IN A SCRAPE OVER POS

Since the closure of Ernest Marples website, how the Royal Mail should share its postcode data has been hotly debated. Should it be free or licensed with added value, asks **Michael Herson**

IWR's November 2009 feature on "Bashing the bootleggers" sparked a broader debate among public interest groups and bloggers on whether public sector data should be given away free – on the basis that taxpayers are 'entitled' to use it. This debate was further stimulated by the government's announcement in December of a consultation into greater access to Ordnance Survey data.

This itelf is part of a wider government campaign under its 'Frontline First' initiative to promote transparency across a whole range of data including NHS Choices, public transport and Met Office public weather service data – all of which can be viewed at www.data.gov.uk.

Issues underlying postcode data, however, are more complex. For a start there are two main files involved; the PAF file (postcode address file) which was conceived by Royal Mail when postcodes were introduced primarily to aid the efficient and effective distribution and delivery of mail. This data is widely used today by a whole raft of businesses to validate addresses for everything from deliveries to insurance quotes. The less well known 'PostZon' file, consisting of

eastings and northings, is more a tool to calculate the 'nearest' (such as restaurant to a tube station etc) and to aid (for example) courier companies in calculating their most effective routing to save fuel and time.

SCRAPED INFORMATION
The Ernest Marples' website, set up by Harry Metcalfe and Richard Pope, has temporarily closed following a challenge from Royal Mail about the source of its postcode data.

Metcalfe maintains that the site was originally set up with no commercial objective – ie with no fee or advertising based model. It was aimed at citizen-focused websites, charities and 'back bedroom' start-ups, rather than the serious business user.

He explains that they never had a copy of the PAF or PostZon files, instead adopting an IT technique known as 'scraping', whereby the data is taken from other mapping websites. Their skill was to write a programme to "extract the co-ordinates of a pin on a map". Metcalfe points out that this was done on a 'request basis' rather than via publishing data and regards it as adding a "translation layer". The issue of who owns the rights to 'scraped' data is a murky area, but at

the end of the day someone is paying for it in the form of a licence fee.

Metcalfe passionately believes that postcode data should be made available free because of "the increased revenue that government will get in taxation as a result of this information being free will exceed the cost of maintaining it". His premise is that businesses in 'UK plc' will ultimately be more efficient.

Metcalfe argues that "this data is too valuable to be locked up behind fees that are impractical if you're trying to innovate with government data".

He maintains Ernest Marples never intended to target fulfilment businesses who would in any case gain limited value out of just the physical co-ordinates of the postcode; it is only full PAF file that provides the address and confirms number of the house.

Metcalfe doesn't believe that PAF re-sellers who add value have anything to fear from the data being made available free – "the ones that really perform a good service for their customers will be fine, they will make more money; the ones that don't will go the way of the dodo".

Metcalfe takes no credit for the government consultation on Ordnance Survey data (which is closer to PostZon), but admits his own plans are now "on hold" until after April.

Of course, what you are buying with a licensed copy of either the PAF

businesses that have opened or closed; change of address etc. PostZon isn't updated quite as often but, nevertheless, it is the updates give the value – and that are the root of the maintenance cost.

Guy Mucklow, managing director of PAF re-seller Postcode Anywhere, points out that Royal Mail employs 80 people to look after the PAF data and estimates it costs £20m a year to maintain; so if the data was made available free of charge, who would bear that cost?

"The UK is one of only three countries in the world that have an address database so refined it allows you to find the final address in under seven key strokes. The strength of the PAF database is that it has spawned a data management software industry in the UK, address management, that is world-leading. We feel that there will always be a need for applications to manage and maintain data," comments Mucklow.

Charles Arthur, technology editor at The Guardian, ran the 'Free our data' campaign run through the newspaper in 2009.

He too is pleased with the outcome on the Ordnance Survey data: "this is a lot of the things we want". He shares Metcalfe's view that more tax will accrue to the Revenue as a result of more businesses having access to the data. As an overseas example of 'free' data, he cites the GPS signals that the US government provides to underpin the sat nav technology.

With only 37,000 businesses in the UK licensed to use the PAF/Postzon files, Arthur believes the price model needs to be reviewed - "you could argue that actually there's a lot of

> ...this data is too valuable to be locked up behind fees that are impractical if you're trying to innovate with government data

file or the PostZon file from one of Royal Mail's re-sellers is the right to receive the updates – PAF has an astonishing 100,000 updates a month, mostly from data collected by 'posties' on the street; new builds;

benefit in just reducing the cost to the businesses of doing this - if you put these costs in front of businesses it's an impediment to them starting up".

Emma Gooderham, managing director of Allies, another PAF re-

OSTCODES
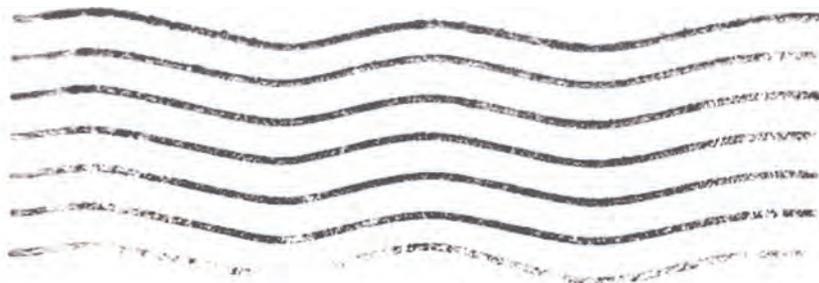
seller, also sees price as a barrier to some market entrants: "There is a perception about PAF that we need to overcome in the industry – people think it is expensive". She believes that although the licence has been quite complicated in the past the simpler version Royal Mail is launching in April will be easier to license for single users.

Arthur says the outcome on the Ordnance Survey data appears to meet that objective – "it lowers costs, increases the potential for activity in the economy, and creates all sorts of side benefits which might not be immediately visible, but the real difficulty is measuring it". Arthur cites Linux as an example of a free operating system that created new business opportunities within the technology space.

IWR columnist Tim Buckley Owen, who wrote the November story, feels "it's part and parcel of a very much bigger issue about the extent to which public sector agencies should be generating revenue from the information that they hold or making it available as a public good." Buckley Owen compares it with the Trading Fund model: "It's better for them to give the stuff away at marginal cost to whoever wants it as this collects more tax."

PAF RE-SELLERS
Terry Hiles, managing director of Capscan, a leading Royal Mail PAF re-seller, explains how to add value: "We make the data usable, so we build in an interface at the front end that allows people to search through the 28 million or so addresses in the Postcode Address File". Capscan also supplies data quality management software that updates files with postcode changes. He admits the firm very rarely sells the raw data because "it is constantly changing", and believes the historical low take-up is because a number of organisations "don't recognise the importance of using tools such as this", so for him it is an awareness issue.

Hiles recognises that the huge PAF file costs a lot to produce and maintain, and questions who pays the piper. "Those that are not contributing to that and therefore to the costs, are in fact leaving those that are paying with a higher fee than they might otherwise have to pay".

He points out that Royal Mail operates in a price controlled environment generating no more than £25m a year from PAF with attendant maintenance costs of £20m a year, with its margin restricted by Postcomm's mandate to 10%, so this is certainly not a 'get rich quick' product. But he concedes that the PostZon file provides limited value to fulfilment businesses unless they are involved in the logistics side of deliveries.

As for scraping, Hiles says: "It rather suggests that simply because you can scrape information off the web, there's no intellectual property rights attached. That's clearly wrong in the same way as illegally downloading songs or copyright works is. I'd have no problem about the data being used for free provided all could access it for free. I do have a problem about a site that basically uses data that others are paying for."

Hiles deals extensively with overseas markets and says they don't obtain data free from the US and Canada; they pay an up-front licensing charge which works out cheaper – "the licensing obligations are less onerous… it works out much, much cheaper in the sense that you pay an upfront fee and you can sell it to 1,000 licensees".

Gooderham of Allies agrees that the same rules are needed for everyone in the industry, and backs the action against Earnest Marples. Her firm's flagship product is 'PostCoder': "It is quite a complex address matching engine that looks at missing and misspelt address elements to really cleanse an address database as quickly as possible in an automated manner" explains Gooderham. PostCoder also offers a 'nearest' service which appends grid references of branches etc when people are looking up the nearest store to where they live. Gooderham feels the current licence is "quite complicated" but is looking forward to the more simplified version for single users available from April this year.

Postcode Anywhere's Mucklow sums it up: "The bottom line is that PAF generates significant efficiencies for the UK – efficiencies we need now more than ever. Those that benefit most are often in the financial and services sectors: government should consider carefully the relative value to these sectors when considering the cost of PAF, especially in the current climate". But he adds, "A fundamental principle of UK fiscal policy is that those who benefit most from national assets pay more. Why would PAF be any different?"

On the issue of the 'scraping' Mucklow feels it is more the question of responsibility for collecting and maintaining the data "…although scraped, copied and derived data is a complex issue, we must remember this data costs real money to collect and maintain and has value to users".

He points out that the UK is not unique and "all overseas databases cost money to varying degrees … the Royal Mail data is an international leader, providing low cost data

> ## It's undeniable there will be innovative use and new users, particularly very small ones, might be enticed

bearing in mind its high quality. It's not free but it saves companies money and is certainly worth paying for."

Ian Beesley, chairman of the PAF Advisory Board which is part of the regulatory environment, feels a 'pay-as-you-go' model might be fairer; suggesting there might be "payments at very low unit cost but which generate enough revenue to maintain the required quality". He is a supporter of increasing usage – "It's undeniable there will be innovative use and new users, particularly very small ones, might be enticed". However any attempt to give it away for free would raise question marks about the quality in Beesley's mind: "Quality is an important driver of costs but it's vital in data and I think the prime concern of the Board will be to maintain the quality of the PAF file for those who, putting it crudely, already pay to use it".

ROYAL MAIL
There is evidence that Royal Mail is listening. Giles Finnemore head of marketing for PAF, feels Royal Mail is now addressing customer concerns. "We've reduced the entry price for using PAF on a website to just £1 per 100 uses. For those doing 400k or more uses on a website the price is capped at £4k". Finnemore adds: "We've also reduced the entry price for use in internal systems. Two years ago it was £150 for unlimited use; from April this will be £75 with a new pay as you go entry price of just £8 per 100 uses. This makes using PAF much more affordable for small organisations and start-ups".

Finnemore believes this will create more opportunities for re-sellers by introducing new solutions targeted at low volume and infrequent users. Effectively, this moves more towards a pay-as-you-go system which meets some customer concerns and enable re-sellers to reach a broader market.

He says: "Royal Mail isn't in this to make money. The price changes are revenue neutral," He claims they just cover the cost of maintaining the file, including gathering and validating the data and keeping it up to date. The regime enforced by Postcomm is sufficient, he says: "PAF is regulated to ensure that we treat all organisations the same, be they commercial or non-commercial… they are all charged exactly the same price for the use of the data".

As for the relationship between Royal Mail and Ordnance Survey: "We licence each other's data. We licence some of OS's data and put it onto PostZon, and it licenses ours and puts it into Code Point."

There is clearly a case for maintaining the data file at the proper level of quality: re-sellers then add value to it and ultimately UK business operates more efficiently – reducing carbon emissions in the process, which no-one has yet quantified. And since the 'pay-as-you-go' system to be introduced by Royal Mail in April does embrace smaller users, is it more a case of 'tweaking the model' by improving awareness and reducing the entry price than going overboard and making everything available free of charge? ∎