



MERSEYSIDE
POLICE

Machine Learning and Topic Modelling

(Understanding Demand and Crime MO)

****S40(2)** & **S40(2)****

Machine Learning and Topic Modelling

- Background & Context
- Understanding Calls Demand - Topic Modelling
- Automating MO

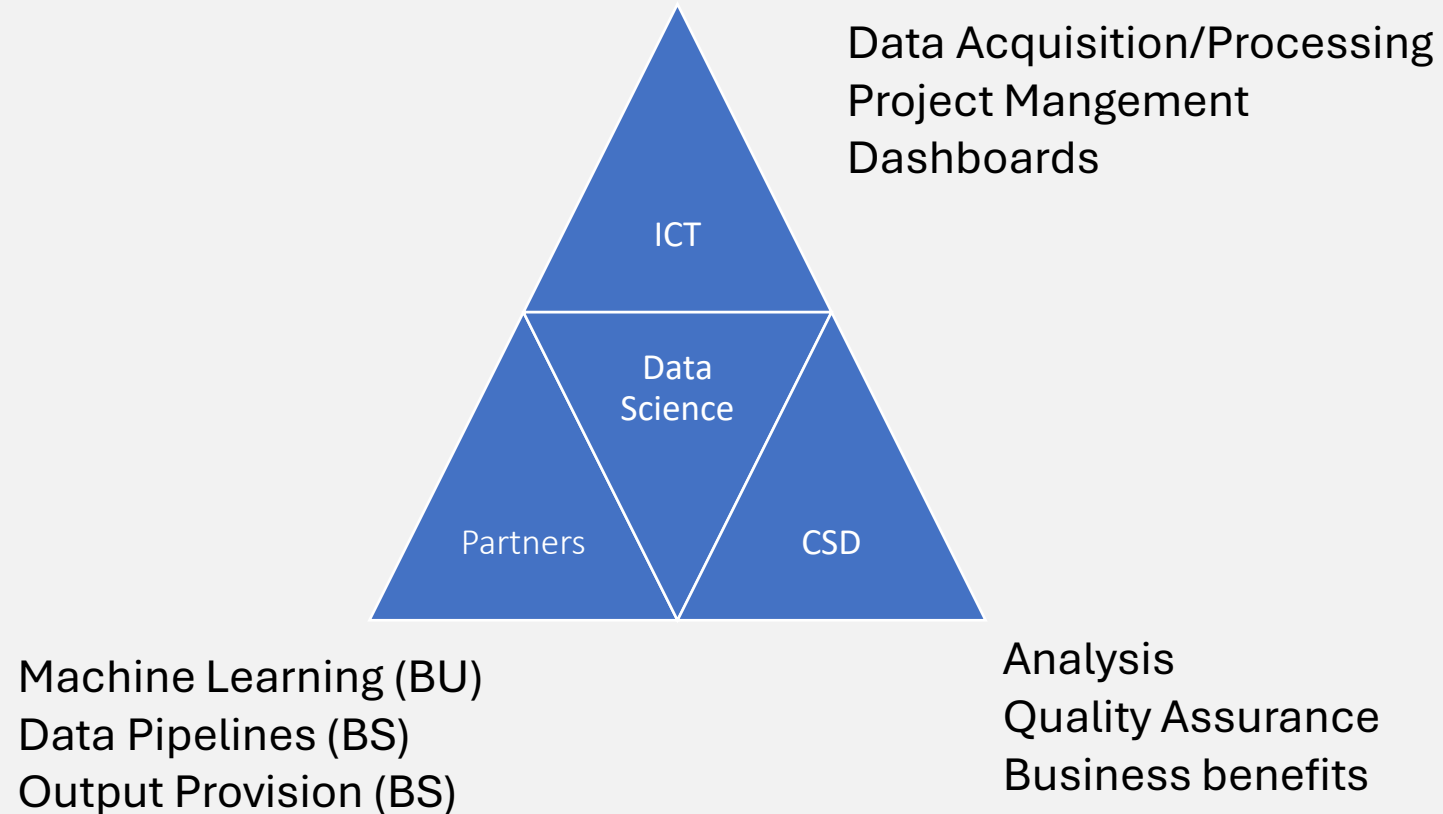
Topic Modelling – Background

- Understanding our demand
- No record of non-incident calls
- Only audio files

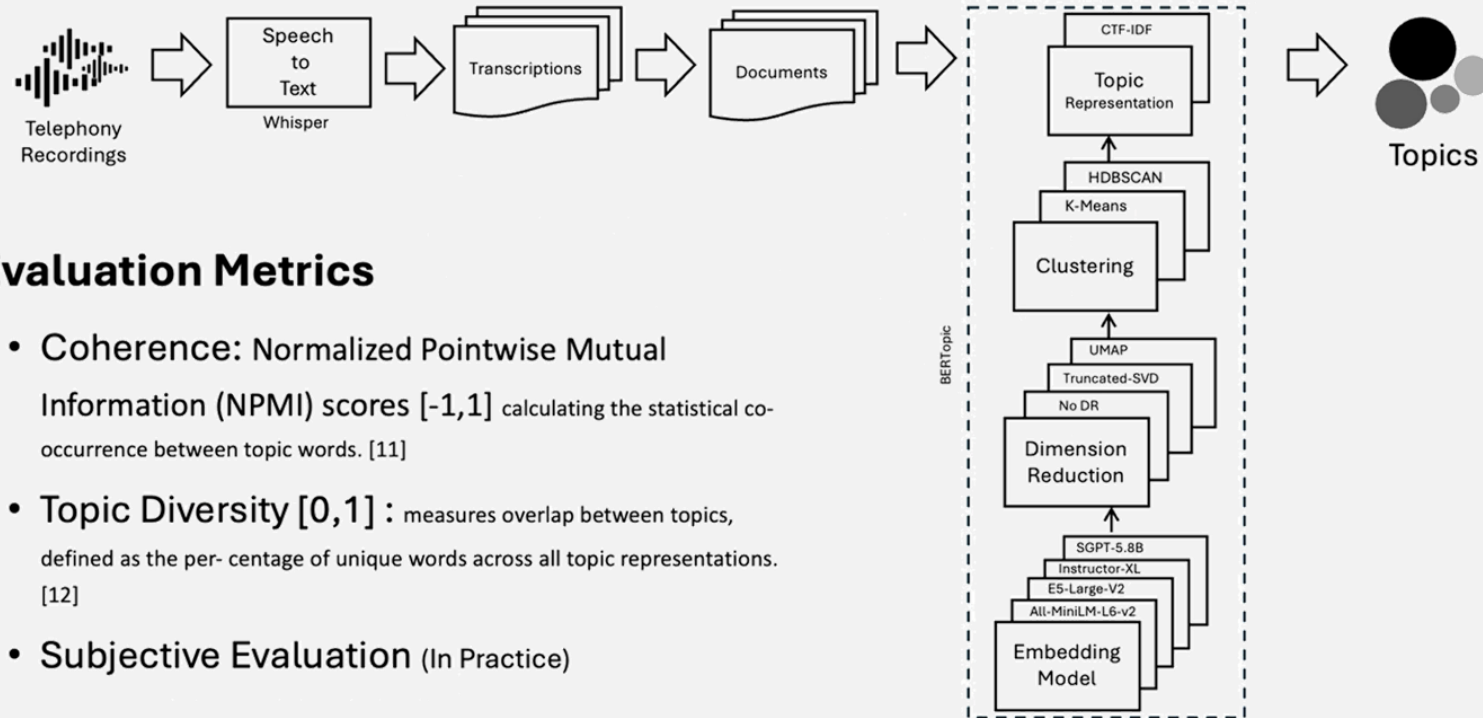
...we don't know what half of our calls are about

Topic Modelling – Approach

- Merseyside Police not capable of full Data Science ...
...but does have skills!
- **Collaboration is key!!**



Topic Modelling – Process



• Evaluation Metrics

- **Coherence:** Normalized Pointwise Mutual Information (NPMI) scores $[-1,1]$ calculating the statistical co-occurrence between topic words. [11]
- **Topic Diversity $[0,1]$:** measures overlap between topics, defined as the percentage of unique words across all topic representations. [12]
- **Subjective Evaluation (In Practice)**

- Data Acquisition
- Transcription Issues
- Unnecessary topics
- Wrong topic labels
- Similar topic labels
- Numerous topics for a call

Topic Modelling – Example

ML-generated Transcription (speech-to-text) 393 words

Call audio file

*Originally
no insight
into reason,
or content, of call.*



Hello there, I'm calling from Whiston Hospital and one of the charge nurses. We've got two patients that are fighting outside. OK, bear with me one minute. Thank you. Whiston Hospital, yeah? Yeah. OK, bear with me. And have you got any weapons? Not that I'm aware of. We don't even know who the patients are. They've just pulled up in cars and two of them are fighting outside and causing a disturbance and they're blocking the entrance. We've rang security, but... Two males fighting, did you say? Yeah. Are we aware of outside the hospital? Literally outside A&E, which is our roundabout to the entrance. Outside A&E? Yeah, outside the A&E department, yeah. I can still hear shouting from there while I'm on the phone too, as well. Any descriptions? There's quite a few people. There's five or six people at the moment. There's mainly two males, but I've just come straight back in to ring you. No worries, it's fine. Bear with me a minute. We've got a sense of an emergency. Thank you. Sorry, security have been notified, but you're unsure if they're there? Yeah, I think they're coming now, but... OK, they are. And the males are blocking, is that right? Yeah, they've been blocking the entrance because they've got the car arguing and... The entrance to A&E? Yeah, the roundabout to A&E where all the patients pull up the drop-off area. OK, it's the drop-off area. No, this is just like, sorry, delaying it if I'm OK. Have you got a description of the white males? There are two white males, yeah. I haven't got any other descriptions of that, though. I think they'll all be on camera, as well, if needed. So if you just speak to security they'll all be on camera. And we don't know any of their names or anything? No, not yet, no. I can try and find out. I'll do some more information before you come and try and find out. Or why they're fighting. Yeah. All right, no worries. What's your contact number? [REDACTED] yeah. [REDACTED] All right, log number 1000 of today's date and we have dispatch on an emergency response. We've got two controls on read. All right, perfect. Thank you very much. All right, thank you. Cheers, thank you. Cheers, bye-bye. Thank you.

**'Whisper' performance:
93% accurate**

LLM-produced Summary (reduced token length) 53 words

Two patients are fighting outside A&E at Whiston Hospital, blocking the entrance. Security has been notified but their presence is uncertain. The fight involves two white males, and there may be additional individuals involved. The hospital has logged the incident (log number 1000) and has two controls on standby for an emergency response.

**'LLama2' performance:
98% accurate**

LLM-generated Topic Label (single topic per call)

Violence

**Topic label
accuracy currently
being tested**



Topic Modelling – Top 10 Topics



Emergency calls
73,578

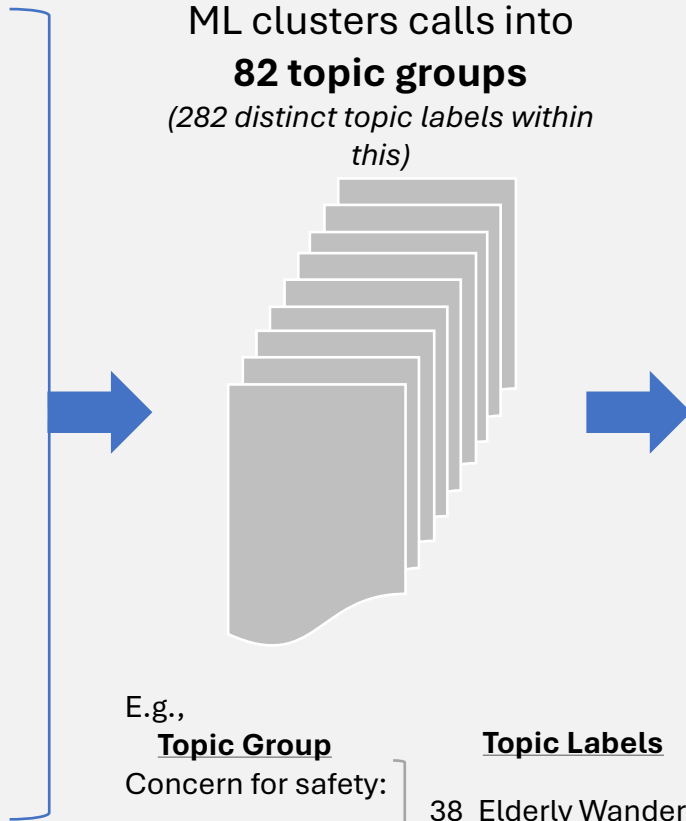


Non-Emergency calls
53,065



Enquiry calls
49,235

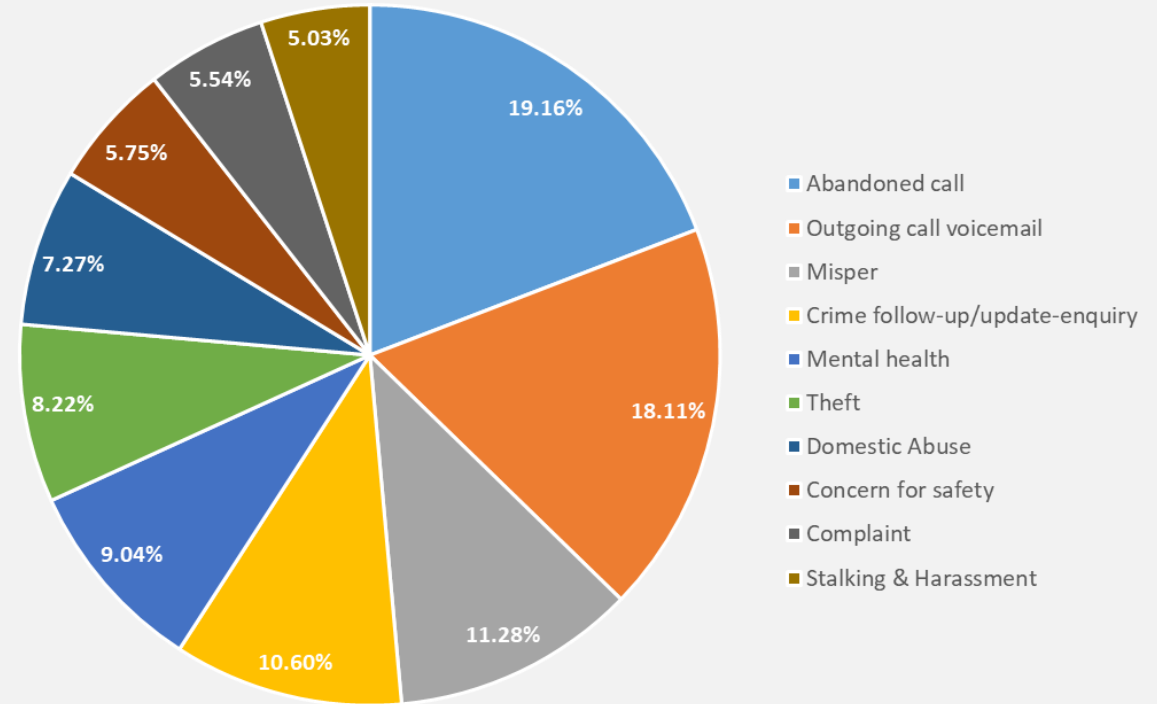
Apr-24 to Jun-24



Topic Labels

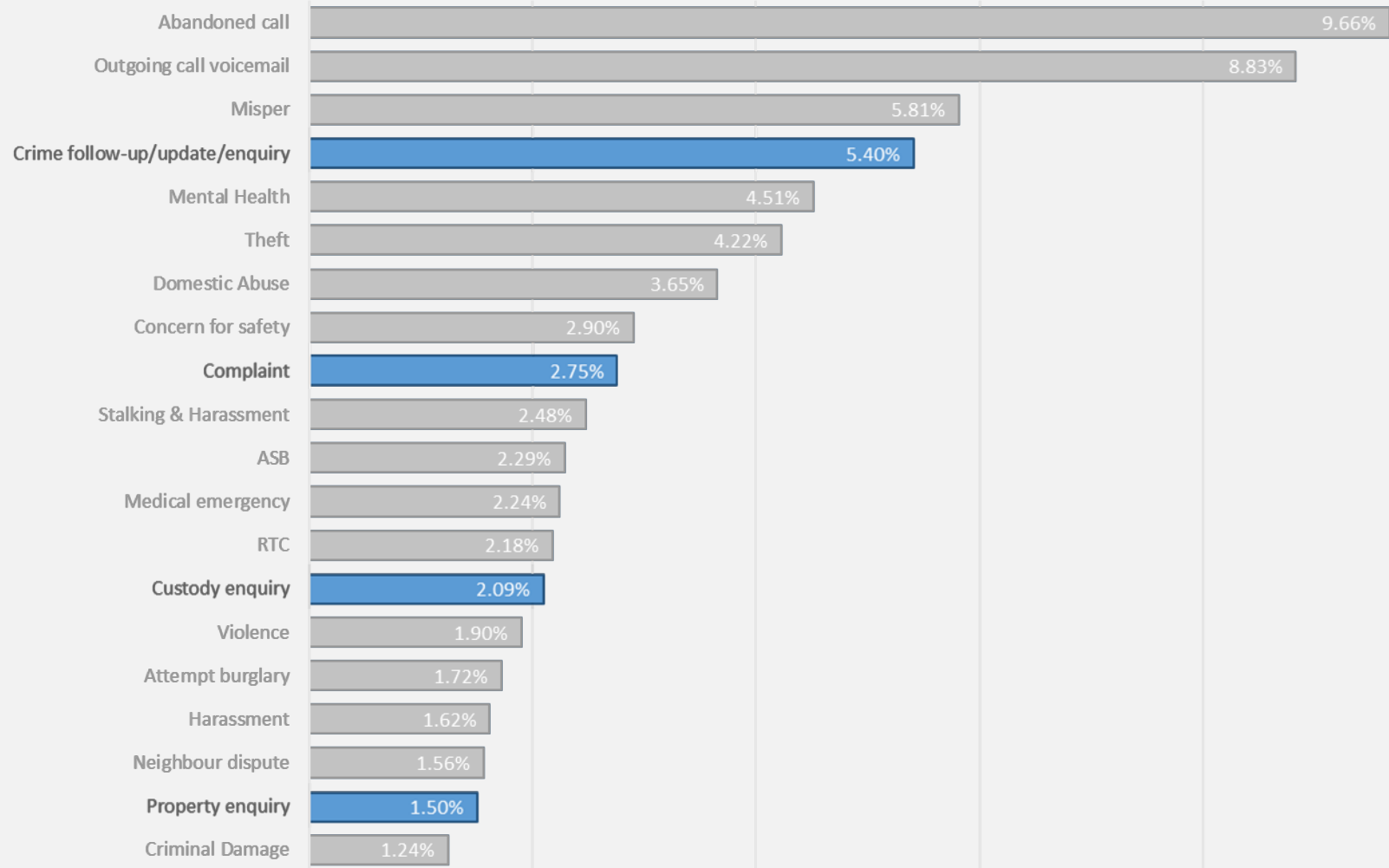
- 38_Elderly Wanderers
- 57_Teenagers Jumping Into Water
- 73_Elderly Stranded
- 150_Welfare Check
- 280_Safeguarding Incident

Top 10 call topic groups



Topic Modelling – Top 20 Topics

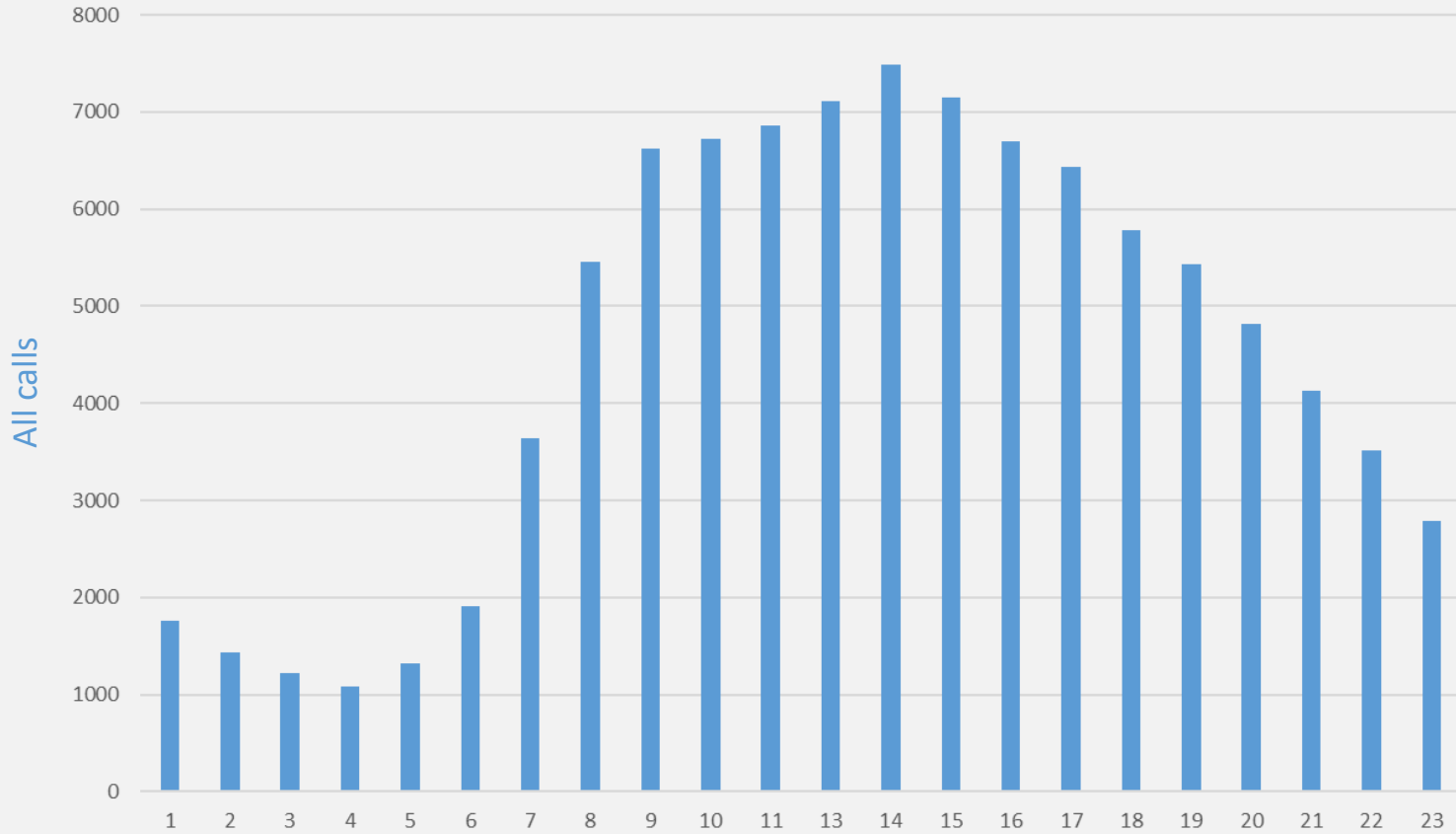
Count of calls by topics
(Top 20 topic groups - 01/04/24 to 14/05/24)



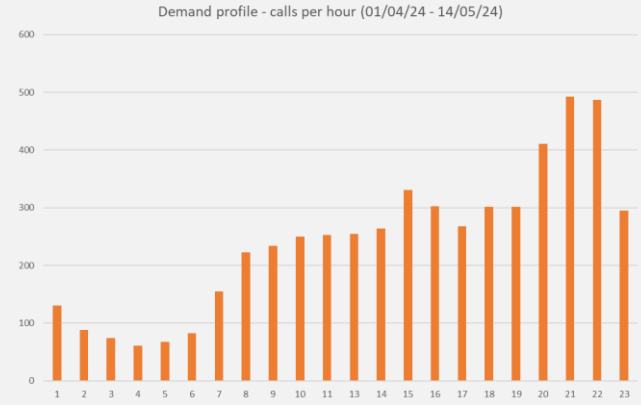
% of grand total

Topic Modelling – Demand Profile

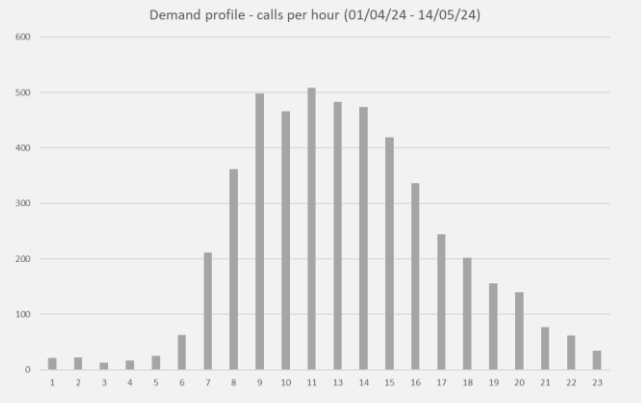
Demand profile - calls per hour (01/04/24 - 14/05/24)



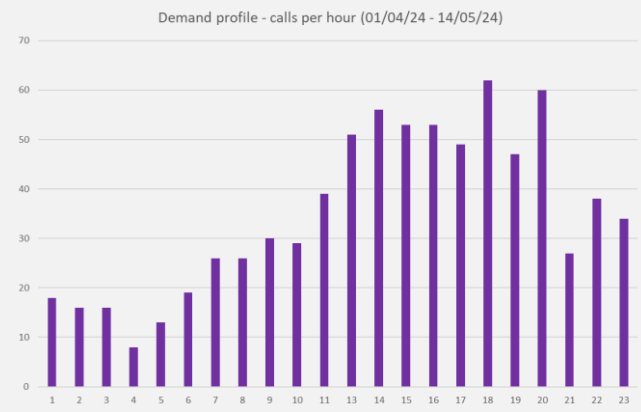
Misper



Crime follow-up/update/enquiry



Chase up deployment

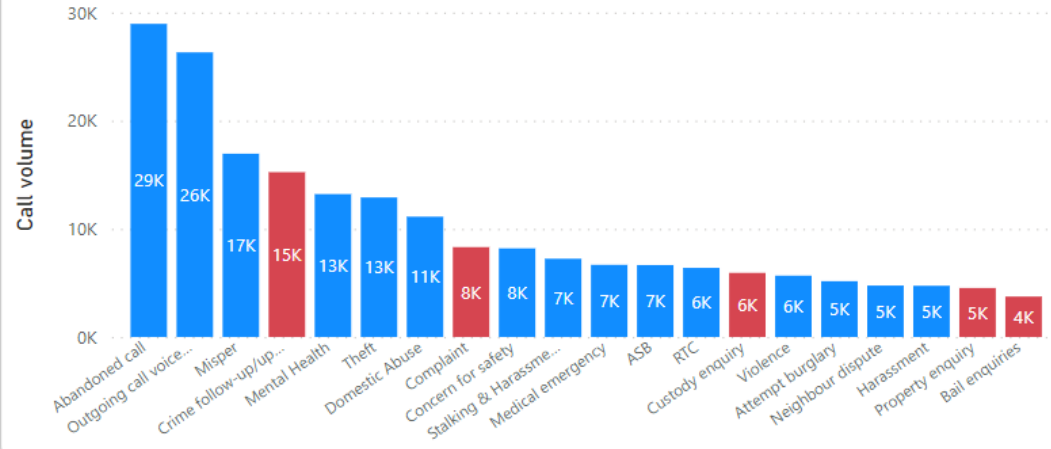


Topic Modelling – Dashboard

Calls by Topic Group

Group	Calls	%GT
Abandoned call	28955	9.79%
Outgoing call voicemail	26318	8.90%
Misper	16944	5.73%
Crime follow-up/update/enquiry	15238	5.15%
Mental Health	13213	4.47%
Theft	12901	4.36%
Domestic Abuse	11123	3.76%
Complaint	8308	2.81%
Concern for safety	8191	2.77%
Stalking & Harassment	7237	2.45%
Medical emergency	6670	2.26%
ASB	6649	2.25%
RTC	6295	2.16%
Total	295661	100.00%

Top 20 Topic Groups



Calls per Hour

Hour	Calls
0	7532
1	5756
2	4856
3	3941
4	3112
5	2741
6	3313
7	5217
8	10159
9	14742
10	17990
11	18787
12	18651
13	19100
14	19569
15	20386
16	19524
17	17849
18	17149
19	16197
20	14952
21	13035
22	11635
23	9468
Total	295661

Calls per Month by group

Year	2024				
	March	April	May	June	July
Abandoned call	5969	6504	7152	7005	2325
Abandoned vehicle	545	569	563	552	195
Alarm activation	483	550	631	619	181
Animal welfare	175	293	382	375	112
Arson	219	263	332	330	116
ASB	1156	1499	1743	1796	455
Assault	360	436	462	402	129
Attempt burglary	1012	1210	1249	1278	408
Bail and court enquiries	353	410	426	405	123
Bail enquiries	809	782	896	918	333
Total	60065	67729	73608	70824	23435

Calls per Day

Week Day	Calls
Wednesday	46734
Tuesday	46492
Friday	43777
Monday	42907
Thursday	41035
Saturday	38779
Sunday	35937
Total	295661

Report drill through

Calls Summary and Transcription

Topic Group: All | Date Range: 01/03/2024 - 31/12/2024

Call Key	Date	Summary	Transcription
12353463	25/04/2024	The owner of an Apple Watch has triggered an emergency alarm after knocking the device, causing it to believe a crash has occurred. The police were mistakenly called, but the issue was quickly resolved and the alarm was turned off.	...latitude longitude minus with an estimated search radius of 51 metres, this method will repeat in 5 seconds. Hello, it's police. Hello? Hello, it's the police. The owner of this Apple Watch has taken... Why is it the police? I'm sorry, I haven't called the police. okay, you've knocked your Apple Watch so it's setting an emergency alarm off. Oh, I'm sorry, I don't know how they've done that. That's okay, no problem. You don't need police fire rambles, no? No. No. I just panicked, I didn't know what it was. Oh, bless you. Yeah, that's is. If you knock your Apple Watch, it thinks you've been in an accident, but that's no problem, we'll close it down, alright. Thank you ver much, bye. Thank you, bye. Thank you, bye.

Topic Modelling – Next Steps

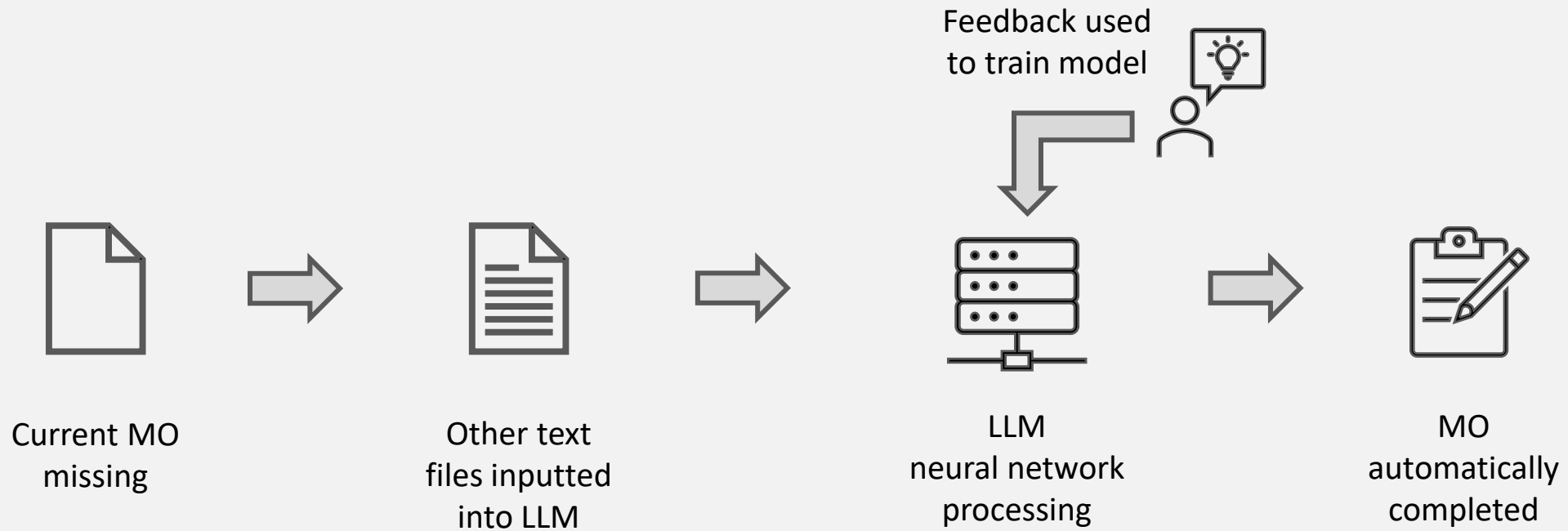
- Analysis of data to inform strategic insight
- Ongoing work to further test, and improve, accuracy
- Additional developments of MI dashboard
- Technical enhancements to process
- Working group in place to consider how data can be utilised further (e.g. sentiment analysis)



Automatic MO – Background

- Unsatisfactory Data Quality
- Find and Fix process
- ...existing 'Get it right second time' process is inefficient

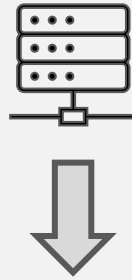
Automated MO – Process



Automated MO – Example

Input

Offender has entered rear garden of property and climbed up to flat roof outside window when he has been seen by victim's daughter who has returned home. Suspect has then shouted "RUN" and fled.



Output

1. The location of the burglary was **Dwelling**.
2. The actual point of entry was **Window - upper rear**.
3. The entry method was **Use code**. This is inferred because the offender opened the window, implying they used a code to unlock it.
4. Damage was caused to the **Rear flat roof door**.
5. **Unknown**. The occupation of the offender is not mentioned in the provided input.
6. Summary: A person attempted to burgle a dwelling by entering through the rear window on the upper floor. They were spotted by the victim's daughter, who returned home unexpectedly, causing the offender to shout "RUN" and flee the scene. The CSI was called to examine the scene, and no obvious forensic opportunities were found, although some objects were noted to be out of place. The offender's description was obtained from the witness statement, and further enquiries are required.

Automated MO – Examples -Structured and Free text MO

Withheld in full S31 S38

Automated Structured MO – Accuracy

Burglary

MO Field / Information	% completed in Niche *	Mixtral accuracy
Location	99%	63%
Entry Point	41%	88%
Entry Method	37%	87%
Attempted Entry Points	22%	75%
Damage Caused	22%	78%
Offender Occupation	21%	33%
Free-text Summary	100%	87% - 99%

* Apr-24 to Jun-24

Sexual Offences

MO Field / Information	% completed in Niche *	Mixtral accuracy
Location	90%	59%
Victim/Offender Relationship	97%	79%
Victim Activities	45%	75%
Act of Violence	6%	77%
Sexual Acts	44%	87%
Offender Behaviour	6%	88%
Offender flee scene	3%	50%
How Victim/Offender Met	N/A	60%
Offender Occupation	38%	48%
Free-text Summary	98%	89% - 100%

* Apr-24 to Jun-24

Violence

Currently being tested



Automated MO – Next steps

- Outputs to be aggregated into groups
- Further testing and accuracy improvements
- MI reports/dashboards
- Analysis of key themes (Intelligence and Strategic Analysis)
- Apply to broader offences