

Report of the First AMiCA Brainstorm

18 October 2010, 10-15u

Lindner Hotel & City Lounge, Antwerp

By Kim Luyckx

About the project

In the AMiCA (Automatic Monitoring for Cyberspace Applications) project, we build a tool that automatically scans messages, videos, and images in social network sites for unwanted content or behaviour by or directed to minors. Potential applications include the detection of cyber-bullying, grooming by paedophiles, sexual harassment, self-mutilation, *etc.*

AMiCA is a cooperation between University of Antwerp (CLiPS and MIOS research groups), University College Ghent (LT3), Ghent University (IBBT-INTEC), and Katholieke Universiteit Leuven (VISICS-ESAT). The one-year pre-project – funded as an IWT-SBO (Agency for Innovation by Science and Technology, Strategic Basic Research programme) project with a primary societal finality – has started on October 1, 2010. By January 2011, we prepare a full project proposal for four years of research in automatic monitoring of social network sites, resulting in a monitoring tool ready to incorporate in existing social network sites. Users already incorporated in the pre-project are:

- * Child Focus (Ellen Stassart)
- * Centrum ter Preventie van Zelfdoding (Grieke Forceville)
- * Federal Computer Crime Unit (Luc Beirens)
- * Mollom (Benjamin Schrauwen)
- * Netlog (Lien Louwagie)
- * Trendwolves (Maarten Leyts)

For legal advice, we consult Patrick Van Eecke, a lawyer specialised in Information Technology law affiliated to DLA Piper Brussels as lawyer and partner, and to the University of Antwerp as a professor. In the course of the project, he will be asked to analyse national and international legislation for privacy issues of monitoring social network data.

Goal of the AMiCA brainstorm sessions

In order to get a good understanding of the potential risks faced by minors using social network sites, we organise brainstorm sessions in the framework of AMiCA. During a brainstorm session, we bring together various types of stakeholders:

- * Government
- * Police
- * Social profit organisations
- * Technology providers and developers
- * Social network site providers
- * Social network site users and their parents

and engage in an open discussion on the valorisation potential of AMiCA and make agreements on transfer of the results of the project between the project team and the relevant stakeholders.

People or organisations that agree to become part of the AMiCA user committee – acting as an advisory board – will be asked for advice and feedback, and will be invited to future user group meetings. Joining the user committee implies commitment to the objectives and outcome of the project, but does not require financial input.

First brainstorm – 18 October 2010

The goal of the first brainstorm session was two-fold. First of all, we wanted to present the project to the various stakeholders and get an idea of their expectations towards an automatic monitoring of social network sites. Secondly, we wanted to discuss the potential applications of automatic monitoring and come to a prioritization, based on

- a. The application's urgency, as assessed by the various social profit organisations, government, and police services
- b. The technical challenges each of the applications poses
- c. The legal and privacy issues of working with social network data and automatic monitoring

The full list of applications:

(Detection of)

- * Grooming by paedophiles
- * Cyber-bullying
- * Sexual harassment
- * Explicit images or videos: nudity, self-mutilation
- * Advertisements for prostitution or escort agencies
- * Breaches of privacy, reputation, personal dignity
- * Incitement to violence, racism, xenophobia
- * Dissemination of totalitarian ideologies
- * Suicidal behaviour

Programme

- 10:00 Introduction, Walter Daelemans (UA, project leader)
Introduction of the participants
- 10:30 Valorization in IWT-SBO, Carine Lucas (IWT)
AMiCA context and objectives, Kim Luyckx (UA)
Legal context of AMiCA, Patrick Van Eecke (UA)
- 11:15 Brainstorm, moderated by Véronique Hoste (HoGent)
- 12:30 Lunch
- 14:00 Closing

Report

This report is presented as a Q&A, based on questions the AMiCA consortium wanted to see answered and observations or questions from the participants.

Q1: Will AMiCA be included in the social network sites or installed as a package on a personal computer?

AMiCA needs to be included in the SNS itself (suggested by CPZ, Mollom, and Netlog) because it allows fast follow-up and can be combined with the other tools (a.o. button for reporting abuse, reporting profiles, removing posts) included in

SNS. Tim Wauters (project partner) indicates that plugging in on the SNS' framework has good effects on the scalability of the automatic monitor. However, Annemarie Ijkema (Microsoft) indicates that the tool can be included in parental control tools, and more specifically in the 'Family Safety' package Microsoft Live offers for free download.

Q2: How will the AMiCA tool be made available and maintained?

Gezinsbond and VIP Jeugd are prepared to advertise the tool among the SNS users and their parents. Carine Lucas (IWT) stresses the importance of maintenance and support for the tool. The more people in the AMiCA community, the higher the chance of long-term success.

Q3: How will the AMiCA tool interact with human monitoring?

The AMiCA tool will be trained on posts/images/videos that have been flagged by human monitors of SNS. When the tool sends a list of new potential risks to the human monitors, the humans report back to the tool on the correct action to be taken (e.g. remove message, block user, remove user, alarm police). That way, the AMiCA tool will become increasingly better with each round. (Such a set-up is referred to as an *active learning* approach). Netlog is prepared to cooperate in such a construction. Both FCCU and Federale Politie are prepared to transfer data from actual offenses.

An important feature to incorporate is an indication of the gravity of the problem. When human monitors receive a list of potential risks from the tool, an indication of the urgency of the risk helps them to react fast to the most urgent ones. In an active learning approach, this will also help the tool deal with borderline cases.

Karolien Poels (project partner) stresses the importance of user feedback – by minors and their parents. There is a difference between the reactions of feelings of various users when facing harmful content or behaviour.

Q4: What is the legal context of a monitoring tool operating on SNS data?

Patrick Van Eecke (project partner) presented a high-level analysis of the legal issues of AMiCA. The project needs to keep the right balance between privacy and freedom of speech on the one hand, and prevention and enforcement on the other hand. He stresses that, for each use case, the exact legal concerns need to be examined (e.g. liability risks, compliance with Data Protection Act, consequences of the monitored behaviour).

Q5: Which applications are most urgent?

The participants in the brainstorm did not present some risks faced by minors on the web (or SNS in particular) as more urgent than others. Self-mutilation in video or images is easy to perceive (according to Jan De Coster and Lien Louwagie) and occurs according to regular patterns and includes similar objects, which makes it easier to train image processing systems for this application.

However, detecting suicidal behaviour in text is a challenge. Cyber-bullying can be subtle, but sometimes it can be obvious. It is essential that the AMiCA tool picks the low-hanging fruit and that extra research effort is invested in the more difficult

applications. IWT-SBO is specifically meant for high-risk technologies, so AMiCA fits in.

Q6: Are there additional applications or features (not in the initial list) that AMiCA can/should include?

- * Scams
- * Phishing sites
- * Self-mutilation
- * Gambling
- * Terrorism

Q7: What is most important: precision (detecting as many *actual* offenses or harmful acts as possible) or recall (detecting all *potentially* harmful posts/images/videos)?

Lien Louwagie (Netlog) says that good recall is important because it limits the search space for the human monitors. However, precision needs to be high as well because it takes a lot of time and money to manually go through large amounts of posts/images/videos. If precision is low, there is only limited added value of an automatic monitoring tool.

European Commissioner for the Digital Agenda (2010-2020) Neelie Kroes, has stated that by 2013, providers will be asked to further develop self-regulatory measures regarding online safety for children.¹ In short, the AMiCA tool needs to reduce the cost of human monitoring and aim at high precision, especially when aiming at increased self-regulation.

Q8: How is follow-up arranged in current SNS?

- * Suicidal behaviour: referred to specialized help in banners or by e-mail (CPZ)
- * Cyber-bullying is reported on by the SNS users themselves (Netlog)
- * Illegal content or behaviour (e.g. paedophilia) is reported to the police directly by the human monitors (Netlog)

AMiCA should suggest a specific follow-up action, depending on the gravity of the risk. A human monitor will then evaluate and correct this suggestion, and give feedback to the tool (cf. *active learning*).

¹ See IP/10/704

<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/10/704>

Participants

Laurent Bounameau	Federal Computer Crime Unit
Walter Daelemans	CLiPS, Universiteit Antwerpen; project leader
Jan De Coster	Centrum ter Preventie van Zelfdoding
Vinciane Goosse	Federale Politie, cel Mensenhandel
Véronique Hoste	LT3, Hogeschool Gent; project partner
Annemarie Ijkema	Microsoft
Lieze Lingier	Vlaams Informatiepunt Jeugd
Lien Louwagie	Netlog
Carine Lucas	IWT
Kim Luyckx	CLiPS, Universiteit Antwerpen; project coordinator
Annelies Mervielde	Gezinsbond
Claudia Peersman	CLiPS, Universiteit Antwerpen
Karolien Poels	MIOS, Universiteit Antwerpen; project partner
Benjamin Schrauwen	Mollom
Heidi Vandebosch	MIOS, Universiteit Antwerpen; project partner
Patrick Van Eecke	Fac. Rechten, Universiteit Antwerpen; legal advice
Luc Van Gool	VISICS, KULeuven; project partner
Bert Van Puyenbroeck	Kind en Gezin
Tim Wauters	IBBT-INTEC, Universiteit Gent; project partner

Apologized

Anita Cautaeers (CAW Federatie)
Filip De Turck (IBBT-INTEC, Universiteit Gent; project partner)
Grieke Forceville (Centrum ter Preventie van Zelfdoding)
Kathleen Goovaerts (Tele-Onthaal)
Maarten Leyts (Trendwolves)
Ellen Stassart (Child Focus)