International Journal of Population Data Science

Journal Website: www.ijpds.org

Making the most of cybercrime and fraud crime report data: a case study of UK Action Fraud

Sara Giro Correia[1,*]

Submission History

Submitted:	25/11/2021
Accepted:	08/02/2022
Published:	11/05/2022

[1]Swansea University, Swansea, SA2 8PP, UK

Abstract

Introduction
Researchers and public authorities are increasingly exploring the potential of administrative data to generate new insights. This includes recent work leveraging the opportunities of the crime report data collected by the UK's national reporting centre Action Fraud (AF). However, the quality of these data and its implications for data users have not been systematically analysed.

Objectives
This paper outlines challenges and opportunities of using AF data in cybercrime and fraud victimisation research and practice and makes recommendations to improve the quality of this dataset.

Methods
The author has undertaken two studies using samples of AF data pertaining to crime reports within the Welsh police forces, between 2014 and 2020. Quality diagnostic checks, reflections and methodological decisions were considered across each study. These were reviewed, key themes were identified and discussed with data users and a broader group of researchers to finalise the recommendations presented.

Results
The strengths and limitations of AF data are discussed and grouped into themes, closely aligned with four quality dimensions widely used by statistical authorities. This includes an assessment of 1) the impact of under-reporting and 2) the purpose and rules of crime recording, on the *relevance* of the data to its users; 3) the *accuracy and reliability* of the data; 4) the consistency of recording and its impact on *coherence and comparability*; and 5) the *accessibility and timeliness* of the data.

Conclusions
Recommendations are made to improve AF data to generate better quality insights across the dimensions of *relevance*, *accuracy & reliability*, *coherence & comparability* and the *accessibility & timeliness* of this dataset. Additionally, a data catalogue would enable frontline officers and researchers to make the most of this dataset, harnessing it to produce key insights for crime prevention, investigation, and victim support.

Keywords
crime data; fraud; cybercrime; police recorded crime; data quality

*Corresponding Author:
Email Address: s.correia@swansea.ac.uk (Sara Giro Correia)

https://doi.org/10.23889/ijpds.v7i1.1721

Introduction

Researchers and public authorities are increasingly exploring the potential of administrative data to generate insights and inform policy and practice. At the same time, fraud and computer misuse (e.g., hacking, or ransomware attacks) constitute a large proportion of crimes experienced by victims, with some evidence suggesting this was aggravated during the COVID pandemic. Based on the Telephone Crime Survey for England and Wales (TCSEW), it is estimated that there were 4.6 million incidents of fraud and 1.7 million incidents of computer misuse (CM) in the year ending March 2021, adding up to 53% of the total UK crime estimate [1]. Furthermore, compared to the pre-pandemic period, while the number of incidents of *other* crime decreased by 19%, fraud and computer misuse (F&CM) increased by 24% and 85% respectively [1].[1] In addition, there was a 28% increase in reports of fraud and a 16% increase in CM reported via the UK's national reporting centre Action Fraud (AF), in the year ending March 2021 [1]. This aligns with work by Kemp and colleagues suggesting increased reporting of certain cybercrime and fraud categories during the COVID pandemic [2]. On the other hand, a US-based survey study found no major change in cyber victimisation, suggesting the greater change in this period may concern reporting behaviour [3]. This is significant as one of the challenges of both measuring and responding to F&CM crimes, remains their relative under-reporting [4].

Nonetheless, in contrast to the localised reporting of other crime types, AF collects reports from across the UK.[2] As such, AF data is a key resource, used by police crime analysts to generate national, regional, and force-level trend analysis, to undertake threat assessments, inform crime prevention campaigns and other activities, as well as being a key basis for investigative and local victim-support responses. Additionally, researchers have leveraged AF data, to better understand reported trends, as well as their implications for policy and practice [2, 7–10]. In particular, this data has an enormous potential for researchers as it provides granular detail on each report and, as the author's previous work demonstrates, allows for the analysis of small geographies and repeat victimisation [8, 11]. However, before conclusions are drawn on substantive matters, it is key to develop a thorough understanding of how data sources were collected, evaluate their quality in relation to wider research aims and prepare them for statistical analysis. This requires the development of 'detectors' and 'metrics' to detect quality issues and 'auditing data sources for quality' [12]. However, the quality challenges, and opportunities associated with using AF data, have not been systematically analysed. As such, this paper addresses the strengths and weaknesses of AF data in facilitating a victim-focused response, particularly with respect to the less developed 'Protect' strand of policing, which aims to increase protection for those who are at risk of (further) victimisation. This is especially timely in the UK, given the government's recent commitments to deliver "an improved national fraud and cybercrime reporting system", and alongside it, to "expand the National Economic Crime Victim Care Unit [NECVCU]" [13].[3]

Building on previous work [4, 15], and work from related fields [16, e.g.], this paper audits the quality of AF data with respect to two samples collected between 2014 and 2020 and examines how its quality impacts on analytical outputs, across four quality dimensions [17]. The first of these is *relevance*, or the extent to which the insights produced meet the needs of users, including law enforcement and researchers. The limitations of using police recorded crime (PRC) as a source of statistical crime data are well documented [18–20] and in line with previous scholarship, this paper firstly considers the quality implications of under-reporting, the original purpose of data collection and the rules which shape the collection of AF data. Secondly, the paper examines whether these data produce *accurate and reliable* results. In other words, the extent to which the data collected by AF accurately portrays the reality of the crime/victimisation experience it is intended to record and does so reliably over time. Thirdly, it turns to the impact of data quality on *coherence and comparability*, i.e., whether it produces insights which are internally consistent, consistent over time and comparable between regions and police force areas. Finally, the paper turns to the *timeliness and clarity* of data insights generated from the AF dataset. Recommendations are made to improve the quality of the dataset across each of these quality dimensions.

Methods

The author undertook two studies using samples of AF data, pertaining to crime reports within Wales, relating to recording periods between 2014 and 2020 (Table 1). From the perspective of evaluating the quality of AF data, the two periods are significant as the data in the second study were collected after several improvements were made to the AF recording system. As such, improvements in data quality were expected between the two studies.

Firstly, each of the studies were summarised and the results of quality diagnostic checks, reflections on limitations and methodological decisions around mitigations were collected.

[1] The methodology of the Crime Survey for England and Wales (CSEW) was changed to account for COVID restrictions, and therefore the results of the replacement Telephone Crime Survey for England and Wales (TCSEW) are not directly comparable to previous years. As such, these are *best estimates* of the percentage change in number of incidents experienced by victims, between May 2020 to March 2021 TCSEW and year ending March 2019 CSEW, calculated by the Office for National Statistics, using comparable sub-sets of data [1].

[2] As the National Lead Force for Fraud, City of London Police operate Action Fraud (AF), the only nationally run crime recording system, as well as the National Fraud Intelligence Bureau (NFIB). While AF collects data for the whole of the UK, City of London lead strategy primarily in England, Wales, and Northern Ireland (NI). Where frauds and cybercrimes are experienced by victims or committed in Scotland, they are usually reported and investigated by Police Scotland, although multiple authorities can have concurrent jurisdiction [5]. Furthermore, AF is not the only source of data on fraud in England, Wales, and NI. Other sources of police recorded crime on fraud include reports from the industry bodies UK Finance and Cifas. In addition, fraud is recorded by other bodies including Trading Standards and the Food Standards Agency. Finally, AF primarily records reports from individuals e.g., 91% of AF records between 1 Jan 2020 and 31 Dec 2020 came from individuals, according to the City of London's online dashboard. While no comparable analysis was possible for Study 1, this figure is consistent with the data collected for Study 2 [6].

[3] While the City of London police has clarified that AF will not be replaced [14], the service is currently being re-commissioned and, based on the author's discussions with stakeholders, there is an interest in ensuring that data quality issues are addressed.

Table 1: Summary of studies

	N cases	N forces	Time-period	Ethics approval	Methodology
Study 1	17,049	4	01/10/2014 to 30/09/2016	03/11/2016	Mixed: Linkage, statistical bivariate analysis, modelling and qualitative thematic analysis
Study 2	11,934	3	01/02/2019 to 30/06/2020	24/09/2020	Descriptive and bivariate analysis

Secondly, the researcher identified the key quality themes across these studies. Thirdly, the early results were shared with practitioners and researchers, who were invited to comment on the findings and recommendations. In addition, the results were shared with representatives from the data provider. This allowed for cross-validation of the findings, as well as of the feasibility of the recommendations made. Finally, data from victimisation surveys were used as comparators where relevant to the analysis. Each of the studies will now be described in turn.

Study 1: Vulnerability & repeat victimisation in wales

This study drew on a sample of crime reports (n = 17,049), made to AF by victims based within the four police forces in Wales, between 1st October 2014, and 30*th* September 2016. A mixed-methods approach was used, encompassing descriptive and bivariate statistics, generalised linear models, deterministic and probabilistic data linkage, as well as qualitative thematic analysis. Key results generated by this study included the unsustainability of an online/offline distinction, patterns of repeat victimisation and an original framework for understanding vulnerability in the context of F&CM victimisation, to better target a victim response [8].

Study 2: COVID and the impact of fraud & computer misuse in wales exploratory study

This brief study focused on analysing F&CM reporting patterns and victim impact, between 1st February 1st and 30th June 2020 (n = 11,934). The analysis was carried out on site, to generate exploratory analytical outputs for the Regional Organised Crime Unit and identify emerging trends and areas for future research, during the COVID pandemic. Descriptive statistics were produced, including those included in this paper.

Results and discussion

Under-reporting

Instead of capturing all crime experienced by victims, PRC captures only those crimes which are both reported to and recorded by the police. The data is therefore limited by under-reporting and shaped by the rules and purpose of recording [18–20]. Under-reporting has a considerable impact on what can be known from recorded crime as F&CM are comparatively under-reported. As shown in Table 2, recent Crime Survey for England and Wales (CSEW) data indicates that at best, 2% of computer crimes and 8% of fraud experienced by individuals were reported to the police via AF in the year ending September 2019 [21]. In comparison, approximately 53% of all theft was reported to the police in the same period. While not directly comparable due to COVID19-related changes in methodology, the Telephone-operated CSEW (TCSEW) suggests a similar ratio with that 9% of fraud and 2% of CM reported in the year ending March 2021 [1].[4] In parallel, under-reporting is also a known issue with respect to corporate victims [22, 23].

Inevitably, the extent of under-reporting demonstrated above, has an impact on the quality of statistical and operational outputs produced using AF data, with respect to the *relevance* dimension of quality i.e., the extent to which the insights produced meet the needs of users [28]. Whether the users of AF data are crime analysts within law enforcement, officers, or researchers in or beyond academia, they must consider whether the questions they have, can be answered using data which relates only incidents reported and given a crime label. As such, questions about the overall extent of victimisation in society are often best answered by victimisation surveys. However, there are also known limitations when using victimisation surveys to understand crime at low geographies, or to measure repeat victimisation [29, 30]. As far back as 2006, the Fraud Review identified the potential for "data matching" being used to identify repeat offenders, prevent repeat offences and address "vulnerability", particularly within the public sector, although it fell short of identifying the need to identify repeat victims [31]. In a more recent report however, Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS, formerly HMIC) identified "the ineffectual use of intelligence products (such as monthly victim lists) given to forces by the National Fraud Intelligence Bureau" [32].

At the same time, it is well established that the 'seriousness' of the crime [33, 34] and police/response perceptions [33, 35, 36] are key factors in reporting behaviour. In fact, the relative lack of seriousness of the crime and/or a cost-benefit rationale were given as a prominent reason for not reporting to Action Fraud by 8% of CSEW experimental statistics respondents in 2017 [37]. As such, one would expect to be able to identify within Action Fraud data significant harms and impacts on the victims who did report, which should enable the selection of key factors in adequately prioritising and responding to victims' needs. However, the most common reasons were never having heard of AF (66%) and assuming fraud was reported by another authority (10%) [37, Table 2].[5]

[4] The calculations by the author are based on the number of offences recorded by AF and referred to NFIB, as proportion of the total number of crimes estimated in the [T]CSEW.

[5] That said, interpreting these results is somewhat complicated given that even if a victim is not aware of AF, they may nonetheless call a

Table 2: Number of ONS estimated crimes, AF recorded crimes and % of recorded crimes for years ending September [21, 24–27]

Crime Type	N Estimated	N Recorded	% Recorded
Fraud			
2015	NA*	219,536	NA
2016	3,617,000**	219,406	6.07
2017	3,236,000	250,821	7.75
2018	3,473,000	271,486	7.82
2019	3,797,000	310,675	8.18
CM			
2015	NA*	14,992	NA
2016	1,966,000**	13,424	0.68
2017	1,505,000	21,516	1.43
2018	1,004,000	24,063	2.40
2019	1,006,000	21,471	2.13
Theft			
2015	3,906,000	1,754,614	44.92
2016	3,552,000	1,784,598	50,24
2017	3,396,000	1,987,091	58.51
2018	3,574,000	1,998,876	55.93
2019	3,578,000	1,880,780***	52.57

*F&CM were not included in CSEW estimates.
**The 2016 estimates were published as experimental statistics. In 2017 these were published as Official Statistics, but since March 2018 received accreditation as 'National Statistics'.
***Excludes Greater Manchester Police as following the implementation of a new IT system in July 2019, the force was unable to supply ONS with data for the quarter July to September 2019.

As such, awareness of AF remains a key issue to be addressed. Nonetheless, subject to a good understanding of the recording rules and quality issues, AF provides a valuable source of data for analysts and researchers on higher-impact and (repeat) victimisation at local level.

Purpose & rules of recording

Previous work has highlighted that the purpose and the rules that shape administrative data collection, and crime recording in particular, have profound quality implications [18, 38, 39]. The purpose and rules of crime recording are set out in the National Crime Recording Standard (NCRS) and the Home Office Counting Rules (HOCR). The NCRS was originally developed by the Association of Chief Police Officers (ACPO) and rolled out from 2002, after a review of crime recording practices by Her Majesty's Inspectorate of Constabulary (HMIC) found large variations in recording rates across forces (between 55 and 82%) [40]. It has the twin aims of promoting consistency of recording across all forces and to take a victim-focussed approach to crime recording. The latter means that recording is based on the victim's account of a crime occurring, rather than the police satisfying themselves that a crime had indeed taken place. While police should keep auditable incident reports for all reports, a crime will be recorded if: "(a) the circumstances of the victim's report amount to a crime defined by law (the police will determine this, based on their knowledge of the law and counting rules); and (b) there is no credible evidence to the contrary immediately available" [41, para. 2.2]. Alongside the NCRS, the HOCR stipulate what type and how many offences should be recorded by police in specific circumstances. In addition, the 'vision' for crime recording across the NCRS and the HOCR is to achieve "the best crime recording system in the world: one that is consistently applied; delivers accurate statistics that are trusted by the public and puts the needs of victims at its core." This is followed by a breakdown of purposes which include enabling crime investigation (the 'Pursue' strand of policing), but also meeting victims' needs (more aligned with the 'Protect' strand). Pursue is focused on investigating, prosecuting, and disrupting fraud and CM, while Protect aims to protect victims and against fraud and CM, including those at risk of repeat victimisation. However, the accurate and consistent 'counting' of crime is not always compatible with meeting victims' needs, a key tension which, it is argued, exists within the NCRS/HOCR. Firstly, this is demonstrated by examining the data that is collected by AF, shared with local forces and the outcomes which are then returned on crime reports. In addition, this can also be seen through the application of the principal crime rule, as well as the "one crime per victim" and the "no victim – no crime" principles to F&CM recording.

Data collected, shared and returned

As shown in Figure 1, not all incidents reported to the police via AF are recorded as crimes. As with all crime types, when a report is made to AF, either via the contact

local police station, at which point they would be referred to AF as the national reporting centre or an officer will record a crime with AF on the victims' behalf. As such, never having heard of AF per se, should not represent a barrier to reporting F&CM to the police.

centre or the online form, it is possible that the situation does not meet the threshold of a crime and is therefore not given a crime label/number i.e., *crimed*.[6] This is unsurprising as the police deal with a significant number of non-crime incidents, as high as "83% of all Command and Control calls" [42]. AF data is then added to the National Fraud Intelligence Bureau (NFIB) 'Known Fraud' database. At this point, there is a distinction between what may be described as the 'Pursue' and 'Protect' strands of the police response. An algorithm scores cases according to whether they have sufficient leads for investigation, for further manual review by NFIB Crime Reviewers, who assemble case bundles which are then 'disseminated' to local forces for further 'Pursue' enforcement action. The extent to which the data disseminated for 'Pursue' activity meets users' needs, is an area for further research.

Protect, on the other hand has, to date, been the responsibility of local forces and as such, all reports made within their respective jurisdictions are separately sent to local forces on a weekly basis, for the purposes of identifying and responding to victims' needs. One exception to this has been the previously mentioned NECVCU, based at City of London Police, which currently services London, the West Midlands and the Greater Manchester areas, but hopes to expand nationally. However, no national-level data is available to paint a picture of the extent of the victim support provided in relation to F&CM. Furthermore, an examination of the sub-selection of AF data which is sent to local police forces for 'Protect' action, as sampled in the author's studies, demonstrates that the recording system is oriented towards traditional 'Pursue', rather than the victim focused 'Protect' strand of activity.

At the time of Study 1, several variables of relevance to 'Protect' activity were collected by AF, but not shared with local forces (Table 3). These variables included type of victim (e.g., individual or organisation), victims' gender, vulnerability indicators (whether vulnerable to financial loss, regularly targeted and/or repeatedly victimised) and a variable on the impact of the crime on the victim's finances, health, and wellbeing. Finally, information on whether the offender was known or unknown to the victim was not collected. By the time of Study 2 however, most of these issues had been addressed with several new 'victim impact assessment' variables collected and shared with local forces, including vulnerability and impact indicators. In addition, a series of questions were asked to determine the victim's guardianship attitudes and awareness of risk, resulting in a final Public Interventions Model (PIM) assessment score.[7] Despite these improvements however, more is needed to ensure that the data collected is optimised to enable local forces to identify vulnerable and repeat victims. This includes ensuring that all variables relevant to a vulnerability and repeat victimisation analysis are shared, but also that the data collected accurately represents what the aspects they were intended to measure, that analysts and researchers understand what vulnerability and impact scores mean and how they should be used. As such, the need for question testing, data validation procedures and the development of a data catalogue are discussed below.

Furthermore, the samples in both studies demonstrate that the outcome data that is systematically recorded by or returned to NFIB by local forces relates primarily to 'Pursue' type activity. In Study 1, only 1.25% of all cases returned an outcome of 'Prevention' or 'Victim Care' [4]. The outcome data in Study 2 was not directly comparable, but it showed that only 0.03% of cases returned 'Protect' or 'Victim Care' outcomes. However, these statistics do not reflect the victim response across the Welsh forces. In Dyfed/Powys for example, every F&CM victim is contacted by their local force to offer support. Rather, these figures illustrate how AF data are not adequate to provide a picture of Protect activity.

The principal crime rule

The principle crime rule is common across many countries and it states that in cases where there is a "sequence of crimes in an incident, or a complex crime, [which] contains more than one type of crime", then "the most serious crime" should be counted [41, Section F]. Generally, violent crimes take precedence over property crimes including F&CM. Furthermore, where a series of property crimes are reported together, the most serious crime is recorded. This is the crime carrying the maximum sentence on conviction or, where the maximum sentences are equivalent, the greatest sentence most likely to be prescribed on conviction. To support the decision of what constitutes the principle, specific guidance is provided on the counting rules for F&CM [45]. Given the relatively higher severity of fraud over most CM offences, this principle tends to favour the recording of fraud over computer misuse, when both are present. While this is aligned with HOCR, quantitative analysis or AF data will therefore under-estimate levels of CM reporting. To correct for the impact of the principal crime rule, data on whether fraud was enabled by a CM offence could be provided to local forces.

One crime per victim

The one crime per victim principle helps to establish who the victim(s) of the crime are and, in order to minimise double counting, establishes that only one crime will be recorded for each specific, indented or identifiable victim. In some circumstances however, minimising double-counting will result in the data collected not being optimised to identify and respond to the needs of victims. This is aggravated in the case of F&CM as there can be multiple victims and, to avoid double-counting, in some situations an individual's report will be recorded as a 'Crime Related Incident' (CRI) rather than a crime, or not at all. In those circumstances, it is unclear whether victim vulnerability is assessed, or their support needs considered.

A victim is defined as "the subject against whom the crime was committed"; for offences against the person this is the specific intended victim (SIV), whereas for property crime this is "the person who had custody/control or proprietary rights in the property at the time the crime was committed" [46].

[6] For an incident to be *crimed* it must be a 'notifiable offence', listed in the Notifiable Offence List (NOL) contained in the HOCR. Notifiable offences include all offences that could possibly be tried by jury plus a few additional closely related summary offences dealt with by magistrates and are listed in the Notifiable Offence List (NOL) contained in the HOCR.

[7] This table was put together based on the data inspected by the author and discussions with stakeholders.

Figure 1: The Victim Journey in Numbers, England and Wales, year ending March 2015 [43, 44]

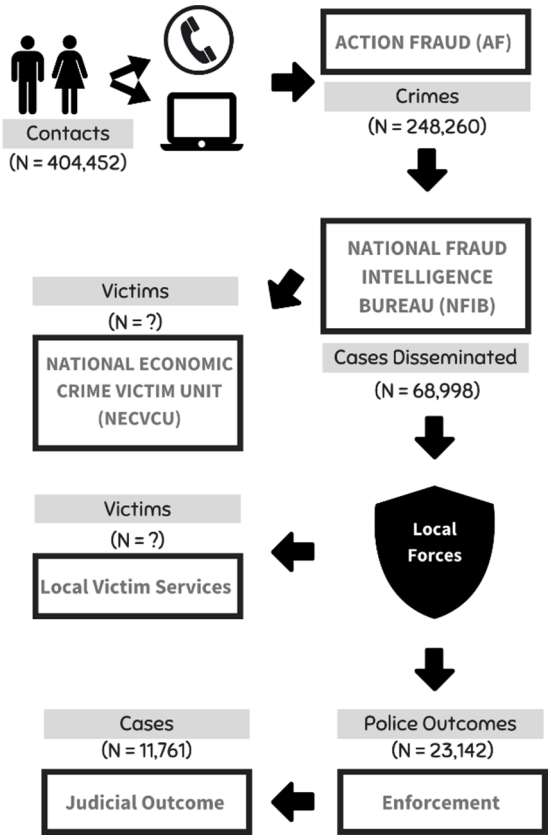


Table 3: Demographic, vulnerability and impact variables collected by AF and shared with local forces

Variable group	**Variable**	**Study 1**	**Study 2**
Demographics	Victim type	Not shared	Shared
	Gender	Not shared	Not shared
	Ethnicity	Shared	Shared
	Age	Shared	Shared
Vulnerability indicators	Known offender	Not collected	Shared
	Disability	Not shared	Shared
	Repeat victim	Not shared	Shared
Impact indicators	Confidence	Not collected	Shared
	Health	Not shared	Shared
	Finances	Not shared	Shared

However, in its F&CM guidance, many fraud subtypes, as well as the category of Computer Virus, Malware and Spyware are subject to the SIV principle, generally intended for crimes against the person. That is because while 'the victim' in relation to these crime types may be better conceptualised as a 'hybrid' of individuals, corporations and digital systems [47], the rules are designed to avoid the double recording of the same crime. As such, where victims report an instance of cheque, plastic card or online bank account fraud, they will be asked to contact their financial institution in the first place and this will not be recorded as CRI or a crime [45]. Furthermore, if an individual is infected by a virus which is affecting machines on a global scale, a CRI will be recorded. These rules are intended to minimise double-counting of crimes which might be reported by multiple victims.[8] Notwithstanding this rule, previous research has shown that F&CM can have impacts beyond financial loss including on victims' health and wellbeing [48], which they may require support with. In addition, support and advice may be required to avoid repeat victimisation. In fact, the accurate measurement of repeat victimisation may also be affected by the one-crime-per-victim principle.

[8]That said, it is still possible for an individual to report a fraud with losses that are subsequently refunded. As such, double counting of recorded crime may happen to some extent between the various sources feeding into the fraud recorded by the police. On the other hand, it is unclear which other organisation would be expected to report widely spread computer viruses, or how AF operators might decide which viruses are indeed sufficiently global in scale.

Where a victim receives a fraudulent cold call, phishing email or letter, or a malicious attachment, but they do not engage with the offenders, they are not considered a SIV, and a CRI will be recorded. However, if the victim acts on any of the information given by the offenders or they are repeatedly targeted, they become the SIV – even if no money/property is lost. As such, in some cases where only one fraud has been recorded, the victim has already been repeatedly targeted. Furthermore, for many categories of advance-fee frauds, where the same victim is contacted multiple times by the same fraudsters and on each occasion the victim is defrauded in the same way (i.e., the same NFIB category applies) then one crime is meant to be recorded when the victim reports these instances together, even where they span years of interaction. However, in such circumstances, arguably the report represents multiple instances of victimisation. In contrast, where an individual is repeatedly targeted by hackers, generally one crime will be recorded for each device and/or account hacked. As such, repeat victims of Hacking will be more easily identified than repeat victims of Advance-fee fraud. As noted, where the SIV principle is not met, a crime related incident (CRI) will be recorded instead of a crime that *counts*. However, as noted above, CRIs are not passed to the police forces within whose jurisdiction the victim falls. At the same time, the new 'repeat victim' variable (Table 3) is now automatically populated by the system and will therefore be highly skewed by these recording rules. As such, using crime records to study and develop operational responses to repeat victimisation, will be limited by the ways in which report data are collected and processed.

"No Victim – No Crime" principle

Section 3.5 of the NCRS elaborates further on the principle of "victim focused recording" and introduces the "No Victim – No Crime" principle. This principle stipulates that if no victim of crime can be immediately identified, "the matter must be recorded as a crime related incident until such time as the victim is located or comes forward to provide an account" (41, para. 3.5i). However, the "recording without victim confirmation" guidance provides two exceptions to the concept of 'no victim - no crime' i) the police believe recording to be appropriate and necessary ii) reports by parents, carers and Professional 'Third' Parties (doctors, nurses, social workers and teachers) where recording must occur regardless of whether the victim has given their permission for the reporting individual to speak to the police and irrespective of whether the victim subsequently confirms that a crime has been committed. Recording without victim confirmation is key to capturing the experiences vulnerable victims of F&CM. Without this principle, no crime would be recorded where vulnerable individuals are identified by third parties (e.g., family members, social services, Trading Standards, or the Citizens Advice), but do not believe themselves to be victims. As previous research has noted, this has been observed in relation to especially vulnerable fraud victims [48]. However, the recording system currently allows for a wide range of "proxy" reports e.g., where a police officer has taken the report at the station and has now entered the information through the AF online portal. As such, it is not possible for local forces to easily identify where recording took place "without victim confirmation" and utilise this as a vulnerability indicator.

Accuracy and reliability of records

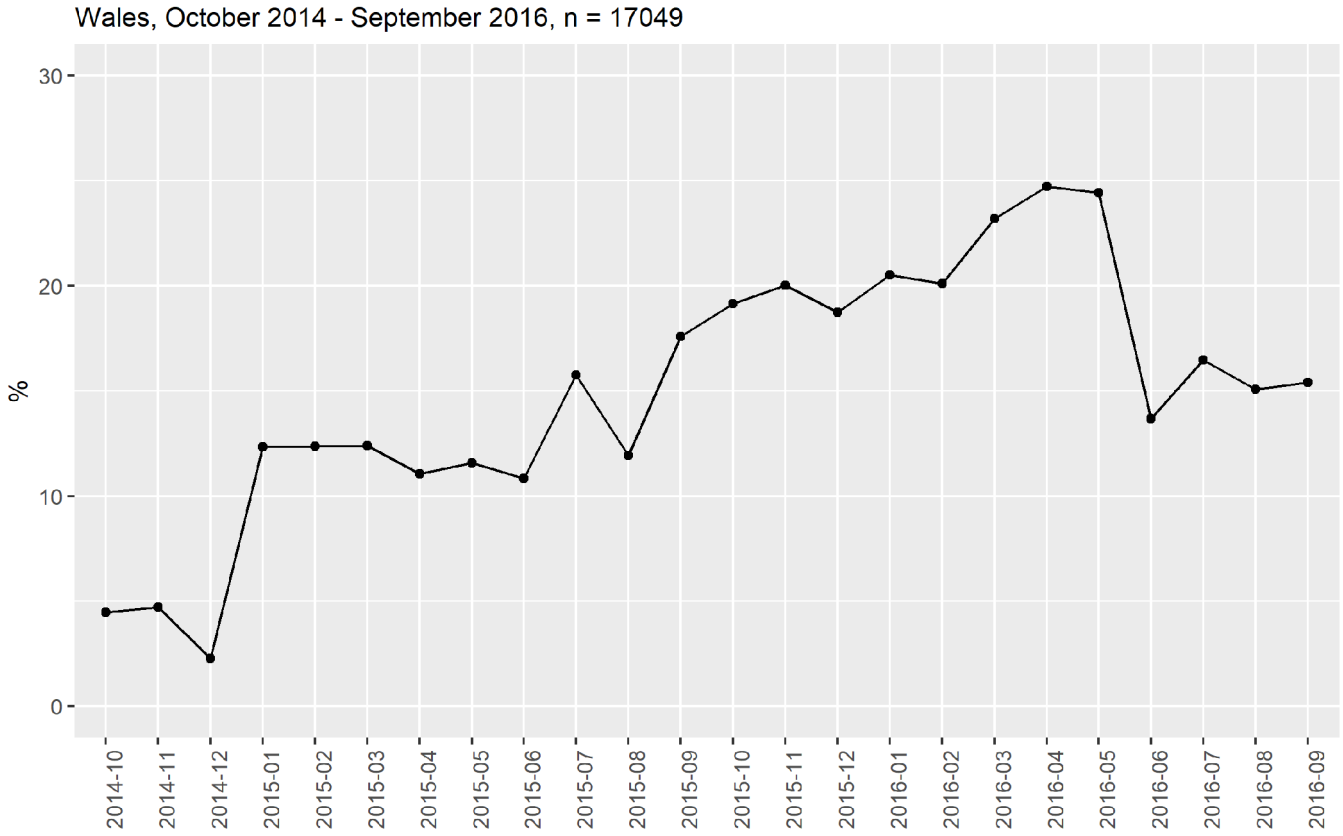
This section examines whether AF data is collected accurately and reliably i.e., accuracy is maintained over time. Accuracy has two sides, one relates to the extent to which the data was collected in accordance with recording rules; the other concerns the extent to which the data collected produces accurate representations of reality, irrespective of these rules. In addition, it should be noted that sources of error related to the accuracy of crime recording affect not only AF data or the UK, but any data which constitutes PRC [49, 50]. With respect to the first, the accuracy of F&CM crime records has doubtlessly improved since the introduction of AF, as a centralised recording system reduces error and increases consistency (see next section). Despite improvements however, Studies 1 and 2 still demonstrated issues of accuracy and reliability, including the presence of duplicate records, outliers and implausible values, recording error and missing values. The last two are considered in greater detail below.

Recording Error and 'Other' Fraud

A key source of uncertainty which will affect the accuracy and reliability of AF data is recording error. Despite the above-mentioned advantages of centralised recording, errors may result from both the quality of the information provided by victims directly, as well as the quality of the recording made by others on their behalf. For example, when a victim reports F&CM, the crime is categorised according to the NFIB crime codes, part of the HOCR. However, it is possible that some reports are mis-recorded, especially given the complexity of the recording rules. At the time of writing, there are 48 NFIB fraud categories relevant to individual and business victims, including the NFIB90 category of 'Other' fraud, plus eight unique computer misuse categories [51]. Given the large number of categories, it is unsurprising that this analysis revealed that the correct crime category was not always recorded. Mis-categorisation may be aggravated by recent concerns over the training and empathy levels of some AF call centre staff [52, e.g.]. Furthermore, when a victim reports directly to AF via the online tool, there is an assumption that they understand and interpret the form as intended. However, unlike survey designs, the questionnaire used to collect AF data has not been subject to cognitive tests and developed accordingly to minimise response error.

Alongside the above, academic, or operational analysis of AF data requires that crime categories be grouped and merged to achieve statistical power, as many have too few cases. Results should also be relatable to previous literature and thus contribute to an evidence base to inform criminal justice policy and practice. However, there is little guidance to crime analysts within police forces or researchers working with AF data, on how best to aggregate and use NFIB categories in research. In the process of the above studies, a working typology including 9 fraud categories and 2 categories of CM [4, 8] was developed as starting point (Table 6). This typology is, to the extent that it was possible, compatible with existing typologies, particularly Levi & Burrows [18],

Figure 2: Proportion of NFIB90 crimes recorded in Study 1, per month



Button, Lewis and Tapley [53], Wall [54] and Yar [55]. It also reflects the aggregation undertaken by the Office for National Statistics to produce official PRC statistics. At the same time, this typology is data-driven, as it was developed to maximise statistical power.

While the above exercise improved the usefulness of the data, a close analysis of crimes categorised as 'NFIB90 - Other Fraud (Not covered elsewhere)', was revealing in terms of the accuracy and reliability of the data. Overall, 15.40% of cases in Study 1 and 23.34% of cases in Study 2 were labelled as 'NFIB90' (14.82% and 23.50% respectively for individual reports only). This is a significant proportion of all crimes recorded, representing the second and first most commonly recorded categories in Studies 1 and 2 respectively. Furthermore, while this proportion tended to increase over the period cover in Study 1 (Figure 2), it has remained consistently high over the period cover in Study 2 (Figure 3).

Given the high levels of prevalence shown above, a closer look was taken at a sub-sample of individual NFIB90 reports in Study 1, to identify and examine any significant patterns. A sub-sample of 160 individual victims who reported 332 incidents between them, were selected for Thematic Analysis (TA). One important limitation of this analysis relates to the reliance on a relatively small sub-sample of incident descriptions (3% of n = 11,841). However, due to the access restrictions, the researcher had to manually verify that personal information was removed from each incident description, before they could be extracted for further analysis at the university and there was a limited time available for this task. At the same time, the use of a combination of random and purposive sampling was intended to maximise the utility of this sample.[9] Of the TA sub-sample, 33 crimes (approximately 10%) reported by 25 victims were classed as NFIB90 and the themes identified within these cases are represented in Figure 4. While further research is necessary to better understand NFIB90, the results of this analysis are indicative.

The most significant theme to emerge was 'Courier Fraud' (36% of incidents coded). While this does not constitute an NFIB category, the AF website states that Courier fraud occurs when a fraudster contacts victims by telephone purporting to be a police officer or bank official [56]. The predominance of Courier Fraud suggests that adding a tag, if not a dedicated NFIB category to reflect this type of fraud, may contribute towards reducing the significance of the 'catch-all' NFIB90 category. While frequently changing NFIB categories would not be desirable in the interest of consistent crime reporting, a revamped recording system would ideally allow the flexibility to create and update crime tags, in order to monitor ongoing and emerging trends. Finally, several of the themes identified highlight that the volume of 'NFIB90' also includes mis-categorised cases. These include cases which, in line with HOCR, should have been categorised as 'romance', 'online shopping', 'consumer' and 'credit card' fraud. In addition, 'telephone (tel) preference' and 'plane ticket' themes might

[9] All reports made by victims who reported three or more incidents were purposively selected for TA (58 victims, 208 incidents). This was done because Study 1 focused on repeat victimisation. At the same time, a random selection of 22 repeat victims who reported two incidents (44 incidents), totalling 252 incidents reported by 80 repeat victims were also selected for TA. An equal number of reports from one-time victims (n = 80) were also randomly selected for TA to ensure a balanced sample of repeat and one-time victims.

Figure 3: Proportion of NFIB90 crimes recorded in Study 2, per month

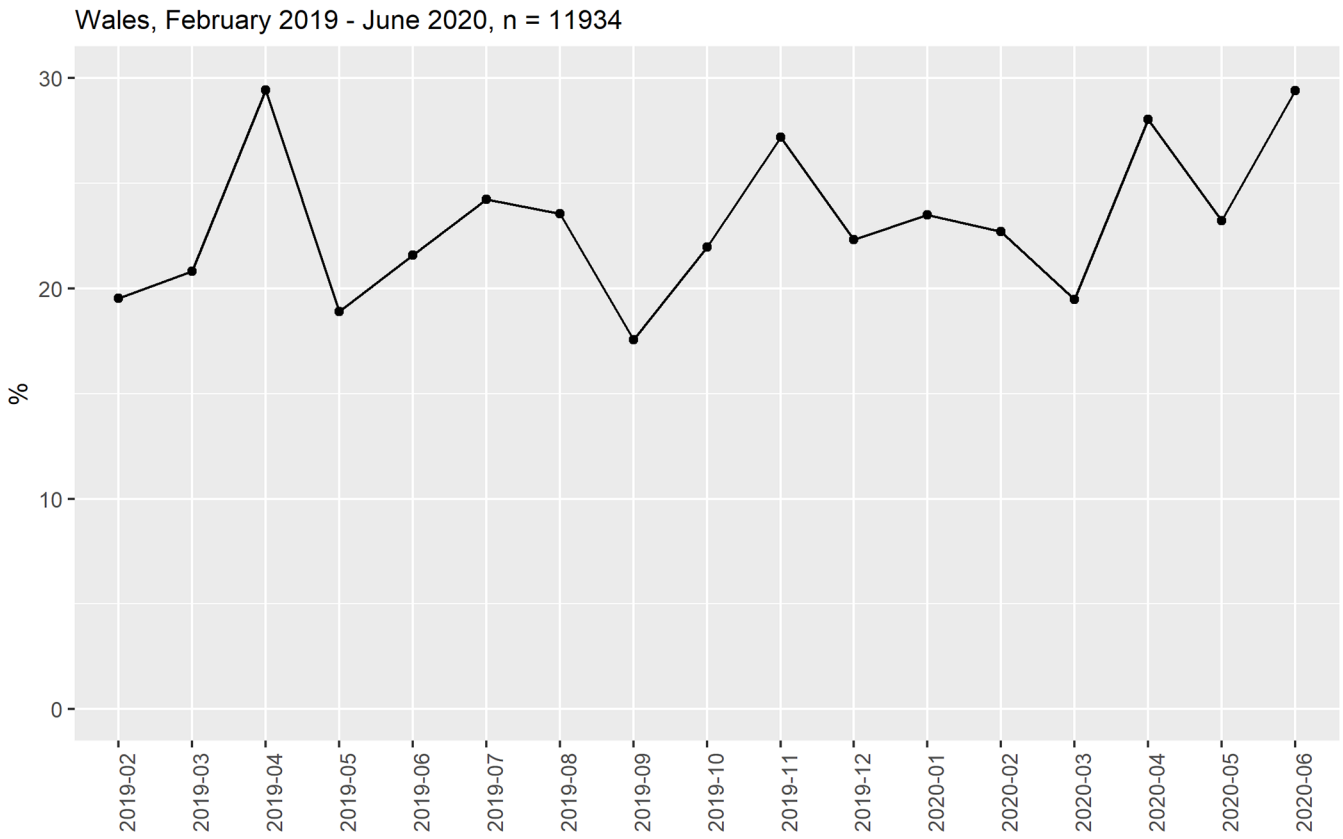


Figure 4: Tree diagram of 'NFIB90' records

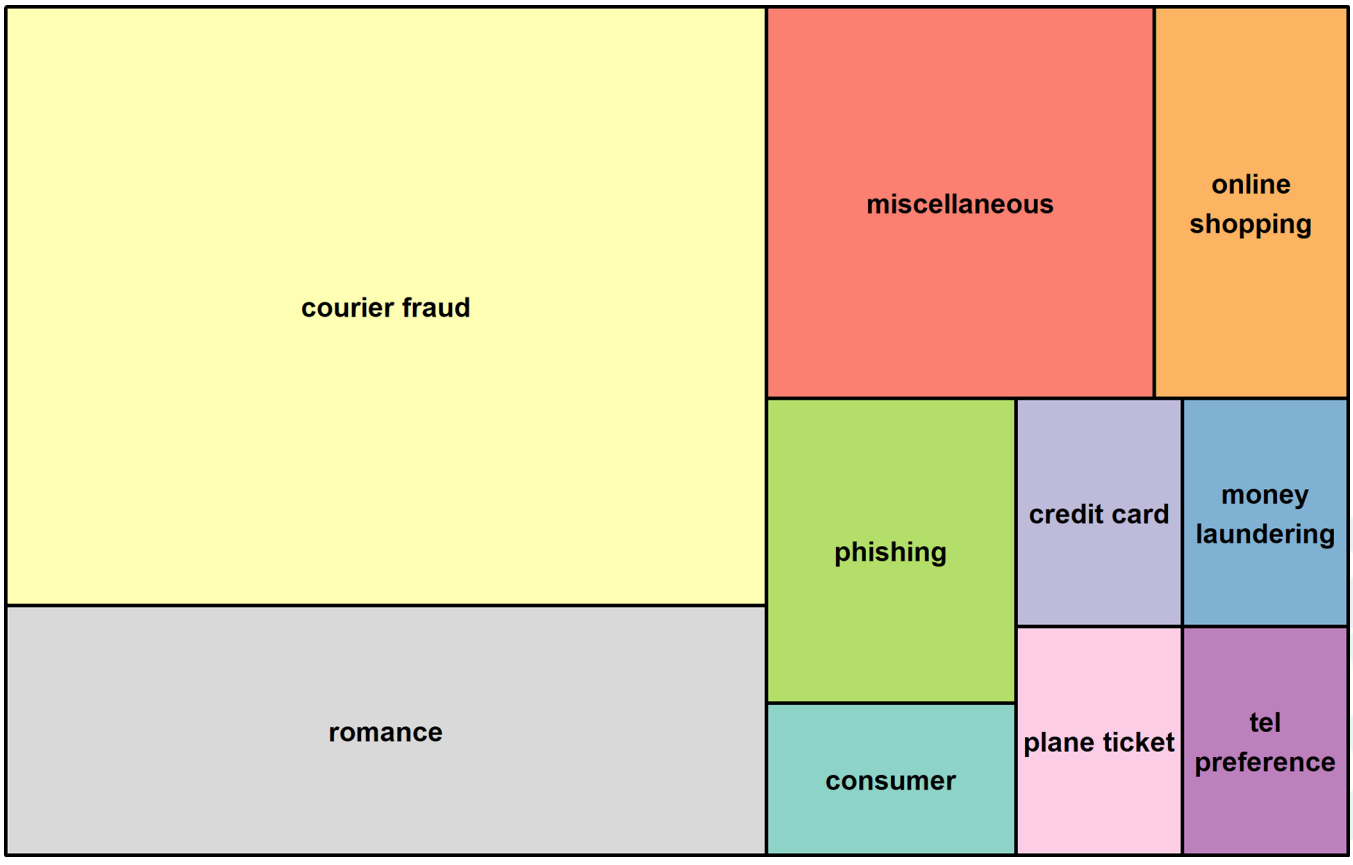


Table 4: Percentage of missing values per key variable, for all victims and for individual victims across studies 1 and 2

	Study 1 – % Missing		Study 2 – % Missing	
	All (n = 17,049)	**Individuals (n = 11,845)**	**All (n = 11,934)**	**Individuals (n = 11,009)**
Force	0	0	0.01	0.01
Date	0	0	0	0
Victim Type*	18.01	0	0	0
Age	29.53	17.99	22.45	15.93
Gender*	5.95	0.17	NA	NA
Ethnicity	42.10	33.47	23.61	17.20
LSOA **	0	0	1.17	1.08
NFIB category	0	0	0	0
Incident description	9.38	2.96	2.10	NK
Direct loss	30.35	30.52	27.12	26.78
Disseminated	92.82	92.41	99.52	99.54
Outcome* **	85.40	85.51	91.6%	92.5%
Referral	98.67	98.67	99.99	99.99
Call for service	94.64	95.87	92.02	92.84
Impact Assessment	NA	NA	31.79	31.96

*Victim Gender was not provided but in Sample 1 could be derived/coded based on other variables. Victim type was derived for sample 1 and directly available with sample 2.
**Lower Super Output Area (LSOA) is a geographical location derived based on victim postcode.
***derived based on the aggregation of several variables, but not directly comparable between Study 1 (which includes variables call, outcome and partner) and Study 2 (where it includes outcome and call).

have been coded as advance fee and ticket fraud respectively. Finally, the cases coded under the theme of 'phishing' should not have been recorded as a crime.

Missing data

Missing data will have a considerable impact on academic research and operational analysis, as it can skew results directly or hinder bias and/or cross-sectional testing. As a rule of thumb, levels of missingness above 5% within a variable are a red flag, particularly if missingness is not random i.e., if significantly prominent among certain victim sub-groups. Furthermore, checking for this kind of bias requires high level of completion within demographic variables. Between Studies 1 and 2, the quality of AF data improved, as measured by the level of missingness within key variables in each sample (Table 4). The proportion of missing values in the variables age, ethnicity, direct financial loss and the free-text description of the crime has decreased between the studies. In addition, as previously noted, several new 'victim impact assessment' variables are now shared with local forces. This assessment is given to individual victims on a voluntary basis, with only 25.74% of individual victims agreeing to take part, while 31.96% did not respond (Table 4). This high level of missing data makes the interpretation of results uncertain.

In addition, among those who agreed to the assessment, the level of missing values within victim impact variables is also not encouraging (Table 5). Alongside the levels of missing data, the lack of consistent coding of missing values also adds uncertainty to what can be known e.g., a missing value for direct financial loss may be indicative of no loss, or of a loss that is unknown at the time of reporting. Similarly, it is unclear how one should interpret missing values within the victim impact assessment variables in Table 5. As such, the quality of AF data would be improved by ensuring that the recording system is designed to reduce levels of missingness. One way to do this is by making questions compulsory, while allowing individuals maximum choice e.g., a 'Prefer not to say' option for ethnicity or crime impact, rather than allowing blanks. In addition, clear rules need to be developed around when data should be coded as missing. For example, a "Yes/No/Not Known" triage question in relation to direct financial losses will help disambiguate whether a missing numeric value for direct financial loss indicates 'no loss' or 'unknown loss' at the time of recording.

Consistency of recording, coherence and comparability

This section considers whether AF data is capable of producing statistical insights that are internally consistent, consistent over time and comparable between regions and police force areas. The increase in the volumes of recorded fraud by approximately 70.5% between 2011 and 2013 [1] suggests an improvement in the consistency of recording across England and Wales with the introduction of AF. Additionally, the volumes of records in Study 1 (i.e., collected from local forces), matched those reported via official statistics.[10] With

[10] A comparison with official statistics of crimes recorded by AF and referred to the NFIB was possible for the year ending September 2016. For this year, there was an exact match between the counts within the sample used in Study 1 and the official counts, as published by ONS, for Dyfed/Powys (n = 1,525), Gwent (n = 1,632) and South Wales (n = 3,803). For North Wales the ONS published figure was n = 1,903 while the sampled figure was n = 1,841 (a different of 3.3%). As the North Wales data was acquired separately (because it falls outside the jurisdiction of the Southern Wales ROCU), this may be due to an error

Table 5: Percentage of missing values for victim impact assessment variable in Study 2 (individual reports only)

	% Missing (n = 2,834)
Disability	0
Repeat Victim	0
Known Offender	0
Health impact	48.41
Financial impact	51.52
Confidence	44.25
PIM assessment	62.39

respect to internal consistency, the key source of uncertainty in relation to AF data is the previously mentioned recording error, aggravated by the multiple layers of interpretation (or hermeneutics) which records may be subjected to. Furthermore, coherence over time will be impacted by external events, as exemplified here by the analysis of the volume of reports over AF's crisis in the summer of 2015 and the impact of COVID on fraud and CM recording. Finally, the comparability of fraud and CM recording figures across police force areas and regions could be improved by the development of a data catalogue to aid crime analysts and researchers in making the most of the data and thus generate insights relevant to policy and practice.

Multiple hermeneutics

The incident descriptions in the samples of Studies 1 and 2 captured the "voice" of several populations, a considerable source of uncertainty with respect to coherence and comparability, which adds to the complexity of working with this dataset. An AF record may be the result of the victim's direct report via the website and therefore be written in the first person, in accordance their understanding of the reporting tool and the questions asked, representing one layer of interpretation. However, it may also be recorded by the call operator or a police officer, based on their interpretation of the victim's report: two layers of interpretation. Furthermore, an operator may record a crime, based on the details taken by an officer, in turn based on their interpretation of the victim's report: three layers of interpretation. Finally, the voice may be that of the AF operator, based on the account of a police officer, where the report was made by someone else, on behalf of the victim. In this situation, four layers of interpretation or hermeneutics are possible. These layers of interpretation constitute a considerable source of uncertainty with respect to the internal coherence of the data. While this is not so problematic with respect to the categorical and numeric variables collected by AF, it can affect the quality of the free-text description. In Study 1, for example, the average number of characters included in the incident description was 607, the median 492, but it varied from 5 to 2033 characters. However, the free text description of the incident is key to identifying new trends. As such, ensuring that it is a compulsory field, and that there is sufficient and accessible guidance on how to complete it is essential. Further, for reports taken by operators and police officers, it is reasonable to expect a minimum level of detail.

Impact of external events

The impact of external events on reporting is evident when the changes in the levels of reporting are considered against the backdrop of key events. In the case of Study 1, the volume of data collected by AF was deeply affected when Broadcasting Support Services (BSS), the not-for-profit organisation which then ran the AF call centre, suddenly went into administration in July 2015 after losing the tender contract for the continued provision of this service to IBM. This led to AF operating with a skeleton staff and is clearly reflected in the fall reports over that period, within the four Welsh police forces covered in Study 1, which then took a significant amount of time to recover (Figure 5). Furthermore, while the period covered in Study 2 does not allow for a clear examination of the impact of the lockdown on volumes of F&CM recording, official statistics show that there was a +28% increase in reports of fraud and a +16% increase in CM reported via Action Fraud [1]. Throughout the pandemic, other data source such as the Telephone-Operated Crime Survey for England and Wales suggest that the increase in reporting is linked to an increase in the overall levels of F&CM being experienced by victims, as criminal took advantage of the COVID-19 crisis. In relation to Study 1 however, the drop in records relates to the availability of the AF service. As such, understanding the impact of external events on the data is essential context to the adequate interpretation of any insights produced using the AF dataset.

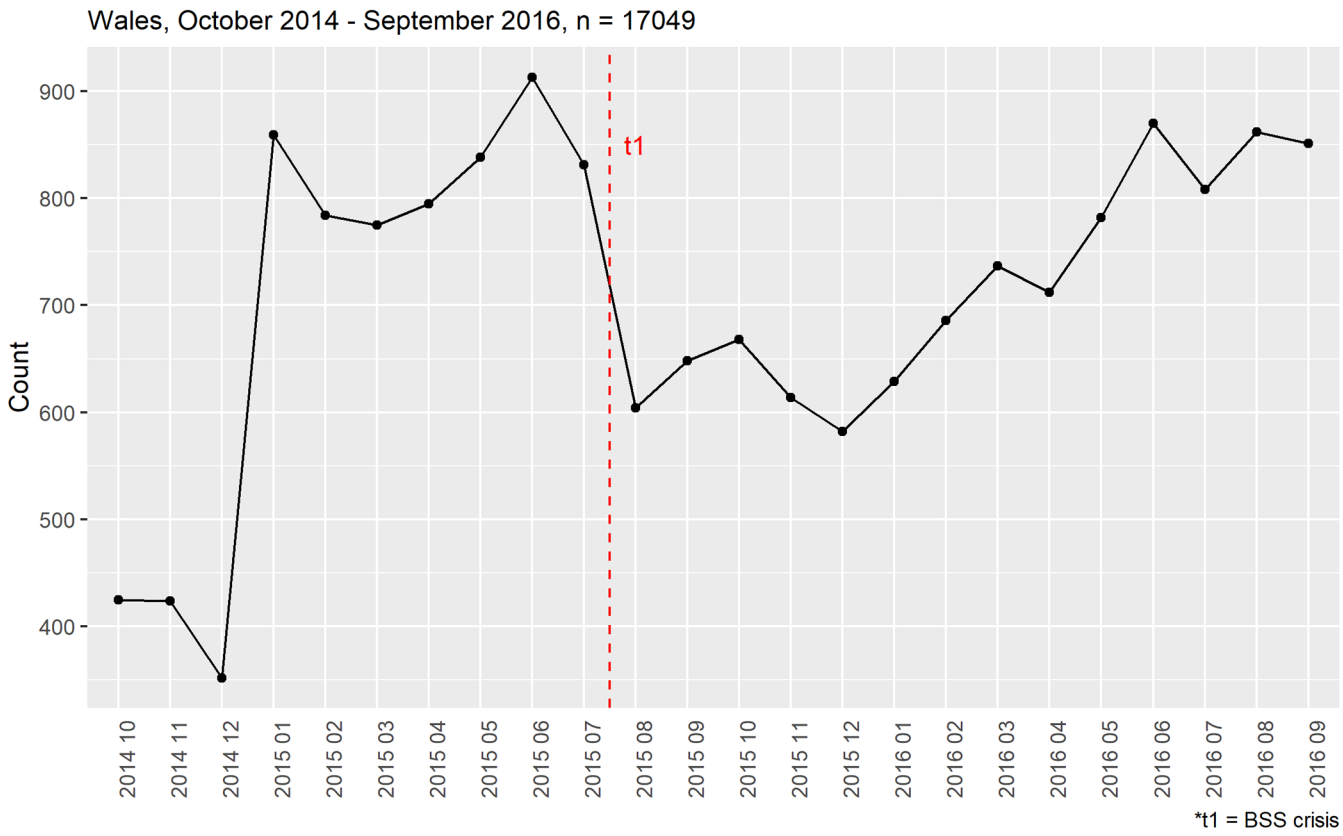
Timeliness & accessibility

With respect to the timeliness and punctuality with which police forces and crime analysis have access to AF data, this has much improved between the two studies covered in this paper. At the time the data for Study 1 was collected, victim reports were shared with local forces on a monthly basis, which meant there was an inevitable delay between victim reports and a local Protect response, where available. Since then, this has improved considerably with victim information now shared with forces on a weekly basis. This has improved the timeliness of AF data with respect to the needs of Protect officers.

In addition, the accessibility of AF data to the wider research community has improved through the online

in the processing of the data before it reached the author. However, this was considered a small error within the overall sample. On the whole therefore, the data on the ground reflects official counts.

Figure 5: Number of crimes recorded in Study 1, per month



publication of the NFIB Fraud and Cyber Crime Dashboard [6]. This interactive dashboard allows researchers to access several AF data variables and apply filters (e.g., by crime category, region, police force area), to explore the previous 12 months of AF records. Researchers have already been able to harvest this data through the portal and use it to make valuable contributions to the field [7, 23]. At the same time, the variables made available are limited and the dashboard not optimised for data download. As such, researchers wishing to investigate aspects such as victim impact of repeat victimisation, might wish to explore alternative data access routes, including research collaborations and entering into Data Processing Agreements with NFIB at the national level, the Regional Organised Crime Units at regional level, and/or specific local police forces.[11]

Closely linked to the above, users of AF data need access to sufficient supporting metadata and guidance about how the data was collected, to enable them to produce and present outputs in a clear, accessible, and impartial basis. Tables 4 and 6 list the variables accessed for Studies 1 and 2, but this does not constitute the full spectrum of variables collected by AF. Furthermore, little information was available to describe variables and their classifications, when these were shared with the researcher in both Studies 1 and 2. The lack of a *data catalogue* is therefore a major limitation, which will hinder analysists and researchers ability to make the most of this dataset and producing insights relevant to research, policy and practice.

[11] In the context of the studies mentioned here, the data was accessed through a partnership with Southern Wales Regional Organised Crime Unit (ROCU).

Recommendations

The above analysis has identified several challenges when using AF data in research and to inform victim-focused and intelligence-led responses. In particular, these data provide a better insight into reporting/recording, than crime patterns or victimisation risk. However, it has a huge potential with respect to identifying serious crimes, vulnerable and repeat victims, generating intelligence and monitoring whether CJS responses are adequate to meet victims' needs. To achieve this, four key areas for improving the quality of AF data are detailed below.

Firstly, to improve the relevance of the outputs produced with AF data, the collection could be optimised to enable a 'Protect' response, e.g., improved impact and vulnerability indicators to identify and respond to victims' support needs. In addition, to correct for the impact of the principal crime rule and improve intelligence, data on whether fraud was enabled by a CM offence could be provided to local forces. These improvements should not come at the detriment of collecting data to the 'Pursue' strand of the police response. However, there has long been a recognition that law enforcement will not be able to 'arrest their way out' of fraud and CM. As such, collecting the right data and sharing it with local forces, will allow CJS agencies and researchers to assess victim impact and vulnerability and to identify high incidence repeat victims. This will enable a victim response even where there are no leads for investigation and arrests to be pursued.

Secondly, the accuracy and reliability of the insights produced using AF data can be improved by ensuring that the levels of missing data are reduced, and data accuracy

Table 6: Fraud and CM working typology based on NFIB categories

Fraud	**Definition**	**Included NFIB categories**
Advance-fee & Romance	Where fraudsters use telephone, mail or e-mail to target victims with a fictitious scenario and persuade them to pay a fee, in advance of either financial (e.g., in the case of fake lottery winnings or lender loan frauds) or emotional recompense (e.g., in the case of romance fraud).	NFIB1A; NFIB1B; NFIB1C; NFIB1D; NFIB1E; NFIB1F; NFIB1G; NFIB1H; NFIB1J
Business compromise	Fraud committed by an employee against their employer, or by a business against another business.	NFIB8A; NFIB8B; NFIB9; NFIB10; NFIB11
Card and banking	Fraudulent use of cheques, plastic card (including credit, debit, prepayment and store cards) and bank accounts.	NFIB5A
Consumer	Fraud in connection with the purchase of goods or services (allegedly) rendered.	NFIB3A; NFIB3B; NFIB3C; NFIB3D; NFIB3E; NFIB3F; NFIB4A; NFIB6B;
Investment	Investments in fraudulent 'opportunities' including investment in pensions, shares, pyramid schemes, time shares and others. Distinguished from advance fee based on the specific nature of the investment pretext used by the fraudsters.	NFIB2A; NFIB2B; NFIB2C; NFIB2D; NFIB2E; NFIB16B; NFIB16C
Other	All other fraud not covered elsewhere.	NFIB17; NFIB18; NFIB19; NFIB90;
Public	Fraud against public authorities including fraudulent applications for grants and driver's licenses, benefit and tax fraud.	NFIB12, NFIB13, NFIB14; NFIB15; NFIB20A
Retail	Fraud committed against retailers that does not involve online sales or cheque, or plastic card sales. It includes refund fraud, label fraud and obtaining goods or services with no intent to pay.	NFIB3G
Services	Fraudulently applying for legitimate services including credit, insurance etc. It also includes fraudulently setting up direct debits from another's account.	NFIB4B; NFIB5B; NFIB5C; NFIB5D; NFIB5E; NFIB6A; NFIB7; NFIB16A
Computer Misuse	**Definition**	**Included NFIB categories**
Hacking	The hacking (i.e., unauthorised access) to a computer system. It includes the hacking of computers, servers, telephone systems, social media and email accounts, with and without blackmail.	NFIB52A; NFIB52B; NFIB52C; NFIB52D; NFIB52E
Malware, Virus & (D)DOS	Criminal acts which impair the operation of a computer system including computer malware, viruses and (Distributed) Denial of Service attacks.	NFIB50A; NFIB51A; NFIB51B

is monitored and tested. In particular, the categories of crime used should not over-rely on the catch-all 'Other' fraud category and improved training is necessary to reduce human error. At the same time, upstream data validation and harmonisation would encourage continuous improvement while maximising efficiency. Presently, there is limited information available to data users documenting whether and how AF data are validated, or the outcomes of any such validation. In line with previous research using PRC [57, 58, e.g.], extensive data cleaning was therefore required to make this data usable for analysis. The time and resource this involved would not be feasible for most crime analysts and could compromise the timeliness of statistical and/or operational outputs. However, with some level of data harmonization and validation introduced 'upstream', the potential of AF data may be fully realised. Drawing on administrative data quality literature [15, 16, 59], key accuracy validation checks for AF data are proposed, which could be conducted at regular intervals, before AF data is sent to the NFIB, or shared with local forces and other researchers (Table 7).

Data harmonisation (e.g., the consistent coding of NA values) and data validation checks i.e., checks to monitor the quality of the data collected, would contribute towards ensuring AF data quality is continuously improved. For greater efficiency, such checks might be carried out 'upstream' from the data being shared with NFIB and onwards with local forces and other researchers. While non-exhaustive, the Table 7 draws on previous research to suggest a set of validation checks which would be relevant in relation to AF data.

Thirdly, the coherence and comparability of AF data can be improved through user research, recording audits and a vision which seeks to align operational and statistical quality. Were the recording tool to be re-developed in the future, user research should be undertaken to test the extent to which victims understand the online reporting tool, before it is deployed. This includes the data collected about the crime itself (e.g., the process of selecting the best-fitting crime

Table 7: Proposed accuracy validation checks for AF data

Validation check	**Description**
Record and variable counts	Overall number of records and variables collected for the reference period.
% Duplicates	Percentage of duplicate records i.e., duplicate unique crime reference and overall record.
Longitudinal sense checks	Check for significant and unexpected changes in variable distributions including top frequencies of crime types and victim characteristics over time.
% Missing and missing bias	Percentage of missing values within each key variable, taking into account any conditionality that applies. In addition, chi-squared tests may be useful to determine whether missing values is significantly more prevalent among population sub-groups (e.g., across gender or ethnicity).
% Out of Range	Percentage of out-of-range values within each key variable i.e., outside the parameters specified by recording rules.
% Implausible	Percentage of 'dubious' cases within variables i.e., outliers, not out of range and not necessarily incorrect, but implausible.
% Incompatible	Percentage of incompatible cases between logically related variables (e.g., date of birth and victim age).

category), but also the victim impact assessment variables. To make the most of AF data, the aim should be to, as far as possible, approximate operational data quality to statistical quality. In this respect, designing robust questionnaires has a long tradition in the social sciences and therefore the assistance of social scientists may prove valuable. From a technical perspective, user testing and human-centred design would be beneficial in future system development. Following this approach, accessibility and user satisfaction testing may be designed-into the tool itself, with metadata collected to continuously evaluate and improve the reporting tool – something which the current 'static' system does not permit. At the same time, regular audits of crime records and recording 'in action' will provide analysts and officers with the confidence they need that the data collected is accurate and reliable.

Finally, developing a data catalogue would enable frontline officers and researchers within academia and beyond, to harness the full potential of this dataset and produce insights needed for crime prevention, investigation, and meeting victims' needs. Crime analysts and researchers need to be aware of the data quality issues noted throughout this paper and any future changes to data collection and processing which may impact the relevance, accuracy and reliability, or the consistency and comparability of analytical outputs produced using AF data. As such, this paper recommends the development of a data catalogue setting out the variables contained in AF data, their data classes (e.g., numeric, discrete categories, date, etc.), the range of acceptable values and any additional notes relevant to their use in producing research and operational insights (e.g., eligibility and/or recording rules). A mechanism to update the catalogue when changes to data collection and processing occur is also key.

Conclusion

A continued conversation within the research community and with data providers is needed to enable researchers to access and utilise the wealth of data collected in relation to cybercrime and fraud by police authorities. The analysis and recommendations made in this paper will help researchers be better prepared to develop adequate research designs that utilise fraud and computer misuse crime records to its full potential and better understand what data is available when applying for access to this dataset. The strengths and weaknesses of AF data must be fully understood, to realise its potential to help tackle substantive research areas and to aid police crime analysts working with AF data, as well as policy makers' working towards improving the response to F&CM. This analysis will also aid frontline officers and crime analysts to make the most of this dataset, harnessing it to produce key insights for crime prevention and meeting victims' needs. Finally, the insights presented here will be valuable to policy makers and practitioners involved in the development and design of crime recording systems in today's data-driven world.

Acknowledgments

Some of the work presented here resulted from a Ph.D. studentship funded by the Economic and Social Research Council, in collaboration with the Southern Wales Regional Organised Crime Unit (known as 'Tarian', the Welsh word for 'shield'). I would like to express my immense gratitude for the fantastic feedback received on earlier versions of this paper from colleagues, research partners, stakeholders and the anonymous reviewers.

Statement on conflicts of Interest

None to be declared.

Ethics statement

Ethical approval for the studies was provided by the Research Ethics and Governance Committee at HRC School of Law, Swansea University (Ref Correia 03/11/2016).

References

1. Elkin M. Crime in England and Wales: Appendix tables - Year ending March 2021. In: ONS. Titchfield, UK. 2021.

2. Kemp S, Buil-Gil D, Moneva A, Miró-Llinares F, Díaz-Castaño N. Empty Streets, Busy Internet: A Time-Series Analysis of Cybercrime and Fraud Trends During COVID-19. Journal of Contemporary Criminal Justice. 2021;37(4):480–501. https://doi.org/10.1177/10439862211027986

3. Hawdon J, Parti K, Dearden TE. Cybercrime in America amid COVID-19: the Initial Results from a Natural Experiment. American Journal of Criminal Justice. 2020;45(4):546–62. https://doi.org/10.1007/s12103-020-09534-4

4. Correia S. Responding to victimisation in a digital world: a case study of fraud and computer. Crime Science. 2019;8(4). https://doi.org/10.1186/s40163-019-0099-7

5. Fraud Advisory Panel. Fraud in Scotland. 2020. Available at: https://www.fraudadvisorypanel.org/wp-content/uploads/2020/07/Fraud-in-Scotland-4th-ed-July2020.pdf

6. City of London Police. NFIB Fraud and Cyber Crime Dashboard - 13 months of data. 2021 [11 November 2021]. Available from: https://colpolice.maps.arcgis.com/apps/opsdashboard/index.html#/60499304565045b0bce05d2ca7e1e56c.

7. Buil-Gil D, Miró-Llinares F, Moneva A, Kemp S, Díaz-Castaño N. Cybercrime and shifts in opportunities during COVID-19: a preliminary analysis in the UK. European Societies. 2020. 11 August 2020.

8. Correia S. Vulnerability and Repeat Victimisation in a Digital World: A Study of Computer Misuse and Fraud Reported in Wales. Swansea: Swansea University; 2021.

9. Levi M, Doig A, Gundur R, Wall D, Williams M. Cyberfraud and the implications for effective risk-based responses: themes from UK research. Crime, Law and Social Change. 2017;67(1):77-96. https://doi.org/10.1007/s10611-016-9648-0

10. Levi M, Doig A, Wall D, Gundur R, Williams M. The implications of economic cybercrime for policing. London: City of London Corporation; 2015.

11. Correia S. Patterns of online repeat victimisation and implications for crime prevention. The Symposium on Electronic Crime Research (eCrime 2020); 2020 November 16 – 19; Boston [virtual conference]: The Institute of Electrical and Electronics Engineers (IEEE).

12. Hand DJ. Statistical challenges of administrative and transaction data. Journal of the Royal Statistical Society. 2018(181):555–605. https://doi.org/10.1111/rssa.12315

13. UK Government. Beating Crime Plan: Fewer victims, peaceful neighbourhoods, safe country. London: UK Government; 2021. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1015382/Crime-plan-v10.pdf

14. Jacques P. 'No intention to scrap Action Fraud', says City of London Police. Police Professional. 4 August 2021.

15. UNECE. Guidelines for measuring the quality of administrative sources for use in censuses. Geneva: 2021 October. Avialable from: https://unece.org/sites/default/files/2021-10/ECECESSTAT20214_WEB.pdf

16. Wallgren A, Wallgren B. Register-based Statistics: Statistical Methods for Administrative Data. 2 ed. Chichester: Wiley; 2014.

17. eurostat. Eurostat Quality Assurance Framework. 2013. Available from: https://ec.europa.eu/eurostat/documents/64157/4372717/Eurostat-Quality-Assurance-Framework-June-2013-ver-1-1-EN.pdf/352234ca-77a0-47ca-93c7-d313d760bbd6.

18. Levi M, Burrows J. Measuring the Impact of Fraud in the UK: A Conceptual and Empirical Journey. British Journal of Criminology. 2008;48(3):293-318. https://doi.org/10.1093/bjc/azn001

19. UKSA. Statistics on Crime in England and Wales. London: UK Statistics Authority, 2014: Assessment Report 268.

20. ONS. Crime in England and Wales: Year ending September 2016. Office for National Statistics, 19 January 2017.

21. ONS. Crime in England and Wales: Appendix tables - Year ending September 2019. 2020.

22. van de Weijer SGA, Leukfeldt R, van der Zee S. Cybercrime Reporting Behaviors Among Small- and Medium-Sized Enterprises in the Netherlands. In: Weulen Kranenbarg M, Leukfeldt R, editors. Cybercrime in Context. Cham: Springer; 2021. p. 303-25.

23. Kemp S, Buil-Gil D, Miró-Llinares F, Lord N. When do businesses report cybercrime? Findings from a UK study. Criminology & Criminal Justice. 2021: p.1–22. https://doi.org/10.1177/17488958211062359

24. ONS. Crime in England and Wales: Appendix tables - Year ending September 2015. Titchfield2016.

25. ONS. Crime in England and Wales: Year ending September 2015. 21 January 2016.

26. ONS. Crime in England & Wales: Appendix tables - Year ending September 2016. ONS, 2017.

27. ONS. Crime in England and Wales: Appendix tables - Year ending September 2018. ONS, 2019.

28. eurostat. Quality Assurance Framework of the European Statistical System V2.0. eurostat; 2019.

29. Brimicombe AJ. Definition of 'repeat victim' and calculation of repeat victim statistics from police recorded crime and incident data. Government Statistical Service, 2014: CSAC(14)12.

30. Brimicombe AJ. Mining Police-Recorded Offence and Incident Data to Inform a Definition of Repeat Domestic Abuse Victimization for Statistical Reporting. Policing. 2016: pp. 1–15. https://doi.org/10.1093/police/paw025

31. The Fraud Review. Fraud Review Final Report. London: Office of the Attorney-General, 2006.

32. HMICFRS. Fraud: Time to Choose, An inspection of the police response to fraud. Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services, 2019.

33. Skogan WG. Reporting Crimes to the Police: The Status of World Research. Journal of Research in Crime and Delinquency. 1984;21(2):113–37. https://doi.org/10.1177/0022427884021002003

34. Tarling R, Morris K. Reporting Crime to the Police. The British Journal of Criminology. 2010;50(3):474–90. https://doi.org/10.1093/bjc/azq011

35. MacDonald Z. Revisiting the Dark Figure: A Microeconometric Analysis of the Under-reporting of Property Crime and Its Implications. The British Journal of Criminology. 2001;41(1):127–49. https://doi.org/10.1093/bjc/41.1.127

36. Zawitz MW, Klaus PA, Bachman R, Bastian LD, DeBerry MM, Rand MR, et al. Highlights from 20 Years of Surveying Crime Victims. The National Crime Victimization Survey 1973–92. US Government Printing Office: Bureau of Justice Statistics, 1993: Special Report NCJ-144525.

37. ONS. Percentage of incidents of fraud and computer misuse reported to Action Fraud, and reasons for not reporting incidents to Action Fraud, year ending September 2016 CSEW (Experimental Statistics). 2017.

38. Daas PJH, Arends-Tóth J, Schouten B, Kuijvenhoven L, editors. Quality Framework for the Evaluation of Administrative Data. European Conference on Quality in Official Statistics; 2008.

39. Furnell S, Dowling S. Cyber crime: a portrait of the landscape. Journal of Criminological Research, Policy and Practice. 2019;5(1):13–26. https://doi.org/10.1108/JCRPP-07-2018-0021

40. HMIC. On The Record. HMIC; 2000.

41. Home Office. Home Office Counting Rules For Recorded Crime: Crime recording General Rules. Home Office, 2020.

42. College of Policing. Estimating demand on the police service. College of Policing, 2015.

43. Commissioner of Police. National Lead Force: First Quarter Performance Report. Commissioner of Police, 2015.

44. Commissioner of Police. National Lead Force: Q4 Performance Report. Commissioner of Police, 2016.

45. Home Office. Home Office Counting Rules For Recorded Crime: Fraud & Computer Misuse. Home Office, 2020.

46. Home Office. Home Office Counting Rules For Recorded Crime: Crime recording General Rules. Home Office, 2018.

47. van der Wagen W, Pieters W. The hybrid victim: Re-conceptualizing high-tech cyber victimization through actor-network theory. European Journal of Criminology. 2020;17(4):480–97. https://doi.org/10.1177/1477370818812016

48. Button M, Cross C. Cyber Frauds, Scams and their Victims: Taylor & Francis; 2017.

49. Skogan WG. The Validity of Official Crime Statistics: An Empirical Investigation. Social Science Quarterly. 1974;55(1):25-38.

50. Tompson L, Johnson S, Ashby M, Perkins C, Edwards P. UK open source crime data: accuracy and possibilities for research. Cartography and Geographic Information Science. 2015;42(2):97–111. https://doi.org/10.1080/15230406.2014.972456

51. Home Office. Home Office Counting Rules For Recorded Crime: Fraud. Home Office, 2021.

52. Morgan-Bentley P. Shamed anti-fraud firm will lose public contracts. The Times. 2019 10 September 2019; Sect. News.

53. Button M, Lewis C, Tapley J. Fraud typologies and victims of fraud Literature review. National Fraud Authority 2009.

54. Wall DS. Cybercrimes: New Wine, No Bottles? In: Davies P, Francis P, Jupp V, editors. Invisible Crimes: Their Victims and their Regulation. London: Macmillan; 1999. p. 105-39.

55. Yar M. Cybercrime and Society. 1st ed. ed. London: SAGE Publications; 2006.

56. Action Fraud. Courier fraud. 2021. Available at: https://www.actionfraud.police.uk/a-z-of-fraud/courier-fraud

57. Brimicombe A. Analysing Police-Recorded Data. Legal Information Management. 2016;16(02):71–7. https://doi.org/10.1017/s1472669616000207

58. Brimicombe AJ, Brimicombe LC, Li Y. Improving geocoding rates in preparation for crime data analysis. International Journal of Police Science & Management. 2007;9(1). https://doi.org/10.1350/ijps.2007.9.1.80

59. Daas PJH, Ossen S, Vis-Visschers R, Arends-Tóth J. Checklist for the quality evaluation of administrative data sources. The Hague: Statistics Netherlands; 2009.

Abbreviations

ACPO: Association of Chief Police Officers
AF: Action Fraud
F&CM: Fraud and Computer Misuse
HMIC: Her Majesty's Inspectorate of Constabulary (now Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services)
HOCR: Home Office Counting Rules
NCRS: National Crime Recording Standard
NECVCU: National Economic Crime Victim Care Unit
NFIB: National Fraud Intelligence Bureau
ONS: Office for National Statistics
UKSA: UK Statistics Authority