

Free Our (Flood) Data

Owen Boswarva for correlatedrisk.com

4 January 2012

Correlated Risk

Summary:

Recent open data initiatives in the UK have focused on the four largest Government trading funds that manage the nation's 'information infrastructure'. However in this article I make an argument for releasing the Environment Agency's flood data assets as open data, in order to support wider re-use of flood information in the insurance industry as well as better public understanding of flood risk and flood protection.

Here in the UK the recent surge of interest in 'open data' policy, and the business case for wider availability and re-use of public sector information, has so far focused on the four largest Government trading funds that manage the nation's 'information infrastructure'.

The Ordnance Survey, Met Office, Companies House and Land Registry have been corralled into a new Public Data Group, and the Cabinet Office is pushing some [new open data releases](#) through channels. However beyond that the Government's direction of travel is somewhat unclear. Although there have been strong [signals of unity from the European Union](#), open data advocates should not be too confident that the political battle has been won in Whitehall.

My background is in the development and analysis of risk data for insurance applications. I broadly support open access to publicly-owned data assets as a matter of good economic sense as well as democratic principle, but I also have a particular interest in the availability of large data sets that describe geographic perils. In the UK that mainly means the [National Flood Risk Assessment](#) (NaFRA) and other high-quality flood data sets maintained by the [Environment Agency](#).

Most of the Environment Agency's flood data and mapping is available for commercial re-use, but only under restrictive licensing terms and at a cost that most small and medium sized businesses will find prohibitive. In this article I make an argument for open data release of most or all of the flood data and mapping that the Environment Agency currently makes available under commercial licence.

The Information Fair Trader Scheme - Still Fit for Purpose?

The Environment Agency (EA) is one of a [list](#) of public sector bodies accredited under the [Information Fair Trader Scheme](#) run by the Office of Public Sector Information (now part of National Archives). This list includes [trading funds](#) that have been set up specifically to generate revenue from supply of public information, but also bodies (including the British Geological Survey and the Coal Authority) that, while their terms of reference do not depend on this type of revenue, have nevertheless sought approval to license some of their data or information to third parties on commercial terms.

IFTS accreditation is subject to an approval process, based on compliance with the [Re-use of Public Sector Information Regulations](#) (RPSIR) as well as a core principle to maximise re-use of public information. In practice however most of the bodies on the IFTS list were approved or re-approved for accreditation years ago, before the economic benefits of open data were fully debated and recognised.

The main practical effect of IFTS is that accredited members are able to set licensing fees for re-use of specific data products themselves, without having to demonstrate "genuine and pressing exceptional reasons" why the data should not be made available for re-use at [marginal cost](#).

Identifying Environment Agency Flood Data Assets

The Environment Agency's [Commercial Licensing](#) Team publishes an [Information for Re-Use Register](#) on its website, which summarises a range of flood mapping and data products as well as other water-themed data sets that would be useful to anyone with a technical interest in flood risk.

Strictly speaking this is not a proper [Information Asset Register](#) as described by the OPSI, because it covers only the assets that the Environment Agency is promoting for re-use and excludes some additional resources that the EA maintains for internal purposes only or that have not yet been 'productised'. For example the latest version of the Register excludes the 'second-generation' [Flood Map for Surface Water](#) that the EA distributed to local authorities in late 2010.

I've listed the main flood-related data sets on the Information for Re-Use Register at the end of this post. The core assets are the NaFRA and Flood Map spatial data sets, which provide an indication of the likelihood of flooding from rivers and/or the sea for any area within England and Wales. In addition the Register includes separate data sets that cover historic flood events, large reservoirs, flood warning areas, and the river network. Some of this information is visualised in the [What's In Your Backyard](#) (WIYBY) mapping facility on the Environment Agency's website.

The Commercial Licensing Team also publishes a quarterly newsletter called Acorn, which provides updates on use of Environment Agency data products and on products in development. This newsletter is distributed to added-value resellers and other 'channel partners' that license EA data on commercial terms.

Many public bodies now record their information assets on the data.gov.uk website, but Environment Agency use of that resource is patchy. Although some of the EA's access-restricted data (such as the LiDAR products) is listed, at the moment metadata for most of the EA's flood data sets is [missing](#) from the data.gov.uk database.

Policy Discussions

Under the current Government the Environment Agency is [no longer allowed](#) to make policy itself or lobby publicly to influence Government policy. The Environment Agency's approach to open data will likely depend on guidance from the [Department for Environment, Food and Rural Affairs](#) (DeFRA). Ultimately whether to free Environment Agency flood data will be a decision for the environment minister [Richard Benyon](#) and his policy advisors.

As in any large organisation there are probably different opinions within the Environment Agency as to the best use of its data assets. Based on the current pricing and licensing strategy the EA's Commercial Licensing Team has a rather business-minded interpretation of the RPSIR framework as it applies to commercial re-use. However there is a counter-argument that opening up access to the EA's flood data would better support the public task of the wider organisation, and also better support public understanding of flood risk and flood protection.

In September DeFRA and the Environment Agency made a joint presentation to the UK Transparency Board. The [minutes](#) of that meeting indicate that the Board had some unanswered questions about the EA's policy for data charging. My own impression is that, at a policy-setting level, DeFRA and the EA have not yet fully understood the difference between making data available to the public and making it open data.

The EA reportedly responds to 44,000 requests for information each year, more than the rest of the public sector put together -- which underlines the interest in and potential for re-use of its data. The EA also does a substantial amount of valuable and commendable work to share environmental information within the public sector, supporting a multitude of important local authority and emergency response functions. Not least the EA has carried much of the responsibility for UK compliance with the EU Floods Directive and INSPIRE Directive, which has driven the availability of public information in the form of online maps and reports.

In the UK policy context however 'open data' means data that:

- is available to the public either for free or at marginal cost of supply,
- is obtainable in bulk, i.e. not just viewable via a particular website or service, and
- is released under an open licence -- normally the [Open Government Licence](#), which allows re-use for both personal and commercial purposes.

DeFRA Working Group on Flood Information Sharing

Last year, as part of wider negotiations with the insurance industry on long-standing issues around flood risk management, DeFRA convened a working group with the stated aim of ensuring that information on flood risk is "transparent and available to all". The group met regularly over several months as a forum for debate between insurers, the Environment Agency, the [National Flood Forum](#) and various others with an interest in flood data and information. I attended most of the meetings as a 'subject matter expert' on use of flood data models for insurance purposes.

The working groups operated under [Chatham House](#) principles intended to promote free discussion, so I won't go into who exactly argued which points of view. However last month DeFRA [published a report](#) with the following summary of the pros and cons of releasing the NaFRA flood data set as open data:

The Working Group also discussed the potential benefits and negative consequences of making the entire NaFRA dataset available to all, free of charge, and not subject to third party licence restrictions. In support of the proposal it was noted that:

- A free NaFRA dataset could be more easy to use with other formats such as Google Maps.
- NaFRA is paid for by the taxpayer, via Defra's grant to the Environment Agency, so the dataset could be made available to the taxpayer for free, including for commercial use.
- Individuals could use the data to produce many different and flexible tools, such as Apps.
- Such Apps or an interface with Google Maps could improve the ordinary consumer's understanding of their flood risk and could be easier to use than the current maps on the Environment Agency's website.
- People might more easily be able to find out about their flood risk online.
- Insurers and other businesses, such as estate agents, could have access to the most up-to-date Environment Agency flood risk assessments, free of charge and in a format that is most useful to their specific use.

- The Environment Agency could save money as it might not have to respond to so many individual requests for detailed flood risk information.

Against the proposal, it was noted that:

- Making the NaFRA dataset available for free could have a negative impact on private companies who specialise in producing and selling flood risk information.
- There are some restrictions on licensing the NaFRA dataset for not charge because some information is provided by third parties under licence.
- It could run counter to the argument that a 'one stop shop' is needed for flood risk information, in order to reduce confusion for members of the public. It could be difficult for people to know where to go to get the 'official' data.
- The Environment Agency would still need to respond to individual requests for detailed local information as it would be difficult to make the detailed local knowledge involved in mapping available automatically.

The NaFRA flood model is of special interest to insurers because the data outputs reflect likelihood categories written into the [Statement of Principles on the Provision of Flood Insurance](#), a market agreement between the industry and Government. However most of the arguments for releasing NaFRA as open data could be applied just as readily to the Flood Map (which is aligned with local planning criteria) and the EA's other flood data assets.

The pros and cons above are a fairly balanced record of the longer discussion in the working group, and most of the points will resonate for anyone familiar with the wider debate over open data. We have the argument that releasing the data would enable more innovative approaches and 'mash-ups' with other technologies, we have the argument that the public should have full access to data already funded from taxation, and conversely we have the fear that free data will disrupt the existing market, and institutional concerns about losing control of the data and its potential misuse or misrepresentation.

In my view there are four plausible reasons for objecting to open data release of Environment Agency flood data, and in each case there are compelling counter-arguments:

1. Competition and disruption to the existing market for flood risk information

Several private companies have developed their own national flood risk models for the UK, which are (at some level) in competition with the Environment Agency's NaFRA and Flood Map data products. [JBA](#), [RMS](#) and [Ambiental](#) all maintain flood models that cover pluvial flood risk in addition to risks of flooding from rivers and the sea. Data sets derived from these models are licensed and used by some insurers and reinsurers, and

in other sectors for environment assessment and property conveyancing reports. Until recently the insurer [Aviva](#) also maintained its own national flood model and licensed its data to third parties.

However I'm sceptical that increasing the availability of EA flood data would have much impact on the market for commercial alternatives. That market is inherently limited and based on the added value that the commercial flood models can provide for specific business purposes: that they include additional estimates of the financial costs of flood damage, that they take into account more sources of flood risk, that they cover Scotland and Northern Ireland in addition to England and Wales, etc. Commercial flood data is usually used in addition to Environment Agency information, rather than as a substitute for it.

As long as the Environment Agency retains statutory responsibility for flood risk management, its flood data will remain authoritative for most purposes. As a function of its public task the EA has the resources to manage a rolling programme of updates to its flood models in response to new flood defence works and local development. Private companies cannot compete effectively in that space because the economics of maintaining an up-to-date national flood model primarily for business use require high licensing costs to recoup a substantial investment. The only way to make the commercial models more competitive would be for the Environment Agency to price itself out of the market for re-use of its flood data, i.e. to waste public resources by introducing an [artificial scarcity](#).

It's worth noting here that the Environment Agency makes extensive use of private sector partners to develop and maintain its flood data models. Most of the work on NaFRA was done by [Halcrow](#) and [HR Wallingford](#), and the EA has recently [published a tender](#) for the Water and Environment Management Framework that includes a £48m spend on mapping, modelling and data services for flood and coastal risk management. Lead Local Flood Authorities also use private contractors (including JBA) to produce local flood risk mapping.

2. Loss of revenue and additional distribution costs

While I don't have figures for the Environment Agency's income from commercial licensing of flood data, I gather anecdotally that it is not a significant revenue stream. An [OPSI report](#) in 2009 reported the Environment Agency's income from information licensing (i.e. all environmental information, not just flood data) as between £2m and 5m in any given year, and set that against a then overall budget of around £1.25bn. It is indicative that income from information licensing is not mentioned specifically in the EA's [annual accounts](#).

Additional distribution costs arising from open release of Environment Agency flood data should be easily manageable. None of the data sets are large enough to require

the type of [special arrangements](#) that the Ordnance Survey had to put in place for its OpenData release in 2010. The Environment Agency has an existing download service called [DataShare](#) and, if that does not scale adequately, peer-to-peer distribution is always an option.

3. Third-party intellectual property

Third-party IP is sometimes an intractable barrier to open data release of public information and may be an issue for certain Environment Agency assets. However in respect of the NaFRA flood data my understanding is that the only third-party interest arises from the use of Ordnance Survey [TOIDs](#) (proprietary address references) in one property-level data set. It should be possible to release the spatial (GIS) version of the NaFRA data, and the postcode-level data set, without engaging OS rights.

On the broader issue, if the Environment Agency did embrace an open data approach to flood information it would follow that any future contractual arrangements with private partners to produce flood information should ideally make sure the IP rights are retained by Government.

4. Potential for public confusion

While I can understand that there might be misgivings about potential public confusion if Environment Agency flood data is available from different sources, particularly on the web, in practice I think this is unlikely to play out as a significant problem. Open data release of flood data need not imply a right to use any Environment Agency logo or trademarks, and recourse for any misrepresentation of authority would be unchanged.

There is still a strong argument for bringing the Environment Agency's flood data together with other, particularly local, sources of information -- but there's no obvious reason to suppose broader availability of the data would prevent that. While open release would encourage alternative interpretation of the EA's flood data, the underlying assumption that re-users would set themselves up to imitate the EA's own public use of the data is almost certainly misplaced. As an existing parallel, Government recorded crime statistics have been widely available for bulk re-use for some time now, without any challenge to the Home Office or the police as the authoritative source of crime statistics.

Next Steps

There are at least three upcoming opportunities in the policy process to press the argument for open data release of Environment Agency flood data:

1. In a recent [statement](#) the environment minister Richard Benyon announced he had agreed with the Treasury to sponsor further work to analyse the options for managing

the future financial risks of flooding, in the hope that Government and the insurance industry can develop a "shared understanding" in the spring. The Association of British Insurers is [lobbying](#) Government to subsidise flood insurance in high-risk areas. The availability of flood information is only one concern within a much wider negotiation. However releasing Environment Agency flood information as open data is a relatively affordable concession that DeFRA could make to the insurance industry.

2. The UK Transparency Board has a wide range of items on its agenda. However flood data has been identified as a priority in public consultations and surveys on open data, so it may be that the Board will follow up on the questions raised during the September meeting.

3. According to its previous accreditation report, the Environment Agency's status as an Information Fair Trader is subject to re-verification by the OPSI no later than November 2012. I hope that the OPSI will take that opportunity to closely examine the EA's approach to data pricing and consider open data licensing as a better approach to maximise re-use of EA data.

Appendix - Environment Agency Flood Resources [Information for Re-Use Register](#)

Following is a brief summary of the flood-related information assets currently available for commercial licensing from the Environment Agency:

NaFRA (National Flood Risk Assessment), a spatial data set that provides a "broad-brush" assessment of the average likelihood of flooding inside each 50m by 50m area of land within the flood plain of any main rivers or the sea. This data set is based on a system of three risk categories (low, moderate and significant) that were agreed some years ago in discussion with the UK insurance industry, and includes the remedial effects of flood defences.

The Flood Map, a spatial data set that also provides an indication of likelihood of flooding within the flood plain of rivers and/or the sea although based on different criteria from those used by NaFRA. The Flood Map has been pieced together from local models, so the level of detail is less consistent than in the NaFRA model. The Flood Map is the source of Flood Zones 2 and 3, which determine when consideration is given to flood risk in the local planning and development process. Information on flood defences and areas benefitting from flood defences is available with the Flood Map as separate layers of data.

Historic Flood Outlines, a spatial data set that shows the extents of flooding to the land from individual river and coastal flood events from 1947 to the present, and Historic Flood Map, an additional spatial layer that provides a combined outline of all areas known to have flooded in the past.

Detailed River Network, a spatial data set that shows the centrelines of rivers in England and Wales, derived from Ordnance Survey mapping but annotated by the Environment Agency staff with additional information on flow direction and path. Additionally Real-time and Near-real-time River Level Data is available as a feed based on information from the Agency's network of river gauging stations and river level monitoring sites, and both Daily Mean River Flows and Monthly Maximum Instantaneous River Flows are available as time series data.

Large Raised Reservoirs, a data set that shows the location of any reservoir capable of holding more than 25,000 cubic metres of water above the adjoining land, and Reservoir Flood Map. Maximum Flood Outline, a spatial data set that shows the "worst case scenario" of flooding if any such reservoir is breached.

Flood Warning Areas, a spatial data set containing outlines of portions of the floodplain that contain communities at risk of flooding, and Flood Alert Areas, a spatial data set containing outlines of larger regions of the floodplain that may be affected concurrently by lower impact flooding. This spatial data can be used in conjunction with status updates from the Environment Agency's Flood Warnings Live Feed.

The Environment Agency has also developed a 'second-generation' Flood Map for Surface Water, which has been distributed to Lead Local Flood Authorities but is not currently available for commercial licensing. (This may be because of potential conflicts with the Preliminary Flood Risk Assessment reports recently produced by those LLFAs, some of which made use of the Agency's surface water mapping but also took into account local records of surface water flood incidents.)