Supporting Document for Public Sentiment on Security Cameras

Anonymous

I. Introduction

This document provides additional supporting materials to the main paper "Public Sentiment on Security Cameras." It includes the list of keywords used for data extraction, text preprocessing code, hypothesis assignment code, word cloud analysis, and detailed visualizations of sentiment trends across the different hypotheses examined in the main study.

Appendix A
List of Keywords Used for Data Collection

The following is the complete list of keywords used in the data extraction process from the Suomi24 forum:

- "public camera privacy"
- "CCTV surveillance"
- "security camera"
- "julkinen kamera yksityisyyttä"
- "kamera valvontaa"
- "public", "camera", "privacy",
- "public", "camera", "yksityisyyttä",
- "public", "camera", "yksity",
- "julkinen", "kamera", "yksityisyyttä",
- "julkinen", "camera", "yksityisyyttä",
- "julkinen", "kamera", "yksity",
- "julkinen", "camera", "yksity",
- "cctv", "surveill",
- "cctv", "valvontaa",
- "cctv", "valvo",
- "camera", "surveill",
- "kamera", "surveill",
- "camera", "valvontaa",
- "kamera", "valvontaa",
- "camera", "valvo",
- "kamera", "valvo",
- "security", "camera",
- "turvatoimet", "camera",
- "turvatoimet", "kamera",
- "public", "safety", "concern",
- "julkinen", "turvallisuus", "koskea",
- "julkinen", "turvallisuus", "huole",
- "public", "security", "concern",
- "julkinen", "turvatoimet", "koskea",
- "julkinen", "turvatoimet", "huole",
- "monitor", "cctv",
- "seurantaa", "cctv",
- "seurant", "cctv",
- "monitor", "camera",
- "seurantaa", "camera",
- "seurant", "camera",
- "monitor", "kamera",
- "seurantaa", "kamera",
- "seurant", "kamera",
- "public", "footage",
- "julkinen", "kuvamateriaali",
- "camera", "footage",
- "camera", "kuvamateriaali",
- "kamera", "kuvamateriaali",
- "street", "camera",
- "katu", "camera",
- "katu", "kamera"

Appendix B
Text Cleaning Code

The following Python code was used to preprocess the text data, including handling HTML entities, removing email addresses and URLs, and expanding contractions.

```python
contraction_dict = {
    "it's": "it is", "aren't": "are not",
    "can't": "cannot", "couldn't": "could not",
    "didn't": "did not", "doesn't": "does not",
    "don't": "do not", "hadn't": "had not",
    "hasn't": "has not", "haven't": "have not",
    "he'd": "he would", "he'll": "he will",
    "he's": "he is", "I'd": "I would",
    "I'll": "I will", "should've": "should have",
    "I've": "I have", "isn't": "is not",
    "let's": "let us", "might've": "might have",
    "must've": "must have","mustn't": "must not",
    "shan't": "shall not", "she'd": "she would",
    "she'll": "she will", "she's": "she is",
    "shouldn't": "should not", "I'm": "I am",
    "that's": "that is", "there's": "there is",
    "they'd":"they would","they'll": "they will",
    "they're": "they are","they've": "they have",
    "we'd": "we would", "we're":
    "we are", "we've": "we have",
    "weren't": "were not", "what'll":
    "what will", "what're":"what are",
    "what's": "what is", "what've":
    "what have", "where's": "where is",
    "who'd": "who would", "who'll":
    "who will", "who's": "who is",
    "who've": "who have", "won't":
```

```
    "will not", "wouldn't": "would not",
    "you'd":"you would", "you'll":
    "you will", "you're": "you are",
    "you've": "you have"
}

def preprocess_text(text):
    if pd.isna(text):
        return text

    if not isinstance(text, str):
        text = str(text)

    html_char_map = {
        "&amp;": "&",
        "&apos;": "'",
        "&#39;": "'",
        "â€™": "'",
    }
    for html_char, replacement in
                 html_char_map.items():
        text = text.replace(html_char,
                           replacement)

    patterns = [
        r'\b[A-Za-z0-9._%+-]+
                 @[A-Za-z0-9.-]+\b',
        r'https?://[^ ]*\.(com|fi)\b',
        r'\b[^ ]+/[^ ]+\b',
        r'\b[^ ]*\.\s*(com|fi)\b'
    ]
    combined_pattern = r'|'.join(patterns)
    text = re.sub(combined_pattern,'',text)

    text = text.lower()

    words = text.split()
    text = " ".join(
            [contraction_dict.get
                (word, word)
                    for word in words]
            )

    text = re.sub(r'â\w*', '', text)

    text = re.sub(r'&[^;\s]*', '', text)

    text = re.sub(r'\s+', ' ', text).strip()

    text = re.sub(r'[^a-zA-Z\s]', '', text)

    return text
```

APPENDIX C
WORD CLOUD ANALYSIS

The word cloud, presented in Figure 1, was generated to visually represent the most frequently mentioned terms within the discussions. The word cloud highlights key terms such as "privacy," "security," "camera," and "surveillance," reflecting the central themes of public discourse on Suomi24. The prominence of these words underscores the focus on balancing privacy concerns with the perceived security benefits of surveillance.

Fig. 1. Word Cloud of Frequently Mentioned Terms in Discussions

APPENDIX D
DETAILED VISUALIZATIONS AND ANALYSIS OF SENTIMENT TRENDS ACROSS DIFFERENT HYPOTHESES

This appendix contains detailed visualizations and analyses that complement the main findings presented in the paper.

A. Yearly Trends of Hypotheses Sentiments

The following figures illustrate the yearly sentiment trends for positive and neutral hypotheses:

Fig. 2. Yearly Trends of Hypotheses with Positive Sentiment

B. Word Clouds of Sentiments

The following word clouds visually represent the most common words associated with each sentiment category:
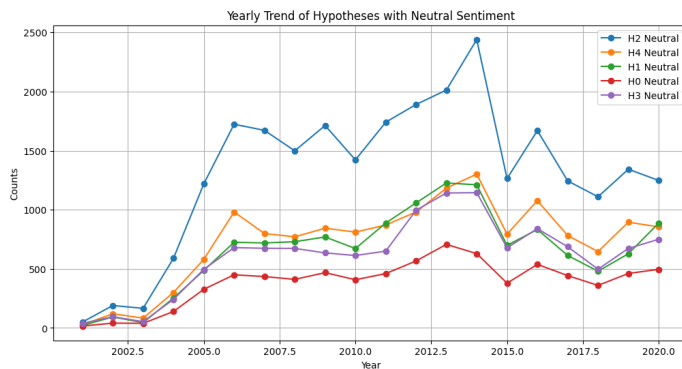
Fig. 3. Yearly Trends of Hypotheses with Neutral Sentiment

Fig. 4. Word Cloud of Positive Sentiments

Fig. 6. Word Cloud of Neutral Sentiments

Fig. 5. Word Cloud of Negative Sentiments