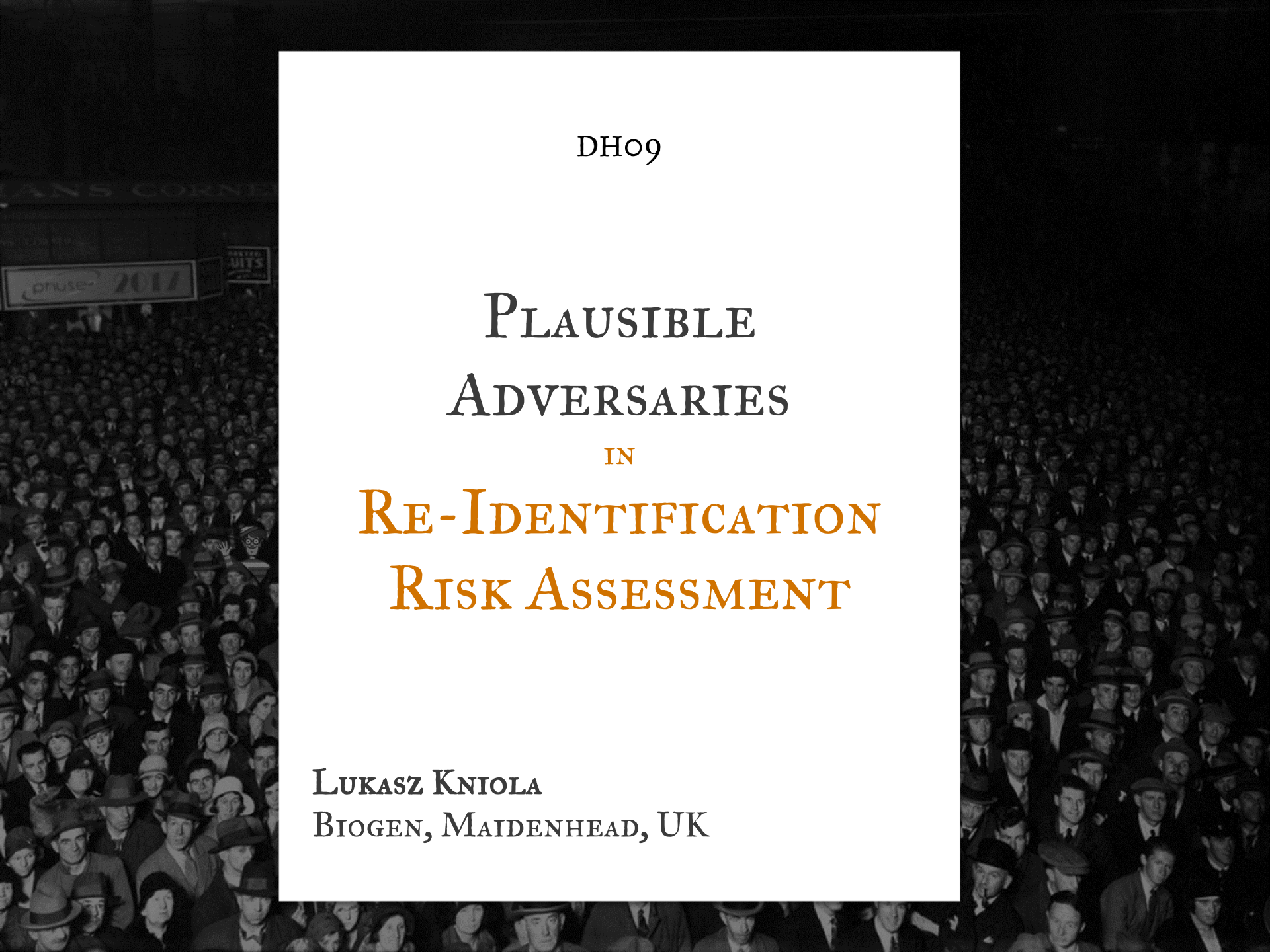
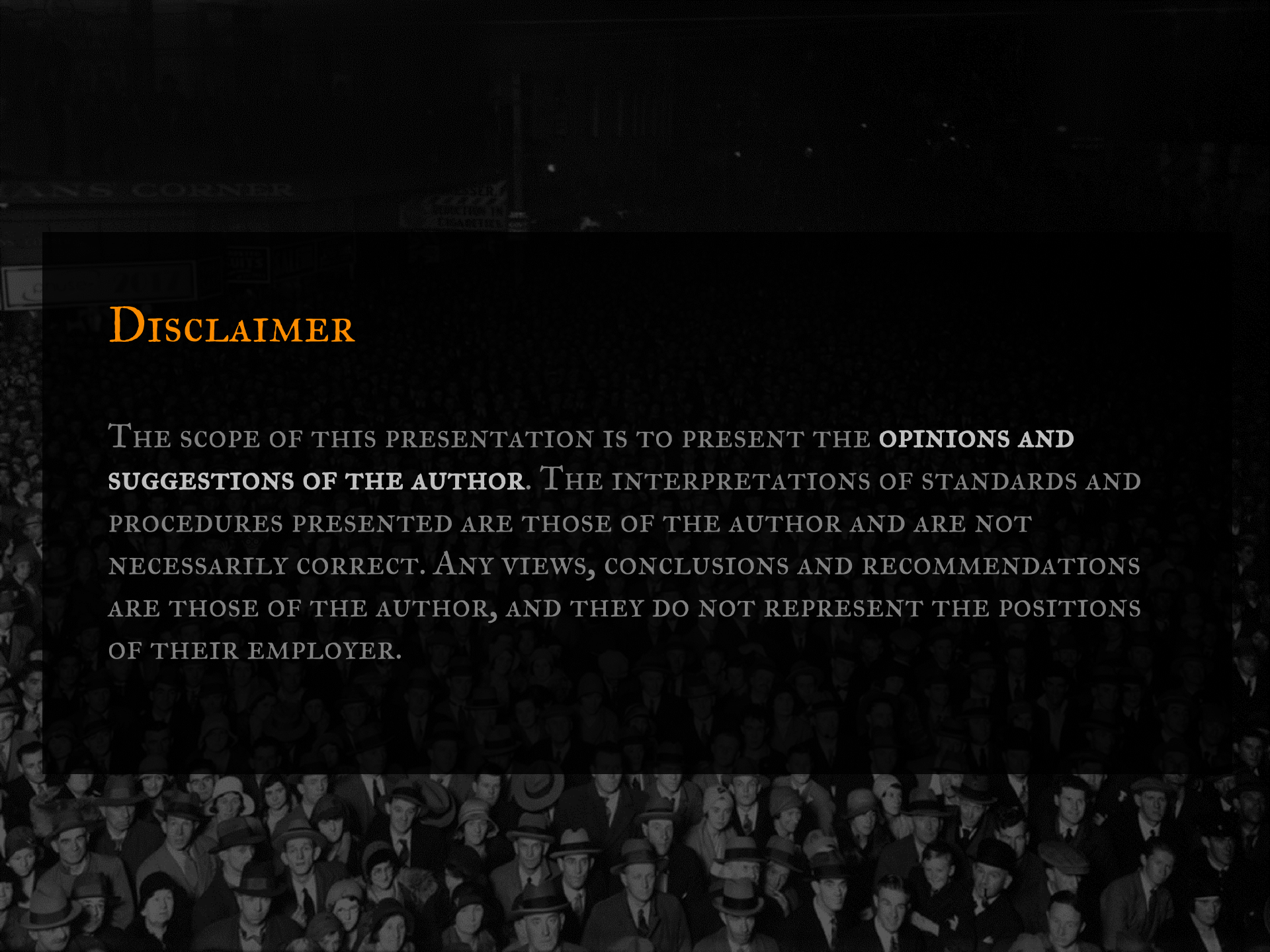




DH09

PLAUSIBLE
ADVERSARIES
IN
RE-IDENTIFICATION
RISK ASSESSMENT

LUKASZ KNIOLA
BIOGEN, MAIDENHEAD, UK



DISCLAIMER

THE SCOPE OF THIS PRESENTATION IS TO PRESENT THE **OPINIONS AND SUGGESTIONS OF THE AUTHOR**. THE INTERPRETATIONS OF STANDARDS AND PROCEDURES PRESENTED ARE THOSE OF THE AUTHOR AND ARE NOT NECESSARILY CORRECT. ANY VIEWS, CONCLUSIONS AND RECOMMENDATIONS ARE THOSE OF THE AUTHOR, AND THEY DO NOT REPRESENT THE POSITIONS OF THEIR EMPLOYER.

AGENDA



. CONTEXT

. METHODOLOGY

. SURVEY

. TRENDS AND OUTLOOKS

ADVERSARIES

POSSIBLE

PLAUSIBLE

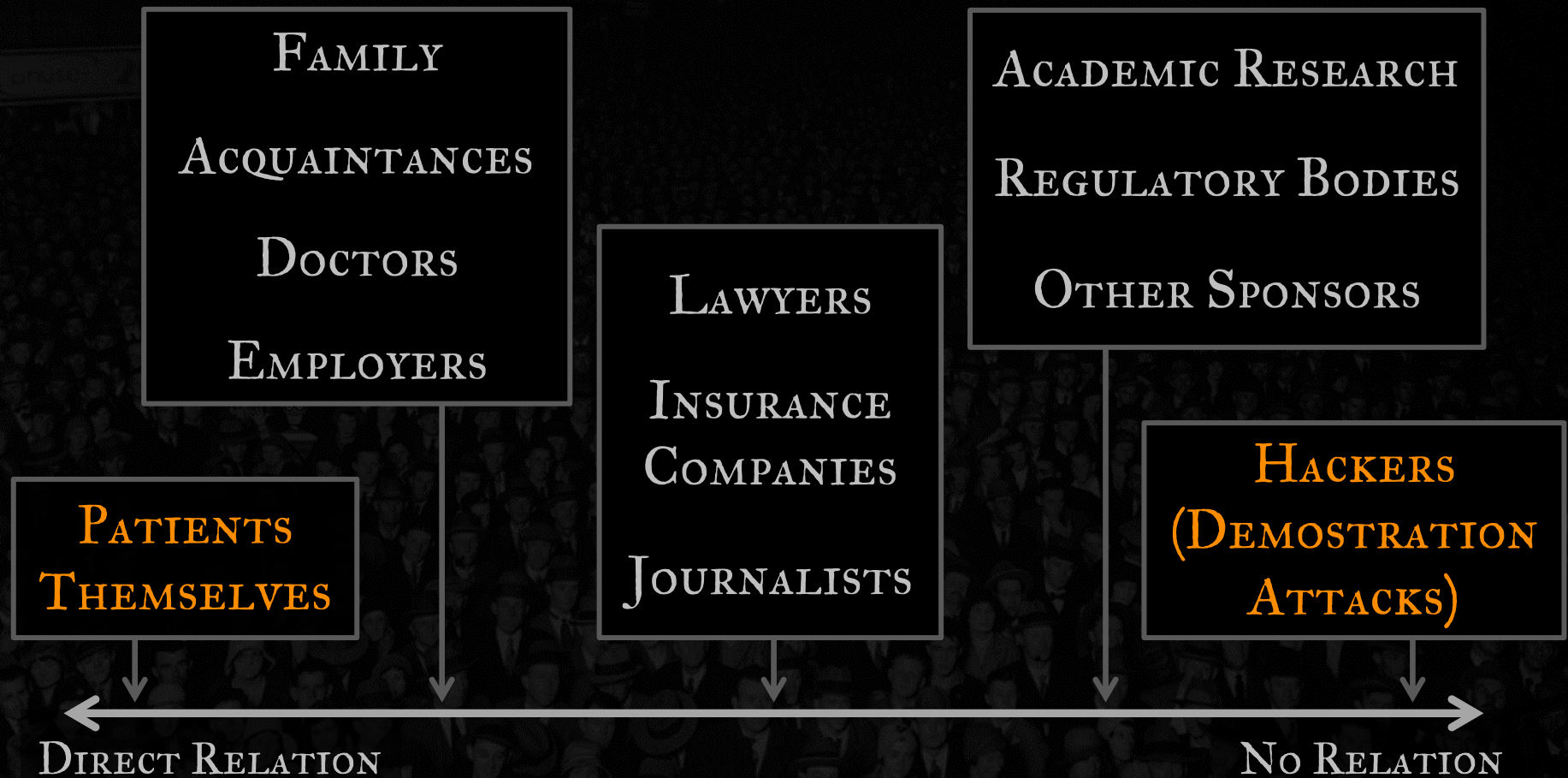
CONTEXT OF DATA RELEASE

INTERNAL
SECONDARY
RESEARCH
(DATA RE-USE)

EXTERNAL
SECONDARY
RESEARCH

PUBLIC
RELEASE

CLASSIFICATION (I/2)



CLASSIFICATION (2/2)

ADVERSARY'S BACKGROUND
KNOWLEDGE

LIKELY / UNLIKELY



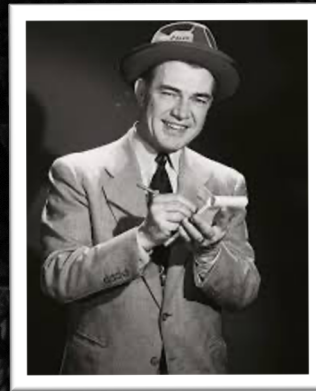
TARGET OF RE-ID:
SPECIFIC / RANDOM

ADVERSARY KNOWS
TARGET IS IN DATA SET?
DEFINITELY / DOUBTFULLY

PROSECUTOR



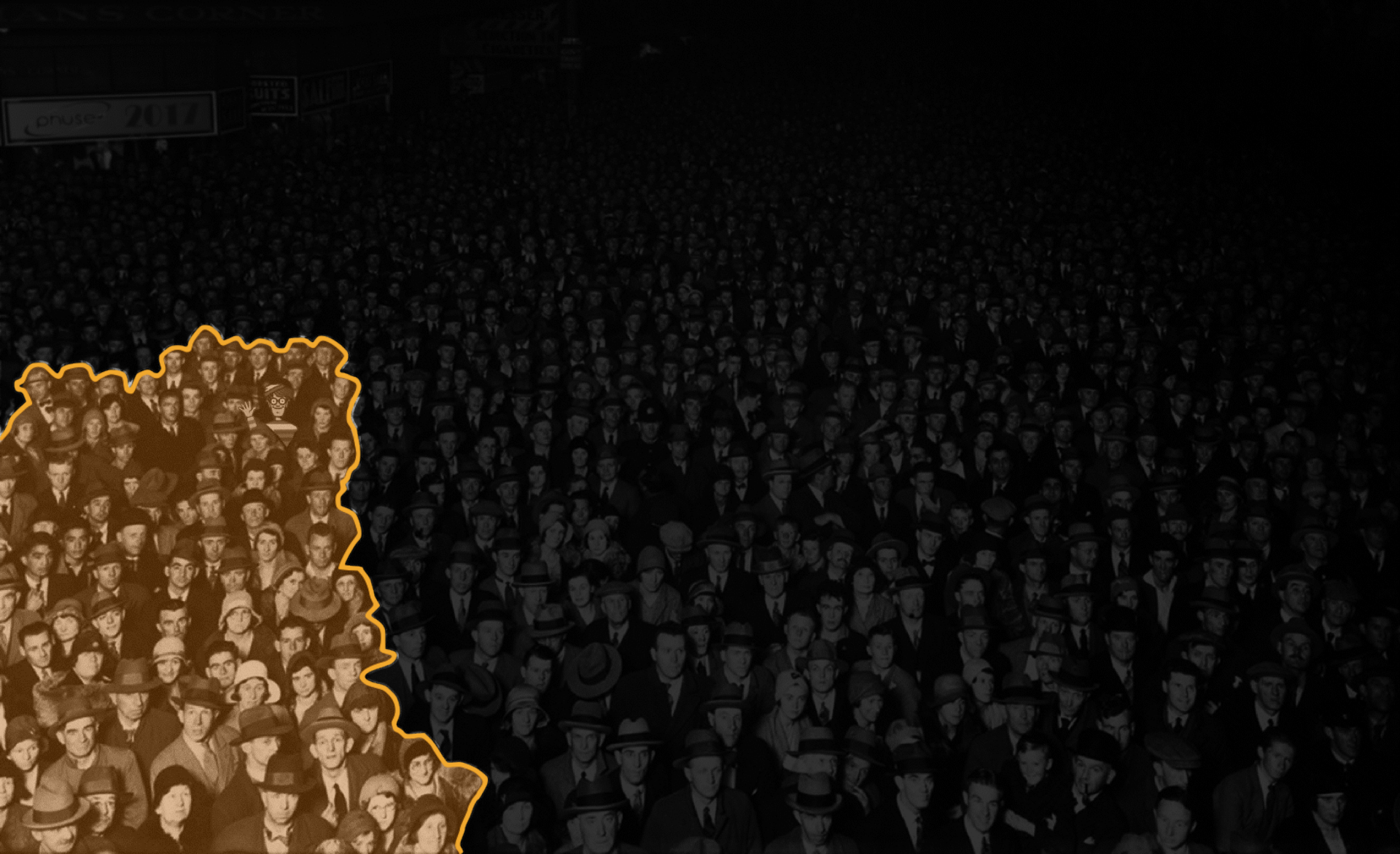
Journalist



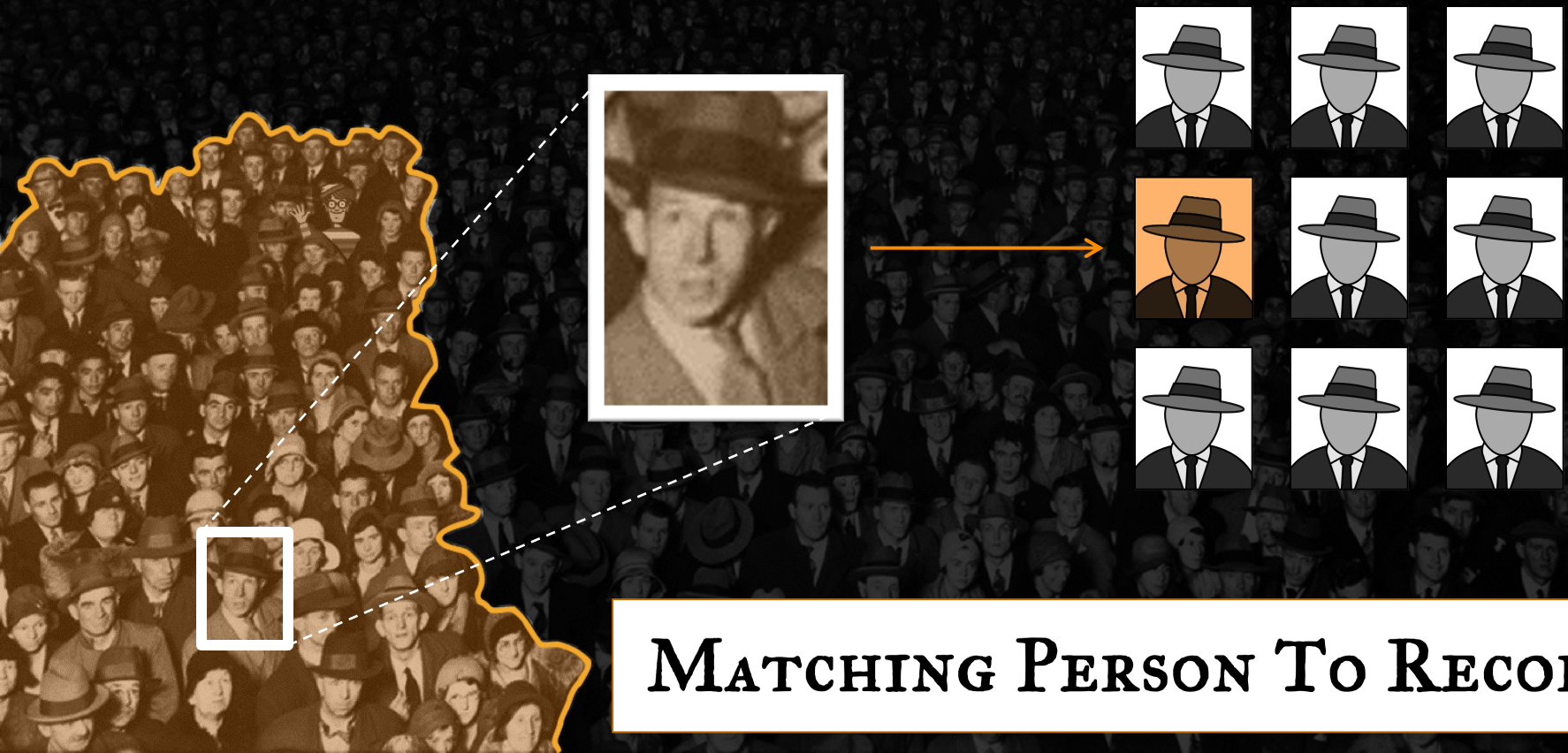
Marketer



PROSECUTOR



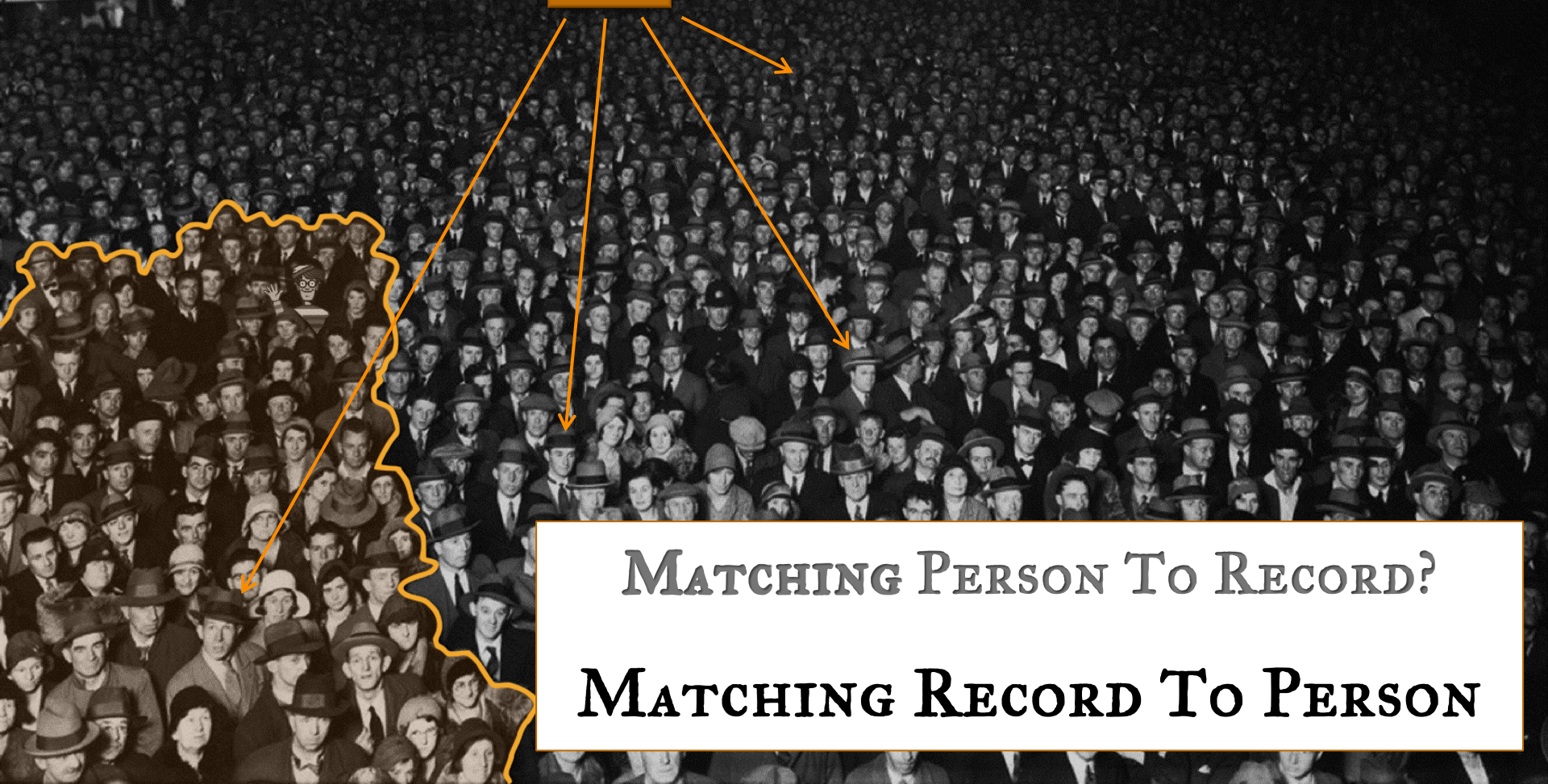
PROSECUTOR



Journalist



Journalist

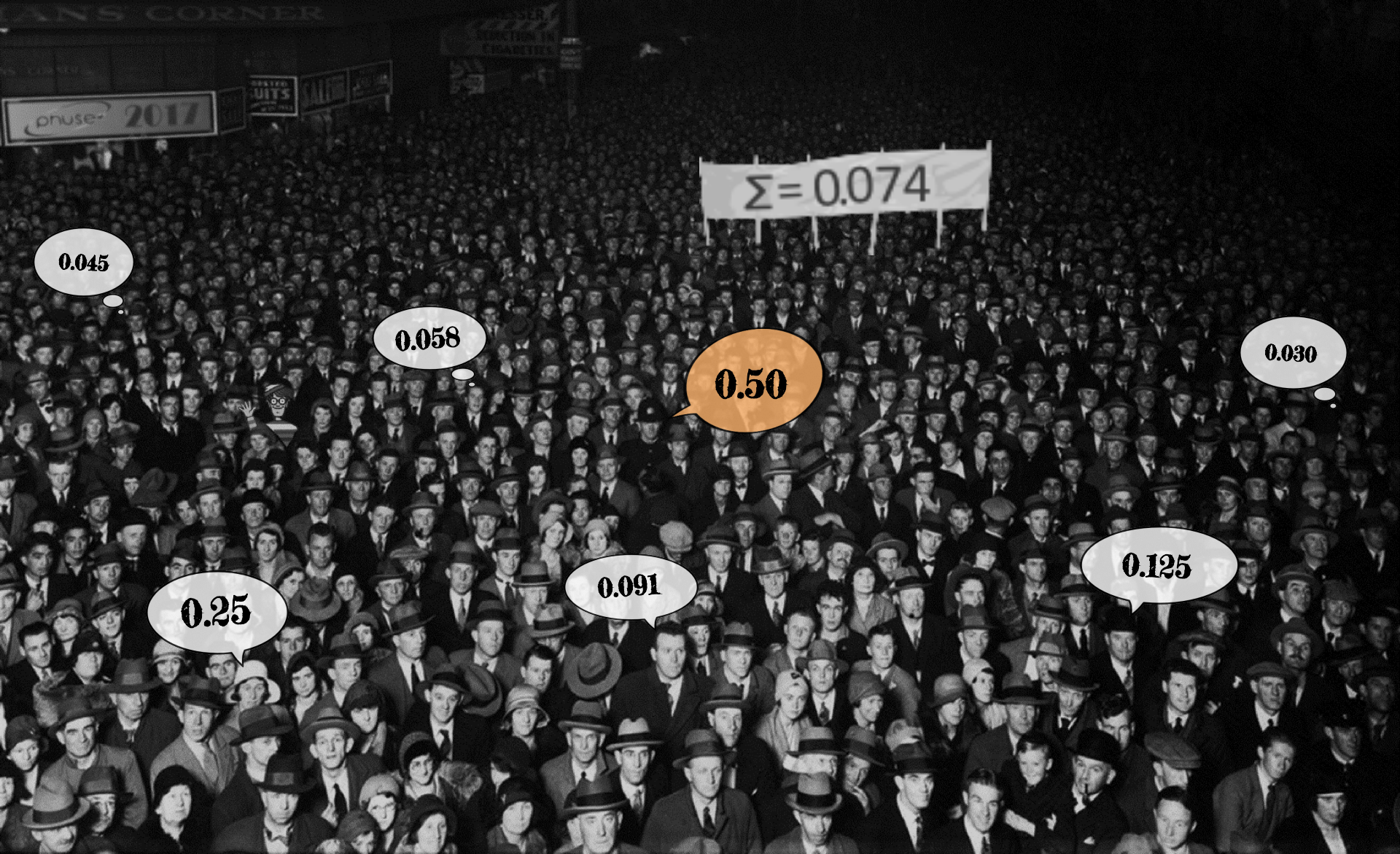


MATCHING PERSON TO RECORD?
MATCHING RECORD TO PERSON

PROSECUTOR $\geq$ Journalist $\geq$ Marketer



MAXIMUM AND AVERAGE RISK



0.045

0.058

0.50

0.030

0.25

0.091

0.125

$\Sigma = 0.074$

UNIQUENESS IN THE DATA SET



HARM $\neq$ RISK



IDENTITY
DISCLOSURE
VS
ATTRIBUTE
DISCLOSURE

INTENT TO HARM
&
POTENTIAL HARM

SURVEY

23 RESPONDERS

6 QUESTIONS

12 ADVERSARY CATEGORIES

TESTING PERCEIVED:

MOTIVATION

MEANS/TOOLS

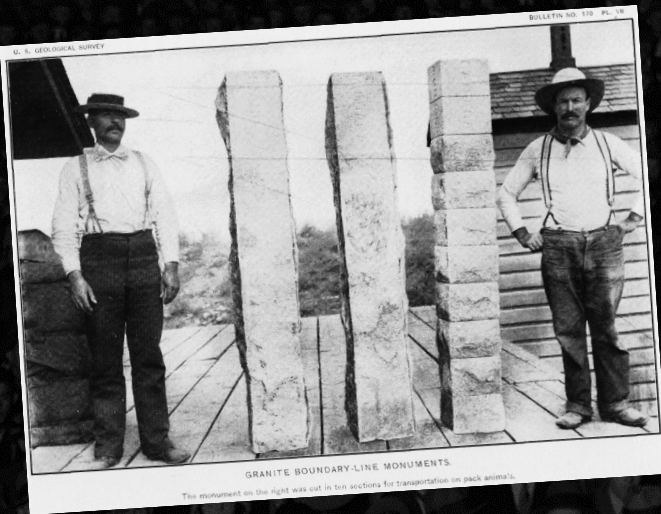
BACKGROUND KNOWLEDGE

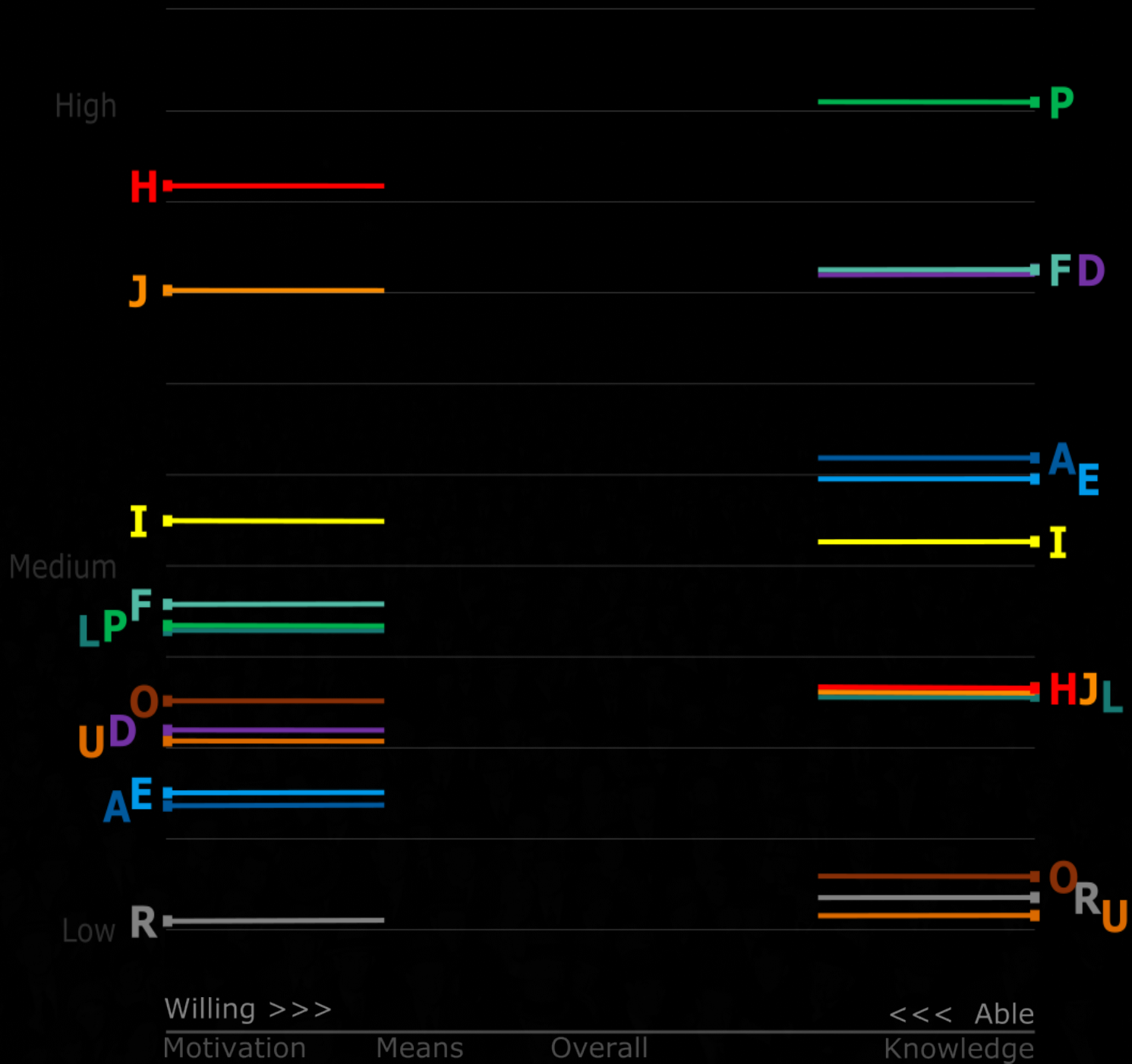
INTENT TO HARM

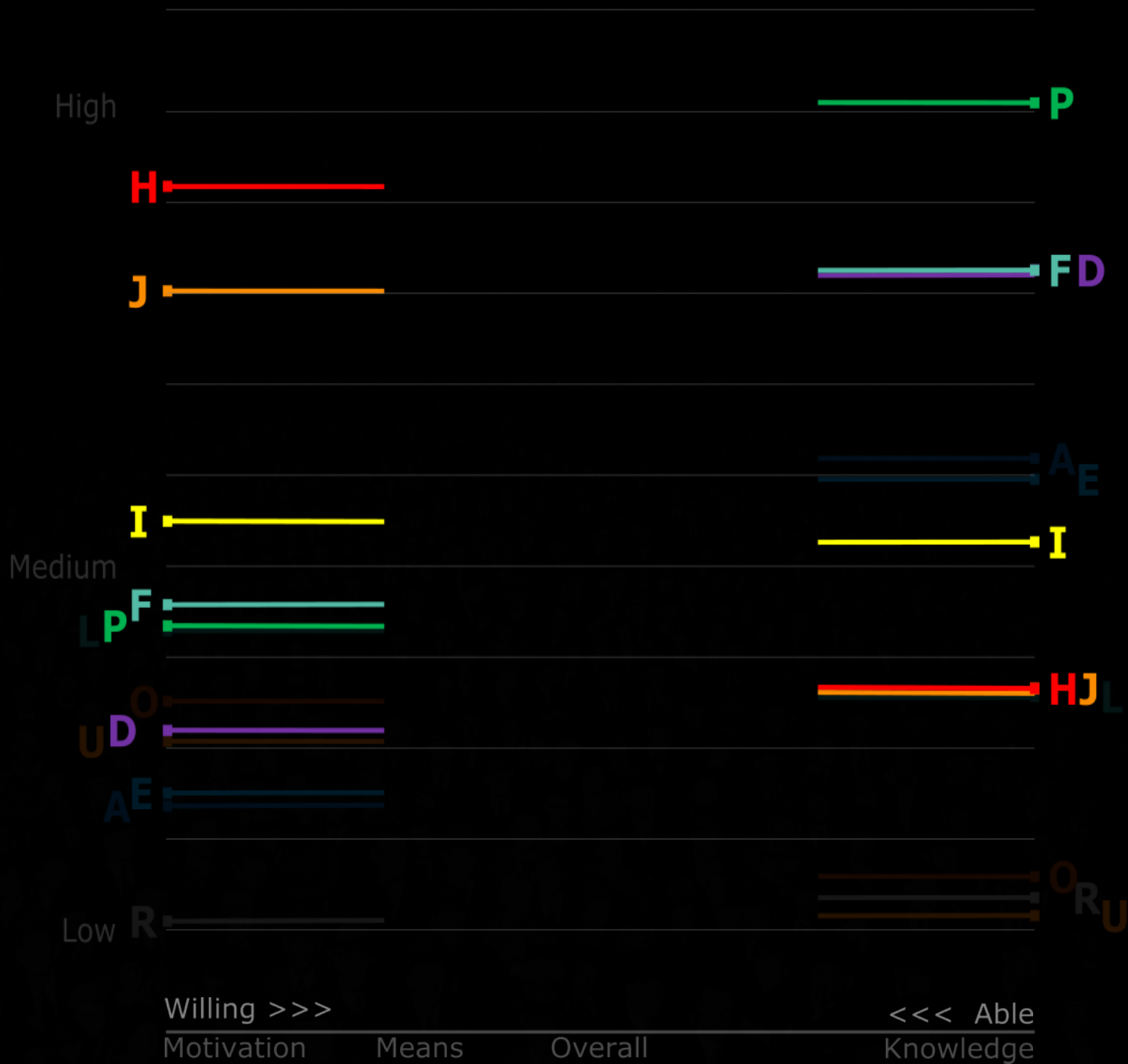
LIKELYHOOD OF KNOWING:

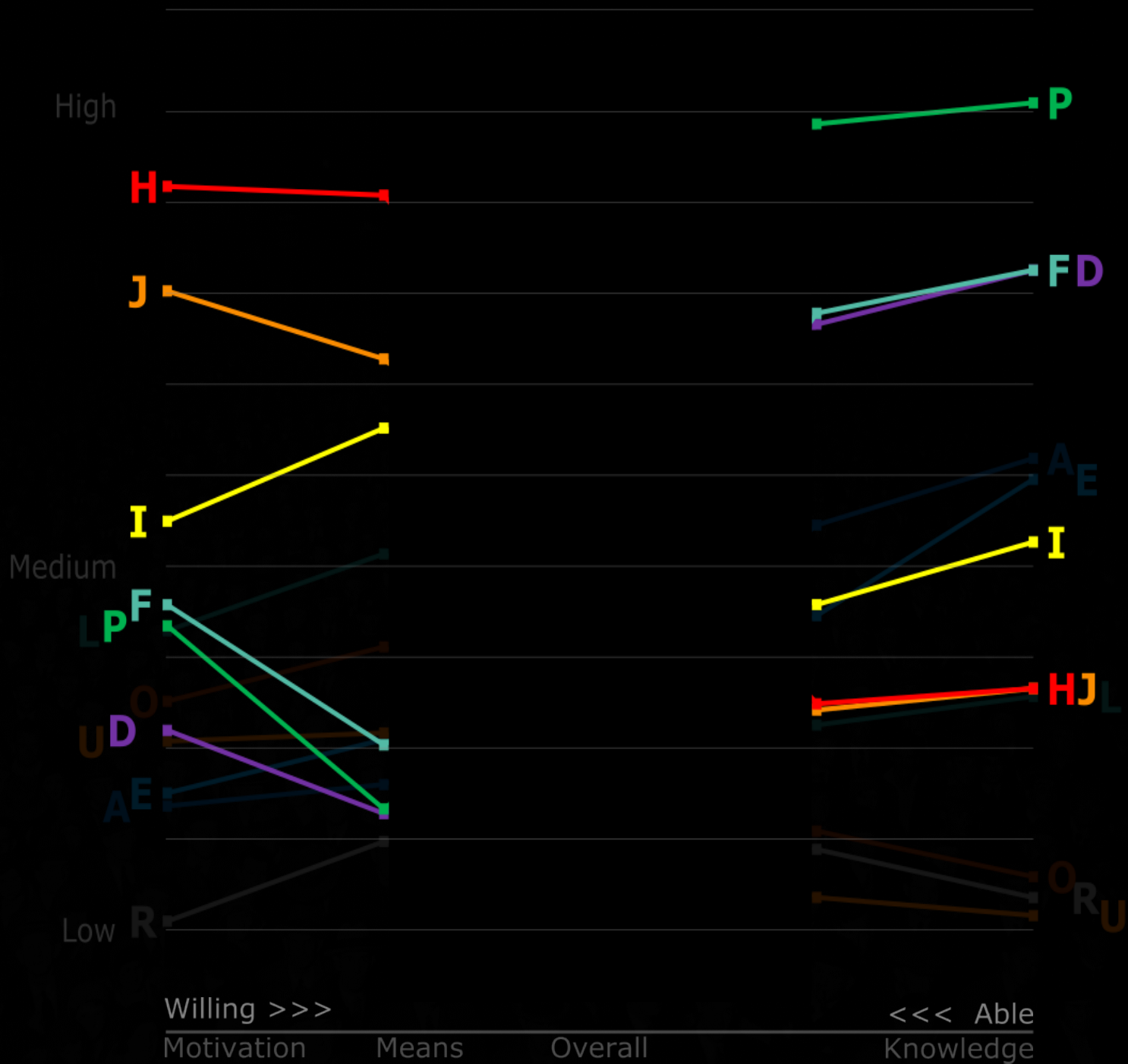
TARGET IN THE DATA SET

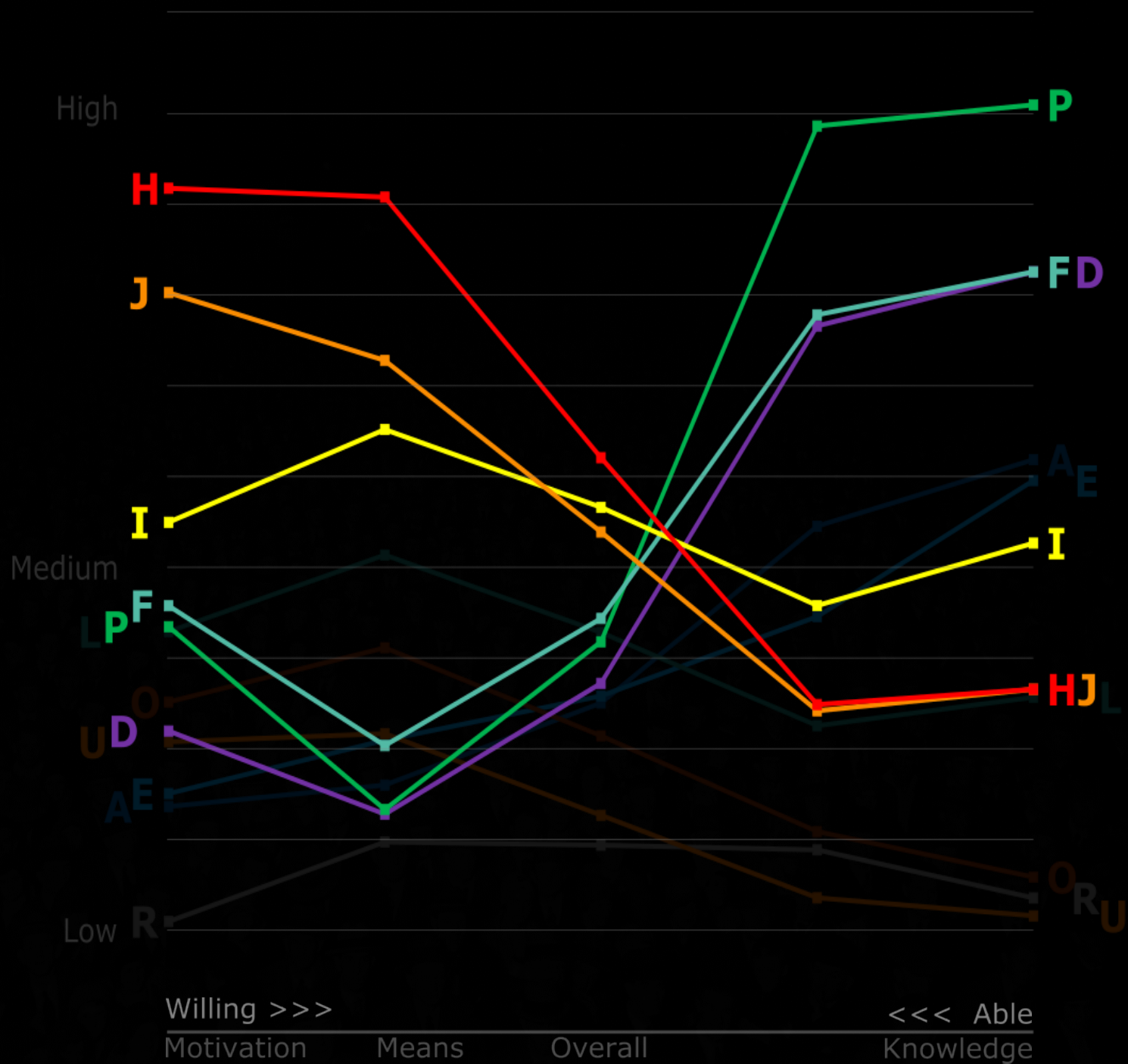
FULL LIST OF TRIAL PARTICIPANTS

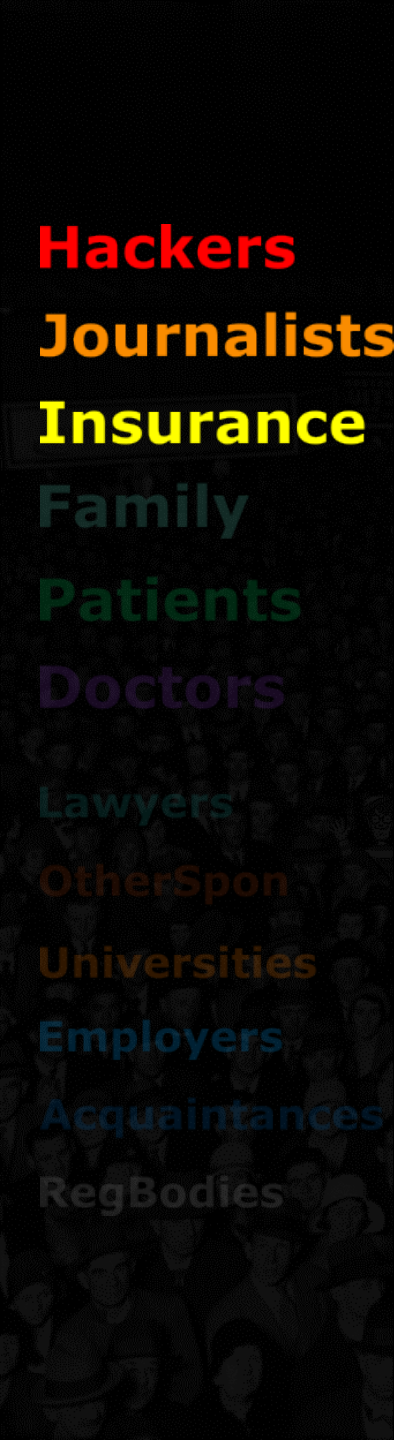




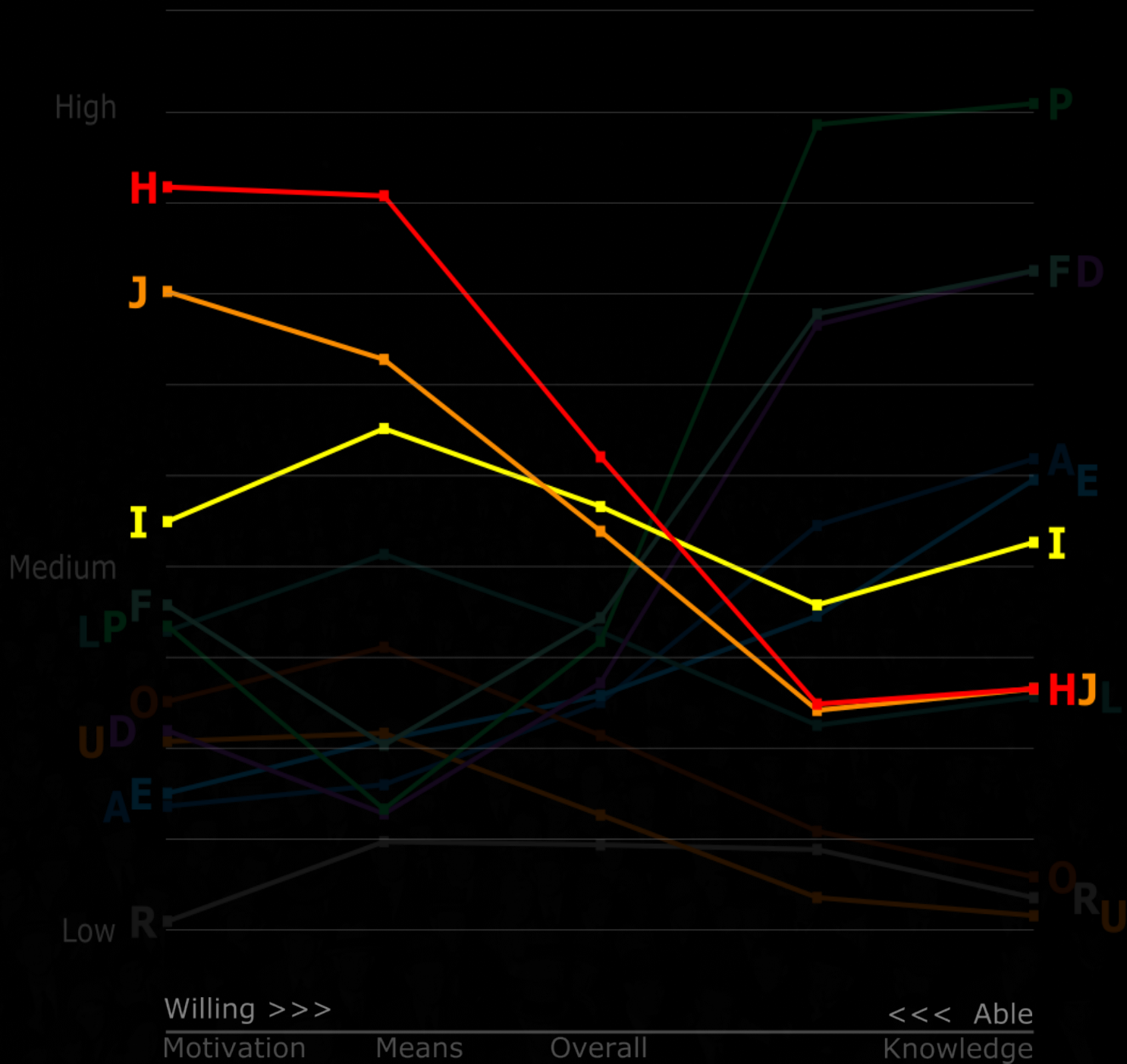


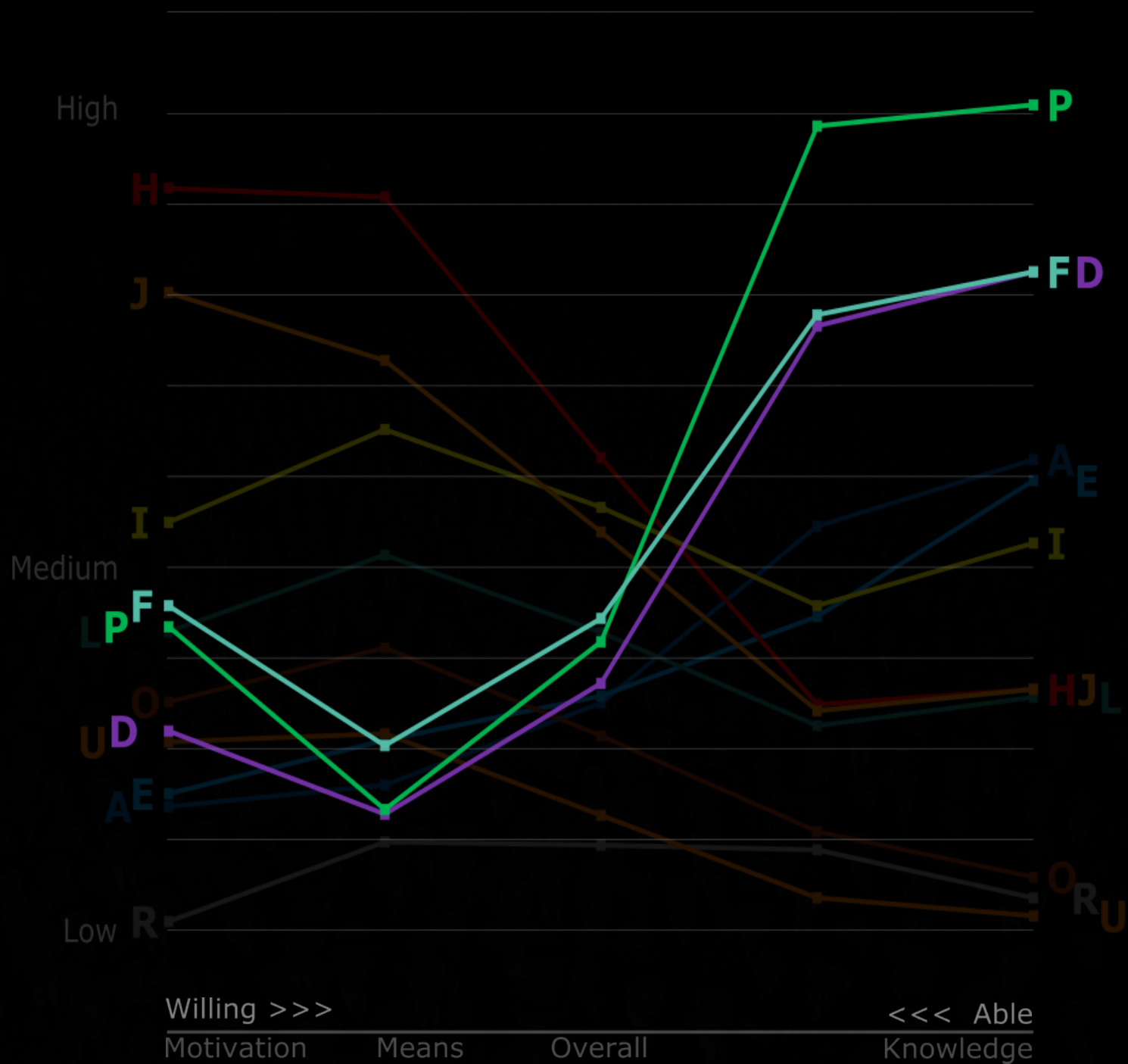






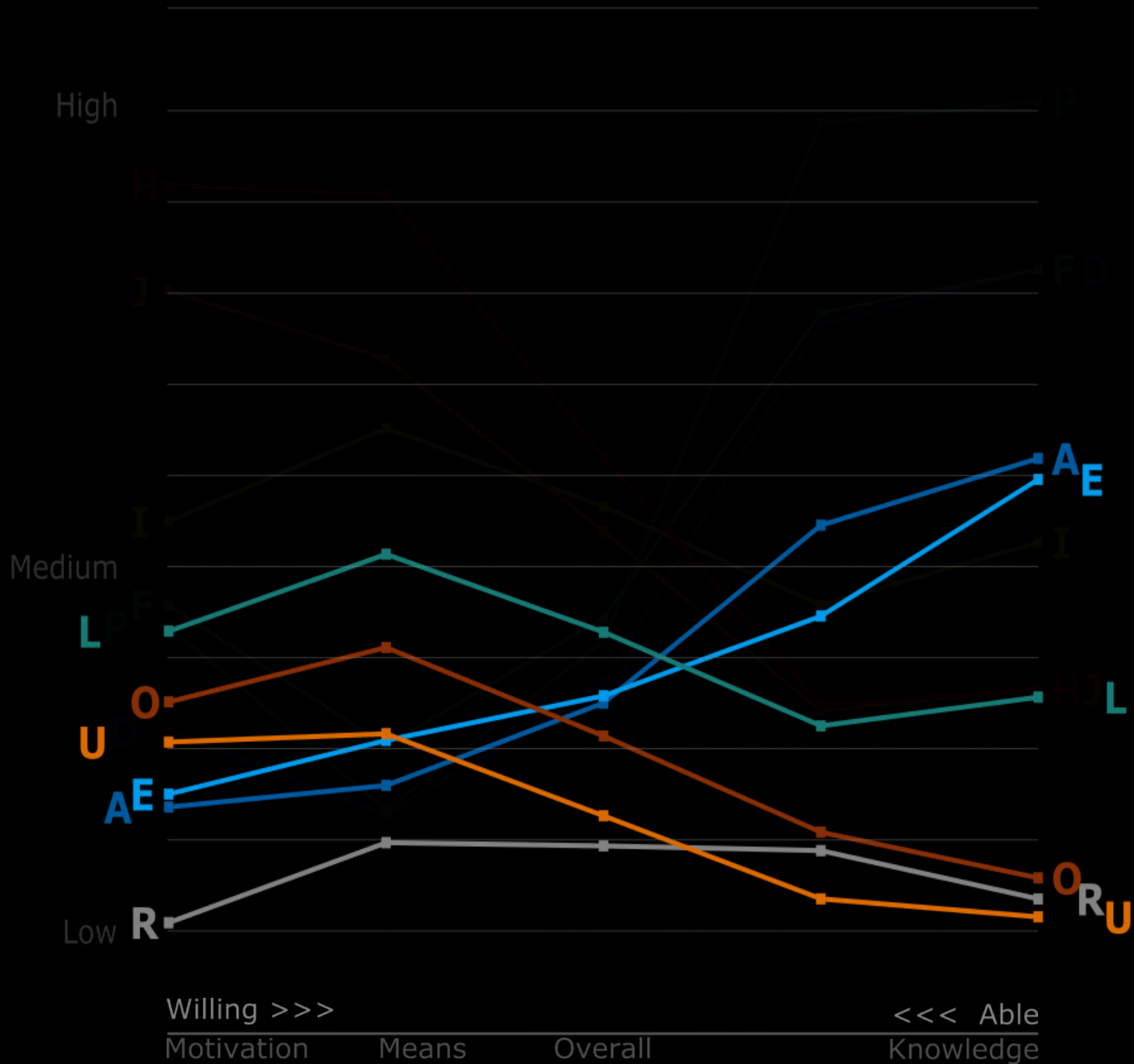
Hackers
Journalists
Insurance
Family
Patients
Doctors
Lawyers
OtherSpon
Universities
Employers
Acquaintances
RegBodies



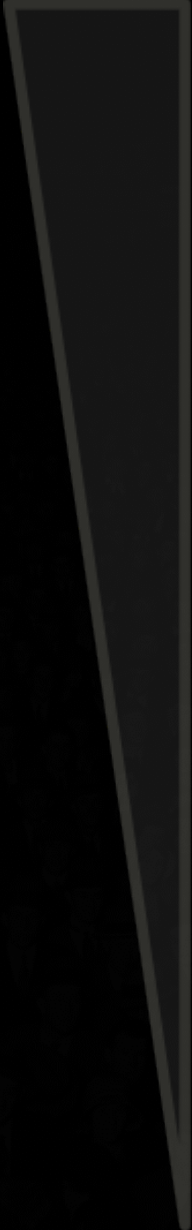




- Hackers
- Journalists
- Insurance
- Family
- Patients
- Doctors
- Lawyers
- OtherSpon
- Universities
- Employers
- Acquaintances
- RegBodies



SURVEY



Hackers
Insurance
Journalists

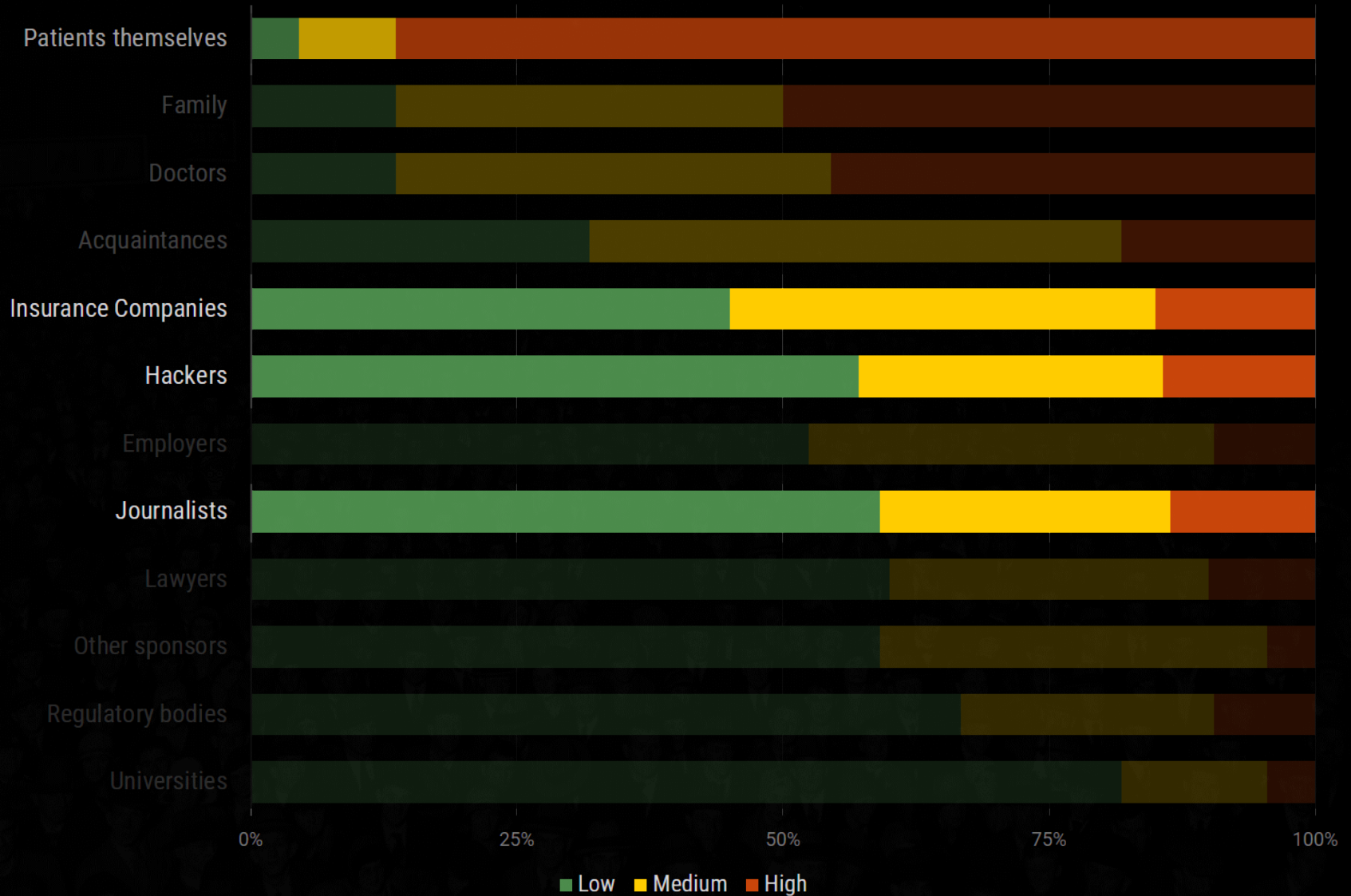
Family
Lawyers
Patients

Doctors
Employers
Acquaintances

OtherSpon
Universities
RegBodies

SURVEY

Likelihood that adversary knows target is in data set



SURVEY

Likelihood that adversary has full list of participants



SURVEY - CONCLUSIONS



Hackers
Insurance
Journalists

JOURNALIST RISK?

Family
Lawyers
Patients

PROSECUTOR RISK

Doctors
Employers
Acquaintances

JOURNALIST RISK

OtherSpon
Universities
RegBodies

JOURNALIST RISK

EMA POLICY 0070

GUIDANCE

DEMONSTRATION ATTACK

PARTICIPANTS OF PUBLIC INTEREST

EMBARRASS DATA CONTROLLER OR PROCESS

ACQUAINTANCE RECOGNITION

VULNERABLE PATIENTS & SENSITIVE INFO

MAXIMUM RISK
ACROSS ALL RECORDS

EMA POLICY 0070

TRENDS

PROSECUTOR



EASIER TO DEFEND

EASIER TO APPLY

MORE CONSERVATIVE

LOWER DATA UTILITY

EMA POLICY 0070 OUTLOOK

JOURNALIST



GREATER DETAIL RETENTION

HIGHER DATA UTILITY

REQUIRES:

BETTER GUIDELINES

DEFINE „SIMILAR STUDIES”

MORE EVIDENCE



THANK YOU

QUESTIONS?

LUKASZ.KNIOLA@BIOGEN.COM