

Crime Data - Guidance

This document outlines how one would apply the lessons learned in previous introductions of transparency to government to the issue of crime data and crime mapping. It contains the recommended best practice for making public data open and transparent.

It is important to note that this refers to the data – the mapping technology is a relatively trivial addition of functionality to the standard provision of data. If the data is correctly presented in a standard format then the functions required to meet the business/political objectives can be developed.

Aims and overriding principles

The reason for releasing crime data is in line with the transparency initiatives in other areas of public service - to enable greater accountability, greater engagement and greater efficiency by allowing inspection by residents, and reuse of the data in innovative applications, mashups, visualisations and analysis. It is an opportunity to describe crime activity in a particular area and allow people to compare their area with others. It will help the police communicate what their priorities are and enable members of the public to identify key issues and crime trends using the facts.

Publishing this data will also enable the public to become more involved in the detection and reduction of crime. Members of the public will be able to give police information on crimes that have been committed and help police as potential witnesses. They will be able to identify local trends themselves and prepare accordingly.

Used effectively crime data will enable communities to become safer, have more information on the issues of crime and antisocial behaviour in their neighbourhood and increase the interaction and enhance relations between residents and the police. This offers a situation where the police can achieve a strong growth in the level of public trust. It will also enable police to introduce smarter, leaner business processes around effective data.

The ability for a resident to view the progress of a crime committed against them through the criminal justice system is something we also want. It is something that will help social cohesion, trust in the criminal justice system and ensure a fair deal for every victim of crime.

Given this it is essential that Police Authorities publish their crime data (and other data too) with a licence and a format that allows reuse, and on the basis that data should be published unless there is an overriding reason not to. This means that Police Authorities should be asking, what's the most they can publish, not what's the least they have to do.

The focus of the guidance is on how, pragmatically, to make the data available quickly rather than seeking to achieve full alignment across every entity. This is consistent with the evolving approach within data.gov.uk. Publishing raw data quickly is an immediate priority, but there are significant benefits to be gained by Police Authorities publishing structured, regularly updated data using open standards, and this guidance helps enable this.

There is, however, no reason why, if they have the resources, authorities should not publish in more ways, or in greater detail than in this guidance. For example, some may start publishing

the information as linked data (in which the data is semantically described using URIs), and we expect to include the successful practices learnt from such experiments in future versions of this guidance.

Licence, Timing & Contacts

Mandatory:

- Data should be published with with a licence that allows open reuse, including commercially. We recommend the new Open Government Licence <http://www.nationalarchives.gov.uk/doc/open-government-licence/>
- There should be a nominated contact person, listed on the page linking to the data files, to oversee the publishing of the data, and this person should also be responsible for answering queries to do with the data, and requests for further information. This will not only increase the skill levels within the police authority and improve relationships with reusers, but also reduce the necessity of going down formal routes (e.g. Freedom of Information requests).
- Data should be published at the most frequent level it is supplied to the Home Office. This data is to be published in individual files, and there should be a single web page that contains links to the individual files.
- Files should be named in a consistent manner with the date included in the file name, and each file name should be unique. You should avoid spaces or characters other than a-z, 0-9, _ and -.
- Authorities may wish to publish the data for longer reporting periods e.g. quarterly, year to date, etc. There is no reason why they should not do this, but this should be in addition to the individual files.
- The Prime Minister has committed to publishing crime data at a level that allows the public to see what is happening on their streets from January 2011. We recommend that where possible data is provided from 1 April 2010. This will add value to the user and demonstrate commitment to complete transparency.

Recommended:

- When errors are discovered, or files are changed for other reasons, rather than 'silently' changing the file it is recommended to publish a revised and differently named version (together with the original files), e.g. 02_2011_v2, 02_2011_v1.
- The data should be published as part of a wider open-data initiative, and there should be a dedicated open-data page/section on the Police Authority's website (the recommended URL for this section is <http://www.yourforce.police.uk/opendata>), as well as listing the files at data.gov.uk.

Scope of Crime Types

Recommended:

For the initial stages of this project we recommend that the following give a range of offences that affect the most people

- Criminal Damage,
- Violent crime (up to and including GBH),
- Vehicle Crime,
- Burglary,
- Robbery,
- Public place antisocial behaviour.

File Format

Mandatory:

The files are to be published in CSV file format. Microsoft Excel files should be converted to CSV. The CSV file must have precisely one header line with field names as suggested. A comma character must separate values, with a new line between separating incidents. Text values that contain a comma must have a “” character at the start and end of the value. This format is the behaviour of Excel when using the “Save As” function selecting CSV as the file type. There should be no comment lines and no blank lines.

Authorities may wish to publish the data in additional formats as well as the CSV files (e.g. linked data, XML, or PDFs for casual browsers). There is no reason why they should not do this, but this is not a substitute for the CSV files, and PDFs in particular do not allow their contents to be reused. Authorities must take full responsibility for the content of their files.

Data content

The content of the published data should match that set out below. Any additional fields that an authority wishes to publish should conform to the general specifications (e.g. quoting values containing commas), and should be added after the fields listed here.

Body

This identifies the police authority to which the data relates, and means that the file is self-describing (i.e. all the information needed is within the file). It should be the URI that represents the police authority. It is a matter of urgency that such URIs are agreed unambiguously. In the case of UK police authorities, they should be of the form of <http://ABC.uk> (where ABC is the id for the authority). You should also add the name of the authority in a Body Name field to aid readability.

Data in the following format is already exchanged between police forces and other partnership and police management information sites.

Crimes

Crime_No	URN	This is the crime number as issued by the police force...
Date_From	Date	Start Date (systems record a start and end date. Some times if the offence is a one time offence either could be completed. We therefore have a rule that uses start date but if not present end date is the start date). The date must be in the UK format DD/MM/YYYY.
Date_To	Date	End Date as the above in reverse
Time_From	Time	As per Date_From in UTC format e.g 11:49
Time_To	Time	As per Time_From
Input_Date	Date	Format as per Date_From
Offence_Classification	From list of 100+offences	Criminal Damage, Violent crime (up to and including GBH), Vehicle Crime, Burglary, Robbery. In addition all incidents of public place anti-social behaviour.
Easting & Northing		The exact location will be "vague-d up" to a geo location roughly equivalent to a minimum 12 homes
Disposal Type	Outcome (Unsolved, Convicted etc)	Latest Disposal Type (i.e. Charged, summons or caution)
Disposal Date	Date of Outcome	Format as per Date_From

Public Place Antisocial Behaviour Incidents

INCIDENT DATE	Date
POSTCODE	6 or 7 digit exact code
SUB_CLASSIFICATION	From a list of types
INCIDENT QUALIFIER	
EASTING	The exact location should be given
NORTHING	
INCIDENT_URN	Unique reference number (cf URN for crime)

Crime Types

Each record may need further appending with a new crime groups and sub groups – such a field should be available even if not used (possibly as below shown as a rough % of total crime in one year for a typical non-metropolitan force)

Violence

1. Serious Violence (As defined under APACS terms) 0.6%
2. Robbery 0.8%
3. Sexual offences (NOT SHOWABLE TO PUBLIC) 1.5%
4. Assault with less serious injury (Defined under APACS terms) 9.7%
5. Common Assaults 7.3%
6. Public order and Harassment 5.2%
7. Other violence 0.9%

Burglary

1. House burglary 3.9%
2. Other burglary 6.2%

Vehicle Crime

1. Theft of motor Vehicle 1.8%
2. Theft from motor vehicle 6.3%

Damage

1. Arson 0.7%
2. Other damage 19.5%

Thefts

1. Shop Thefts 9.5%
2. Other Thefts 17%

Drugs 4.4%

Other Crimes 4.7%