

the GISPro interview



GiSPro: Steve – what is your background?

SC: I studied physics at university but never finished my degree. I then worked for a US company for 3-4 months before starting my own company with two friends. I had lots of ideas and there were interesting things going on. We decided to quit our jobs but I was the only one who took the plunge. It was hairy at times but worked out OK and later I was joined by Nick Black, my current business partner.

GiSPro: How did OpenStreetMap (OSM) start? Were you passionate about maps or was it simply an IT challenge?

SC: The latter. I was living in Central London and had an old laptop and a new and exciting GPS unit. It seemed an interesting thing to do – to look and see where you are on a map; but you couldn't, because there was no map data. I figured out it could be done as a collaborative effort, like Wikipedia. Lots of people had the idea but no one had done anything with it.

GiSPro: Did you imagine then what OSM has become now?

poured in; but they focused on the technology side and ignored the social part. I never cared about getting the technology perfect. A big aspect of getting OSM off the ground was the mapping parties: getting drunk and arguing with people.

GiSPro: Is that a model that still works as numbers have grown or has it been hard to maintain the fun element?

SC: As far as I'm concerned, it's still a lot of fun. Where it starts to break down is where an area is totally mapped. What do you do then? But it's a big planet – there's a lot to be mapped, even in the UK. Mapping parties are still the model. In fact they are essential. It's not about mapping an area, i.e. mapping parties don't do "what they say on the tin". We just map a small area and clear off leaving a few local converts with the incentive to finish it off.

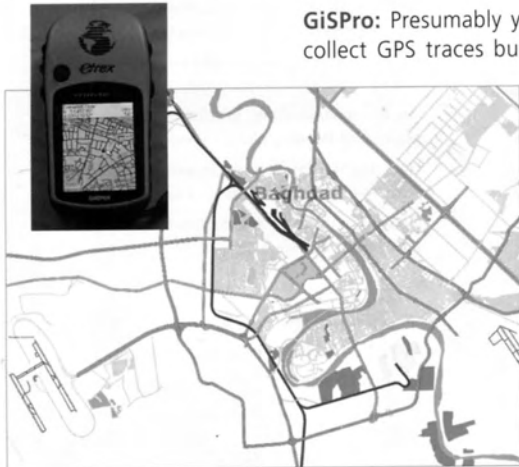
GiSPro: Your users are growing at an enormous rate. Are you concerned that the majority will submit one trace and move on?

SC: What you want is a histogram of the number of users against the amount of mapping they do. There's

Concentrating on the mechanics

OpenStreetMap, the community mapping project, is coming of age. It began with a few enthusiasts collecting mapdata using handheld GPS units – aiming to create a free, editable map of the world. Today, users (contributors) are growing exponentially, it has large corporate data donors and recently held its first AGM. *GISPro* talks to OSM founder, **Steve Coast**.

It all began with a few enthusiasts with handheld GPS but one got bored and went off to map Baghdad! Image below rendered and served using Osmarender and tiles@home (<http://wiki.openstreetmap.org/images/2/2b/Baghdad.png>)



SC: I had an intuition the idea was correct and would take off. There were many similar attempts – collaborative ways of breaking up monopolies. But, no, I can't believe how big it is now: 11,500 users, growing by just over 1000 per month. Until last month it was mainly me running things – the servers, writing server side-code, going to most conferences, organising stuff – but now we're a large enough group with those at the top wanting to help run it. That can be a problem with an open source project: lots of opinions but very little action.

GiSPro: Presumably you get lots of folk wanting to collect GPS traces but fewer to take on the coding and administration.

SC: The latter is much more difficult. . . and boring. I happen to enjoy it. It's not sexy writing server-side SQL code but it's what makes the whole project run. That's why similar projects, such as WikiMap and LondonFreeMap failed – they concentrated on the flowery stuff rather than the basic mechanics. There's a long story of failures; money

definitely a sense that people ramp up to a huge amount then exhaust themselves. But they carry on. Dave Groom lives on the Isle of Wight and finished off that project. He eventually got bored and did Baghdad. Now he is concentrating on woodland in southeast England. There are some very dedicated – and addicted – people. A relative few do an awful lot. But that's fine. It's about local knowledge...and maybe getting someone to notice just one mistake in a place they know.

GiSPro: Is the lure similar to attempting to complete a vast jigsaw and, if the analogy stands, are there bits like the sky that no one wants to do?

SC: Nobody wants to do council estates. But apart from those socio-economic barriers – for places people aren't that interested in visiting anyway – nowhere else gets missed.

GiSPro: Do you think that interest might drop once the emphasis changes from new mapping to updates, and that OSM data will become a snapshot in time?

SC: It's taken us three and a bit years to get here but with the lessons learnt and the software developed you could probably do it all over again in 12-18 months. But in terms of maintenance, so far, it has worked out really well. As people finish their areas they notice new detail that can be added. I said we'd never get building outlines but adding them from imagery is easy and lots of people are doing it. They are not perfect but they will

get better with time. As soon as an area is seen to be complete the definition of "complete" changes.

GiSPro: Is that another problem in itself – the data will never be nationally consistent or complete because there is no control over releasing a particular data layer?

SC: It's important to let go of the concept of completeness. A lot of GIS people and traditionalists in the industry treat completeness as the gold standard. But no map dataset is complete. Not only that; all map providers put things on that are not meant to be there. One street per 100sq km doesn't even exist! The general public assumes a map is complete. I certainly did. But it's nonsense – things change so quickly. Detail is good on the A-Z (let's not go into the cartography) but we capture more detail of things like footpaths, railways, routeable rivers and so on. The same thing happened with Encyclopedia Britannica. Compared to Wikipedia, the accuracy was shown to be about the same. One costs a millionth of the price and destroyed the other's business model.

There's also completeness related to coverage. There's an assumption you cannot sell a dataset of the UK unless you have the entirety but there's a move towards paying per transaction. If 95% of the time users are viewing London data then they can get it for free, only paying when they need to view Winchester, for example, where OSM coverage is currently incomplete. There's going to be a phase change, especially when services are put on top of OSM too.

GiSPro: The Ordnance Survey show little sense of a threat to their business model, from OSM serving the "cheap and cheerful" market if you like. What's your take on that?

SC: The OS has some really good people. But it's like an oil tanker: it takes a while to turn. Inevitably some will get it and some won't. But the whole concept of OSM is not anti-OS. LondonFreeMap was more about being anti-OS rather than being pro-data. It doesn't matter what OS think, as they're not our target audience. Many geodata companies do get it: AND¹ just donated a whole country of data [the Netherlands].

GiSPro: What do AND get out of the deal?

SC: Publicity. They learn a lot from us – in one day they find out what it took us three years to learn. They get to play with concepts. Also, maybe they are not making money out of the Netherlands dataset – it's a market saturated with three big players: Navteq, TeleAtlas and AND. The Dutch government is also going to release all their data in 2009 (although it's perpetually being put back). By the way, they've released their own application, MAP20.org, using the same dataset and an OSM lookalike interface where people can go in and fix things.

GiSPro: Aside from the Netherlands situation, the UK leads the way with OSM. Is there something unique in the national psyche that sits easily with the values and vision of OSM?

SC: Sometimes there is a lack of willingness to run mapping parties. For instance, in Germany there is a huge emphasis on code; they've made an enormous investment both in server-side code and our main editor, JOSM. But very little happens on the social side – which is the key bit.

GiSPro: What do you think your average user would be doing if not contributing to OSM? Hacking into the Pentagon, going off geo-caching?

SC: There really isn't a typical user. Some run large software firms; some are unemployed people doing it day in, day out; some are politically motivated; some are retired and looking for a practical outlet; others just want an excuse to get to know their local area. We are a broad community and hard to classify.

GiSPro: Are data volumes becoming a problem? Do you need to consider using a professional service to manage the data, providing 24-hour cover, and so on?

SC: Where would the money come from? Money goes on essentials. There are breakthrough moments in our evolution. The current configuration operated well for 1000-10,000 users, so now we are having to scale up again. Wikipedia went through exactly the same process.

GiSPro: Does the human infrastructure have to evolve along side the technology with growing pressure on your own time and resource?

SC: My strategy is: when someone is doing good things, get out of their way. Lots of people have opinions but when you see some effort then just let them get on. So far it has worked well as a strategy. I'm not sure about the future, as I'm not so much in control to make those decisions.

GiSPro: Is it hard to relinquish that control?

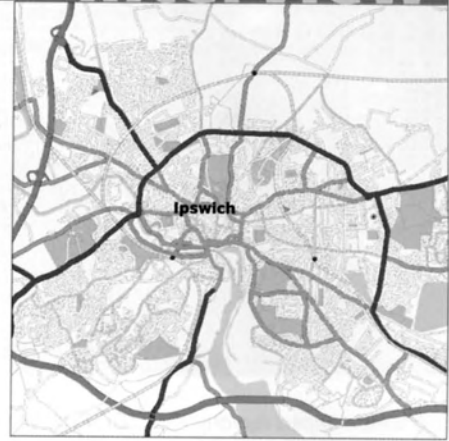
SC: It's strange because I'm so used to running it. But it's also good to see people taking it on for themselves and seeing them adopt the same the approach that I have – even though I wasn't trying to choose yes-men.

GiSPro: Do you have contact with Wikipedia and try to learn lessons from them?

SC: I've spoken to Jimbo Wells a couple of times. He's only interested if we become part of the MediaWiki Foundation but I don't like bureaucracy. They came up with a great idea, built Media-Wiki software then nothing happened. They built a hammer and now everything's a nail. They've lost the willingness to do anything new. I use Wikipedia every day and it's free. Now move on and do more interesting things. That's why I like OSM. There's constant evolution to meet new requirements.

GiSPro: What's your relationship with the Free the Postcode campaign?

SC: I started Free the Postcode (FTP) and when the number of emails got annoying I handed it to the



Above: Ipswich – a typical example of a well mapped town.

(<http://wiki.openstreetmap.org/images/e/ed/Ipswich.png>)

“
A big aspect of getting OSM off the ground was the mapping parties: getting drunk and arguing with people.
 ”

the GISPro interview



A well mapped region around Amsterdam using AND's data.

guys at NPEmaps (New Popular Edition Maps). What they do is really cool. FTP did well to stir up people, especially in Royal Mail. It's achieved its aims. A shame they don't want to make any money out of it.

So I designed into the system that everything is tagged to the user – a bit like Wikipedia – but to a lower level. To bring a copyright claim, someone would have to prove we caused financial harm and had based our maps substantially on their data. We would need money, which we don't have, to pay lawyers to answer that claim. Therefore, the model has been to ask them which roads they think are derived from their map. That's incredibly hard to prove. But if they say we have copied a street that could only have come from their map, we can look back to who input it. And even if they are a large contributor and swear blind they didn't copy it, we can delete the data and ban the user. Then it is case closed. They get to stall us by three months and give us an enormous amount of publicity – which incidentally, I would like!

GISPro: What about your finances? Do you have a steady income stream?

SC: Yes and no. We've had some very large anonymous donations recently. And we have great sponsors in the form of Multimap, Nestoria and Rightmove.

GISPro: What is Multimap's motivation?

SC: It's partly being cool, partly publicity. And it's a good investment for them; they don't want to pay for map data in the future and won't have to. Monetary motivation shouldn't be over-rated though – it's only part of what is going on.

GISPro: You had your "month of OSM" when you asked for donations to support a month of full-time work on the project. Did it leave you frustrated at the slow progress outside of that period?

SC: That came about because I lost the contract I was working on and was brainstorming what to do on the way home from work one day. A lot was achieved but it wasn't sustainable. It has to be a community effort, not a one-person show so that ideas percolate. But it got us over the hump of moving from the old Ruby code base and was a proof that such things can work.

GISPro: Does the "month of OSM" go against the grain of the volunteering model?

SC: That's a bit of a myth. In the beginning, it was only me volunteering anyway – apart from the massive contribution by Imi [Immanuel Scholz] in writing JOSM. Also it was never about the notion of freedom for freedom's sake. The economics of peer production models have been studied and are posited to be more efficient than markets or firms. I always wanted us to make money. I see what Wikipedia does every six months: grand appeals for server resources, the thermometer bar creeping up to a million dollar target. It's horrible. If they sold the data they wouldn't have that problem.

GISPro: What about Yahoo and all their imagery? What's the deal there?

SC: Their API came out after Google's (everyone's did). It gives Yahoo prestige and needed little effort on their part. They helped us and we are grateful for it but we still need to go and figure out the street names.

GISPro: Is it likely that without control of users you will get copyright issues, especially over street names?

SC: When I designed OSM, I knew from my knowledge of digital rights, and take down notices from DMCA² etc, that we could get a spurious claim at any moment.

GISPro: But with tagging would someone go after OSM or the individual user?

SC: Depends on your interpretation of who the publisher is and copyright law. They should go after the OSM foundation, which is a limited company. Some time back I was sent a very large number of postcodes by someone who said he had been collecting them for ten years as a courier. I thought it was remarkably prescient. Someone ran a comparison for us with the real data. Unlike our 10-50m accuracy, his were within 30cm (ie clearly stolen) but with a systematic offset that plotted out as a map of the UK – like a digital watermark. There were also some fictitious postcodes. I wrote and asked him where he got the data from and never heard back. I talked to some IP lawyers who said it was almost certainly attempted entrapment by a large company trying to nip things in the bud.

GISPro: How do you protect against other malicious intent?

SC: The same way as Wikipedia. There are many tools to find out what's happening in your area. Ours aren't as sophisticated but we don't have the problem of spam or differing views on whether Jesus exists, i.e. edit wars. We may have a problem with the border of N. and S. Korea, in which case we'd fall back to the UN definition like everyone else. We haven't had anyone trying to delete things for twisted reasons, there's no motivation.

One thing we do have is the occasional mess up in classification, for example, deciding between unclassified, private and service roads.

GISPro: Is classification a big issue particularly as you are creating a "Planet Map" rather than a "UK Map"?

SC: I knew I wouldn't be able to design an ontology that would scale up well for the UK, let alone beyond. Del.icio.us (Firefox's linking website) proved tagging works as a way of allowing communities to design things bottom up. Our own tagging system has evolved and works well.

Our first pass used the class tag: everything was class_road, class_rail and so on but it wouldn't scale.

“

... it was almost certainly attempted entrapment by a large company trying to nip things in the bud...

”

Then we started tagging with highway in general and split out from there. For example, it used to be that we would use highway_footpath for a footpath. But now there are many types of footpaths so that value highway_footpath splits out and becomes its own key. It is a big aspect of why OSM works and other systems haven't – they've not been able to design a good enough ontology. I originally intended to develop tag equivalencies, e.g. tagging highway_motorway for the UK would be the same as highway_freeway for the States. It needed a higher level data structure but didn't happen as people didn't seem to need it. There are a few edge cases: unlike here, there is no concept of bridleway in Europe. Inevitably we get differences of opinion and endless discussion, which is a good thing as it creates our community.

GiSPro: What's the current thinking about licensing? I noticed a big debate took place at your recent State of the Map conference.

SC: For any open source project there are three definitions of freedom. Public domain – you can do anything with the data; a BSD licence, i.e. do anything with the data (mash-it etc) as long as you attribute it; and a GPL licence, i.e. do all of above so long as the stuff you are doing is licensed in the same way (a viral licence). We have all three camps but at the outset I decided on the Creative Commons GPL licence.

A second issue, and maybe just as big, is whether the Creative Commons licence refers to data. It took 20 years to test GPL in court but Creative Commons hasn't been tested – not on data (it was originally designed to protect works of art). It turns out after a lot of beer and talking that it may not be the right one. So far most people are happy although some threaten to quit the project over it.

GiSPro: Have you had other issues where people get that heated?

SC: First, understand the reason – people pour their lives into it. It's their baby. If it changes they want to know why. In November we changed to a new renderer: Mapnik. Previously, we had white lines on a crappy Landsat image. It was good for rapid mapping, nearly live, but had no infrastructure for continuous rendering. Now people complain if they can't see their changes quickly enough, however good the cartography (and Mapnik is as good as any renderer around). It reflects the passion and sense of ownership. Licensing though touches a raw nerve unlike anything else. Adverts on mapdata is another issue – lots of people don't want it. Also occasionally, the "evil Steve" conspiracy theory – me running the foundation in a way they don't like!

GiSPro: Is it still seen as your project?

SC: Maybe externally, but that's changing. The Foundation began in order to separate OSM from me. Now it's an amalgam of a funding body, an entity that owns things, a branch of responsibility if we get sued and an organisational body.

GiSPro: How do you make spending decisions?

SC: We have a treasurer and quite a few committee members. There is an overall mandate that splits our money three ways: servers, publicity and legal costs. With about £25 (or a claim of financial hardship) anyone can become a voting Foundation member. And the AGM goes a long way to maintain an air of accountability and responsibility.

GiSPro: Do you have a wish list for OSM in the future such as routing?

SC: The main thing is removing segments from the data model. We evolved to have "nodes" linked by "segments" to create "ways". The community found we didn't need segments. It'll be a simplification and time saving in processing the vast amounts of data involved.

Routing will happen very soon. The data model was designed for routing, that's why it's fully connected. Look at any standard GIS format and it's not designed for topology. In OSM things have logical IDs – they are either connected or not. The tagging already allows a wide variety of routing attributes such as width restrictions and height restrictions. I'm surprised it's not happened yet. It just needs someone to write the software – the data's all there. And as soon as people start routing there'll be another phase change. It'll expose flaws in the data. If I route from here to Reigate it'll probably take me down the Thames!

GiSPro: I hear confident predictions from outside of OSM that your data will rival OS Mastermap in a couple of years. Are you that confident and how does the future look?

SC: It won't be that long. It's not a matter of confidence – it's just inevitable. Our data will only get more complex, more detailed. Eventually we will have an editor for squaring buildings, probably before GPS accuracy achieves that directly. And people's attention will continually turn to new objects to capture that would be ludicrous (i.e. unaffordable) for companies like TeleAtlas to match.

The world is open for anyone to take our data – any Opendataset – and add data services on top. But it's not just the data. The tools built by NPENet, FTP and OSM are all OpenSource. You can use them to build your own map – and to a better accuracy, if you have better imagery.

GiSPro: Steve, thanks for your time.

For the latest on OpenStreetMap turn to our report on AGI'07, page 12.

Notes Nick Black joined the conversation towards the end and for simplicity occasional comments of his have been included as if from Steve Coast.

- 1 Automotive Navigation Data
- 2 Digital Millennium Copyright Act



Above: Ljubljana Mapping Party – not all mapping parties are on bicycle. (http://wiki.openstreetmap.org/images/ele3/LjubljanaMappingParty_HorsesEqualsYes.jpg)

“

Inevitably we get differences of opinion and endless discussion, which is a good thing as it creates our community.

”