Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

Police investigations in Internet open sources: Procedural-law issues

Bert-Jaap Koops

TILT – Tilburg Institute for Law, Technology, and Society, Tilburg University, The Netherlands

Keywords:
OSINT
Criminal investigation
Open sources
Intelligence-led policing
Social networking
Cybercrime Convention
Dutch law
Digital investigation

ABSTRACT

Analysing large amounts of data goes to the heart of the challenges confronting intelligence and law enforcement professionals today. Increasingly, this involves Internet data that are 'open source' or 'publicly available'. Projects such as the European FP7 VIRTUOSO are developing platforms for open-source intelligence by law enforcement and public security, which open up opportunities for large-scale, automated data gathering and analysis. However, the mere fact that data are publicly available does not imply an absence of restrictions to researching them. This paper investigates one area of legal constraints, namely criminal-procedure law in relation to open-source data gathering by the police. What is the legal basis for this activity? And under what conditions can domestic and foreign open sources be investigated?

These questions are addressed from the perspectives of European and Dutch law. First, the international legal context for gathering data from openly accessible and semi-open sources is analysed, including the issue of cross-border gathering of data. In particular, article 32 of the Cybercrime Convention and some national implementations are discussed, as well as data protection requirements from European Union law. Next, the paper zooms in on the Dutch legal context for open-source investigations, to illustrate how the issues of a legal basis and other legal requirements are addressed in a specific legal framework.

The paper draws the conclusion that technology-facilitated investigations of open sources by the police often constitute an interference with the right to privacy; hence, they require a legal, statutory basis that is sufficiently clear for citizens to understand what the police are doing. Moreover, open-source investigation tools and practices used must meet general data-protection requirements and forensic reliability standards. The discussion also shows that interpreting existing legal provisions to accommodate open-source investigation tools can lead to convoluted interpretations, suggesting that legal frameworks of investigation powers with a focus on physical-space investigations may need to be revised to accommodate the particularities of open-source Internet investigations.

1. Introduction

Data mining large amounts of data from a wide variety of sources has always been a key method of intelligence services. Since a few decades, also law-enforcement agencies have started to look at data mining as a means of acquiring information for investigating crimes. Particularly in the last decade, 'intelligence-led policing' has arisen as a new

0267-3649/$ – see front matter
http://dx.doi.org/10.1016/j.clsr.2013.09.004

approach to law enforcement, with a significant stress on prevention and pre-emption alongside, or even in place of, repression of crime.[1]

Law-enforcement agencies are primarily responsible for crime-fighting. This implies, traditionally, a specific approach, based on concrete suspicions of specific individuals or groups being involved in a crime, and a repressive strategy, aiming to gather evidence that can be used to prosecute and convict criminals for crimes committed. Increasingly, however, in the paradigm of intelligence-led policing, LEAs also focus on data gathering in more preventative approaches, preferably trying to intervene before crimes are committed. Such approaches typically fall under the banner of intelligence-led policing and often require a legal mandate to enable lawful data collection, except where data collection does not interfere with citizen rights such as the right to privacy.

With the explosive growth in data that are generated and made available through the Internet, not the least in web 2.0 applications such as blogs, Twitter, and social-networking sites, data in open sources have become an attractive target of the police, not only in intelligence-led policing but also in classic crime-fighting. The fact that data are available in open sources would at first sight seem to suggest that the police can collect and process these data without restraint; after all, anyone can access and use them, so why not the police? Upon closer scrutiny, however, open-source data investigations by the police are not as straightforward as they seem.

People have a certain legitimate expectation of privacy even 'in public' – they do not expect to be completely exposed to the public's eye when they move around in publicly accessible places.[2] Analogously, when people upload something somewhere on the Internet, even though they (ought to) realise it is publicly available, they do not necessarily expect the entire world to actually view it. There may be a certain legitimate expectation of privacy, then, even on the Internet.[3] This also applies to police investigations: Internet users may not generally expect the police to scrutinise everything that roams on the Internet, particularly not if the investigations rely on sophisticated tools for mining open sources.[4] Increasingly, police are using systems which facilitate automated selection, analysis, combination, and visualisation of search results. For example, the European VIRTUOSO project has developed:

> *a technical framework for the integration of tools for collection, processing, analysis and communication of open source information. This middleware framework will enable "plug and play" functionalities that improve the ability of border control, security and law enforcement professionals to use data from across the source / format spectrum in support of the decision making process.*[5]

Similar platforms and tools are being built for use by national police forces. For example, in the Netherlands, the iColumbo system is being developed, an 'intelligent, automated, "near" real-time Internet monitoring service' for government investigations.[6] Platforms such as the VIRTUOSO middleware and infrastructures such as iColumbo can be equipped with all kinds of plug-ins that enhance the search and analysis capacities of Internet searches, for example through entity recognition, image-to-text conversion, and automated translation.

In this paper, I will analyse what are the legal basis and conditions for the police to use such systems for automated investigation of open Internet sources, or OSINT[7] systems for short. Since criminal law is to a large extent still a matter of national law and European harmonisation, although greatly extended over the past decades, is still relatively small, the answer to this question will depend on the particularities of the national legal system. In order to provide a somewhat generic as well as sufficiently detailed answer, I will use a two-prong approach. First, I will discuss the issue in more general terms on the basis of European legal instruments regulating the collection, processing,[8] and use as evidence of open-source data (Section 2). Second, a more in-depth discussion will follow of the legal conditions for open-source investigations by focussing on a national legal system. I will discuss Dutch law (Section 3), since the development of iColumbo has triggered academic debate about Dutch criminal-procedural law[9] which demonstrates the complexities of the legal basis and conditions for open-source police investigations.

2. International legal context: European law

Although there are hardly any specific provisions in European law regulating open-source police investigations, various legal instruments touch upon what the police can do to collect,

[1] BE Harcourt, *Against prediction: profiling, policing, and punishing in an actuarial age* (University of Chicago Press, Chicago 2007); BJ Koops, 'Technology and the Crime Society: Rethinking Legal Protection' [2009] 1 Law, Innovation & Technology 93.

[2] Cf ECtHR 24 June 2004, *Von Hannover* v. *Germany*, App.no. 59320/00, §77; ECtHR 12 January 2010, *Gillan and Quinton* v. *The United Kingdom*, App.no. 4158/05, §61.

[3] Although the 'reasonable expectation of privacy' doctrine is not explicitly part of European privacy law, the reasoning associated with it is visible in many European Court of Human Rights cases, and the formulation sometimes explicitly refers to the doctrine. See, e.g., ECtHR 25 June 1997, *Halford* v. *The United Kingdom*, App.no. 20605/92, §45 and ECtHR 12 January 2010, *Gillan and Quinton* v. *The United Kingdom*, App.no. 4158/05, §61.

[4] Micheal O'Floinn and David Ormerod, 'Social networking sites, RIPA and criminal investigations' [2011] 10 Criminal Law Review 766, 775–777, 789.

[5] BJ Koops, CMKC Cuijpers and MHM Schellekens, *D3.2. Analysis of the Legal and Ethical Framework in Open Source Intelligence* (2011).

[6] 'Deelprojectvoorstel, Ontwikkeling Real Time Analyse Framework voor het iRN Open Internet Monitor Network', 'iColumbo'', available at http://www.nctv.nl/Images/deel-projectvoorstel-ontwikkeling-icolumbo-alternatief_tcm126-444133.pdf.

[7] OSINT stands for open-source intelligence.

[8] In data protection law, the term 'processing' is very broad, comprising also collection and use of data. For analytic purposes it is useful to distinguish between the stages of collecting, further processing, and using as evidence of open-source data, and I use the term 'processing' in this article to refer to the second stage.

[9] JJ Oerlemans and BJ Koops, 'Surveilleren en opsporen in een internetomgeving' [2012] 38 Justitiële verkenningen 35; see also BJ Koops, 'Politieonderzoek in open bronnen op internet. Strafvorderlijke aspecten' [2012] 11 Tijdschrift voor veiligheid 30.