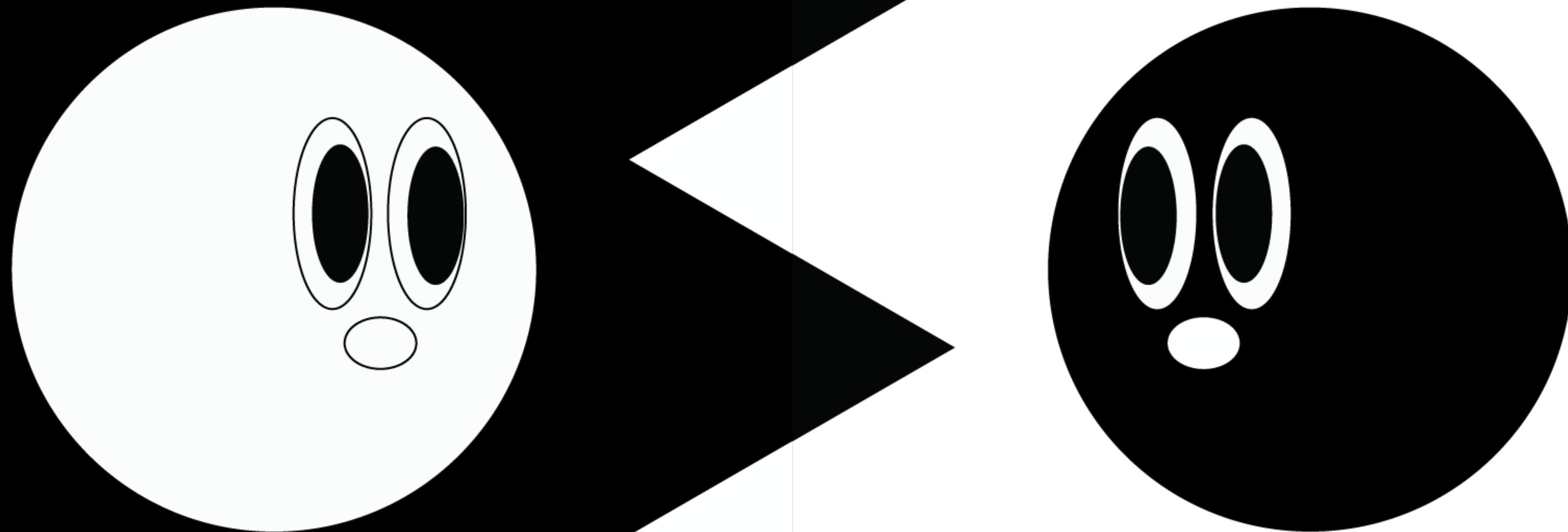
MSc Data Science and AI for the Creative Industries
2022 Graduate Thesis

**Visualising User Embedding of an NCF Recommender Model
– A Creative Approach on Improving Media and Information Literacy**

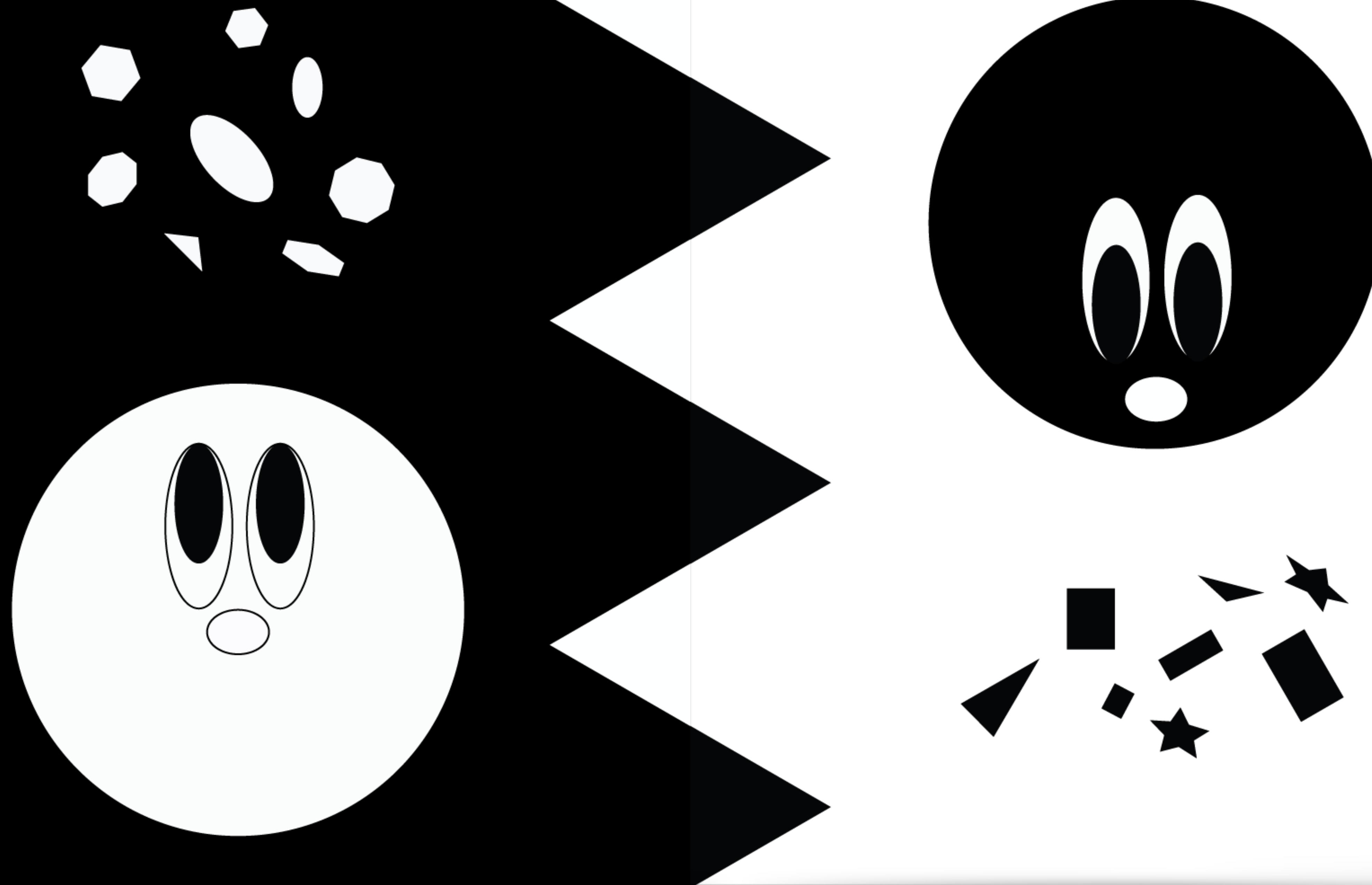
Hsi Fang Chen

Introduction

Filter Bubble | Echo Chamber



Filter Bubble | Echo Chamber



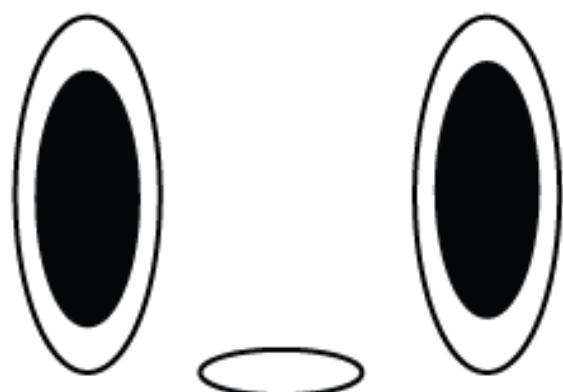
Interdisciplinary Method

MIL

Framework



NCF Model



Visualisation

Media and Information Literacy

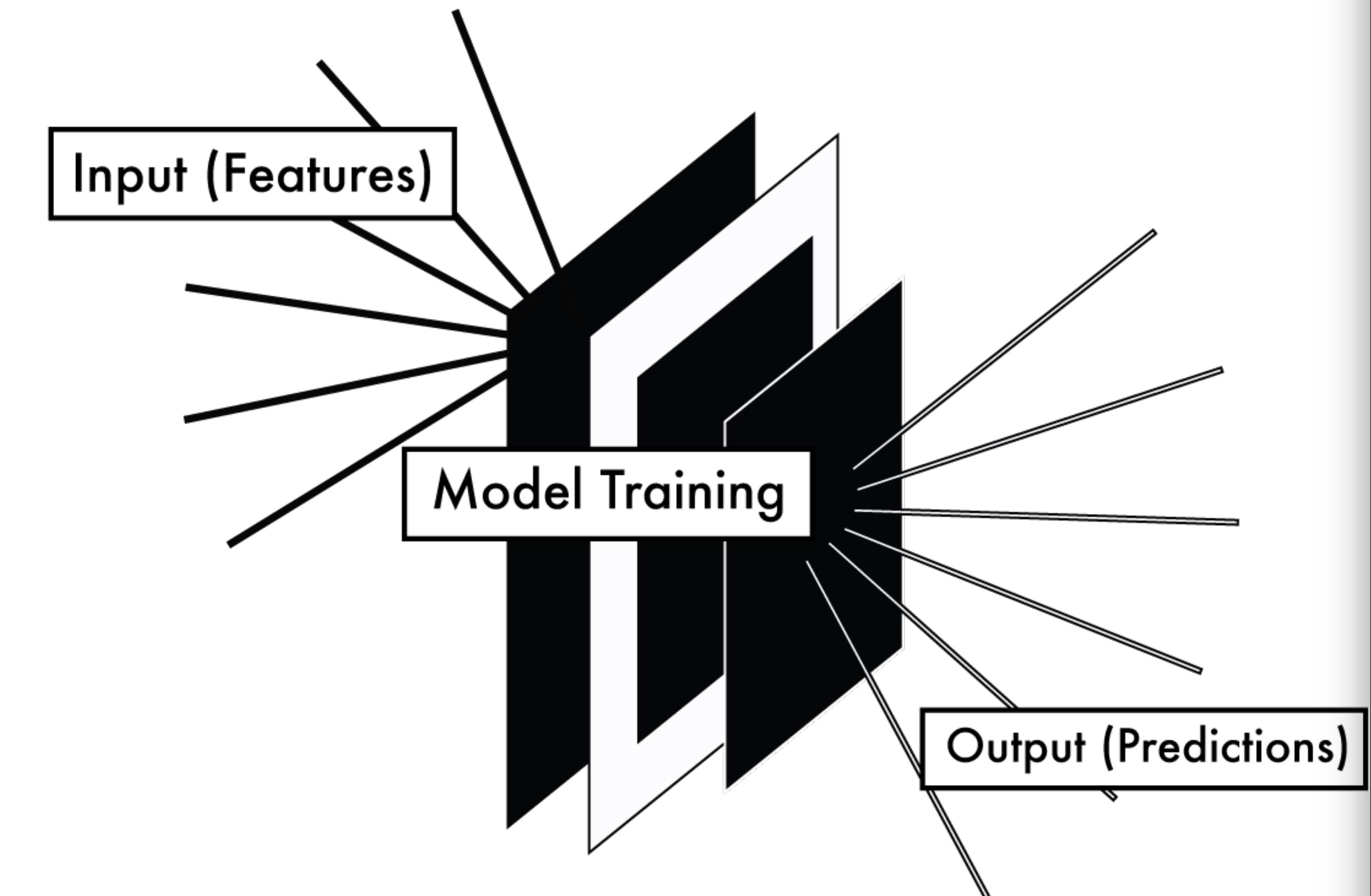
**Media Literacy
Algorithm Literacy
Digital Literacy**



*Critical Thinking
Resilience
Awareness*

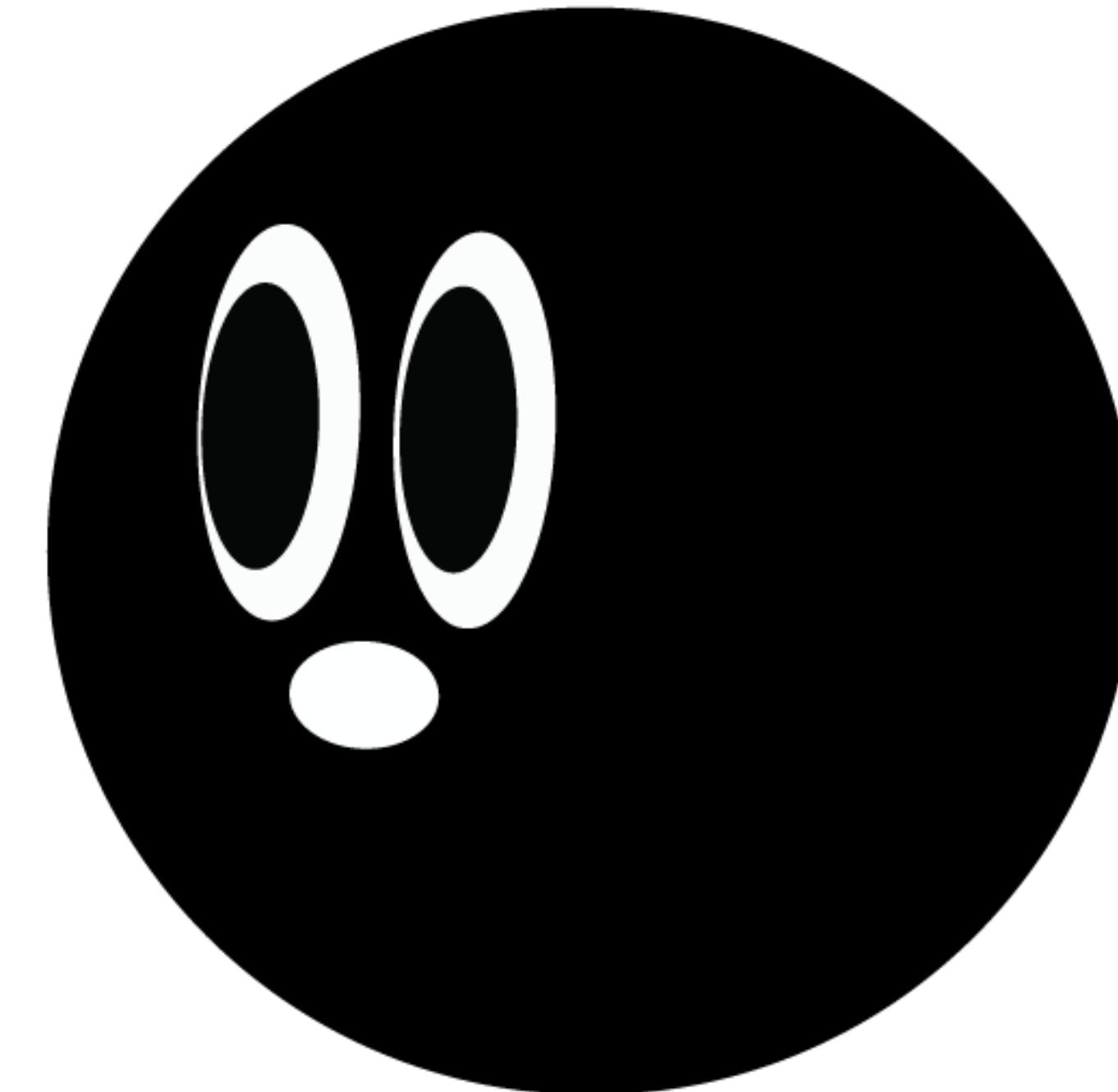
Algorithmic Personalisation

Collaborative Filtering

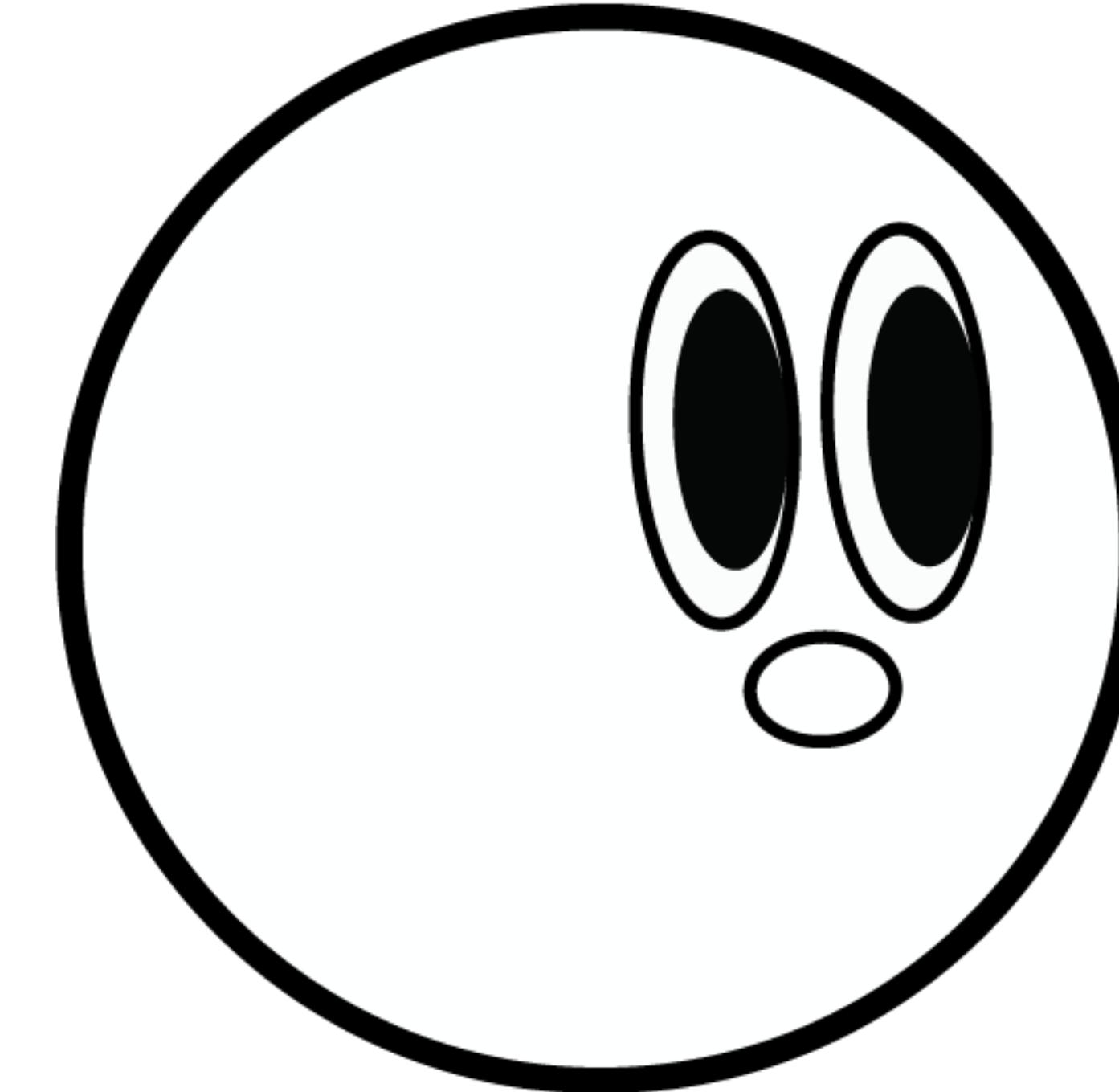


User Recruitment

DSML

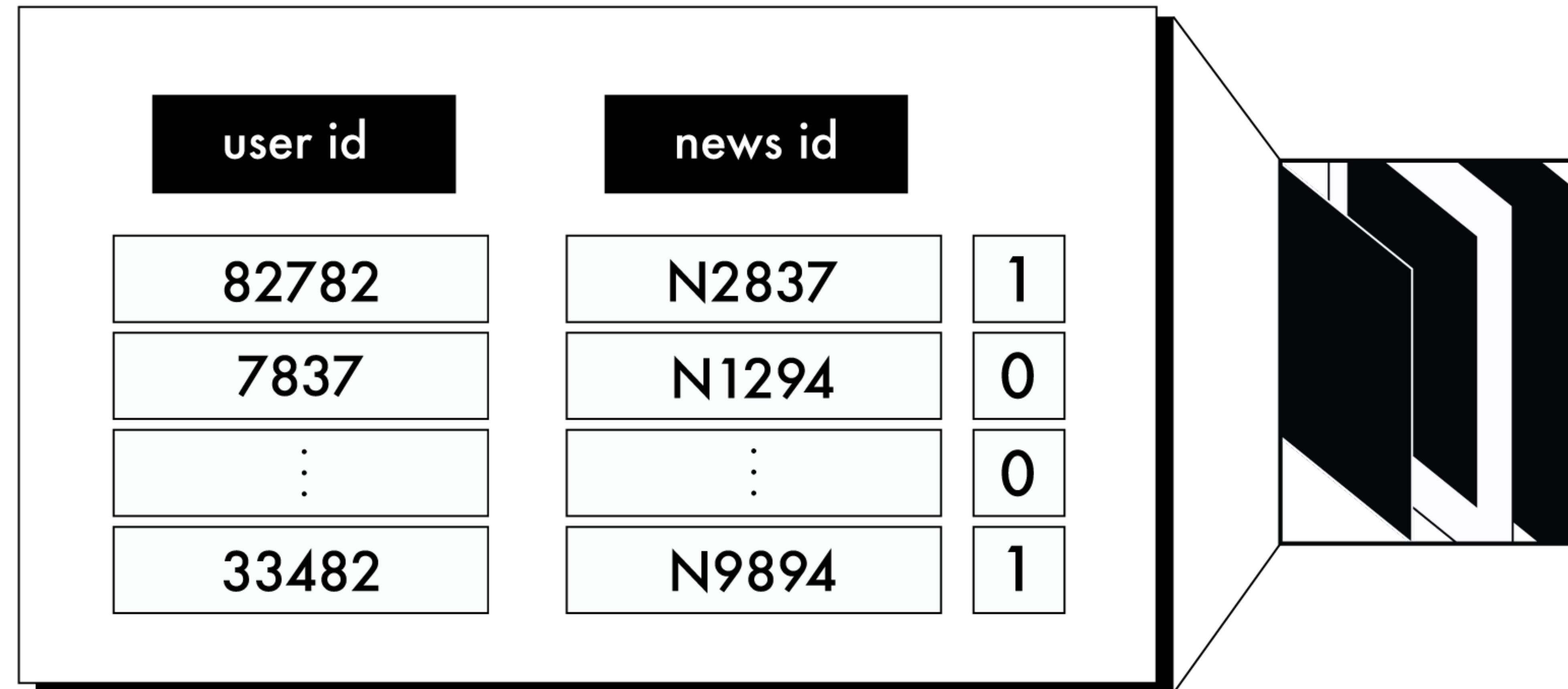


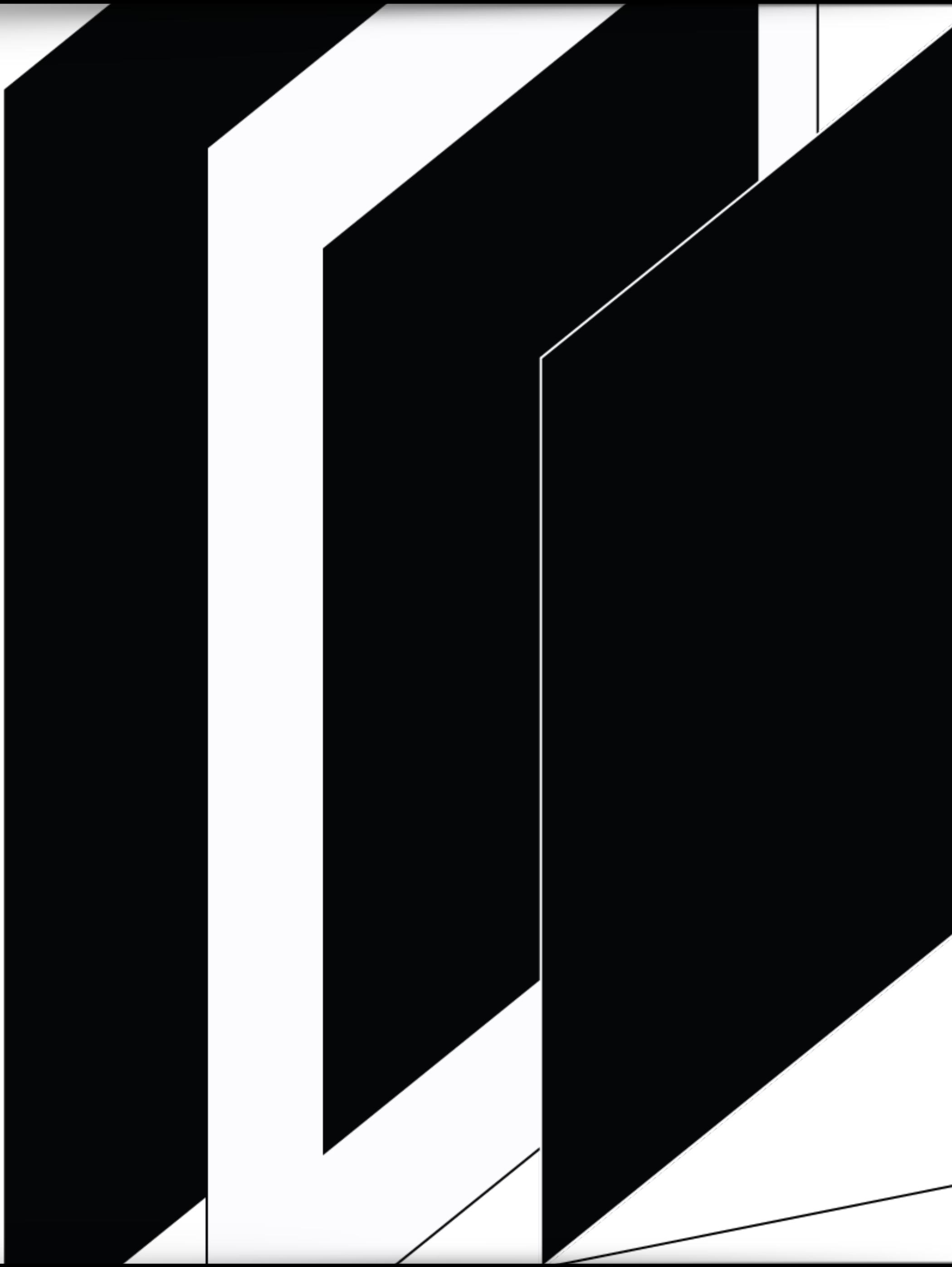
non-DSML



Impression Log

NEWS 1		====	1
NEWS 2		====	1
NEWS 3		====	0
NEWS 4		====	1
NEWS 5		====	0





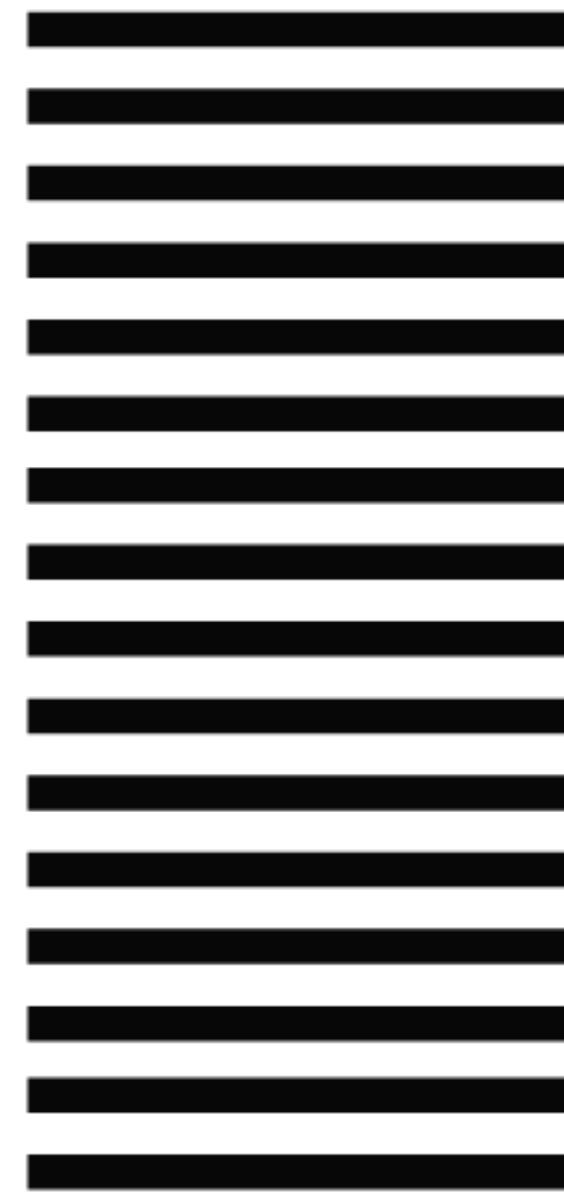
16 - dimensional embedding

0.-23955868
-0.23865387
-0.11432469
...
-0.27637732
0.31800118
-0.2393788

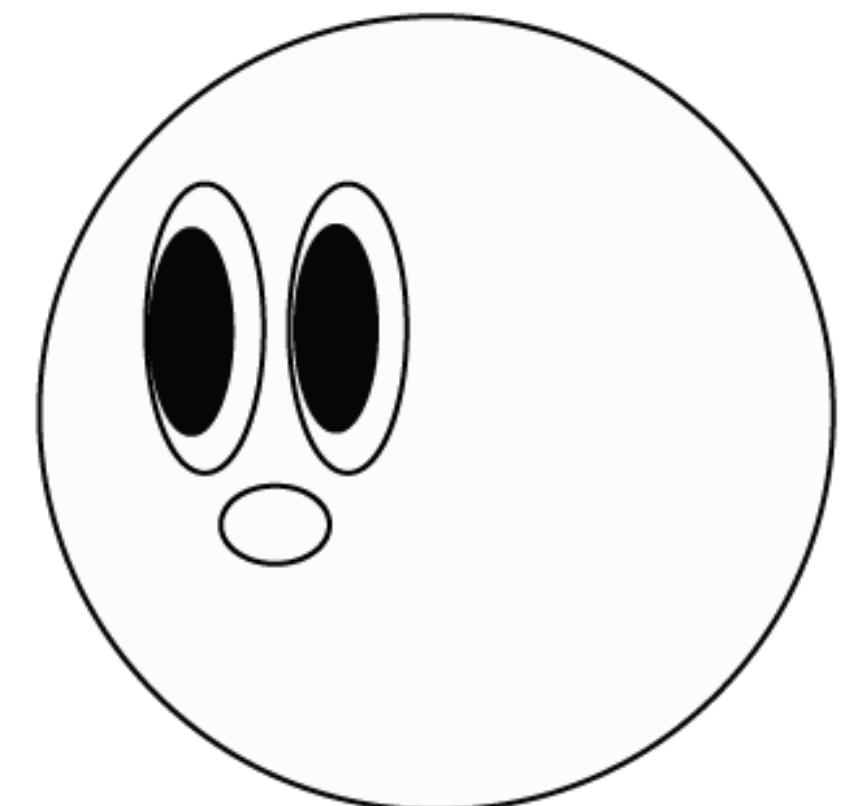
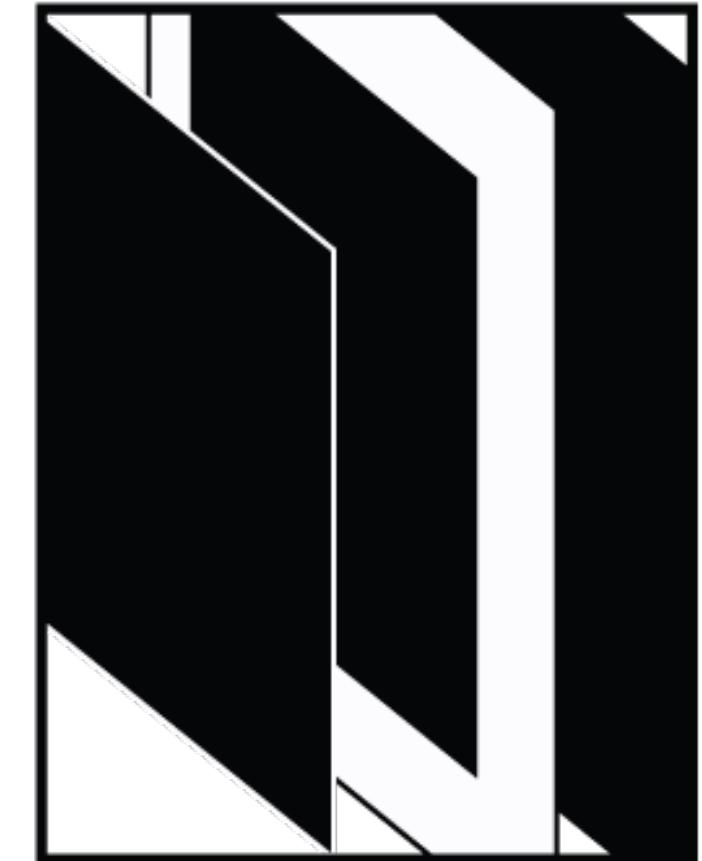
Principal Component Analysis

Dimensional Reduction

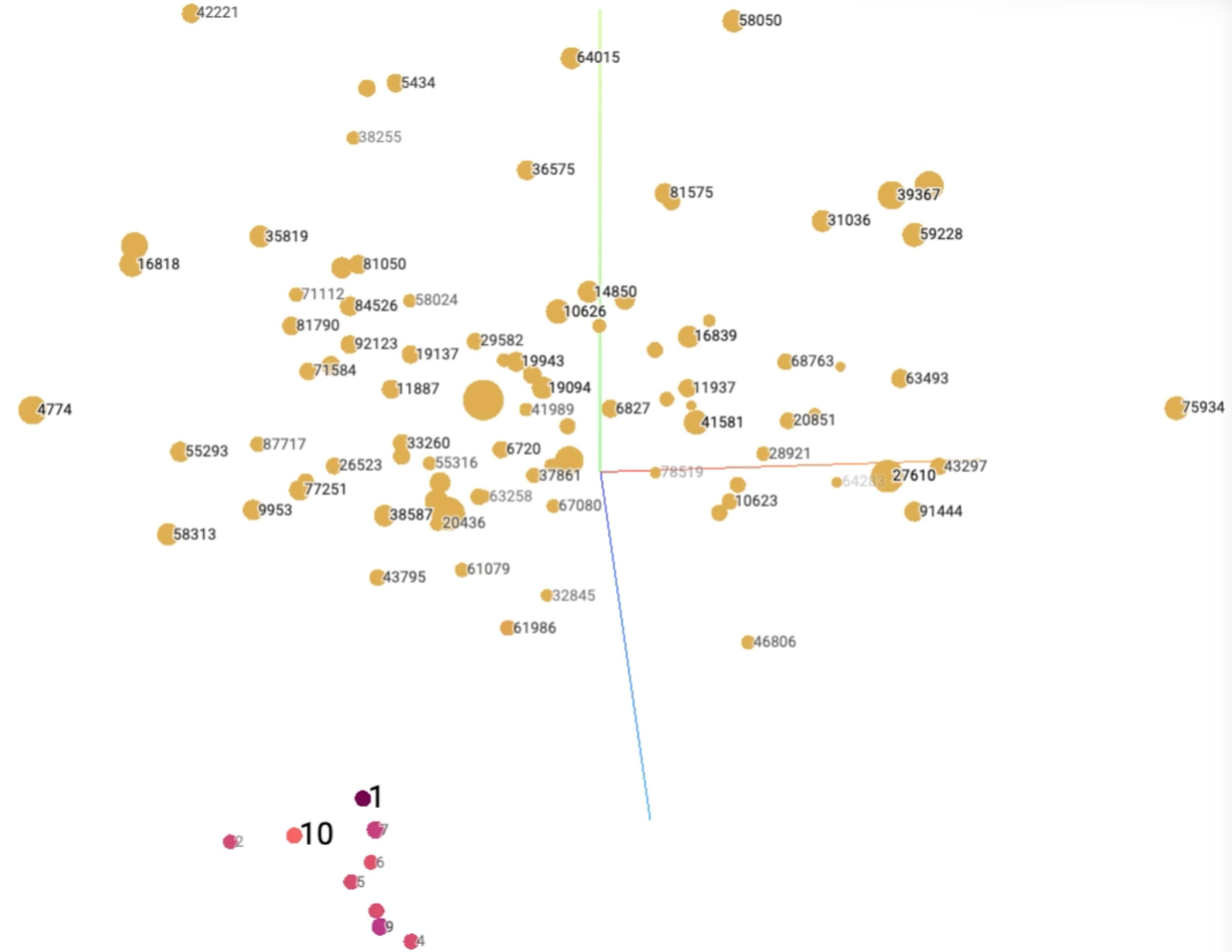
16 - dimension



3 - dimension



Google Embedding Projector



Data Visualisation 3D Graph..

points & numbers

20436

each point
represents a
user

each point is
linked with an
unique user id

meaning of the colours

participants in
this project

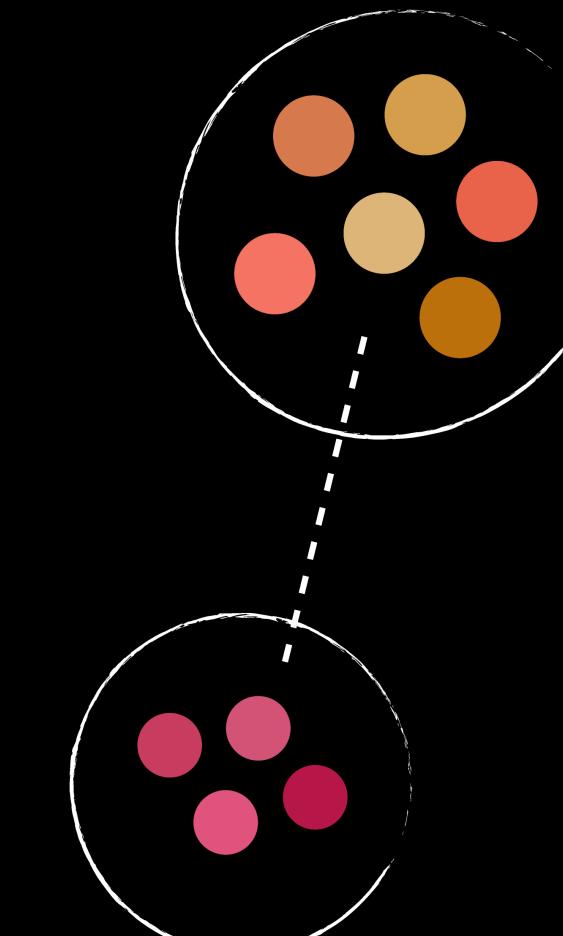
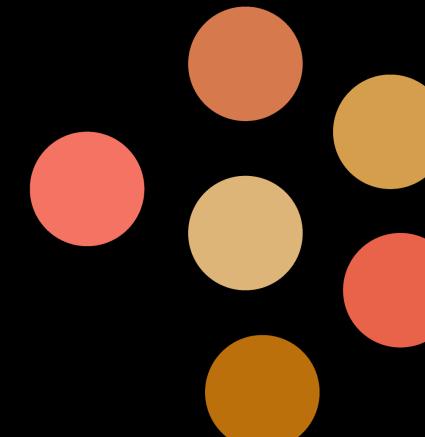
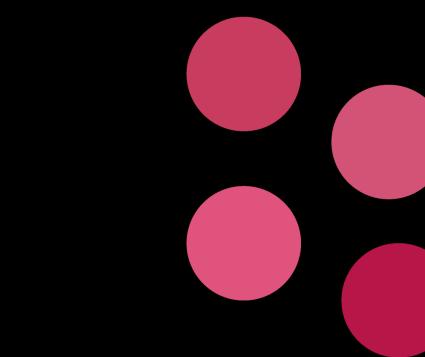
anonymous
users in the
Microsoft news
database

participants in
the project re-
ceived a limited
selection of
news articles

anonymous users
have clicked on the
news on the Microsoft
news platform accord-
ing to their interest

The participants
and anonymous
users are apart be-
cause they have dif-
ferent click history

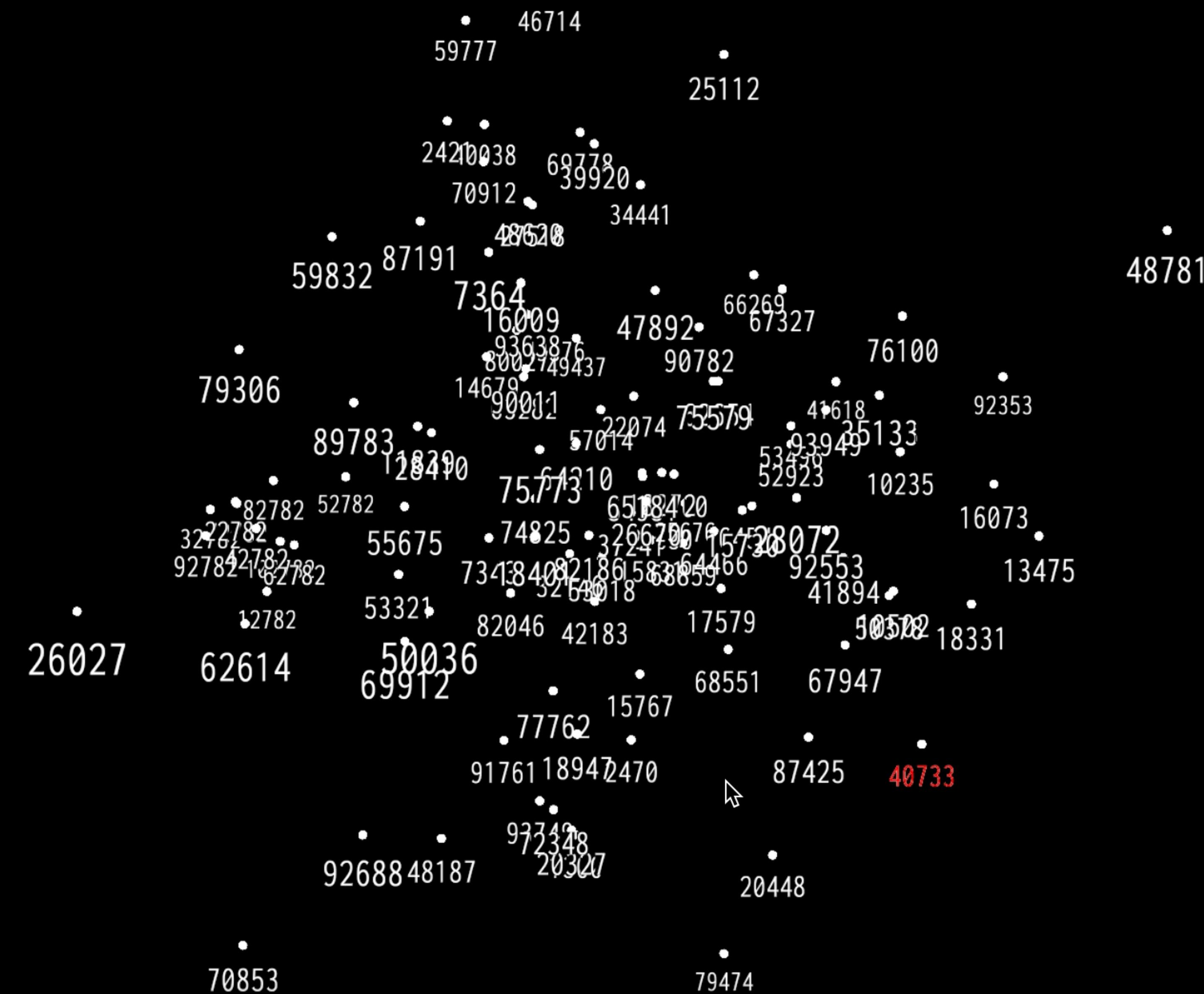
Participants that re-
ceived similar con-
tent formed a dense
cluster, i.e. echo
chamber/ filter
bubble



Outcome

p5.js Demo

46175



type user id below

40733

59832 87191

70912 39920

4865208 34441

79306

89783 82782 52782
3222782 1027128490
92782.

1600938576 66269
3849437 14679 67327
900282 47892
57022074 41618 76100
62654 92353

55675 75773 3828868 16073
55321 18401 42183 15730
64210 1285726 329293949
74825 65184 1035133
326670 64544

15767 68551 9255378 13475
91761 2470 1050218331

7716947

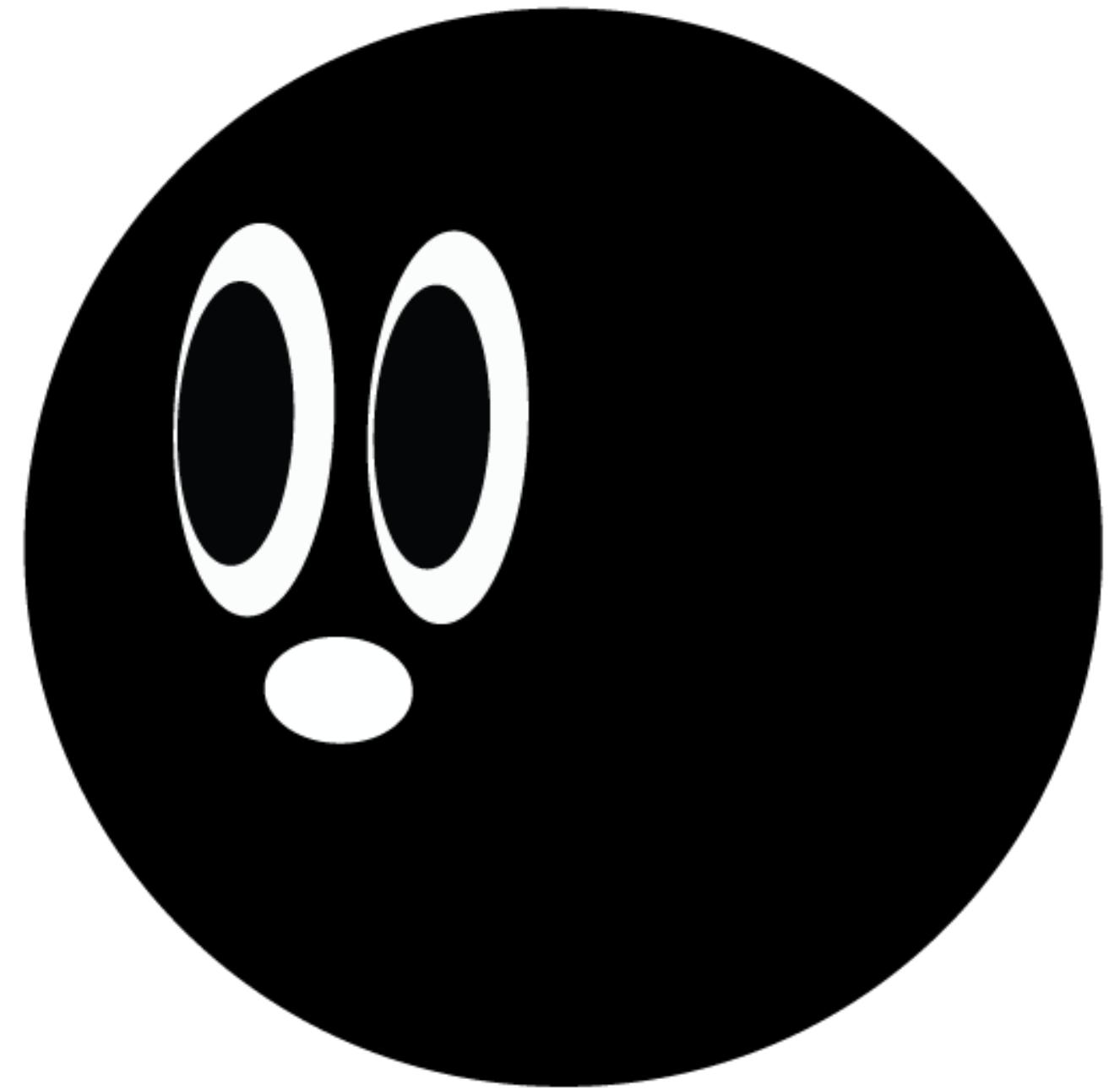
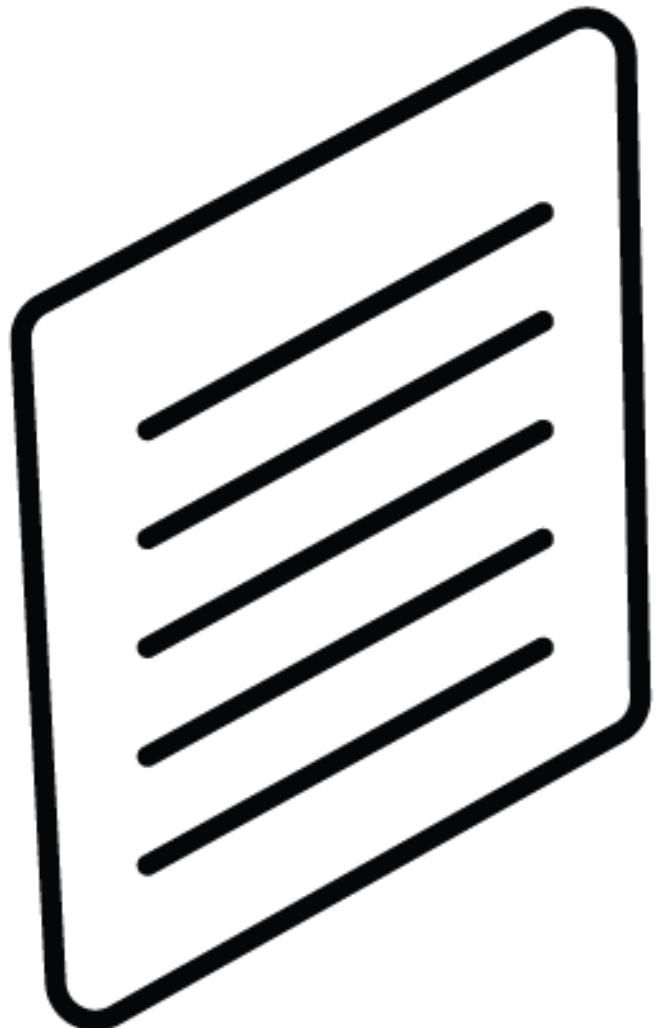
48187 723480

6794733

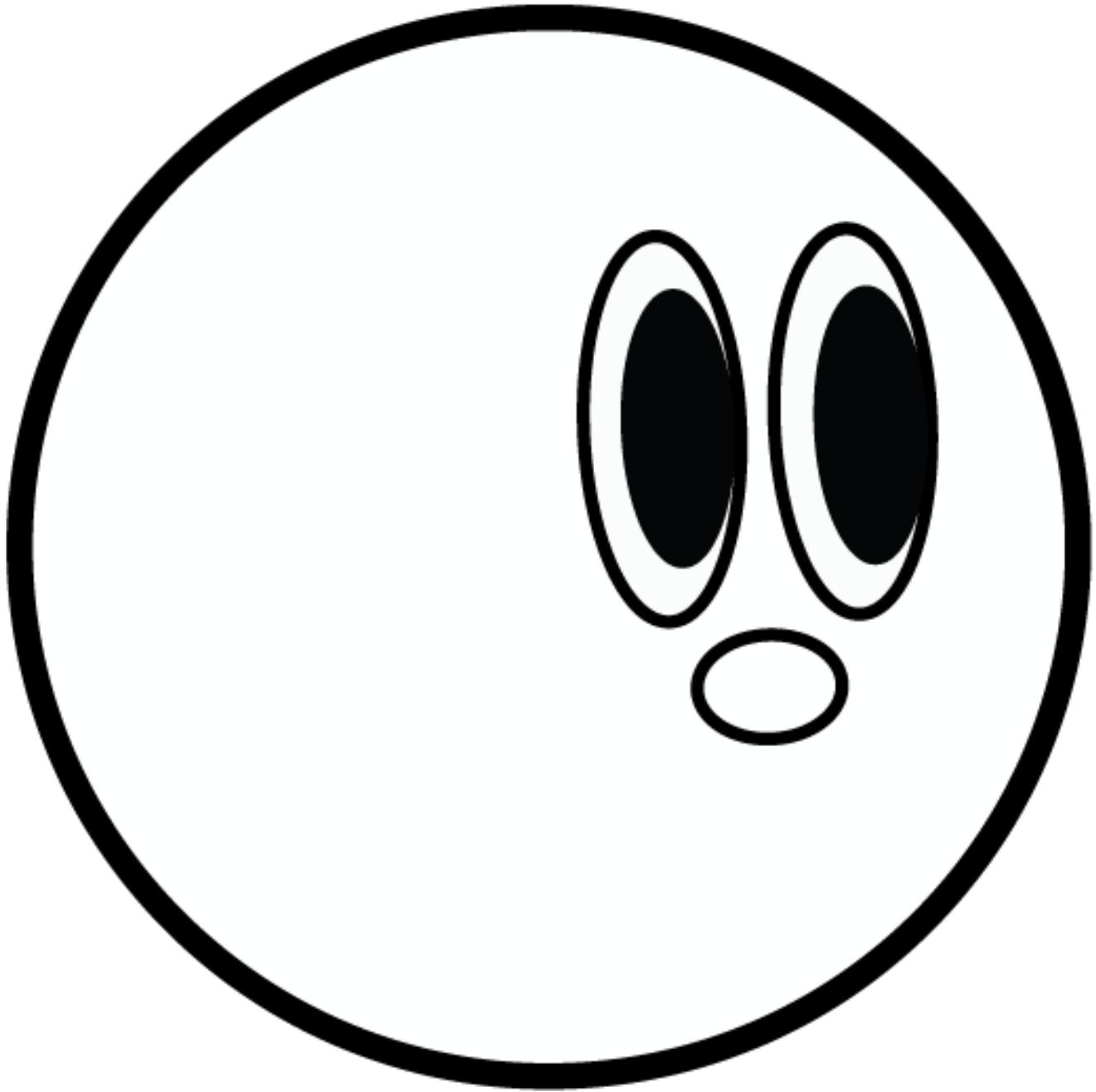
87125

User Feedback

DSML



non-DSML



Evaluation Criteria

- ☒ Topic Familiarity
- ☒ Usefulness | Importance
- ☒ Informative
- ☒ Change of Attitude

only 23 % of all the statements
reject null hypothesis

A few decades ago, when the television replaced traditional newspapers and radio, it became the dominant mass media. There was criticism towards the invention and that even emerged an academic field on “television criticism”.

Today, the younger generation rarely get deceived by TV commercials, and many of them understand that different news channels have different political inclinations.

I believe the same with algorithmic recommenders.

With education and more diverse presentation on the personalisation, which I had proposed in this thesis as an example, would help us to establish a healthier and safer environment to interact with algorithms and personalisation technologies.