Local to European SDI – "Mash Up" or Professional, Industry Strength Infrastructure?

Keith MURRAY and Andrew TRIGG, United Kingdom

Key words: Spatial Data Infrastructures (SDI), Standards, Spatial Information Management (SIM), Digital National Framework (DNF), European SDI (ESDI), INSPIRE & ICT.

SUMMARY

Access and the use of geographic data and information it is growing fast across all parts of the globe. The launch of Google Earth, making mapping and imagery possible on every Internet connected computer has introduced many new users, encouraged by an easy to use and very effective user interface. This has stimulated third parties to build on the Google model by incorporating their own information in a variety of combinations or "mash ups". At the same time the take up of sat-nav units by the general public is currently doubling year on year. Never before has the general public had so much digital mapping at their disposal.

At the same time the European Community is developing legislation and implementing rules to promote better integration of information across Europe, driven by environmental needs. This is acknowledged as an ambitious task given the diversity of member states, public organisations, data types and structures as well as maintenance regimes and quality levels in evidence across Europe today. At the national, regional and local levels digital mapping, and increasingly geographic information, is extensively used in the transfer land and property, life-critical applications in emergencies, property taxation, asset management, the recording of social events of all kinds (crime, health, births, housing conditions etc) to support investment in deprived or needy areas and so on.

By drawing parallels with new developments in the use geographic information with mainstream Information & Communications Technology (ICT) best practice, the paper concludes that the data provided by Mash-Ups today appear to meet the needs of several user communities. At the same time professionally engineered geographic information (GI) is also emerging and is increasingly underpinning business and life critical systems and services.

However it may be possible that there might be cross-over at some stage in the future, should the limitations in data integrity in Mash-Ups become intolerable. As we note in the ubiquitous services of banking, supermarket checkouts, mobile telecoms, the user interface has to be very simple. To make mainstream GI more accessible to a wider range of professional users such an interface may well benefit these users as well.

Local to European SDI – "Mash Up" or Professional, Industry Strength Infrastructure?

Keith MURRAY and Andrew TRIGG, United Kingdom

1. WE ALL START FROM A DIFFERENT PLACE

Over the past decade the movement towards the creation of comprehensive Spatial Data Infrastructures has continued to gain momentum. There are now many initiatives at the global, national and local level geared towards the creation of infrastructures designed to provide a comprehensive, integrated resource for public and commercial use within services and applications.

The Global Spatial Data Infrastructure (GSDI) Association is "an inclusive organization of organizations, agencies, firms, and individuals from around the world. The purpose of the organization is to promote international cooperation and collaboration in support of local, national and international spatial data infrastructure developments that will allow nations to better address social, economic, and environmental issues of pressing importance" (www.gsdi.org).

As such it represents the aspirations of many subsidiary organisations around the world who recognise the advantages and benefits to be realised from the combination and organisation of data, from different providers, in such a way that allows their re-use in 'value-added' products.

The practical realisation of these spatial data infrastructures has proven to be very complex with variable results. The path to a truly integrated, harmonised Spatial Data Infrastructure has many obstacles to be overcome – incompatible standards, data frameworks, resolution, currency, Intellectual Property claims notwithstanding.

In Europe the drive towards the creation of a European Spatial Data Infrastructure has been articulated and led by the European Union's Infrastructure for Spatial Information in the Community (INSPIRE) Directive, which seeks to create a harmonised environment for the provision of interoperable spatial data. A very inclusive approach has seen many organisations across Europe become involved in the definition of the fundamental components of the required infrastructure, viz: metadata; data specifications; network services; data and service sharing; and monitoring and reporting.

Recent initiatives to provide spatial data to an ever-hungry audience, whether as consumer or business, have tended to provide their services through one of two distinct routes – the rapidly emerging mash-up or a more formal, engineered infrastructure. The background and relative benefits of these two approaches is explored in this paper, along with recommendations for the most beneficial and pragmatic solution needed to ensure the maximum benefit to the intended audiences

2. MASH-UPS

The term 'Mash-up' has become a standard part of the vocabulary of the new e-enabled world, but the term is used to represent a variety of data sets and services. Wikipedia defines a mash-up as "a website or web application that uses content from more than one source to create a completely new service". Mash-ups are creating a fundamental change in the way that services are delivered to the consumer. The process of combining data and information from multiple sources has become faster and easier through the availability of standard and simple to understand APIs. Web authors with little development experience are now able to create apparently complex and novel services, using a variety of data from multiple providers, with comparative ease.

Prominent amongst mash-up innovators have been those which provide the ability to combine data through a common geographic location. The most obvious and well known frameworks for map mashups are provided by Google and Microsoft. The appearance of these major league players signals the potential value and significance of such services. Google Earth and Live Local offer unprecedented access to global data, especially large scale imagery, along with the ability to mash other data sets against the base data provided on their sites. In addition to the obvious standard services which are offered to find my nearest bank, theatre, restaurant etc, there are thousands of mash-ups of genre specific data targetted at specific interest groups.

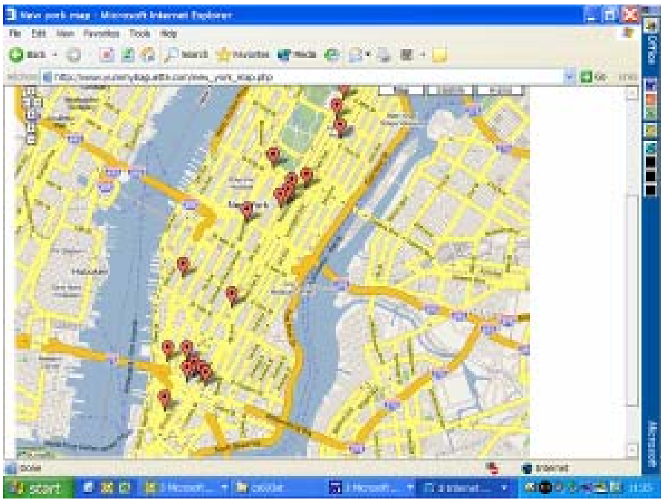

Figure 1: Mash-up example.
http://www.yummybaguette.com/new_york_map.php

http://googlemapsmania.blogspot.com/ and http://www.passthepoi.com/ both track recent innovations using this technology.

A quick scan of the sites will reveal mash ups dedicated to subjects as diverse as tracking the current location of the mash up's author, e.g. http://www.timhibbard.com/wherestim/,

and

sweet shops in New York (see Figure 1).

Whatever the application or interest group served by these map mash-ups, they all have one thing in common – the absolute and relative accuracy of the data used is not important. As long as a feature or event can be geo-located accurately enough to be meaningful, the absolute positioning is of subsidiary importance. Indeed some of the data provided within the Google 'base data' is not internally consistent.

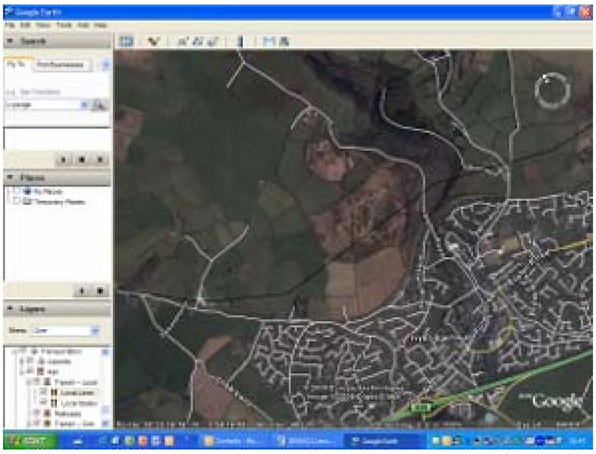

Figure 2: Google Earth – Ivybridge, Devon, England.

Try overlaying road or rail data onto the imagery in a rural part of England and the chances are that the two data sets will not coincide. The example shown, from Ivybridge in Devon, UK shows the natural curve of the railway line on the imagery, limiting development to the north of the settlement. The vector version of the same route follows a path across the tors of neighbouring Dartmoor National Park.

If someone wishes to reference an event on the railway, do they use the image or the transportation alignment? [2006, Barr].

However, data harmonisation and quality is not important to the predominantly consumer and leisure markets that these sites and services are aimed at. It is apparent that the data standrads required for a commercial or critical service would not be met by the current infrastructure provided by these applications.

3. INFORMATION INFRASTRUCTURES

3.1 How do other Information Industries Operate?

How does the GI industry compare with other information industries and will these new innovations exploit GI and attain the significant benefits it promises but often just fails to deliver its full potential? Some traditional industries have evolved over time such as banking while others have adopted ICT to radically transform themselves such as the big supermarkets. Others are children of the digital age e.g. mobile telecoms.

It is often stated that GI is "going mainstream", what does that mean? and do Mash-Ups help us get there? How do we measure "industrial strength" in ICT?

While it is not easy to compare different services and technologies we can take some broad characteristics as a guide to help us. Different ***service industries*** are underpinned by some form of ***application(s)*** and some level of ***investment.*** They generally require ***hardware***, ***software***, ***information content*** and ***training*** to meet the needs of a ***user community***. By taking two traditional industries, banking (evolution), supermarkets and EPOS (revolution) and both private sector; along with emergency services (slower evolution, public sector) and telecoms (new, private sector) we might better understand and answer these questions.

3.1.1 Banking Services

We all rely on our pay finding its way into our bank account each month, sadly our bills seem to empty this reservoir all too rapidly. All of this is executed with 100% reliability. Banking is a long established industry; it has always processed information and delivered multiple services. These have now evolved rapidly from paper transactions (i.e. cheque transfers) to electronic funds transfers, to credit cards, ATMs and more recently digital purse and the use of mobile phones as payment devices. These services are nearly all provided by the private sector but are generally regulated or governed by some form of public administration. Just about every citizen is dependent on them to manage their financial affairs, increasingly by online banking services.

Standards have been evolved by the banking industry in cooperation to enable cross organisational funds transfers. Clearly the industry could not survive without some form of data and system interoperability.

Figure 3. Example of a UK IBAN structure

Picture: APACS – the UK payments association
http://www.apacs.org.uk/payments_industry/ibans_1.html

Harmonisation of bank/branch and account number has not been possible due to the operation of many well established (and different) national implementations; but as we all know, interoperability clearly works since we all pay bills and draw money on our travels. This infrastructure is supported by organisations such as the European Committee for Banking Standards [ECBS] who maintain industry standards such as the Register of European Bank Account Numbers and the IBAN standard. These standards define "keys" required to uniquely identify banks, branches and account numbers across many countries and they do this in a ***100% reliable way.*** This is achieved via interoperability – since cross border harmonisation was not possible.

Plastic card technology is also evolving significantly as well. Organisations such as EMVCo [2006, EMVCo web address] have emerged [founded by Europay International, MasterCard International and Visa International] to develop card technology standards [2006: EMVCo]. Based on elements of ISO standards, organisations have collaborated to extend and develop protocols to implement and maintain the growing infrastructure.

Picture: www.chipandpin.co.uk

Clearly some form of training is required but most users (i.e. the general public) clearly adapt relatively quickly as the methods of use are sufficiently simplified to enable widespread adoption by all members of society. A key characteristic of this industry is the network of fully automated yet entirely reliable services. These are very simple where the user has to interact with the system (ATM, Credit card etc).

3.1.2 Retail & Electronic Point of Sale [EPOS]

How we buy goods such as groceries, clothes or even cars has undergone a revolution over the past 10-15 years. Technology has enabled supermarkets to massively reinvent the entire end-to-end process of buying and selling over this time. By deploying very large scale operations, founded on major network ICT investments a revolution has been possible in the way we shop. The system works by tagging all retail items with a unique identifier (barcode); this is supported by a networked computer technology and very responsive distribution system. It is then possible for the big chains to completely remove the traditional wholesaler and replace this with their own network of warehouses.

Ubiquitous bar coding.

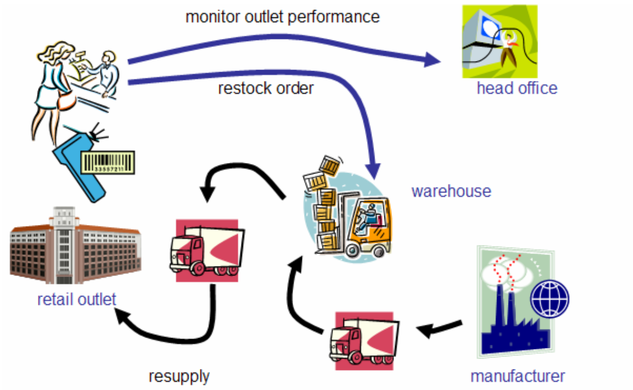

Figure 4: Barcoding and its role in the end to end retail process revolution.

As goods are sold, not only does the customer receive an itemised list, so too is sales information recorded to monitor store performance and stock replenishment. Each night this information is aggregated, analysed and used to restock the supply chain. Deliveries are then made from warehouse to outlet, to provide "just in time" stock replenishment and where required goods are restocked in the warehouse from the manufacturer. The warehouses are strategically located to supply the maximum number of stores in the shortest travelling time. Overall this minimises the space required for stock (i.e. idle funds), high-cost storage space and ensures the vendor is closely in tune with buying trends (often dictated by weather conditions). It gives the supermarket retailers significant buying power over the manufacturers; this in turn helps drive their profit margins. This system is highly dependent on networked technology linking distributed systems and automated information processing.

3.1.3 Emergency Services

The way that emergency services (police, fire, ambulance) operate can be different from country to country, depending on the political and administrative systems. In common they are usually located in the public sector and as organisations they are well established. A common goal is the need to reach locations in the shortest time possible where they often have to be prepared to manage difficult circumstances. Their ICT infrastructure can often vary from force to force, but increasingly the need to share information and co-operate and this is changing the way information systems and technologies are deployed. For example the task of determining or verifying the location of a caller in an emergency call can now be sourced by one of several methods as shown below:

Call From	Location Method	Accuracy
Fixed Landline	From the phone number trace the address and hence the coordinates of that georeferenced address.	1-5m
Mobile	From handset, trilaterate position from known transmitters	100m
Marine	Requires GPS input by user handset. GPS used as .navigation unit in most vessels.	30m

Figure 5: Methods of auto-locating the caller of an emergency call.

3.1.4 Mobile Telecoms

The mobile telecoms industry is clearly a newcomer in comparison with banking and grocery retail. By the mid-1980s, many of the telecoms companies in Europe had developed their own systems and protocols which eventually had to be resolved. The European Union intervened and the countries engaged at that time decided to standardise on a common model which was agreed via a competition. This saw the birth of GSM (Global System for Mobile communications).

This step was major milestone as it enabled and contributed to the massive growth in mobile phones in the 1990's. The approach has now of course been extended into new mobile technologies such as 3G and beyond and demonstrates the need for cross-industry co-operation.

The user is rarely aware of the significant information flows behind a simple mobile call, the cell structure that is constantly tracking the location of the phone and then handling calls, checking authorisation, messages, voicemail and billing. In this industry reliability is balanced with functionality, coverage is never 100% but the technology has developed ways of bridging these gaps (e.g. voicemail, text messages) easy redial etc.

3.2 Summary

In reviewing these four industrial domains there are several common threads that emerge.

- Reliability is essential, life/death situations can depend on it.
- Automation is a key characteristic

- ICT has been exploited to obtain a business advantage; *ICT is a means to an end.*
- Complexity is hidden from the user
- This has required diverse communities to work together and often required competitors to work together to agree standards (e.g. mobile comms).
- Information and data is often distributed, hence system interoperability is essential.

The following section contrasts these developments with the role of geographic information and what it has to do to contribute more effectively to enable our society in the future.

4. LOOKING FORWARD – 2010 AND BEYOND

4.1 New Developments

4.1.1 Mash-ups

As an immature technology, Mash Ups have come a long way, very quickly. In a short period of time they have served to radically increase mass awareness of the advantages of geo spatial data as the fulcrum around which many services can operate. The technology will undoubtedly continue to mature to offer increasingly complex and sophisticated functionality. However, if it is to ever become a serious tool, the key issue of data harmonisation and quality will have to be addressed.

4.1.2 Innovative web enabled applications and old fashioned business models

There is no doubt that the web has enabled innovation and introduced new ways of working. While the ideas have continued to flow, the business models have not always maintained pace, as we witnessed with the demise of the dot.com boom. Creating and maintaining a presence while continuing to meet ever-increasing user expectations will be a challenge for the foreseeable future and in the consumer space this can only be comprehensively serviced by the likes of Microsoft and Google.
The cost of developing and maintaining content is often overlooked and this economic tipping point often forms the boundary of what someone might aspire to and what they might actually end up using instead. Maintenance of information can cost 20-30% of the creation costs, *year on year.*

4.1.3 European Spatial Data Infrastructure [ESDI]

While in many respects it can be said that each country has today some form of spatial data infrastructure, as we noted earlier, there are few formal examples we can refer to. Many of those SDIs listed on the GSDI website [2006, GSDI] are at different stage of development or focus on particular application areas. Generally this is encapsulated in some form of geoportal.

Much is happening within Europe at present and this is being advanced by the INSPIRE [2006, EC] legislation which should be published by late 2006. Work has also started on the "Implementing Rules", i.e. regulations that will underpin the legislation. These have been structured within five areas in the legislation and these topics effectively define the top level components of the ESDI:

- Metadata
- Data specifications (and interoperability)
- Network Services
- Data Sharing (policy)
- Monitoring (of the directive)

It is too early to say what the European SDI will look like, since the task of understanding the current baseline and moving member states to the starting grid is a significant challenge in itself. The table below outlines some of these practicalities. This is the component structure adopted by the Data Specifications Drafting Team. Work is in progress to define the priority components at a largely conceptual level. This is intended to gain early adoption while introducing increasing levels of coherence.

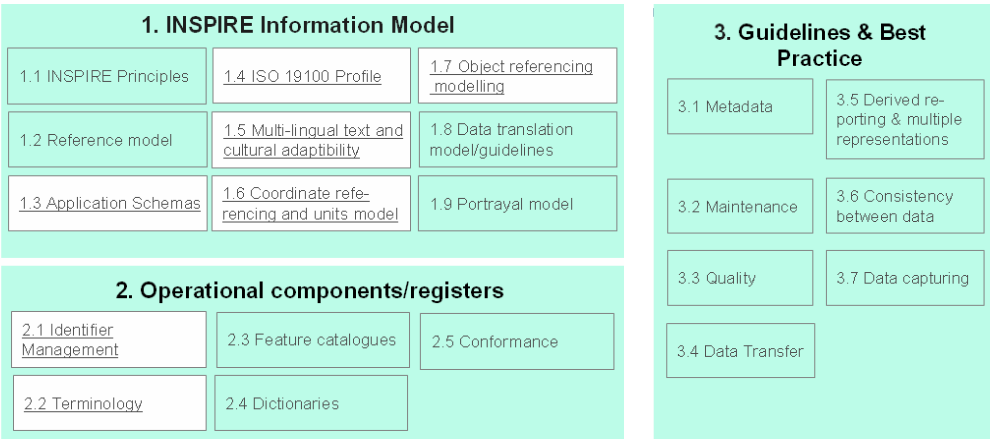

Figure 6. INSPIRE- ESDI: Data Specifications Components (priority items in white)

Geoportals are also in scope of INSPIRE, but from the ICT examples earlier it is clear that there is a need to go beyond this and embed/integrate GI in the wider mainstream information services and flows. This is relevant for reporting (e.g. river quality) and other day to day activities in managing the environment (e.g. impact of transport on the environment). In time we can expect this infrastructure to support wider forms of ICT transactions such as property sales, mortgages etc across member states. Hence there will be an increasing need for *geographic information* over traditional forms such as digital mapping, and specifically geographic information embedded in mainstream ICT.

4.2 Where does GI fit in?

The table below now attempts to compare the industrial sectors reviewed earlier with GI to as a way of identifying any similarities or differences.

	Banking	**EPOS**	**Mobile Telecoms**	**Emergency Services**	**GI**
Users (% population)	95%+	90%+	75%+	10%	10%
Criticality	Absolute	Essential	Now seen as essential	Life/death	Enabler Incr. essential?
Current Process Reliability	100%	95-100%	70-90%	100%	10-20%
Investment	€billions	€100millions	€billions	€10millions	€10millions
Standards	ISO and industry	Industry	Industry	Industry	ISO, OGC and industry
Information Content	ASCII based	ASCII based	Text/voice/ Images/video	ASCII & images	Complex data types
Hardware	COTS & Specialist	COTS & Specialist Accessories	Specialist phone sets	COTS & specialist	COTS
Software	Database driven with significant redundancy. ATMs simplified	Database driven with redundancy. Specialist at point of sale and back office.	Database driven with significant redundancy & Basic phone – req. some skill.	Operational systems complex	GIS software complex
Training	Evolution of well established processes	Checkouts (but users can now scan)	User – self teach	Specialist	Specialist systems tho. Google Earth has changed that.

Table 1. Comparative measures across the example domains observed in section 3.
Note: The assessment is indicative only and based on typical national profiles.

From the table it is clear that banking, EPOS and mobile communications enjoy mass-markets and thereby commensurate investment has been available to assist the process of development. Not all these services must have complete reliability; though in emergencies live depend on it. In banking (and other areas including some involving GI) liability issues arise if services fail to meet given parameters.

GI is then a much smaller player in this world, and one that is more of an enabler to other services, rather than and end in itself. It is therefore moreover dependent on other industries paving the way and breaking new ground for example Geography Mark-Up Language had to wait for XML to emerge in the mainstream initially. Nevertheless the GI industry has work to do to identify where it needs to be and proactively position itself for the future.

4.3 User Communities & Applications

The number of people using geography is both growing and transforming. Internet mapping and latterly Google Earth have brought new people in, while the use of sat-nav in cars is growing rapidly after several years of high expectations. It is clear users range from the professional to the consumer, and that these audiences will tolerate different levels of data integrity. This is also associated with cost and access mechanisms.

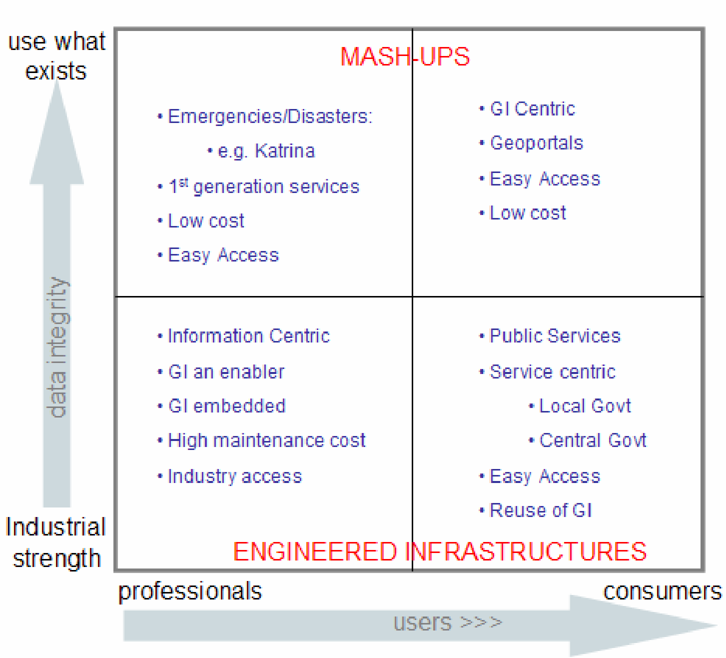

Figure 7: User vs Data Integrity Application Map - 2006

Figure 7 illustrates the current position, but the more interesting question to emerge is how this balance might change in the future? Will highly engineered data remain prohibitively expensive or will new processes provide a sustainable way of maintaining reliable information rich database(s) or will the data disconnects currently evident in mash-ups become the barrier to their penetration into serious application area?

4.4 Joined-up Geography

There is evidence that the current methods of “digital mapping” have failed to move the industry on into the mainstream. It is often used as a dumb backdrop. Software vendors have been reasonably happy with this since users require spatial analysis tools and skills to patch up the gaps in the data sources or worse still duplicate what exists …. because it can be easier to do that than resolve the differences between existing datasets. There is growing evidence that this is changing. Market research has shown that users are frequently spending 20-23% of their project costs in cleaning up data from different sources just to get it ready for use - and this overhead is no longer tolerable.

Several countries in Europe are re-engineering their data, moving to object-based data, underpinned by unique identifiers [2006, Giljohann]. Examples are the AAA model in Germany which is integrating the topographic database (ATKIS), the cadastre/land register (ALK/ALB) and national control system (AFIS), new databases have emerged in Denmark (TOP10DK) and more recently in the Netherlands (TOP10NL & NEN3610) and the INTERLIS developments in Switzerland continue. Other countries such as Sweden, Ireland and Northern Ireland are looking to redevelop their databases in the near future [2006, EuroSDR].

In Great Britain the Digital National Framework has been established to promote improvements in data integrity by using OS MasterMap as the underpinning reference base by cross referencing users "geographic views" to this common base [2006, DNF] using unique identifiers. (See www.dnf.org for case studies).

Integrating these capabilities with new navigation technologies such as Galileo, Rfid technology and many other data forms indicates a promising future.

5. CONCLUSIONS

5.1 2006: A revolution in access to Geographic Information

This is a very dynamic period in the take up of geographic information and public access to information (maps, terrain, images) around the globe. It is a significant step forward for society in general offering the citizen tools to increasingly make decisions and understand the world about them.

The capability to reach new audiences via simple Internet tools has been well demonstrated and such developments are not constrained by artificial boundaries. The tools are effectively available to people no matter whether they are an industrialised nation or in the developing world. These tools and resources are an aid to development whichever continent they are used on.

5.2 GI Breaks out of the Back-room

At this point in time (2006) we can conclude that

- the new web-based innovations (such as Google Earth, Microsoft Live Local, IGN-France Geoportail) significantly extend the reach of GI,
- in general the geoportal is useful resource, especially for consumers and in emergencies, the simplicity of the interface is a significant step forward

- mash-ups are fun (and this is a good thing!) but already expose data problems and this can be expected to grow over time, this may in turn lead to improvements in the interoperability of the content (though this will cost).

5.3 GI goes Mainstream

- the geoportal is less suited for embedded ICT applications.
- for such applications the information content that GI provides is primarily "an application enabler" (not an end in itself)
- to meet this level of need, data integrity is essential; GI has to support automated processes – and to achieve this wholly reliably.
- to attain this future position GI needs to further adopt the characteristics of mainstream ICT as demonstrated by banking, EPOS, telecoms etc by using information chaining and object referencing and ***this is a key emerging development.***

REFERENCES

Barr, R. 2006. Where's the Point? July/August edition of GEOconnexion International Magazine. P26

Digital National Framework [DNF]: 2006. http://www.dnf.org [last accessed 14 July 2006]

EMVCo. 2006. EMV Integrated Circuit Card. Specifications for Payment Systems Common Payment Application Specification. http://www.emvco.com/cgi_bin/0000_welcome.pl [last accessed 14 July 2006]

European Committee for Banking Standards [ECBS] 2006. Register of European Account Numbers. http://www.ecbs.org/download/TR201v3.19.pdf [last accessed 14 July 2006]

European Committee for Banking Standards [ECBS] 2003. IBAN: International Bank Account Number. http://www.ecbs.org/download/EBS204_V3.2.pdf [last accessed 14 July 2006]

EuroSDR. 2006. Features/Objects Workshop, Munich 24-5 April 2006. Proceedings. www.eurosdr.net [last accessed 14 July 2006]

Giljohann., M. 2006. National Data Integrity Initiatives within Spatial Data Infrastructures. MSc dissertation in preparation. City University London

Google Earth. 2006 http://earth.google.com/ [last accessed 14 July 2006]

Global Spatial Data Infrastructure Association: http://www.gsdi.org [last accessed 14 July 2006]

IGN-France. 2006. Geoportal: www.ign.fr/ [last accessed 14 July 2006]

Microsoft Local.Live 2006. http://local.live.com/ [last accessed 14 July 2006]

Wikipedia http://en.wikipedia.org [last accessed 14 July 2006]

BIOGRAPHICAL NOTES

Keith Murray is Head of Geographic Information Strategy at the Ordnance Survey based in Southampton. He has worked as a surveyor and photogrammetrist. He has extensive experience in the use of imagery both in production and in research following a Masters Degree from University College London. He is an immediate past President of EuroSDR, and a member of the INSPIRE Data Specifications Drafting Team and is Chair of the DNF Expert Group. He has published numerous articles on the use of geographic information, new developments in GI strategy and research. He is a member of the Royal Institution of Chartered Surveyors (RICS), the Remote Sensing and Photogrammetry Society (RSPSoc) and a Fellow of the Royal Geographical Society [RGS].

Andrew Trigg is Chief Geographic Information Officer at Land Registry, based in Plymouth. He has spent his entire working career in the Geospatial Data Management industry. He is currently responsible for the maintenance and integrity of Land Registry's national property, mapping and address data sets.

Prior to joining Land Registry in January this year, he was Head of Products and Consultancy at Ordnance Survey, where amongst other things he was responsible for the introduction of OS MasterMap. He has also worked for Wiltshire County Council, Laser-Scan and NERC since graduating with a PhD in GIS in 1987.

CONTACTS

Keith Murray
Ordnance Survey, Great Britain
Romsey Road,
Southampton SO16 4GU
UNITED KINGDOM

Tel: + 44 23-8079-2729
Fax: + 44 23-8079-2660
Email: Keith.Murray@ordnancesurvey.co.uk
http://www.ordnancesurvey.co.uk

Andrew Trigg
Land Registry,
William Prance Rd
Plymouth PL6 5WS
UNITED KINGDOM
Tel: + 44 1752 644549
Fax: + 44 1752 644712
Email: Andrew.Trigg@landregistry.gsi.gov.uk
http://www.landregistry.gov.uk

v1.00g