

# Collaborative Research: EAGER: SaTC-EDU: Just-in-Time Artificial Intelligence-Driven Cyber Abuse Education in Social Networks

PI (FIU): Bogdan Carbutar, Peter Clarke

PI (UCF): Ladislau Boloni

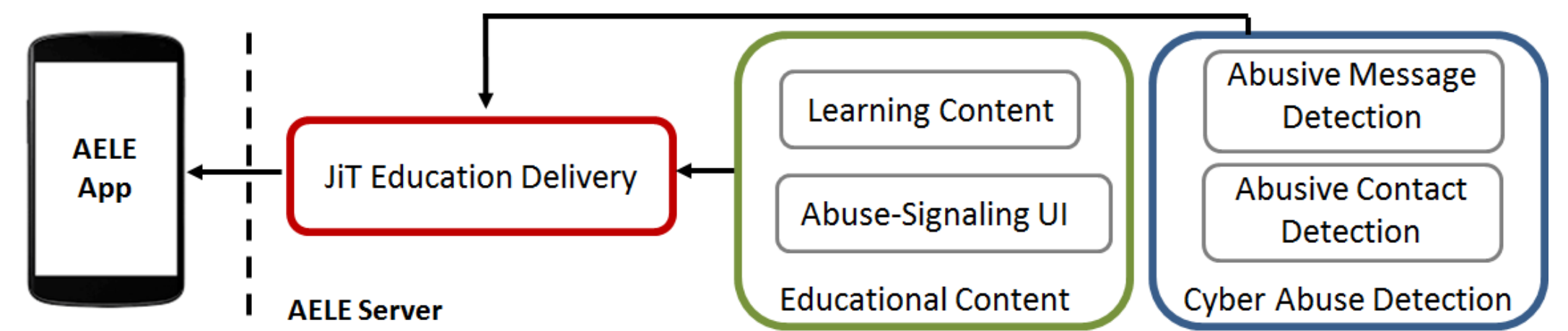


## Objective: Protect and Educate Social Network Users on Cyberabuse

### The Problem: Cyber Abuse in Social Networks

- Social networks encourage casual relations
  - Expose users to cyber abuse
  - Doxing
  - Sextorsion
  - Cyberstalking
  - Name-calling
- Only a fraction (e.g., 40%) of cyber abuse victims adopt self-protective behaviors (SPBs)

### The AELE System



### Research Questions

**RQ1:** How can interaction footprints captured in social networks identify abusive contacts and instances of cyber abuse?

**RQ2:** Can integration of educational content into social networking interactions on increase user adoption of self-protective behaviors and reducing exposure to cyber abuse?

### Scientific Impact

- Automatically detect social network abuse
- Develop learning content for social network safety
- Just-in-time education delivery

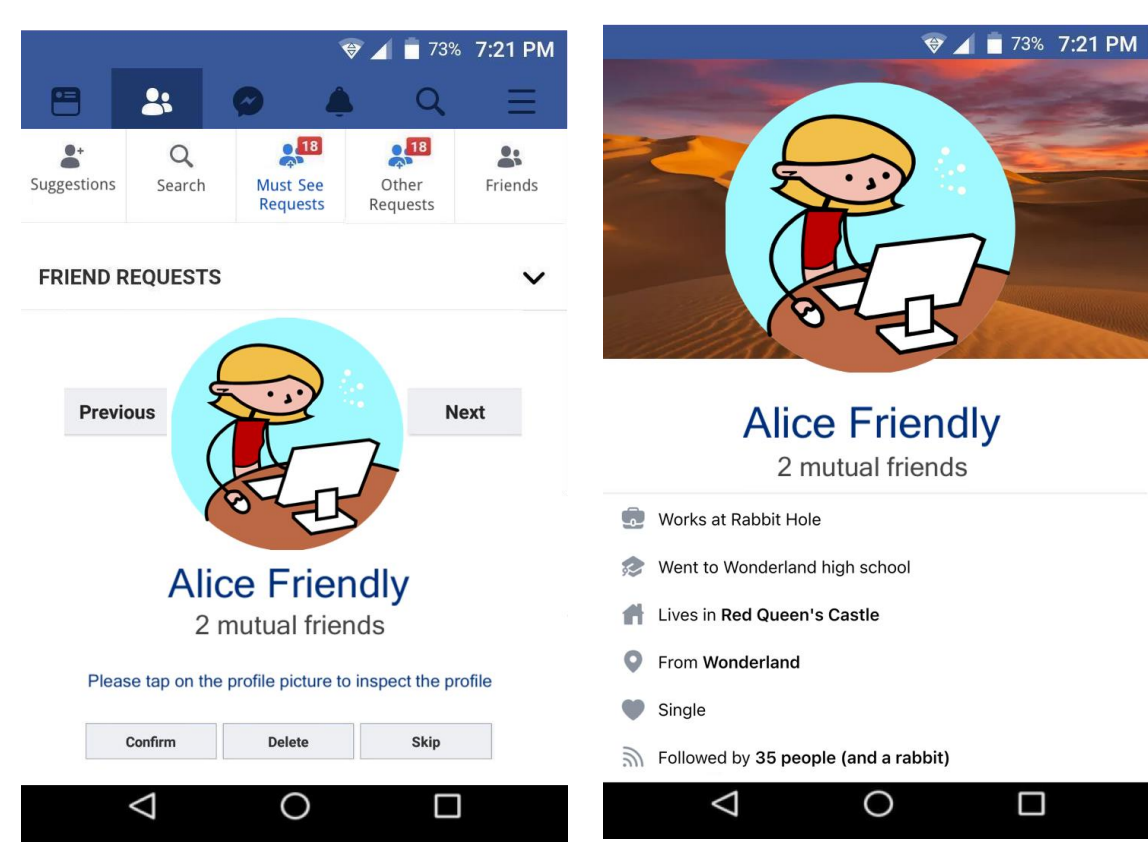
### Approach

- Qualitative study (n = 35) with Facebook users : confirm, delete or skip decisions for a mix of their real pending friends in Facebook and five synthetic friends
- Quantitative user studies (n = 145)

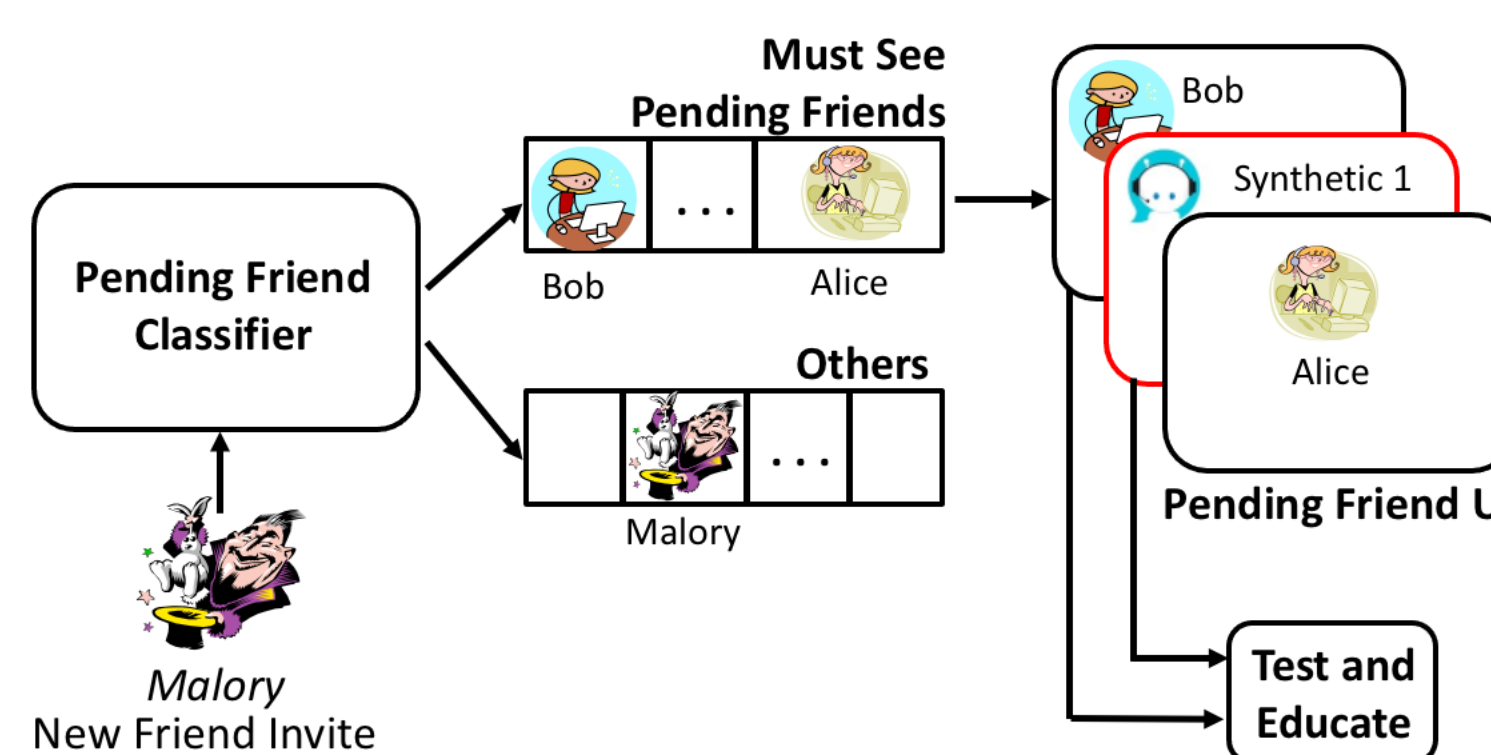
### Data collected:

- Confirm, delete or skip decisions for a mix of real pending friends and five *synthetic* friends
- Percentage of confirmed synthetic friends
- Number of inspected synthetic friend profiles

### Pending Friend UI

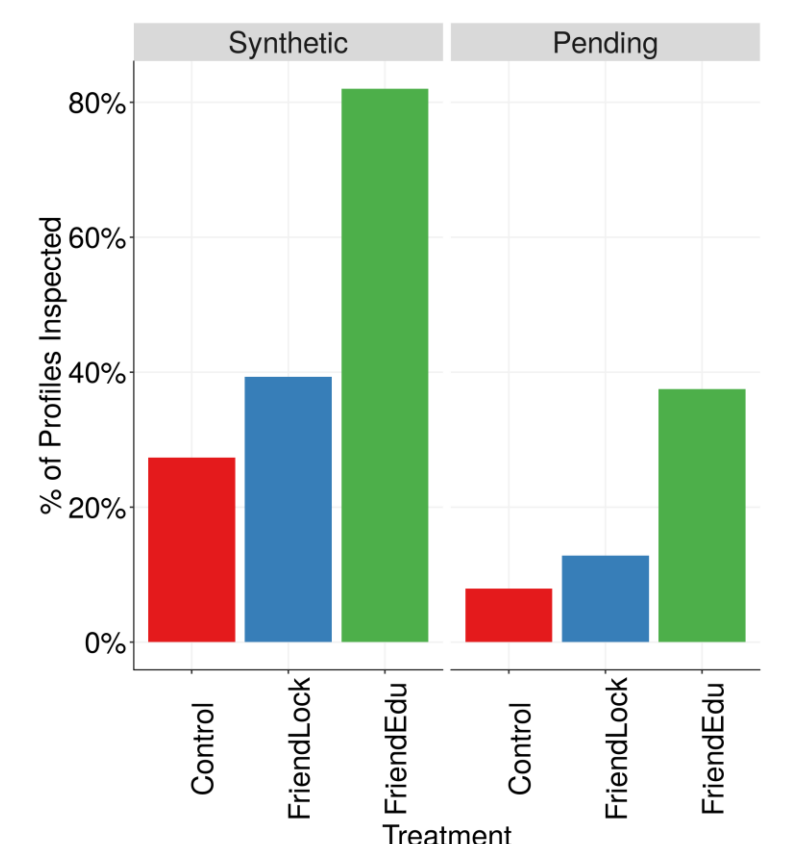


### FriendLock



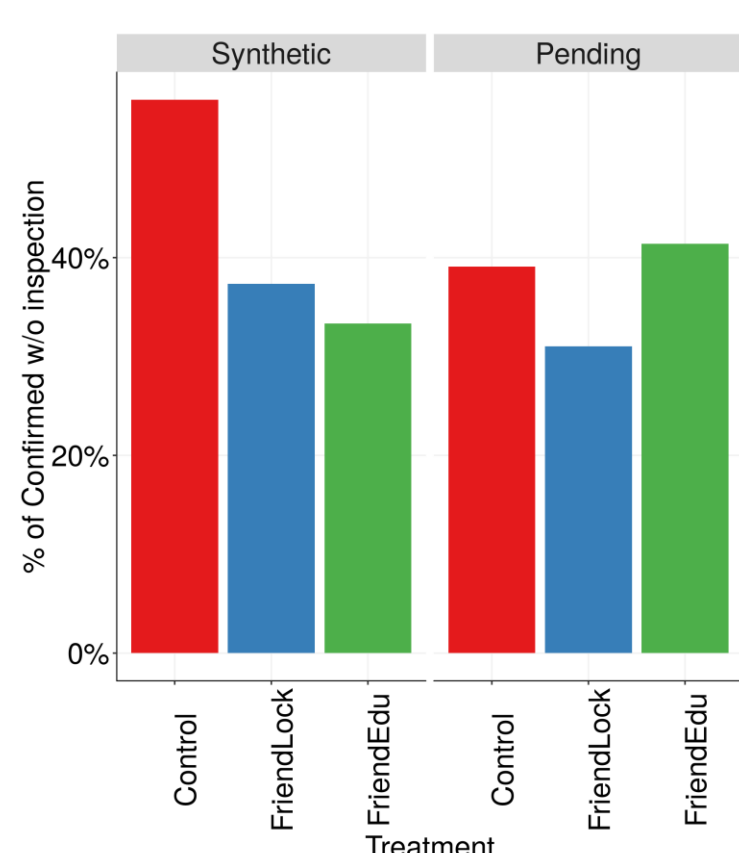
Users can be trained to avoid making automatic decisions with important security implications

### Inspected Profiles



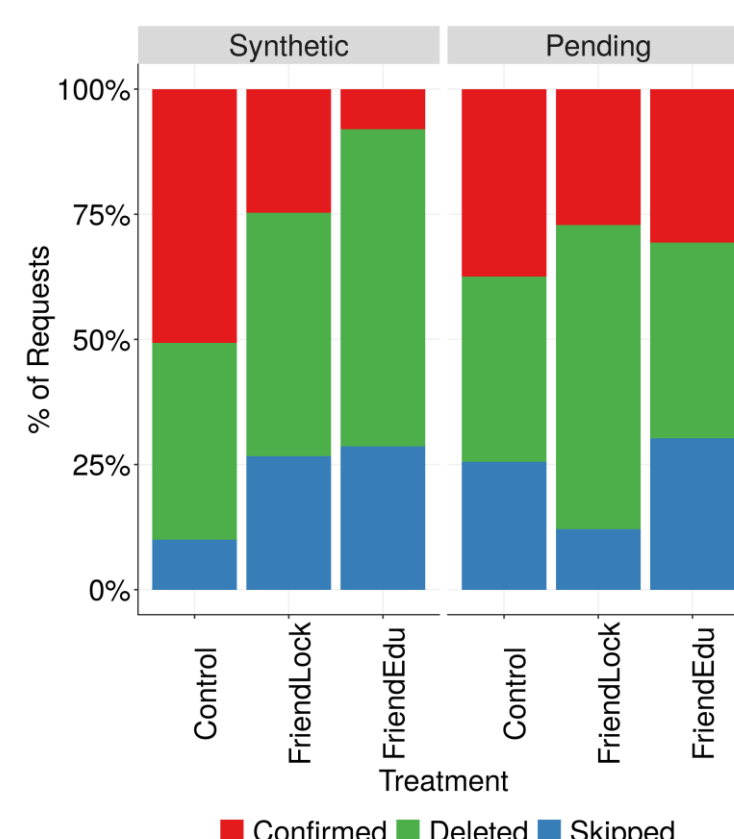
82% inspected synthetic friend profile vs. 27.33% for control

### Blind Confirms



33.3% synthetic friends blindly confirmed vs. 55.9% for control

### Blind Confirms



8% confirmed synthetic friends vs. 50.6% for control

### Pending Friend Decision Classifier

- Predict user decisions
- Features (Facebook data):
  - Gender, age, region, occupation, education, frequency of invitation reception, number of friends and pending friends, past decisions
  - Pending friend gender, age, mutual friends

Algorithm	k	Precision	Recall	F1
GBM	5	81.47%	78.75%	80.08%
	24	96.75%	96.29%	96.52%
RF	5	85.49%	82.72%	84.08%
	25	92.22%	88.88%	90.52%
NB	5	81.02%	79.67%	80.34%
	24	94.23%	92.59%	93.40%
MLR	5	70.91%	70.66%	70.79%
	24	86.17%	85.18%	85.67%
KNN	5	65.96%	59.78%	62.72%
	22	77.77%	77.77%	77.77%

