TRECVID 2021 VIDEO SUMMARIZATION INTRODUCTION AND TASK OVERVIEW

Keith Curtis
National Institute of Standards and Technology

Gareth J. F. Jones
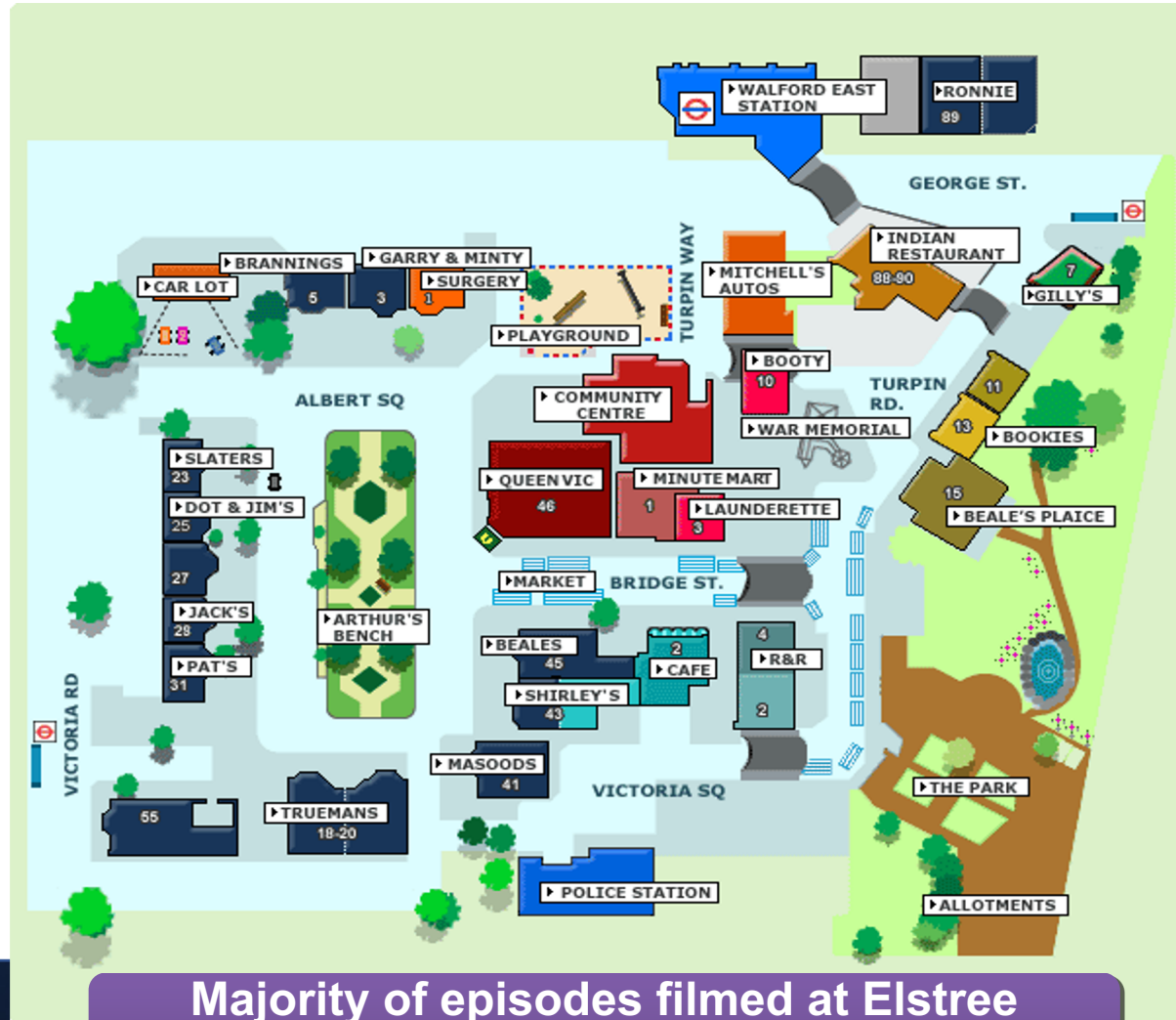Dublin City University

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
  - 5 characters with 4 images at different angles for each character
  - A time frame of the series for which to use for summarization of 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 5 characters who played a major role in the series within a 10-12 week timeframe.

- Watched full episodes from selected timeframes 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 organizers for final 5 major events for each character to be asked in content questions during evaluation.
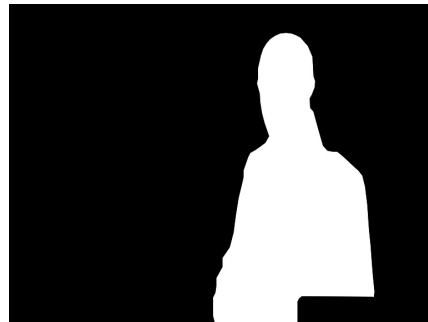
NIST

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, 50 seconds length.

- Run 2 – max 10 shots, 100 seconds length etc.

Submission Requirements

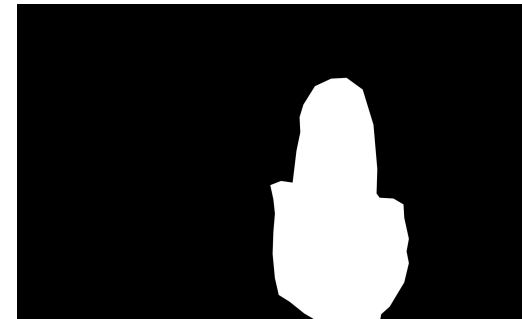| Character	Max	Jack	Tanya	Archie	Peggy
Start Shot #	60_1	60_1	60_1	79_1	79_1
End Shot #	70_2040	70_2040	70_2040	89_2036	89_2036
Max Shots Run 1	5	5	5	5	5
Max Length Run 1	50 seconds	50 seconds	50 seconds	50 seconds	50 seconds
Max Shots Run 2	10	10	10	10	10
Max Length Run 2	100 seconds	100 seconds	100 seconds	100 seconds	100 seconds
Max Shots Run 3	15	15	15	15	15
Max Length Run 3	150 seconds	150 seconds	150 seconds	150 seconds	150 seconds
Max Shots Run 4	20	20	20	20	20
Max Length Run 4	200 seconds	200 seconds	200 seconds	200 seconds	200 seconds

NIST
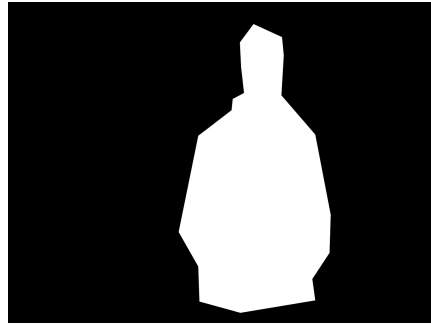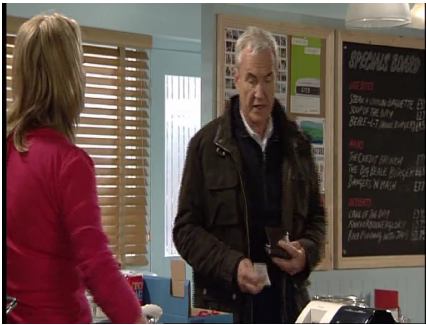
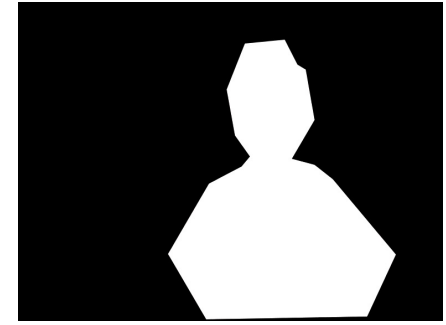# Queries – Character Example Images

Max

Jack

Tanya

Queries – Character Example Images

Archie

Peggy

NIST

VSUM 2021: 3 Finishers (out of 11)

Team	Organization
Adapt_team	ADAPT Research Centre
EURECOM	EURECOM
NII_UIT	National Institute of Informatics, Japan University of Information Technology, VNU-HCMC, Vietnam

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.

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

Max

1. What was the cause of Max's serious injuries which left him in hospital?

2. What is/was the relationship between Max and Tanya?

3. What kind of weapon does Max obtain from Phil?

4. Where are Max and Jack during the voilent confrontation between them when a gun is drawn?

5. Who is responsible, or who does Max believe is responsible, for the serious injuries which left him in hospital?

Evaluation – Content Questions

Jack

1. What happens when police breaks in the door of Jack and Tanya's home?

2. Where are Max and Jack during the voilent confrontation between them when a gun is drawn?

3. Who does Jack offer to pay in order to withdraw their statement to the police?

4. Why is Jack a suspect in the hit and run on Max?

5. What does Jack reveal to Tanya about his dodgy past?

Evaluation – Content Questions

Tanya

1. What does Tanya reveal to the police while being interviewed at the station?

2. What is/was the relationship between Max and Tanya?

3. What does Jack reveal to Tanya about his dodgy past?

4. What does Tanya discover in the sink and on Jack's clothes?

5. What big move were Tanya and Jack planning for the future?
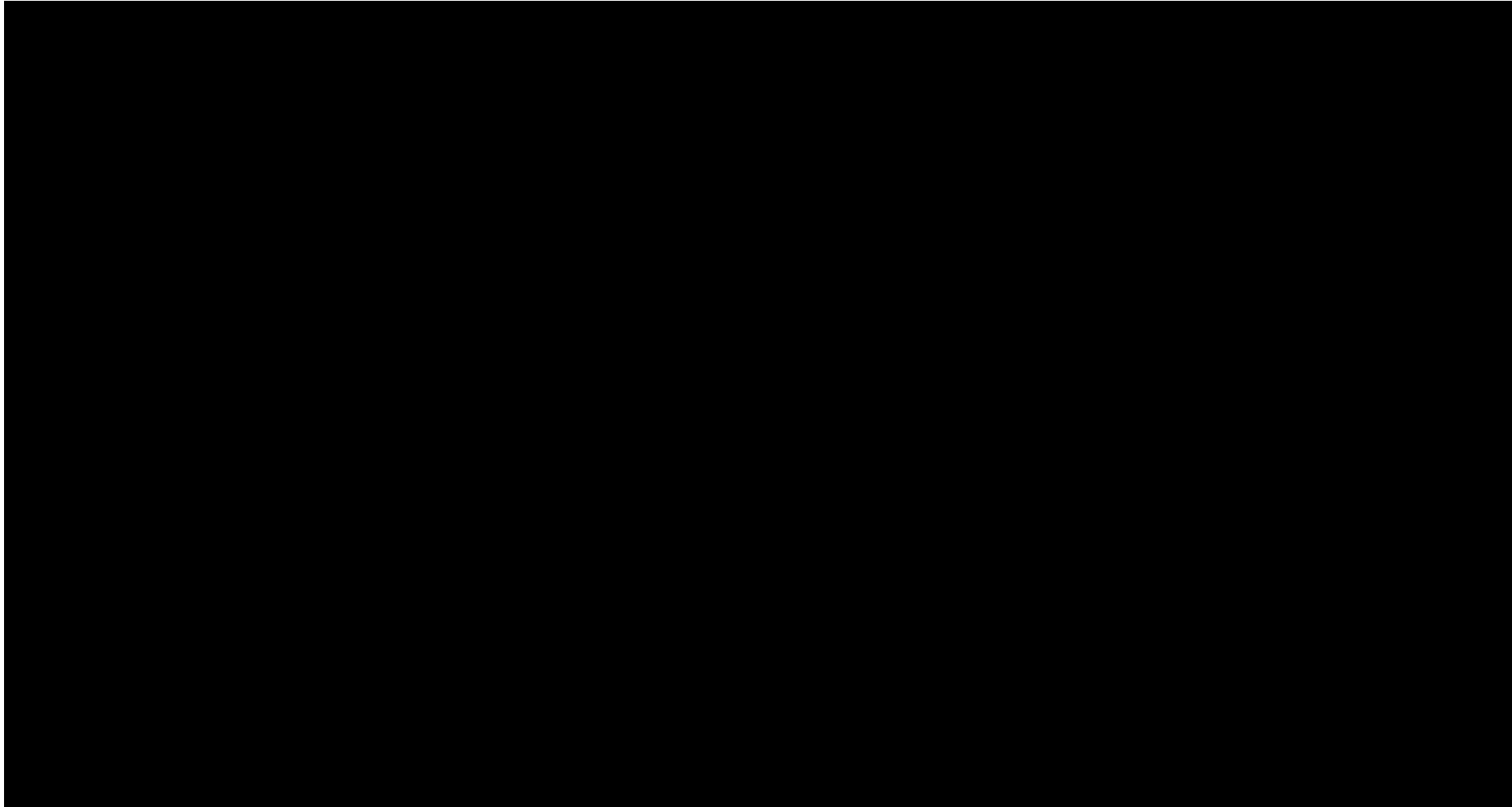
Evaluation – Content Questions

Archie

1. What happens when Phil throws Archie in to a pit?

2. What happens after Danielle reveals to Archie that Ronnie is her mother?

3. Where do Peggy and Archie get married?

4. What happens when Archie arrives at the pub after Peggy invited him?

5. What happens when Archie is kidnapped?
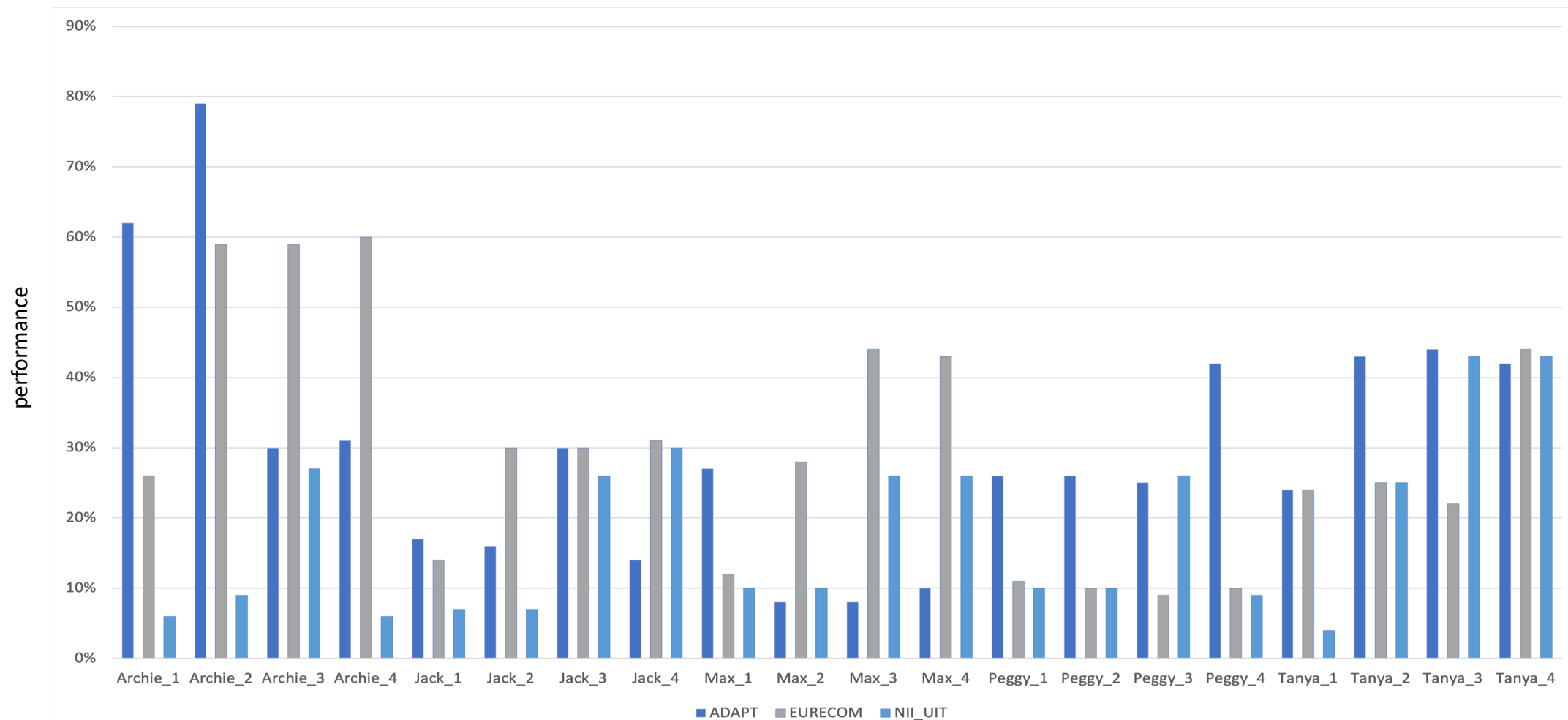
Evaluation – Content Questions

Peggy

1. Who does Peggy ask to kill Archie?

2. Where do Peggy and Archie get married?

3. Show one of the challenges which Peggy faces in her election run.

4. What does Peggy overhear Archie saying, which causes their marriage to be over?

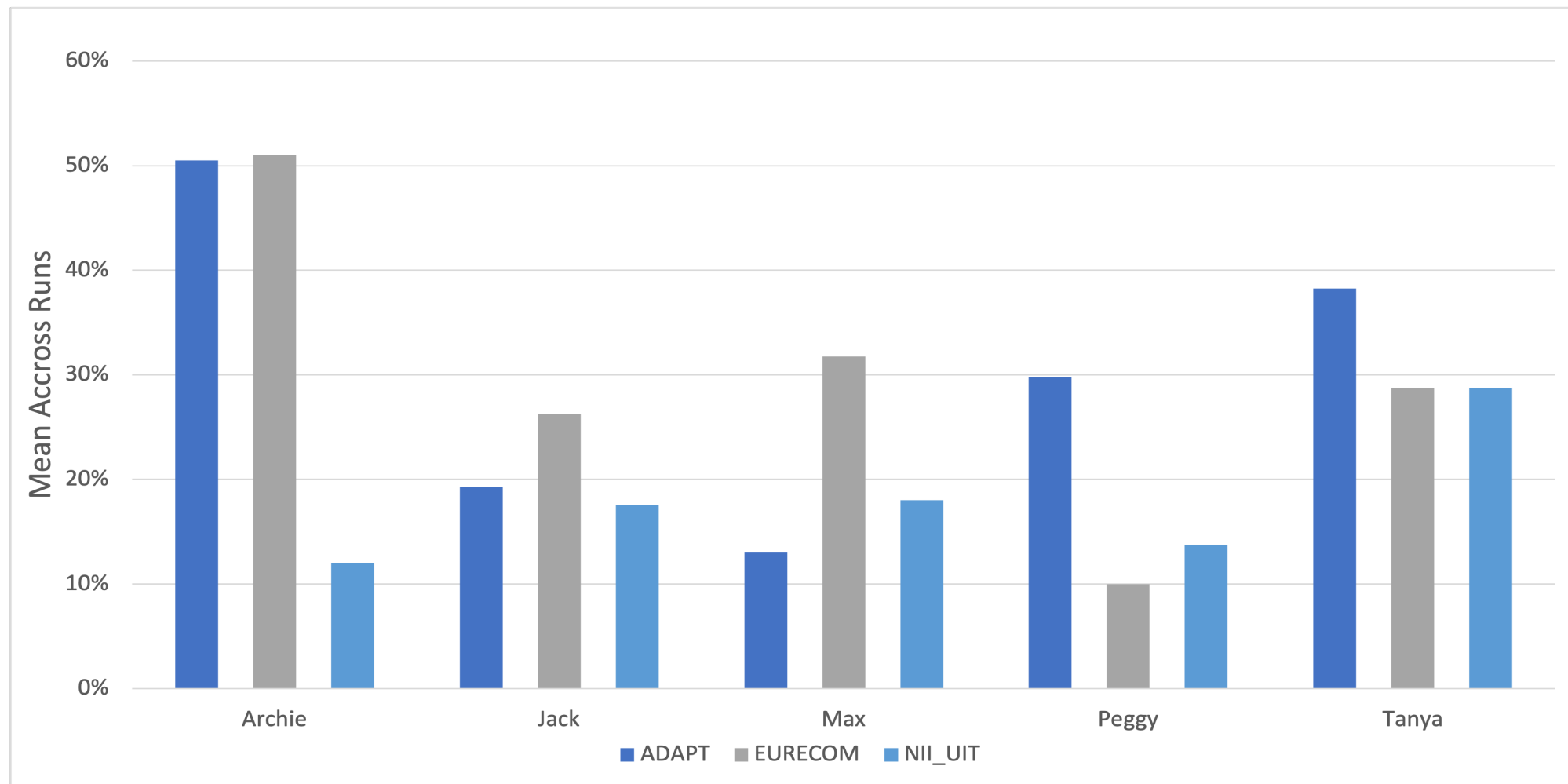5. What is Janine doing to irritate or anger Peggy?
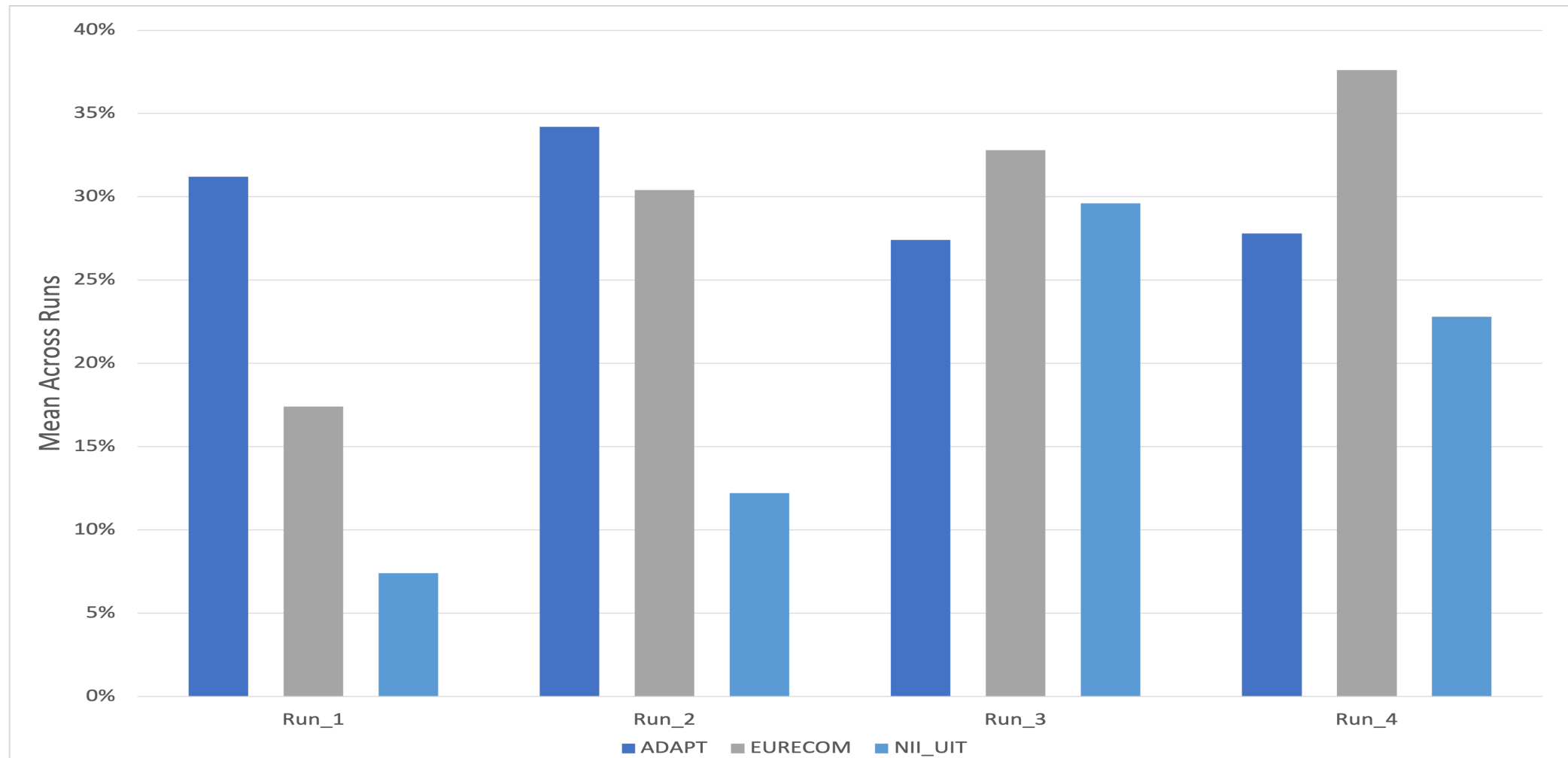
Sample Summary Submission – Max, Run 2

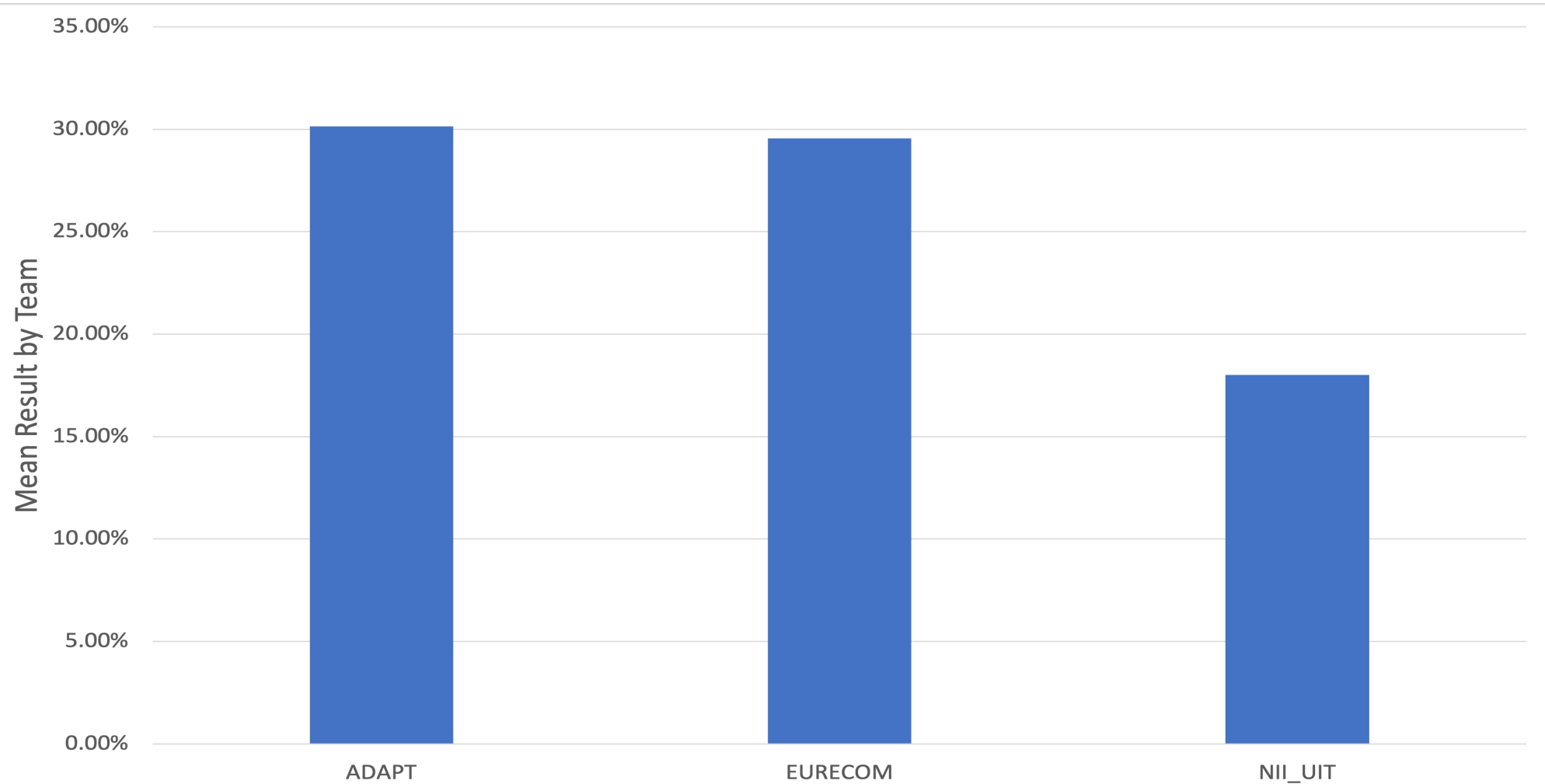# Main Task Results – Query + Run

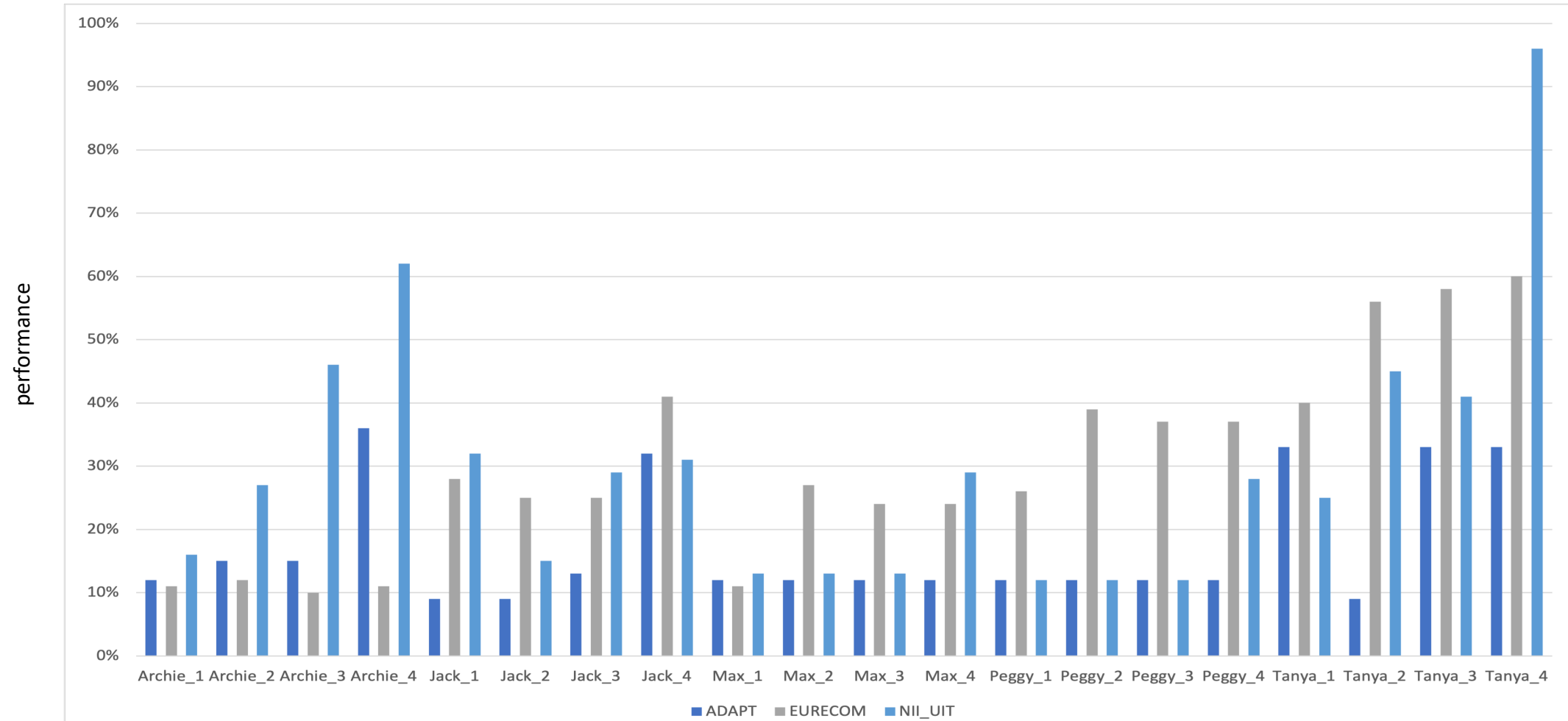# Main Task Results – Query

Main Task Results – Run
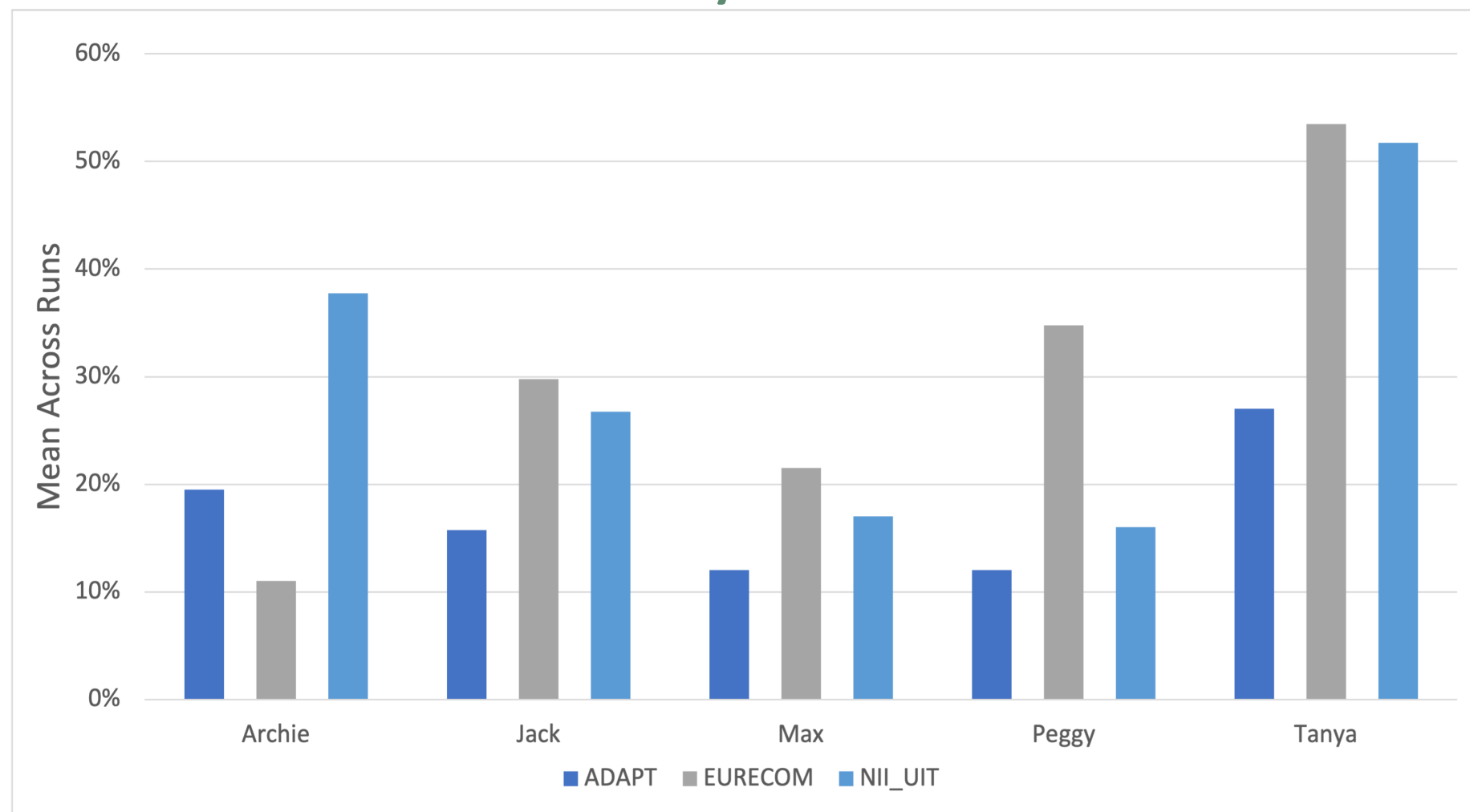
Main Task Results – Team
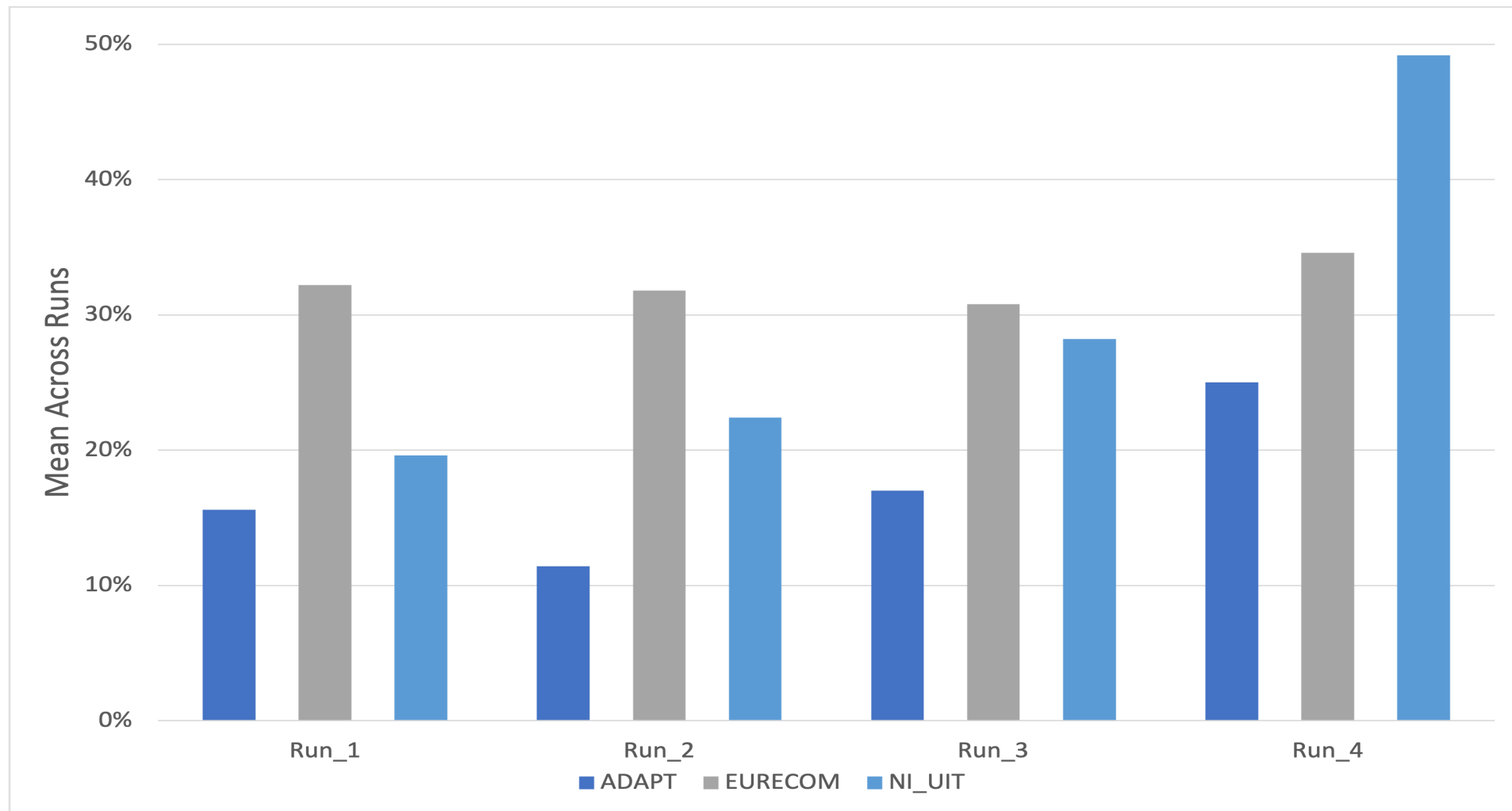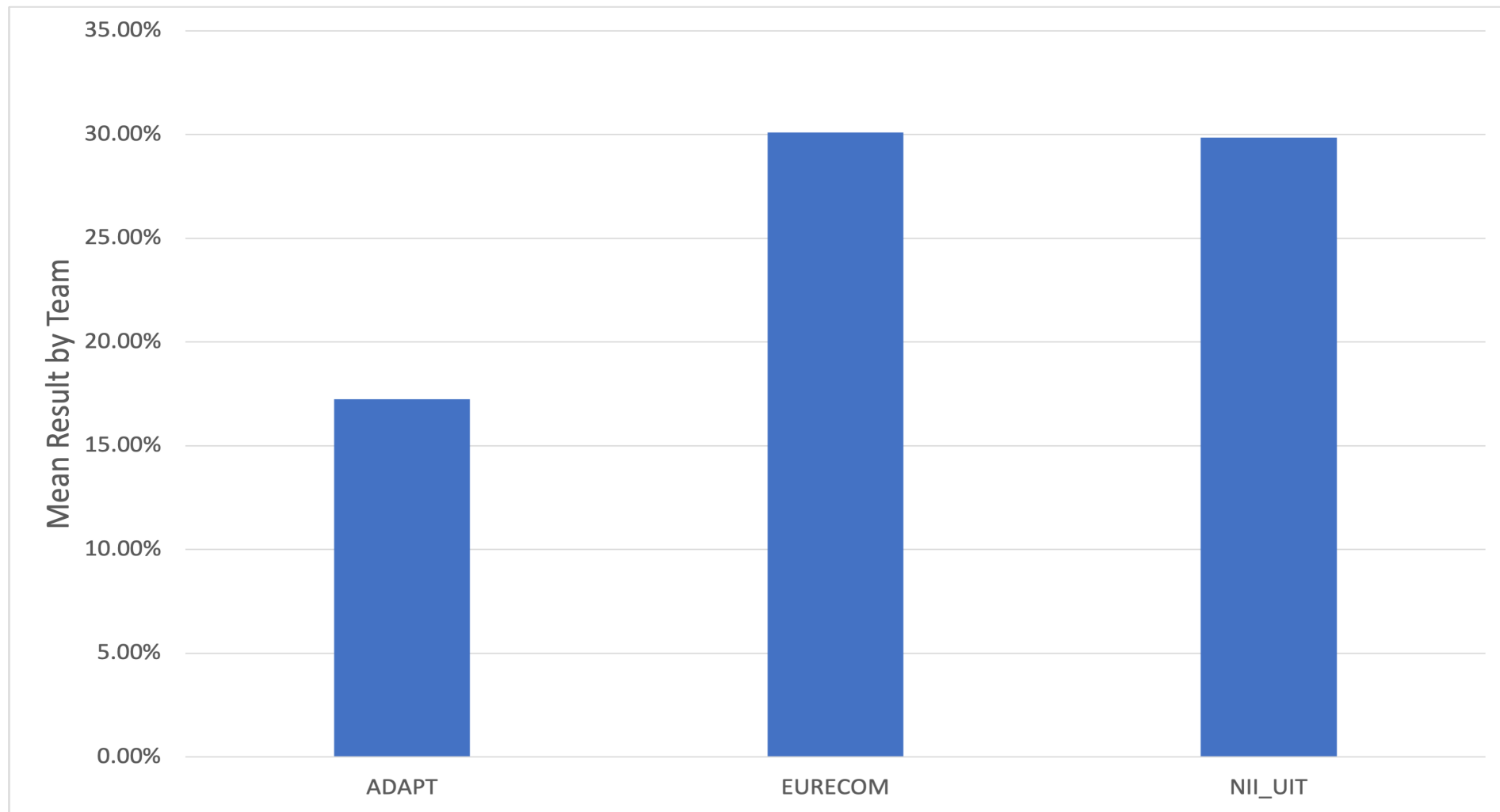
Sub Task Results – Query + Run

Sub Task Results – Query

Sub Task Results – Run

Sub Task Results – Team

Conclusions

1. Run 3 conditions (max 15 shots, 150 seconds) give best results for the main task. This is consistent with the previous year of the task. However for the subtask, run 4 conditions give best results.

2. For the main task of unknown questions, Archie was found to be the easier character / query to summarize major life event of. For the sub task of known questions, Tanya was the easiest character / query to summarize.

3. For the main task of unknown questions, Peggy was the hardest and for the subtask Max was hardest.

4. Across the tasks, Max and Peggy were consistently found to be hard queries and Tanya was consistently found to be easier.

Conclusions

5. 11 teams registered for this years task. Out of these 3 teams submitted runs. An improvement on last year (2 out of 12), but still not enough.

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.