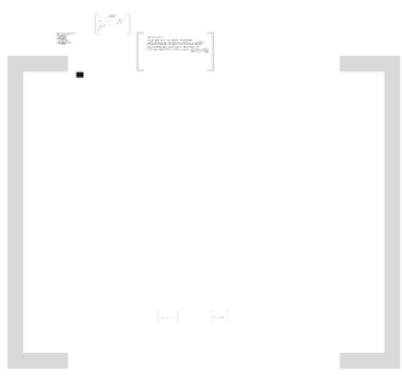




Automating Surveillance  
Dan Marshall - University of Bath  
dan.marshall@bath.ac.uk



Opening title - background for opening gambit -  
links to:  
Unique Selling proposition - background for Point B -  
links to:  
Overview - Preview of entire presentation - background  
for presentation and forecast of time  
Drill down  
Summary slide - summarize and restate Point B  
Name slide

Opening title: Madrid Train Bombing  
Point A: How to catch bombers ethically  
USP: Ethical dangers in automation.  
Overview: Manual - Auto - Partial  
Drill down:  
Summary: Manual has problems, auto has  
problems, partial is soln but has problems  
(albeit less so than man or auto)  
Name slide

10 coordinated bombings on Cercanías (commuter train), Madrid,  
11 March 2004 (three days before general elections),  
killed 191  
wounded 1,800

- Point A: How to catch bombers without sacrificing ethics
- Point B: Partial automation is the solution
- WIIFY: Resolution to problem
- Audience: Computer Scientists/Philosophers
- Structure: Problem/Solution
- Opening Gambit: Anecdote - Madrid Bombings
- USP: Ethical concerns regarding automation of surveillance
- Link forward:
- Forecast time: Next 10 minutes looking at these in more detail
- Call to action: Partial is best in terms of efficacy; also best in terms of ethics (although there is a cost to pay)





# Surveillance

## Manual

manually operate their algorithms and decide if true



## Full Automation

full algorithms interpret files and decide if true



## Partial Automation

partial algorithms interpret files and operators decide if true



# Manual

manual (operator filters information and decides alone)

# Manual

manual (operator filters information and decides alone)

## Processing Capacity



Information overload  
Inattention blindness  
Boredom  
Filter by stereotyping



## Prejudice



Differential prejudice  
- based on social "stereotype"  
- race is making in error  
- automatic prejudice



## False Positives/ False Negatives

False positive:  
Innocent incorrectly identified as threat

False negative:  
Guilty incorrectly identified as non-threat

# Processing Capacity



Information overload

Inattention blindness

Boredom

Filter by stereotyping





Information overload

Inattentional blindness

Boredom

Filter by stereotyping

# Prejudice

Norris and Armstrong (1999)

- 39% of targeted were teenagers

65% "homophobic" were teenagers

## Norris and Armstrong (1999)

- 39% of targeted were teenagers
- 65% "no obvious reason" were teenagers
- 15% population teenagers
- 1% prison population teenagers

Similar findings for black and male

Two problems:

Stigmatization of group

Harassment of innocent

# Behavioural profiling?

- Hard to define "suspicious"
- Poor training to date
- Mask for group profiling

# Prejudice

Norris and Armstrong (1999)

- 39% of targeted were teenagers
- 65% "no obvious reason" were teenagers
- 15% population teenagers
- 1% prison population teenagers

Similar findings for black and male

Two problems:  
Stigmatization of group  
Harassment of innocent

Behavioural profiling?

- Hard to define "suspicious"
- Poor training to date
- Mask for group profiling

# False Positives/ False Negatives

False positive:

Innocent incorrectly identified as threat

# False Positives/ False Negatives

False positive:

Innocent incorrectly identified as threat

False negative:

Guilty incorrectly identified as non-threat



# Manual

manual (operator filters information and decides alone)

## Processing Capacity



Information overload  
Inattention blindness  
Boredom  
Filter by stereotyping



## Prejudice



Differential prejudice:  
- based on skin "color"  
- race training in the  
- automatic prejudice



## False Positives/ False Negatives

False positive:  
Innocent incorrectly identified as threat

False negative:  
Guilty incorrectly identified as non-threat

# Full Automation

full automation (computer filters and decides alone)

# Full Automation

full automation (computer filters and decides alone)

## Processing Capacity



"Highway police issuing speeding tickets, being harassed, are unable to be completely objective and impartial. Their decisions may be affected by the race, sex, class, age, appearance, and manner of the people they pull over. Machines that detect speed, identify license plates, and issue tickets accordingly will be unaffected by such things." Weisauer (2012)

No information overload = no error?

What computers do, they do well.  
They just can't do very much.

Is this always the case?

## Prejudice



## False Positives/ False Negatives

False positives:

- Computer cannot filter as much as human brain

False negatives

- Difficult to identify every possible eventuality

Problem: Associated People

Solution: Social Force Model

- attractive force provided by the goal of the individual
- attractive force keeping those in the same group together
- repulsive force of the walls
- repulsive force of individuals not in the same group



**SUBITO**

© 2012 University of Michigan  
All rights reserved. This work is licensed under a Creative Commons  
Attribution-NonCommercial-ShareAlike 4.0 International License.

# Processing Capacity

No information overload = no error?

What computers do, they do well.  
They just can't do very much.

# Prejudice

“Highway police issuing speeding tickets, being human, are unlikely to be completely consistent and impartial. Their decisions may be affected by the race, sex, class, age, appearance, and manner of the people they pull over. Machines that clock speeds, identify license plates, and issue tickets accordingly will be unaffected by such things.”

Westacott (2003)

Is this always the case?

# SUBBITO

Surveillance of Unattended Baggage and the Identification and Tracking of the Owner



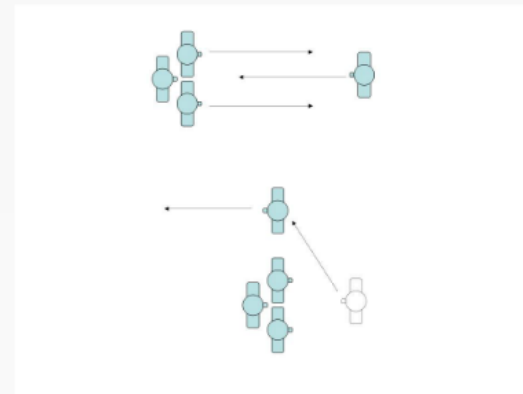
# Problem: Associated People

## Solution: Social Force Model

- + attractive force provided by the goal of the individual,
- + attractive force keeping those in the same group together
- repulsive force of the walls,
- repulsive force of individuals not in the same group

# SUBITO

Surveillance of Unattended Baggage and the Identification and Tracking of the Owner



Problem: not everyone walks the same distance apart  
Determined by culture, age and sex  
Could lead to culture/age/sex bias in software

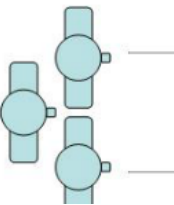
Possibility of prejudice entering  
system through the software

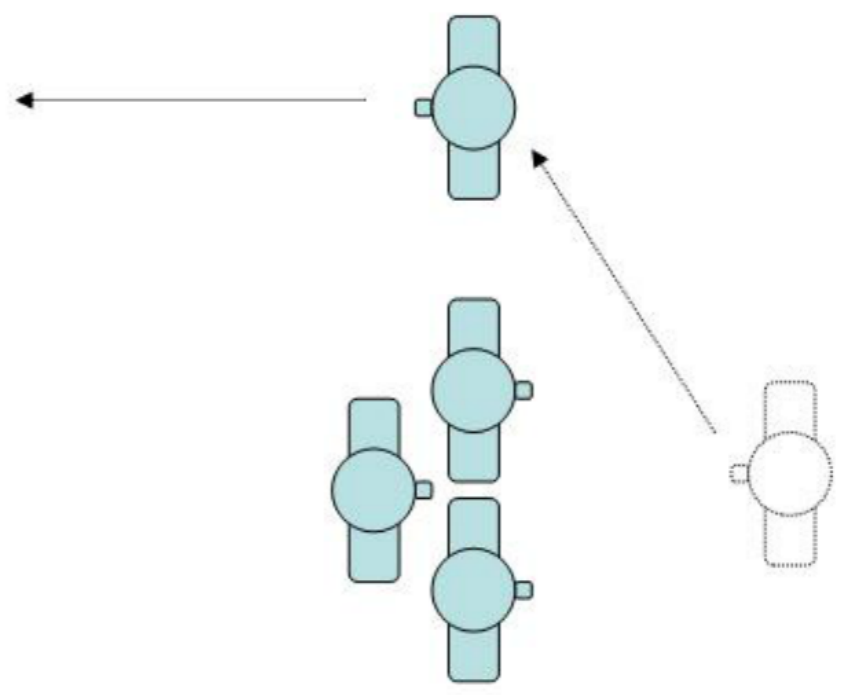
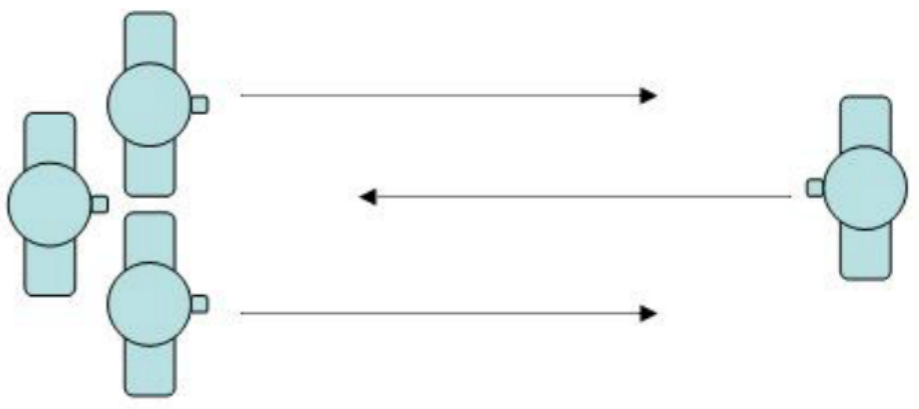
Worse, could be  
institutionalised

# Problem: Associated Peo

## Solution: Social Force Model

- + attractive force provided by the goal of the individual,
- + attractive force keeping those in the same group together
- repulsive force of the walls,
- repulsive force of individuals not in the same group





Problem: not everyone walks the same distance apart  
Determined by culture, age and sex  
Could lead to culture/age/sex bias in software

Possibility of prejudice entering  
system through the software

Worse, could be  
institutionalised

But:

Unintentional prejudice may be less insulting

Possibly more easily rectified throughout system

# Prejudice

"Highway police issuing speeding tickets, being human, are unlikely to be completely consistent and impartial. Their decisions may be affected by the race, sex, class, age, appearance, and manner of the people they pull over. Machines that clock speeds, identify license plates, and issue tickets accordingly will be unaffected by such things."

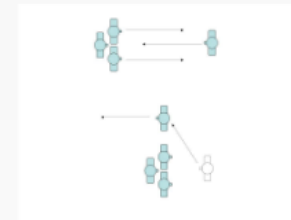
Westacott (2003)

Is this always the case?

Problem: Associated People

Solution: Social Force Model

- + attractive force provided by the goal of the individual,
- + attractive force keeping those in the same group together
- repulsive force of the walls,
- repulsive force of individuals not in the same group



## SUBITO

Semantics of Unintended Biases and the Identification and Testing of its Origin

Problem can be solved by using the same force as the one that caused it, but in a different direction.

Quality of people is not the same as the quality of the system.

It is not the quality of the system.

*But:  
Unintentional prejudice may be less insidious  
Possibly more easily rectified through our system*



# False Positives/ False Negatives

False positives:

- Computer cannot filter as much as human brain

False negatives

- Difficult to identify every possible eventuality

# Full Automation

full automation (computer filters and decides alone)

## Processing Capacity



"Highway police issuing speeding tickets, being human, are subject to the complexity constraints and heuristics. Their decisions may be affected by the race, sex, class, age, appearance, and names of the people they pull over. Machines that detect, identify license plates, and issue tickets accordingly will be unaffected by such things." Weickert (2012)

No information overload = no error?

What computers do, they do well.  
They just can't do very much.

Is this always the case?

## Prejudice



## False Positives/ False Negatives

False positives:

- Computer cannot filter as much as human brain

False negatives:

- Difficult to identify every possible eventuality

Problem: Associated People

Solution: Social Force Model

- attractive force provided by the goal of the individual
- attractive force keeping those in the same group together
- repulsive force of the walls
- repulsive force of individuals not in the same group



SUBITO

The "Subito" logo is a stylized representation of the word "Subito" in a bold, sans-serif font. It is positioned at the bottom right of the slide, below the Social Force Model diagram.

# Partial Automation

partial automation (computer filters and operator decides)

# Partial Automation

partial automation (computer filters and operator decides)

## Processing Capacity



## Prejudice



## False Positives/ False Negatives

Best of Both?  
Computer excels at consistent recognition and simple filtering  
Computer is worst at advanced filtering  
Human is worst at consistent recognition and simple filtering  
Human excels at advanced filtering

May be recognized by operator and flagged up  
or  
May play to operator's own prejudices so no action  
or  
May choose to overrule correct interpretation by software

Reduces false positives (common sense in addition to software)  
Reduces false negatives (computer sees what he might miss)

# Processing Capacity

Best of Both?

# Processing Capacity

Best of Both?

Computer excels at consistent recognition and simple filtering  
Computer is worst at advanced filtering

Human is worst at consistent recognition and simple filtering  
Human excels at advanced filtering

# Prejudice

May be recognized by operator and flagged up  
or

May play to operator's own prejudices so no action  
or

May choose to overrule correct interpretation by software

# False Positives/ False Negatives

Reduces false positives (common sense in addition to software)

Reduces false negatives (computer sees what he might miss)



# Partial Automation

partial automation (computer filters and operator decides)

## Processing Capacity



## Prejudice



## False Positives/ False Negatives

Best of Both?  
Computer excels at consistent recognition and simple filtering  
Computer is worst at advanced filtering  
Human is worst at consistent recognition and simple filtering  
Human excels at advanced filtering

May be recognized by operator and flagged up  
or  
May play to operator's own prejudices so no action  
or  
May choose to overrule correct interpretation by software

Reduces false positives (common sense in addition to software)  
Reduces false negatives (computer sees what he might miss)

# Surveillance

## Manual

manually operate their algorithms and decide if true



## Full Automation

full algorithms interpret data and decide if true



## Partial Automation

partial algorithms interpret data and operators decide if true



# Remaining Concerns

## Complacency

Operator + Computer  
Too much faith in computer  
Reduced processing  
• not by capacity but by choice  
Problem with aircraft pilots



## Prejudice

Unresolved  
Prejudice more easily remedied in automated  
More easily remedied in Operator as Blinkered  
Re-introduction of operator = re-introduction of human prejudice



## Privacy

Problem in manual  
Largely resolved in automated  
Less scope in Operator as Blinkered



## Function Creep

Problem with most technology  
Underscore future use of technology  
More problematic in automated  
More problematic in Operator as Blinkered

## Filtering

Am I not just arguing for two (vice one) layers of filtering?  
Yes, but:  
Vendors seeking cheapest (single layer) surveillance  
Demonstrates that human element is crucial (at least for now)  
So can't remove human without significant cost  
But can improve on human with some cost

# Complacency

Operator + Computer

Too much faith in computer

Reduced processing

- not by capacity but by choice

Problem with aircraft pilots

# Automating Surveillance

Kevin Macnish - University of Leeds

[kevin.macnish@gmail.com](mailto:kevin.macnish@gmail.com)