Integration of social media

Can social media be used in counterterrorism training?

Much of real world events around the world can be followed on various social media, as individuals share their experiences and thoughts.

SAAB Training & Simulation wishes to use this to construct better simulators for training military and emergency services. The first step would be to integrate social media to a database, from which the data can be analyzed and utilized.
But what information can be found on social media and how is it accessed?

Here are a few social media and a summary of what information can be found on each of them:

Facebook
A social network that might be considered to be the most well known social media.

- User information, habits and interests
- News and opinions

Instagram
A content sharing community focused around sharing photos and short videos.

- Photos and videos
- Very little text based data

Twitter
A microblog allowing users to post 140 character messages.

- News and opinions
- Streaming APIs available for live coverage

Google+
A social network with a slightly more focus around the user's interests.

- User information, habits and interests
- News and opinions

Common API features
These social media have some common features when it comes to their APIs:

- REST APIs exists and are used most requests
- Responses return objects as JSON-strings
- Successful requests returns whole objects, not just sought after information.
- Enterprise APIs are available for expanded functionalities

These similarities will make it easier to integrate several social media to the same database. It is however necessary to sort out unnecessary information before storing entries in the database.

Challenges
There are of course some challenges when data mining social media. For instance:

1. accessibility: can the desired data be reached?
2. Will data collection be in violation of any laws or policies?

These are just some of the considerations that must be made. Different implementations will have different answers to these questions, it is therefore important to analyze the needs of an application in development. One implementation that can do everything may be tempting but will most likely waste not only time spent in development, but also storage space.

Proof of concept
To prove that is can be done, a simple implementation was constructed that reads Tweets from Twitter's Sample Stream, and saves them in a database.

The stream returns JSON-strings which are parsed into .Net objects, which are then used as templates when creating entries in the database.

The Connection class retrieves the strings, converts them to object using JSONparser. The resulting objects are sent to WISEhandler, which handles interactions with the database.

Author: Michael Johansson