

THE DIGITAL NATIONAL FRAMEWORK - BRIDGING INFORMATION THROUGH GEOGRAPHY

K.J.Murray^a, G. Hart^b & P Allan^c

^{a-c} Ordnance Survey, Romsey Road, Southampton, SO16 4GU
(Keith.Murray, Glen.Hart, Peat.Allan)@ordnancesurvey.co.uk

Commission IV, WG IV/4

KEY WORDS: Interoperability, GIS, Integration, Federated, database, Spatial Infrastructure, Acquisition

ABSTRACT:

We are currently witnessing massive growth in the use of geographic information around the world. Governments are seizing on the power of geography, location based services are now finally starting to develop and the citizen has more access to geographic information today than ever before. This paper describes how, to support this greater adoption, geography is gradually being transformed in Great Britain to better support and service all users and “join up” the many disparate forms of information. The concept underlying this “joined up” approach in Great Britain is an initiative called the Digital National Framework (DNF). Elements of the DNF are being developed in conjunction with other mapping agencies and evolved in conjunction with a wide number of users in the government, utility and private sectors. The paper will also demonstrate how a whole range of technologies such as remote sensing, photogrammetry, laser scanning, real time kinematic GPS, land survey and data in 3rd party databases are being combined to take geographic information to new levels of integration which offers potential for a powerful national capability. At the heart of the DNF model features such as buildings, land parcels, roads etc, are *each* referenced by a unique identifier (known as a TOID) to enable data linking & data sharing – the TOID represents an entity at a physical location. The structure of DNF will be described in more detail in the paper but the scope also incorporates the national coordinate system and the relationships with GPS, reference and application data models, metadata, feature catalogues, referencing relationships, data quality, compliance and terminology. Fundamental principles include reuse of information (capture once and use many times), the ability to apply common methods of publishing derived information, enabling flexible visualisation and support for web services of many kinds. The DNF is being implemented in Great Britain and is being underpinned by Ordnance Survey’s new layered database known as OS MasterMap. OS MasterMap is an integrated database, which will incorporate detailed topographic data, imagery, addresses, transport networks and in time a new height model, land related information, and more. The DNF concept is an evolutionary inclusive framework, to develop a coherent and consistent approach to georeferencing and *integration of disparate datasets that are underpinned and bridged by geography*.

1. INTRODUCTION

1.1 The Information Age

Where ever we are in the world, most of us now live in a society driven increasingly by information. Information is collected to determine the population and the characteristics of that population. In turn this information is used to determine the economic and social fabric of a national economy. For example “do we have sufficient housing?”, “does housing meet the emerging social characteristics – e.g. single parent families”, “is the housing located where it is needed”, “is the services infrastructure sufficient to support growth i.e. water, sewerage, power etc?” and so on. In many other fields too, information is sought, collected, processed and used. The explosion in the use of the Internet alone has generated a whole new economy around information. Never before has technology made it so easy to collect information of all kinds, store it and use it in many different ways.

1.2 From Data to Wisdom

It is a well established thesis that raw **DATA** is valuable but has limitations. With processing and some additional data we can often turn this raw material into **INFORMATION**. For

example a simple aerial photograph can be developed into a map-like orthoimage from which measurements and interpretation applied to yield information about something of interest. Thereon an analysis and comparison of two or more orthoimages from different periods, or several epochs and integrated with other evidence may then yield **KNOWLEDGE** - for example the nature of the retreat of glaciers in a mountainous area over several decades. Further information from other studies combined with this knowledge may conclude a general trend about natural cycles in the Earth’s climate change. Other information may prove that this is also being affected by pollution by Man and his vehicles, industry and lifestyle.

WISDOM gained from such a level of understanding and the integration of many disparate forms of information such as this might then lead to the direct actions required to mitigate the risks to Man and the future of mankind. Such data and information in several other disciplines e.g. research into deadly diseases, requires the integration of several disparate kinds of data, information sharing, information processing and the sharing of knowledge from many sources.

It would be regrettable then, in our rapidly advancing technological society, if in the way we acquire use, and process

geographic information, we lost the ability to connect it up and thereby derive the knowledge and wisdom to address the major issues that face us; not just at international level, but equally importantly right down to the benefit of the individual citizen.

2. GEOGRAPHIC INFORMATION

2.1 Geography – the common denominator

It is a recognised fact that underpinning most information is location. People live somewhere, people travel from one location to another; crimes occur in buildings, in the street, in all kinds of places, medical services are available at specific locations and so on. Therefore geography has the potential to be the common denominator to the (apparently) disparate information we collect and handle each day.

2.2 Linking disparate information about a location

Information will be collected by an organisation for a specific purpose e.g. education records, health records and so on. When this information is combined with that of other organisations; new information may emerge and establish new levels of knowledge. For example information about health combined with living conditions or the local industrial infrastructure may reveal trends such as the prevalence of disease in certain conditions and locations. Likewise information about education attainment coupled with employment success or benefit claims can highlight areas where groups or individuals are becoming socially excluded. Once this knowledge is determined, some form of investment may be appropriate to assist regeneration to improve the social fabric of the community. Similar information collected over time would demonstrate whether the condition is improving, or not, and therefore whether investment has been successful.

Geography alone does not make the connection, since the close proximity of two coordinates, of unknown pedigree, may only imply a relationship. A more reliable link is formed when objects are given identities e.g. buildings or addresses and information is then linked to these objects via the identifier to establish *an explicit connection*.

3. INFORMATION FROM MULTIPLE SOURCES

3.1 Data Collection

The way map or geographic information is collected and has grown over the past decade has changed the entire complexion of the geographic information [GI] industry. Until recently photogrammetrists and surveyors went about processes which would have been recognised 50 years, if not 100 years previously – over the last 10-15 years that has changed and continues to change.

3.2 The mapping or cadastral agency today

Today a database (of some kind) is central to an organisation concerned with developing and maintaining geographic information. Typically different people undertaking different jobs will update this database. A surveyor in the field for example, will now probably use a mobile computer with an extract of the database to be updated contained on it. He is

likely to use GPS to collect the coordinates and a coding system to assign attributes and edit facilities and create objects e.g. building features, street information or add attributes. Measuring devices may use infra-red or other technologies while instrumental observations will still be incorporated on some occasions. The photogrammetrist is now likely to use digital images from space or aerial images, in either stereo or mono to create new data or update existing data. *The acquisition of data is becoming cheaper and simpler - but data management more complex.*

Other information is also increasingly available. For example electronic builders plans can be used to update the data and offer new services commonly referred to as “pre-build” information. The utility company (gas, electric, water, cable/telecoms) need such information to plan their distribution network and link it into their existing records, before the building takes place. The integration of pre-build information into a database can also save costs and improve currency of data, thereby providing a better service all round.

3.3 The wider GI community

Of course, the national mapping/cadastral agency [NMCA] is not the only player. There are several organisations that collect similar definitive levels of base data used for referencing, often across government, and there are many others who collect data for all kinds of applications (and normally reference it to the NMCA data).



Figure 1. Example of reference information.

Topographic information and integrated postal addresses and street centreline information.

Reference information

Framework information (also known as reference data) is generally taken to cover that data or information which underpins or is used as a reference base for applications data. For example a user may record a scene of crime record [application data] to a building or an address [reference information]. While this is mainly the domain of the NMCA, the state or region, local authorities or municipalities, the hydrographic survey and sometime others contribute to reference information, the exact model differs greatly from country to country.

Application Information

People and organisations have collected information for thousands of years. All through that time some of this information has been georeferenced - in some way. From the early census information in the Middle East, historical battles and event and records such as the Domesday Book can often be referenced to a location and often very accurately.

Today there is a greater perception of the need to georeference information arising from the growing recognition, largely at government level, of the real value of geographic information in delivering solutions and establishing more effective ways of analysing information. Such information will be collected to satisfy an application such as property values for taxation, river quality samples, crop health, highway traffic statistics etc. As such this information may or may not be suited to serve in other applications. However when conflated with other datasets the quality and usefulness of the data will generally improve, albeit at a cost of resolving many discrepancies that will inevitably be found.

3.4 Are we moving towards coherence or chaos?

As technology improves one would expect maturity and the integration across datasets to emerge, but there are several recognised barriers.

Business Priorities

Organisations, as we have seen, collect information to meet a business need whether they are in government or the private sector. Resources are often scarce and therefore business priorities are the major objectives to be met. As a consequence the cost of extending a datasets capability for perhaps the sole benefit of the third party is rarely an attractive proposition.

Boundaries of the enterprise

While there may be a coming of minds at international levels over key issues such as global warming, this then transforms when we return home to manage national priorities. This “silo mentality” often surfaces and will strengthen down to the state/region/department levels of government. Further this continues to local government/municipality level where the remit of elected officers is their local area and in many respects quite rightly so. Hence each of these “levels” of administration tends to establish barriers to information sharing and exchange. Differences in specification and practice are evident across different parts of government in almost every country. In several European countries data maintenance can be distributed across the same level of government e.g. detailed information is mapped by municipalities in Denmark and Italy. However very significant differences in the nature and the content of the data will often be found from one municipality to another. In larger countries e.g. Germany the federal states and the mapping agency have achieved a common standard (ATKIS) for some datasets.

New Technology for all

New technology is a boon to the commercial and public sectors alike. The cost of entry into operational data collection is now much lower than 10 years ago. New technologies such as GPS, digital imagery and mobile PCs make it easier to get in a car and go out and collect data such as street centrelines and attributes. LIDAR and airborne radar require more investment but are increasingly being operated and are providing new and

different sources of data. The growth in the availability of new datasets is significant and in many countries now offers alternatives to the traditional reference data.

Little standardisation

The ISO 191xx family of GI standards [www.isotc211.org] are now emerging, and are starting to provide a common model. However the ability of the majority of these standards to influence real data standardisation is limited to the provision of a simple profile at this stage, but this is a start. Elsewhere, such as the UK the British Standards for geographic information provide a tighter model for some datasets e.g. street information. However these standards fall short in terms of providing a generic model to support several application areas and have so far encouraged divergence rather than reuse of definitive information.

Generally there are few models to support a coherent structure and form of georeferencing and data collection. It's not uncommon for organisations to adopt new digital methods as a direct replication of the tried and tested paper methods processes – it's perceived as the line of least risk. Hence data such as environmentally sensitive sites may be collected at small scales of say 1:25,000 while the cadastre may be maintained at 1:2,000. When a third party attempts to develop an automated process to determine whether any properties fall within the environmental area, the reliability of the results would be limited and liability on the provider potentially significant.

Summary

The combination of new technology, the limits and priorities of the business enterprise and the lack of effective standards provides barriers to seamless information exchange and hence automated GI based services.

4. THE DIGITAL NATIONAL FRAMEWORK

4.1 Joined-up geography

The need to provide a vision for better integration of geographic information has been emerging for several years in many countries. Great Britain has enjoyed a robust mapping infrastructure for over 200 years and this has been reinforced at regular intervals, for example in 1938 through a fundamental review [Ordnance Survey, 1938] that established the core foundations for today's GI infrastructure in Great Britain. While the growth in the use of GI in Great Britain has been significant, for all the reasons explained earlier, it remains difficult to easily link or share applications and some reference information and thereby automate analysis processes. The concept of the Digital National Framework [DNF] was born in 1999 following a period of consultation in GI and the publication of several reports on the benefits of GI [e.g. OXERA, 1999].

OS MasterMap

From the DNF concept OS MasterMap quickly evolved. OS MasterMap is as a layered database of reference information of which the first deliverable, in November 2001, was the detailed and definitive topographic framework for England, Wales and Scotland. There are over 430,000,000 features in the

topographic database and 5,000 updates are applied every working day. This is provided as the foundation for all other information and is based on the National Grid coordinate system. A transformation [OSTN02] now defines the OSGB36 National Grid in conjunction with the ETRS89 positions of the National GPS Network stations.

Since then, other layers have been added: the roads components of an Integrated Transport Network, colour orthoimagery and all 26 million postal addresses georeferenced to better than 1m. All of these, except imagery at the moment, cover the entire area of England, Wales and Scotland [234,410 Km²]. Further layers are being developed or will be migrated in the near future.

A consultation paper on the concept of DNF and the methods to support usage was published in early 2000 [Ordnance Survey, 2000]. Since OS MasterMap was released take up of the concept has been steady and the adoption and evolution of the methods has been progressing. Several successful projects have been completed, such as the development of a new national database recording and showing all land open to public access. In the light of this experience and further work at national and international levels, a further White Paper on DNF will be published in 2004.

4.2 The Concept

“The Digital National Framework (DNF) provides a permanent, maintained and definitive geographic base to which information with a geospatial content can be referenced” [Ordnance Survey, 2000] and incorporates:

“..a set of enabling principles and operational rules that underpin and facilitate the integration of geo-referenced information from multiple sources”

DNF is broader than the national mapping agencies data, other organisations contribute to the framework and all kinds of users can create their information on the maintained base with greater confidence and assurance. Clearly change in process and practise will be required to achieve this and a stepped approach to adoption (the “levels” described later) has been developed to support this.

4.3 DNF principles

DNF is supported by four guiding principles:

1. The detailed ground rules shall be driven by the strategic needs of the wider GI community.
2. Data should be collected only once and then re-used.
3. Reference information/data should be captured at the highest resolution. Information may then be published via a) different representations at scale b) multi-resolutions across scales - to meet different application needs.
4. DNF will incorporate and adopt existing de facto e.g. OpenGIS and de jure standards e.g. ISO, wherever they are proven and robust.

Therefore DNF is not an Ordnance Survey strategy; it is aimed at providing a better framework for all users of GI in Great Britain. While it will not necessarily be appropriate everywhere its adoption, partial or otherwise, would be welcomed and the

experience could help with current European initiatives such as INSPIRE [European Commission 2004] and EuroSpec [Luzet & Land, 2004].

4.4 Reference data and Applications data

Already this paper has recognised that reference information is not the sole domain of NMCA’s and there will normally be several other organisations that might contribute to the framework. The methods will be published in the public domain and maintained. Therefore the framework is inclusive in nature and open in the way it can be used by the industry at large.

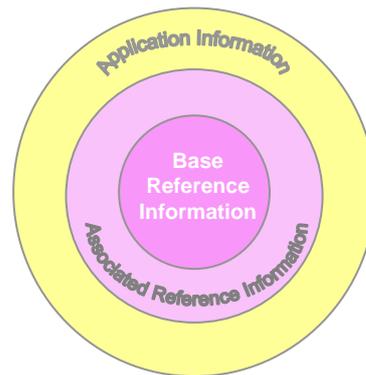


Figure 2. The relationship of Framework and Application Information

The world is not so black and white in terms of Reference and Application data. There will be shades of grey in the boundary. For example a base topographic (or cadastral) map will underpin a land registration system. The parcels defined in that system are therefore at the “application” level. However this information is itself often used as a reference feature for other information that may be shared across organisations. Therefore Reference Information will evolve and will be defined by usage by several parties for linking their Application data or information. In the DNF model, any further information which is definitive, and is a) used widely for referencing and b) associated with the Base Reference Information is known as “Associated Reference Information”. Table 1 below illustrates the different types of information that broadly fall into these three categories.

4.5 Georeferencing

The heart of DNF is the provision of a coherent georeferencing system. One that is both extensible and inclusive. Inclusive in the sense that many organisations may contribute to the referencing framework and it is also open to all those who wish to reference their application information to it, and then share and exchange their information with third parties or exploit it commercially. All parties who have invested in creating and maintaining such data are likely to wish to secure that and future investments by protecting their information in some way to ensure that its future is sustained.

The georeferencing framework is based on giving each object in the landscape a unique identifier. This is known as a TOID and is a sixteen digit unique number, bearing no intelligence. Each building in OS MasterMap is referenced by a TOID as is every other feature such as fields, woodland, street segments etc.

Application Information	
Census records	
Crop records	
Education records	
Health records	
Meteorological records	
Property ownership/occupancy records	
Scene of crime information	
Traffic flow information	
Water quality samples	
Utility records	
...and so on	
Associated Reference Information [ARI]	
Administrative and Electoral boundaries	
Land & Property: • addresses & • property parcels	
Boundary datasets [postcode, statistics, health, planning etc]	
Geology	
Land Use/Cover	
Networks : • transport/paths • water	
Points of Interest	
Base Reference Information [BRI]	
Imagery	
Height	
Marine - coastal zone	
Detailed topographic data including distinctive names	
Geodetic framework and transformations	

Table 1. Examples of Reference and Application information. Note these lists are not necessarily exhaustive and attributes may be added to information in any one of the three categories. (This model is currently under development).

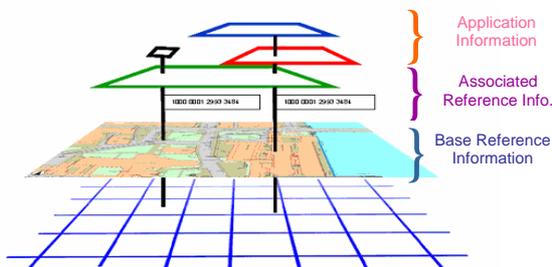


Figure 3. DNF model structure

The TOID can provide a common geographic-link across different user views of the real world and assist in the assembly of individual features into integrated user-defined objects.

Building user-defined features

Features such as buildings, land parcels, property boundaries etc can be aggregated to create a users view of the real world – whatever their application and need. Each business will take a different view of the real world and interpret it differently to meet their business need (land registration, grounds maintenance, utility, mail delivery, navigation etc). Each application is different and will require a different selection of components, but based on a common geography. If someone wished to create a school feature they could build this from all the existing features within the perimeter of the school, incorporate other components such as addresses and classifications/attributes or link in other records. This is illustrated in Figure 4 below.

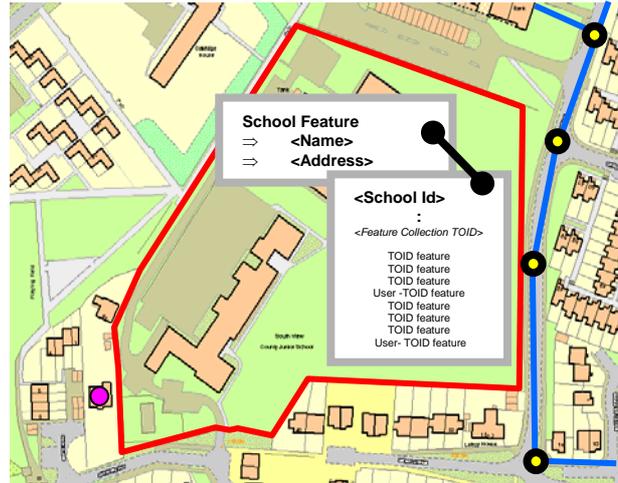


Figure 4. School Feature formed by referencing several objects. Point, line and area geography can be modeled.

Alternatively if one only wished to model just the property boundary above e.g. for land registration purposes, one could simply reference the bounding features. Where geometry is user-defined, i.e. does not relate to any topographic feature, then the user assigns this a TOID from their own list and uses that feature in the same way as the base reference data. Therefore DNF is flexible and allows users to choose the appropriate approach for themselves – providing it is consistent with the overall model to support data exchange.

Feature classification

It is clear from the above that when one wishes to integrate information it is necessary to link with one or more third party datasets. This is rarely easy, since each organisation will have its own coding system, descriptions and records structure. Again this presents a major barrier since organisations need to continue to function within their core business remit. Harmonization can therefore take several years to achieve. This is also a common problem for the mapping agencies across Europe when combining data into a common form for cross border customers [Murray, Bray & Steenson, 2001].

One way of resolving such a problem is to adopt a “global schema” and to relate each organisation’s features to that schema. The schema should be based on a Feature Catalogue, which both registers the features to be used nationally and describes them so that everyone has a common understanding of the meaning of “Railway Station/Building” or “Railway Station/Car park” for example.

5. PROGRESS TOWARDS THE DNF CONCEPT

5.1 An evolutionary concept

DNF is an evolving concept. It requires a change in business practice. Such changes are inevitable as we move on from digital mapping to the more demanding world of geographic information. To provide a “stairway” to linking information three levels have emerged. This is proving successful as some technologies are developing faster than others and the take up of and maturity of several technologies is far from uniform:

Level 1: This is a simple approach, which cross references the identifier in the user's database with an identifier in the reference information, ideally on a 1:1 basis. No attempt to manage any associated geometry is made. This approach ensures that the user information is geographically linked and the reference, about the same object, can be shared with others who may also hold information about that object. For example crimes committed in a particular building or address. This also provides a step towards level 2.

Level 2: Geometry is introduced in this level. This is referenced either to a single feature or collection of features, as in Figure 4 (above). The resulting composite feature is stored in the users system along with references to these components. The user again is able to share information with others, ease data management, improve data integrity and will benefit from updates to the underlying features to maintain a consistent level of currency in their own data.

Level 3: It is possible for the user not to hold any geometry/geography at all (apart perhaps, from any they might create of their own). Again using the same approach in Figure 4, the user would simply hold the references and their geographic feature would be constructed only when required. This effectively requires a database driven solution and ideally some form of reliable Web Feature Serving (WFS) if local reference data holding is to be dispensed with. As a technically advanced solution it potentially offers the greatest gains in terms of data management, data sharing and process automation.

5.2 Testing the Concept

Several organisations across government and the private sector have built applications around the TOID. Currently several further testbeds are being undertaken to develop, refine and document the methods. It is clear that feature lifecycles need to be better managed by all stakeholders (reference information providers and application users) and rules determined and evolved to ensure changes are made when agreed conditions are reached. For reliability these should be as automated as possible.

5.3 Benefits

While we are still in the early day's development of DNF, which is recognised as a medium to long-term development, several benefits can be identified:

Savings

- Reduced data collection and maintenance costs
- Reduced data duplication
- Reduced data integration costs and times

Data Quality

- Improved data integrity across all adopting datasets
- Better data quality (currency, consistency, accuracy, completeness)

Automation

- Easier and greater reliability in information/data interoperability.
- Easier user data exchange and sharing
- "Plug and play information" offering faster solutions

Better solutions

- Greater clarity and emergence of definitive core datasets
- Inclusive & open – many organisations may contribute
- Improved intra and inter-business communication
- Greater reuse of information and thereby higher value of that information.

Ultimately this will contribute to a stronger knowledge economy at the local, regional, national and international levels.

6. CONCLUSIONS

In conclusion:

- we are in the midst of a revolution in the growth, acquisition, use, dissemination and use of information. This is impacting on the GI community and us all in particular in terms of deriving knowledge and decision-making.
- this should be an opportunity to capitalise on the knowledge, experience and skills to maximise the investment in GI and to reaches it's full potential. However the rush to collect data, easy to use new technologies and organisational barriers may significantly undermine this goal.
- The DNF concept is an evolutionary inclusive framework to develop a coherent and consistent approach to georeferencing and integration of disparate datasets that are **underpinned and bridged by geography.**

REFERENCES

Literature:

Luzet, C., & Land N., (2004). EuroSpec – a cornerstone for the building of the European Spatial Data Infrastructure. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Istanbul, Turkey, [In preparation]

Murray, K.J., Bray C., and Steenson T. (2001). Better connected – the three Ordnance Surveys improve georeferencing links. Proceedings of the Association of Geographic Information Conference 2001 (section t2.7) [CD-ROM], London, UK.

Ordnance Survey, (1938). Final Report of the departmental Committee on the Ordnance Survey. London: HMSO.

Ordnance Survey, (2000). The Digital National Framework - Consultation paper 2000/1 [no longer available online – but can be obtained from the authors]

Websites:

European Commission (2004). Homepage of the INSPIRE website. Available from: <http://inspire.jrc.it/home.html> (accessed 30 April 2004)

OXERA (1999). The economic contribution of Ordnance Survey GB. Oxford Economic Research Associates Ltd. Oxford.

<http://www.ordnancesurvey.co.uk/oswebsite/aboutus/reports/oxera/index.html> (accessed 30 April 2004).