TRECVID 2020 VIDEO SUMMARIZATION INTRODUCTION AND TASK OVERVIEW
*Pilot Task

Keith Curtis
National Institute of Standards and Technology

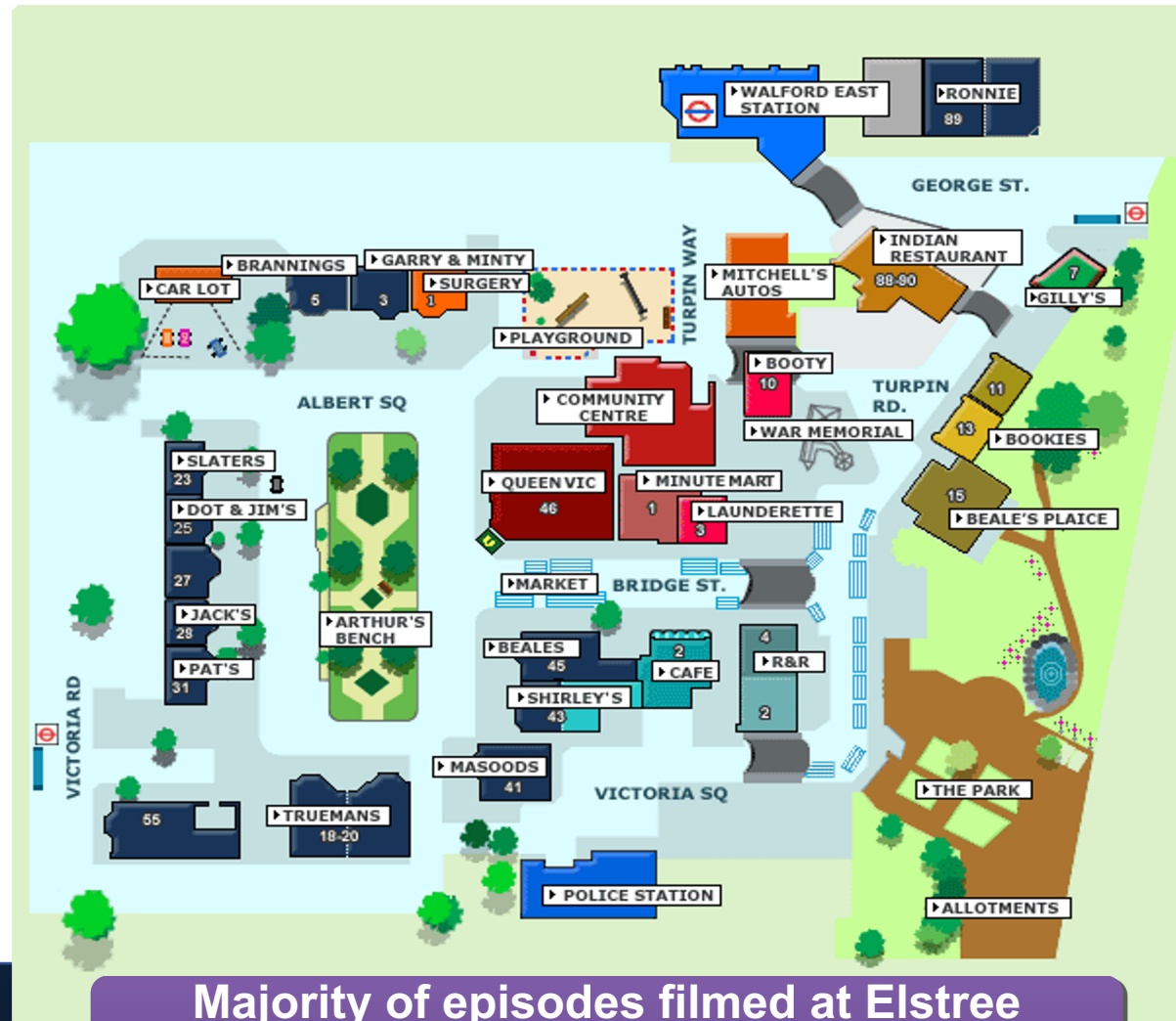Gareth J. F. Jones
Dublin City University

NIST

Table of Contents

- Task Definition

- Data

- Topics (Queries)

- Participating Teams

- Evaluation and Results

- General Observation

Data

- The British Broadcasting Corporation (BBC) and the Access to Audiovisual Archives (AXES) project made **464 h** of the BBC soap opera EastEnders available for research
  - 244 weekly "omnibus" files (MPEG-4) from 5 years of broadcasts
  - 471,527 shots
  - Average shot length: 3.5 seconds
  - Transcripts from BBC
  - Per-file metadata

- Represents a "small world" with a slowly changing set of:
  - People (several dozen)
  - Locales: homes, workplaces, pubs, cafes, open-air market, clubs
  - Objects: clothes, cars, household goods, personal possessions, pets, etc
  - Views: various camera positions, times of year, times of day,
  - Use of fan community metadata allowed, if documented

EastEnders World

Majority of episodes filmed at Elstree studios. Sometimes filmed on 'location'.

Task Definition

System task:

- Given:
  - A collection of BBC Eastenders test videos with a master shot boundary reference
  - A time frame of the series for which to use for summarization
  - 3 characters with 4 images at different angles for each character
- Return a summary of the major life events of each character within the specified time frame of the series (**4 runs for each character requested).**

Topic + Content Question Procedure at NIST

- Viewed omnibus meta-data to isolate 3 characters who played a major role in the series within a 10-12 week timeframe.

- Watched full episodes from that timeframe of series, took a list of major event that include target characters.

- Iteratively refined list of major events down to 10 for each character.

- Refined list of events voted on by TRECVID team for final 5 major events for each character to be asked in content questions during evaluation.

NIST

Run Types

- Plan long term: to allow teams to submit multiple runs by using different techniques / weightings.

- For this year, we first want to gauge optimal number for the maximum number of shots and maximum length for each summary. Runs types used to gauge this.

- Run 1 – max 5 shots, 150 seconds length.

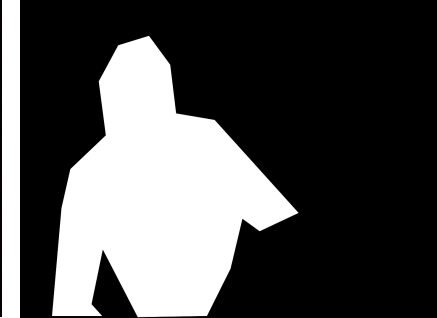- Run 2 – max 10 shots, 300 seconds length etc.

Submission Requirements

Character	Janine	Ryan	Stacey
Start Shot #	175_1	175_1	175_1
End Shot #	185_1736	185_1736	185_1736
Max Shots Run 1	5	5	5
Max Length Run 1	150 seconds	150 seconds	150 seconds
Max Shots Run 2	10	10	10
Max Length Run 2	300 seconds	300 seconds	300 seconds
Max Shots Run 3	15	15	15
Max Length Run 3	450 seconds	450 seconds	450 seconds
Max Shots Run 4	20	20	20
Max Length Run 4	600 seconds	600 seconds	600 seconds

Queries – Character Example Images

Janine

Ryan

Stacey

VSUM 2020: 2 Finishers (out of 12)

Team	Organization
MeMAD	Eurecom and Aalto University for MeMAD
NII_UIT	National Institute of Informatics, Japan (NII); University of Information Technology, VNU-HCM

NIST

Evaluation – Subjective Measures

- 5 Binary (yes or no) questions on content.

- 3 Likert scale ratings (1 - 7) on Tempo, Contextuality, and Redundancy.

1. Tempo / Rhythm (1 - 7). 7 is Best.

2. Contextuality (1 - 7). 7 is Best.

3. Redundancy (1 - 7). 1 is Best.

NIST

Metrics

- Want to rate content more important, but still want to give credit for the other factors. 7 * 3 = 21 for best possible scores for Tempo, Contextuality, and Redundancy (Redundancy is best at 1 so this would be flipped). 100 - 21 = 79 available for content. 79 / 5 content questions = 15.8 possible (rounded) per question.

- This gives a perfect summary 100 points. A summary with no relevant content but all perfect scores for the other factors would get 21 points. Overall this would give summaries a maximum score of 100 down to a minimum score of 3.

Evaluation – Content Questions

Janine

1. What is causing Ryan to be sick in bed?

2. How does Janine attempt to kill Ryan while he is in the hospital?

3. What happens when Janine attempts to play recording of Stacey?

4. Who stabbed Janine?

5. Who gives Janine the recording of Stacey?
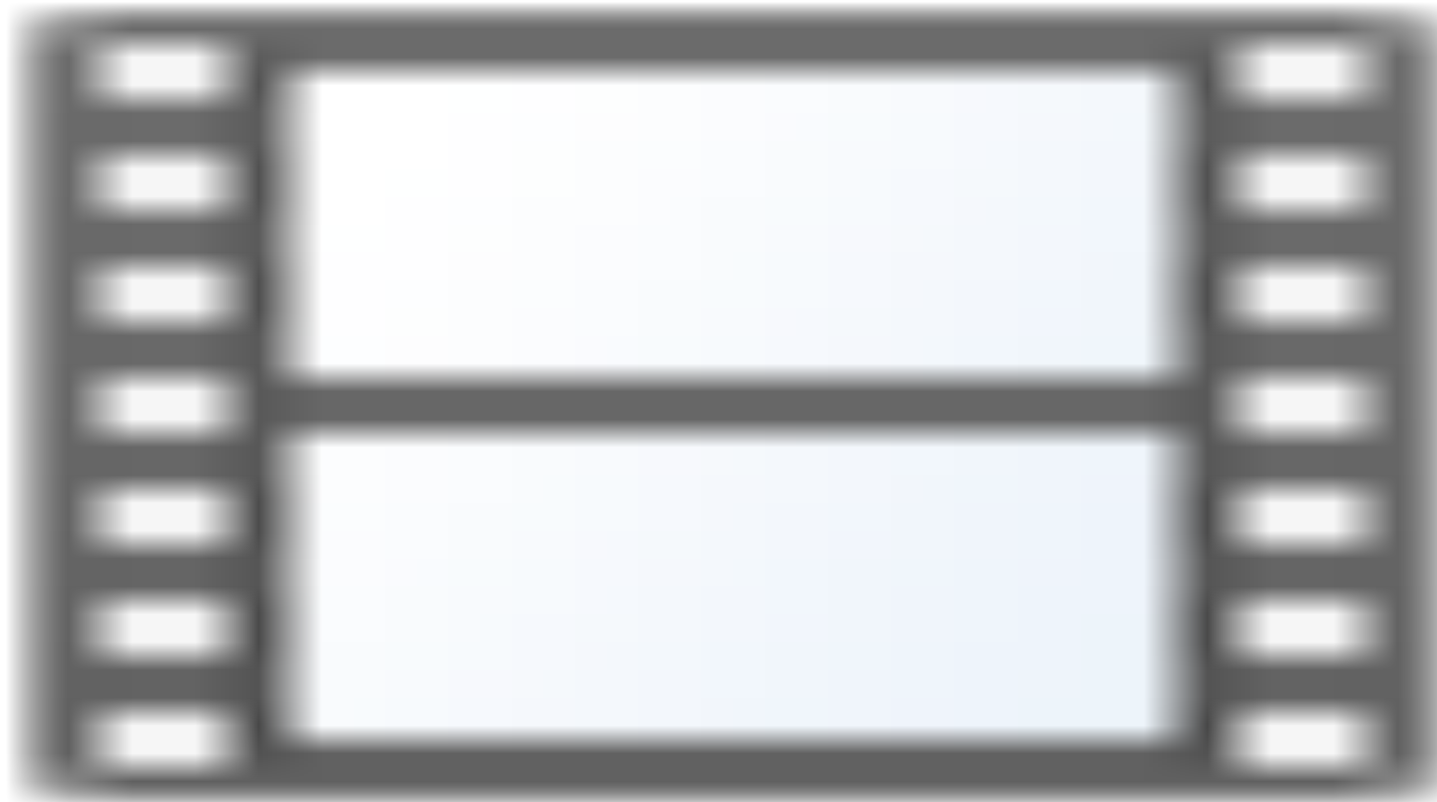
NIST

Evaluation – Content Questions

Ryan

1. How does Janine attempt to kill Ryan while he is in the hospital?

2. What does Ryan do when Janine is lying in the hospital?

3. Where is Ryan trapped?

4. What does Ryan tell Phil he can do for him?

5. Who is Ryan with when going to put his name on the baby's birth cert?
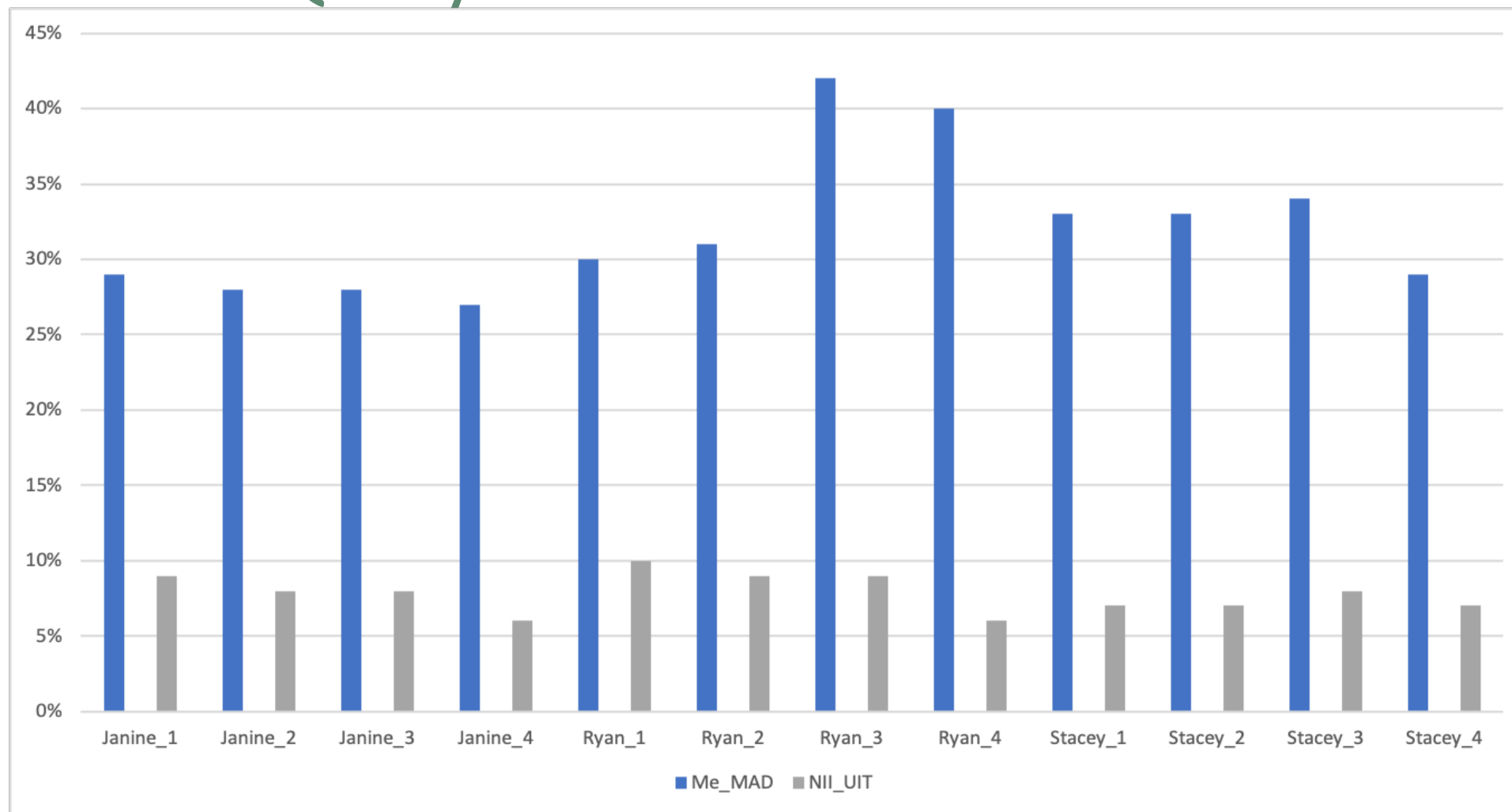
Evaluation – Content Questions

Stacey

1. Who climbs up the roof to talk Stacey out of jumping off?

2. What does Stacey reveal when in a cell with Janine, Kat, and Pat?

3. What does Stacey admit to her mum in bedroom when mum is upset?

4. Who confronts Stacey in restroom where Stacey finally admits to killing Archie?

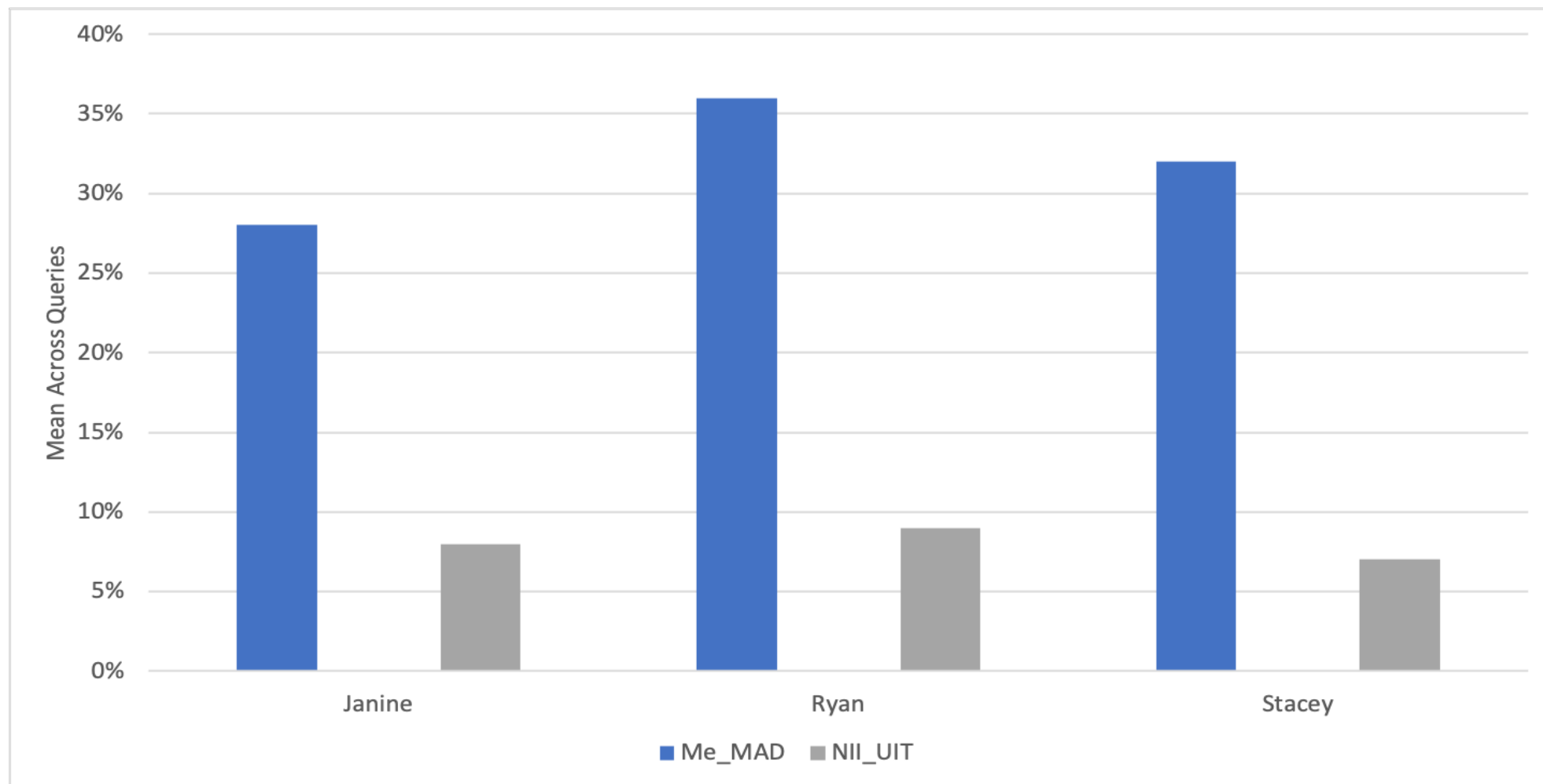5. Who calls to Stacey's door to tell her to get her stuff and go after Stacey's mum had called the police?

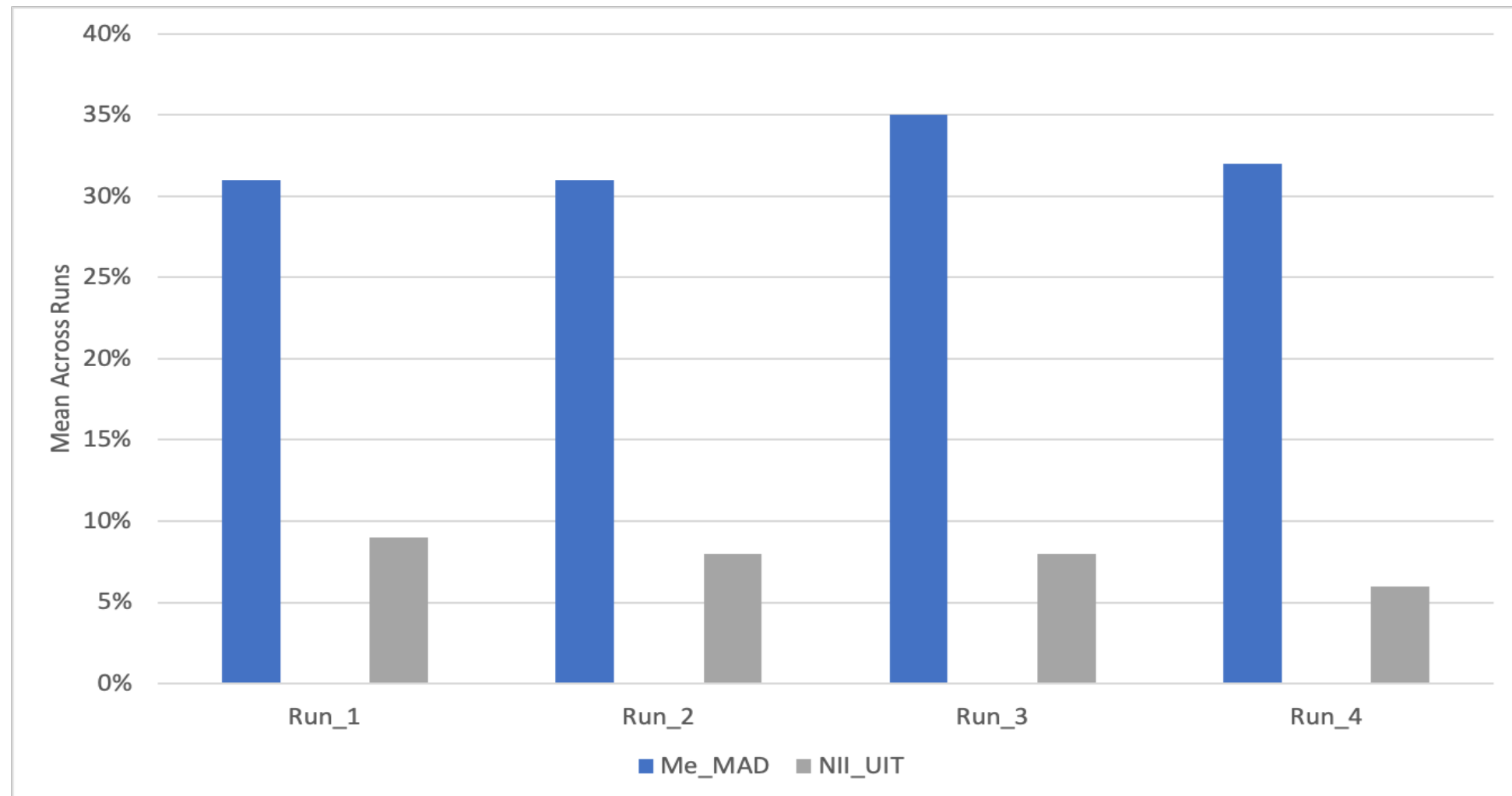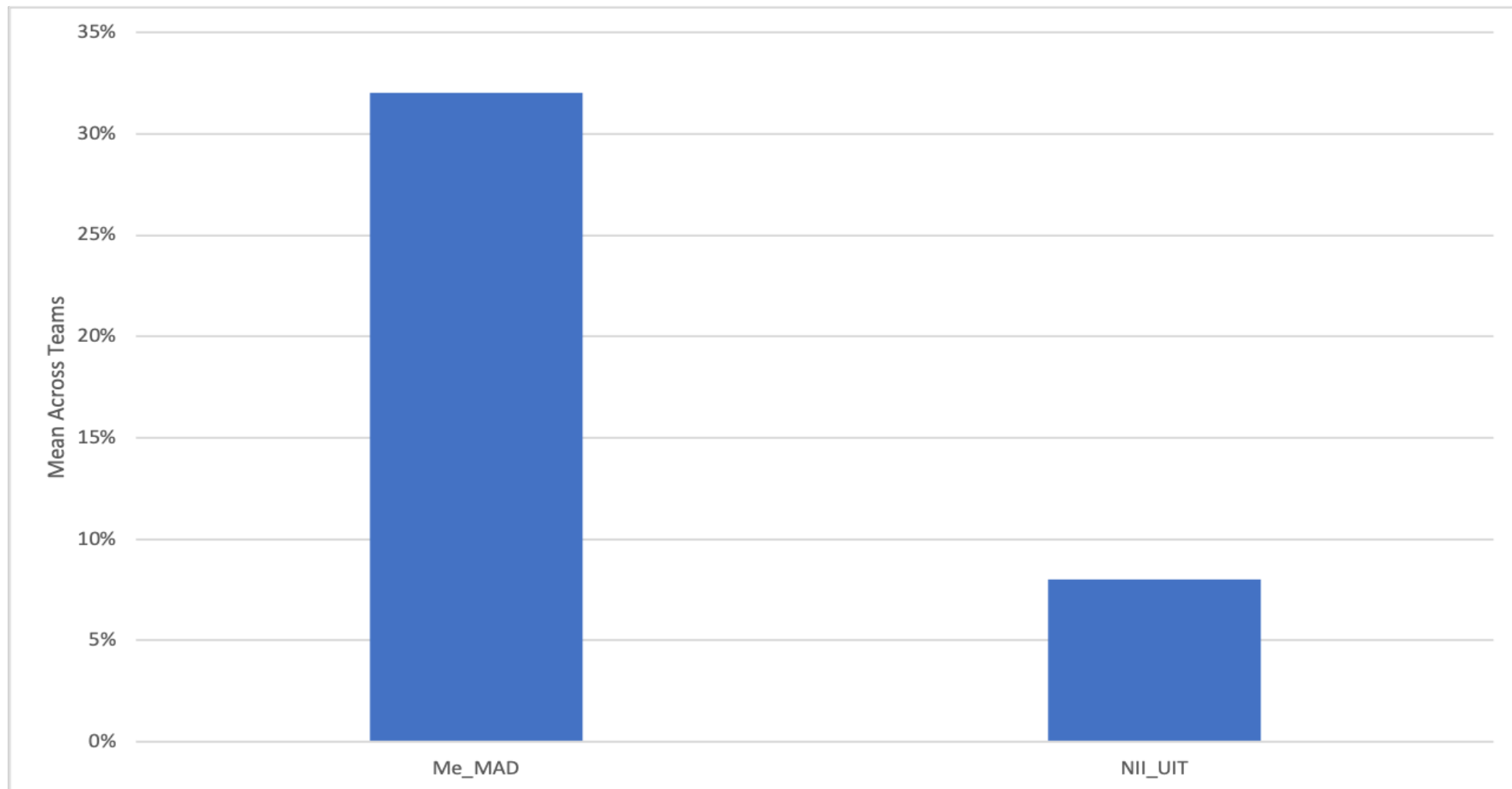# Sample Summary Submission – Ryan, Run 2

Results – Query + Run

Results – Query

Results – Run

Results – Team

Conclusions

1. Run 3 conditions (max 15 shots, 450 seconds) appear to give best results.

2. That being said, maximum summary lengths were very long, with no submission needing the full allowed time. Plan to reduce this for future years.

3. Ryan found to be the easier character / query to summarize major life event of. Janine most difficult.

4. Highest results for Ryan. Lowest results for Janine.

Conclusions

5. Only two out of 12 registered teams submitted runs. Not enough, we need to encourage more teams to participate.

6. Due to small number of submissions, this year's task was evaluated in house by TRECVID team. Best to be evaluated by outside annotators not already familiarized with the data.

7. Assuming more teams participate next year, plan for submissions to be assessed by students at Dublin City University.