# Building bridges between social science, grid, and geospatial communities: a reflection on practice.

Max Craglia[1], Bridgette Wessels[2], Mike Griffiths[2], and Andrew Costello [2]
[1]Joint Research Centre of the European Commission.
[2]University of Sheffield.
Email address of corresponding author: Massimo.craglia@jrc.it

**Abstract.** This paper reports on one of the 11 pilot projects funded by the ESRC as part of its e-social science research programme. The project studied the feasibility of "Collaborative analysis of individual and area-based social exclusion", and had two main objectives: firstly, to address a major social science issue with both theoretical and policy implications, namely the extent to which individual and neighbourhood effects are able to account for the geographical variation of crime patterns; and secondly, to explore the opportunities and challenges offered by the Grid from a socio-technical perspective, i.e. how the different disciplines and theoretical traditions (criminologists, sociologists, planners, spatial analysis, and computer scientists) can engage with this new way of working, and how they shape emerging technology. The paper presents the model of the spatial distribution of offenders developed for the whole of England at one hectare level, and the contribution of the Grid in enabling the data analysis and visualisation. It considers the lessons learned through the interaction of Grid technology and the multi-disciplinary research team working on the project, and reflects on the research opportunities and challenges coming from the convergence of the social sciences, Grid, and spatial data communities.

## Introduction

This paper reports on one of the 11 pilot projects funded by the ESRC in 2003-04 as part of its e-social science research programme. The substantive social science research question that the project addressed is the extent to which individual and neighbourhood effects are able to account for the geographical variation of crime patterns. This issue has been at the core of environmental criminology since the early work of the Chicago School of Sociology in the 1920s on the effects of immigration, urbanization, and rapid industrialization on delinquency. The observation that delinquency rates where highest in the ghetto communities closest to the expanding industrial districts and lowest in those further away lead Shaw and McKay (1942) to formulate the theory of social disorganization, which provided a major impetus to urban studies and criminology but was also heavily criticized from those emphasizing the key role of individual factors in accounting for criminal behaviour on the one hand, and those placing greater importance on societal structures on the other. The theoretical debate between the relative importance of individual, family, school, social ties, and neighbourhoods factors on crime patterns continues to date in spite of major work in this field over the last 20 years (e.g. Wilson 1987; Friedrichs and Balsius 2003; Sampson et al. 1997). As the evidence regarding the influence of area effects remains unclear and difficult to quantify, the opportunity offered by the Grid to undertake large scale data analysis over distributed information resources promised to provide new insights into this key research issue, which has both theoretical and policy implications. With these considerations in mind, the project has the following two main objectives:

1. to explore, quantify, and model the spatial distribution of crime in relation to socio-economic and neighbourhood characteristics based on user-driven applications of Grid technology,
2. to critically reflect on the evolving relationship between social scientists, technologists, and the Grid as input to the development of training material and the further deployment of Grid technology in the social sciences.

An additional objectives was to prototype a web-services platform for sharing the results of the model with the stakeholders in the region of South Yorkshire, namely local Authorities, youth services and the police. This part of the project is discussed in Craglia et al. (2005). The paper is organized as follows: the next two sections discuss the data sources, analysis, and data modelling. This is followed by a discussion on the Grid application and the relationship between technology and research practices in the project, leading to broader lessons for the further development of Grid computing in the social sciences, and the synergies that my be foreseen between the Grid and the spatial data community. The concluding section summarizes the key achievements and directions for further work. A more detailed discussion of the project from an ethnographic perspective, and of the lessons for e science development is included in Wessels et al. (2005).

# Data Analysis

## Data Sources

For the purpose of this one year project the research team, which comprised researchers form Town & Regional Planning, Sociology, Criminology, and Computer Science, took the view that it was necessary to focus only on data already available and be theory-led. With this in mind, and following an extensive review of the literature in the fields of environmental criminology and social exclusion, the project team decided that the most promising sources of available data (i.e. not requiring new collections) would be the following:

*Census data:*
Data was downloaded via CasWeb at Output Area (OA) and Super Output Area (SOA) for the whole of England for 120 variables suggested as having potential from review of literature in the following domains:

- Ethnicity and Age
- Economic Activity and Occupation
- Socio-economic classification and Qualifications
- Household Characteristics including vacancies and overcrowding
- Tenure
- Car Ownership
- Migration

*Crime Data:*
The data made available by South Yorkshire Police following the signing of a written protocol with the University of Sheffield, included:

- 371,000 reported victims of crime

- 46,800 offenders who have committed 118,000 offences
- 17,000 young offenders who have committed 45,000 offences
- 70,000 thefts form cars, 63,000 burglaries, 28,000 damage to dwellings

The data was provided with X,Y coordinates at 10 metre resolution or better (1 metre in some instances), and covered the time period 1998-2003. A unique feature of this data collection is that it provided the opportunity to link the locations of known offenders, offences, and victims in a way that has not been reported before in the literature.

*Youth Data:*
Connexions South Yorkshire made available data for November 2003 and March 2004 for all 16-18 years old in South Yorkshire (approximately 30,000) including unit postcode, age, sex, ethnicity, and whether in education, employment or training or not.

*Mosaic Classification:*
The MOSAIC classification of consumers for the whole of Great Britain at unit postcode level (over 1.5 million records) was made available by Experian, a leading geomarketing organizations, to the University of Sheffield especially for this research. MOSAIC classification segments consumers into 11 major groups and 61 detailed types based on socio-economic characteristics and lifestyle surveys. Used mostly for marketing, it is becoming increasingly used in public policy to target resources to particular segments of the population.

*Index of Multiple Deprivation 2004:*
This Index is published by the Office of the Deputy Prime Minister at Super Output Area, and contains seven domains which relate to Income deprivation, Employment deprivation, Health deprivation and disability, Education, skills and training deprivation, Barriers to Housing and Services, Living environment deprivation and Crime (ODPM, 2004).

## Results of the Analysis

The data analysis was divided into three stages:

1. Analysis of individual variables for the crime and youth dataset, including aggregation at census geography level, calculation of counts, rates, and standardised rates, and identification of outliers and extreme values;
2. Analysis of each variable in relation to key census variables and MOSAIC classification;
3. Correlation analysis among the key variables to identify statistically significant relationship supported by the review of the literature on environmental criminology.

For the statistical analysis of the data, a partition was made between Sheffield and the other three local authorities (labelled "the Coalfield") given that Sheffield has a disproportionate number of students, which skew the type and patterns of crime in the region. The key findings from the analysis undertaken show that:

1. over 70% of the variance of victimisation is accounted for by the proximity of the residential location of offenders (68% in Sheffield, 79% in the surrounding local authorities "the coalfields")
2. Offences such as Domestic burglary and criminal damage are also strongly related to the location of offenders (64% and 76%), with again stronger correlations in the coalfields than Sheffield.
3. The geographic distribution of offenders appears to have strong correlations with Census data, whilst Offences do not.

Given that very little data has been available in the past linking victims and offenders, these are important quantitative findings, which confirm qualitative interviews with young offenders (Wiles and Costello, 2000).

# Data Modelling

In the light of the findings above, it was decided to concentrate our attention on Offenders on three grounds: little work has been done in this area because of lack of data, strong relationship with census data, and opportunity to develop subsequently risk maps of victimisation based on the result of the modelling of offenders. We also decided to focus on a statistical modelling approach rather than a compositional approach based on the MOSAIC classification. The modelling process took two stages:

1. Using stepwise regressions to identify the most significant variables in accounting for the variance of offenders, and reflection on the findings in the light of the literature review
2. Experimenting with different types of models to see which would yield better results.

The final model selected is a General Linear Model of the Poisson family in which the response variable (counts of offenders) is transformed using a logarithmic function. Poisson models are particularly used in the case of discrete count data that are not normally distributed (Haining, 2003). This is the case of the counts of offenders, which have a particularly skewed distribution. The criteria used to arrive at the specification of the model have been:

1) Use of nationally available data (i.e. not the specific datasets available for South Yorkshire that may not be as easily available elsewhere);
2) Good fit of the model;
3) Independent Variables that are not only statistically significant but also stand up to the literature, and with coefficients of the right sign (i.e. the relationship is in the expected direction);
4) Parsimony: i.e. trying to have as few variables as possible to keep the model simple without sacrificing too much in terms of predicting power.

The final model takes the following form:

$$\log(\text{Observed / Expected values}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6$$

Where:
$X_1$ = Percent Economically Active Unemployed
$X_2$ = Percent of Households Renting from Other (hostels, secure accommodation, prisons, boarding houses hotels and other communal establishments)
$X_3$ = Percent of Households with Lone parents with dependant children
$X_4$ = Percent of residential spaces vacant
$X_5$ = Index of Multiple Deprivation 2004- Health Domain score
$X_6$ = Index of multiple deprivation 2004 Crime Domain score

The Expected Value for each area i is the regional rate (i.e. Total Offenders/total population) multiplied by the population of the ith region. In a log form this becomes the log of the total population for each region plus a constant that is then incorporated in the $\beta 0$ coefficient. This is very useful for fitting the model nationally, as all is needed is the population of each SOA (known), rather than the expected number of offenders (non known). The overall fitness of the model is measured by the Cox and Snell R2 = 0.8215. Figure 3 below shows the scatterplot between observed and predicted values for South Yorkshire.

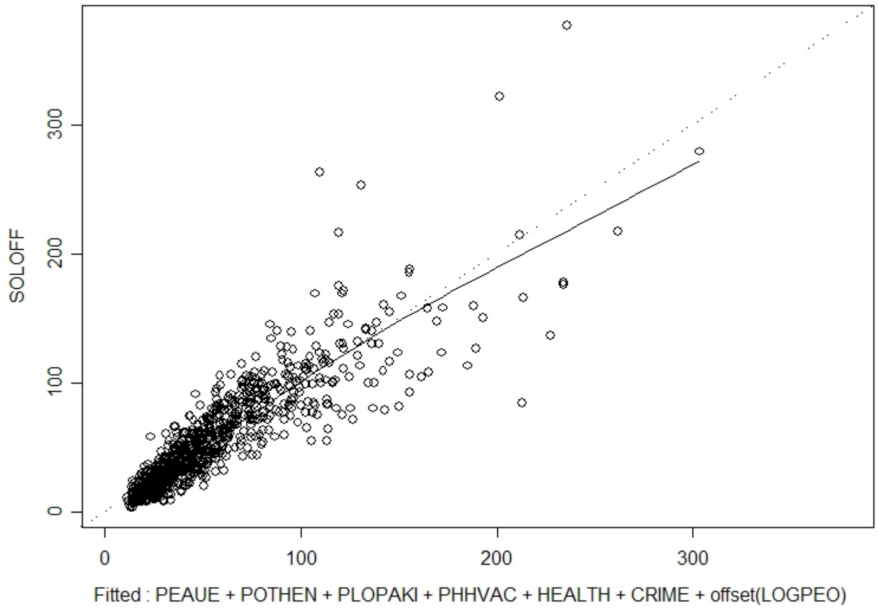


**Figure1*:* Scatter plot of observed and fitted values.**

The model was developed at the level of Super Output Areas. The advantage of using this level of aggregation is that it provides a convenient operational definition of at least the “near” neighbourhood, although there are clearly multiple definitions and layering of what constitutes a neighbourhood. In addition, the SOA level aggregation allows to overcome problems of small numbers when using ratios, and is also useful because a key variable used in the model specification, the Index of Multiple Deprivation, is released at this scale. On the other hand the disadvantage of using census polygons is that they are space filling, and that the smaller, or the more dispersed, the population the larger the polygons for reasons of statistical confidentiality. Therefore, areas with relatively few people stand out disproportionately on the map. To overcome this representational problem, the polygon files were converted to a grid, and then filtered through a 1 hectare cell grid based on the postal address file, which contains the location and number of residential addresses at 100 meter accuracy. Figure 2 and 3 below show the observed and fitted values of the geographical distribution of offenders in South Yorkshire at 1 hectare cell.

Having developed the model at the regional level, we were able to run it for the whole of England. Again the results were filtered through the 1 hectare cell Grid which also enables a smoothing process at different scales which can more easily identify patterns free from the constraints of administrative or census geographies. As an example Figures 4 shows England smoothed at 5 kilometres.

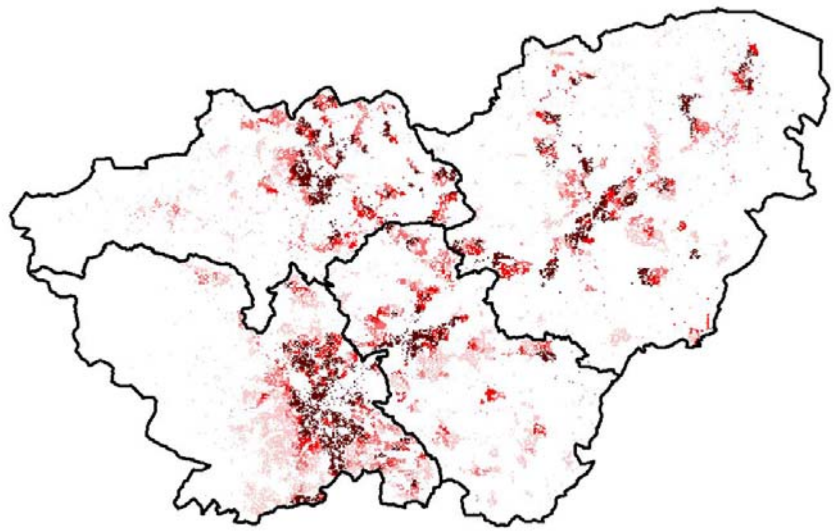

**Figure 2: Observed number of Offenders at 1 hectare level.**

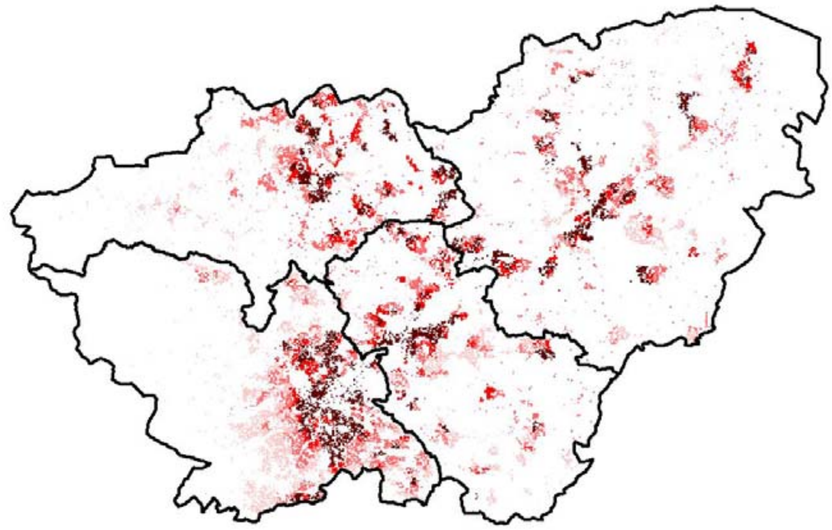

**Figure 3: Modelled number of Offenders at 1 hectare level.**

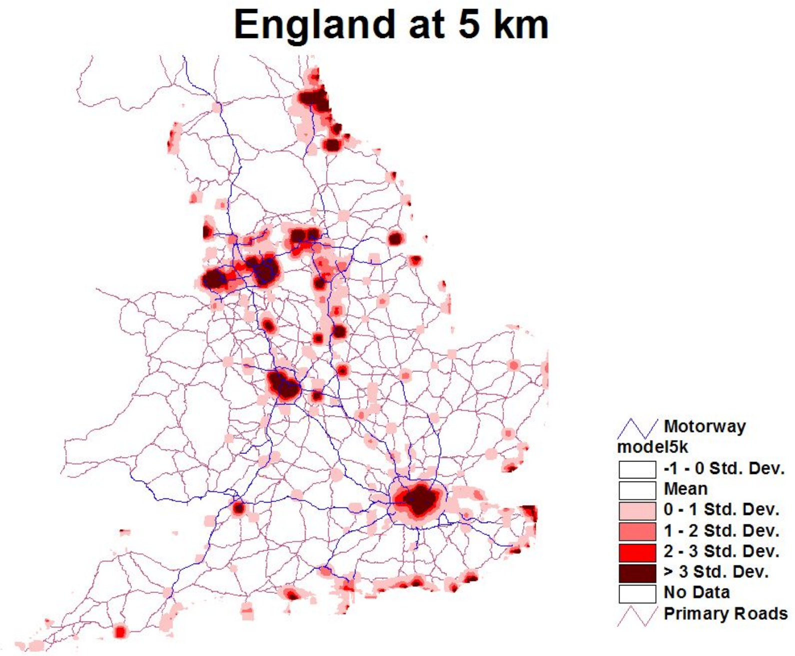


**Figure 4: Modelled distribution Offenders, England, 5km.**

# Development of Grid Applications

The configuration used at Sheffield involves a cluster of 10 Sun machines, connected to two other clusters at Leeds and York forming the White Rose Grid. Application software with which users were familiar (statistical software, SPlus, and Geographic Information System, ArcView) were installed on the Unix cluster but access to the distributed processing capabilities of the Grid it became necessary to write bespoke applications.

A first application developed allows users to retrieve Census data from Mimas, and store it locally to perform subsequent analysis. This first application tested the development toolkit of EASA (Enterprise Accessible Software Application http://www.easa.aeat.com/) an off-the-shelf application environment, which then enables developers to publish the application on a portal for easy access by users. Having proved the value of this solution, the second main application was developed to enable the smoothing of the model results for England at different scales. This was particularly important for the project, because smoothing the data at 5 kilometers (Figure 4) requires calculating for each of the 35 million hectare cells covering England, the mean value across the neighbouring 50 cells, and returning the result for display. Figure 5 shows the portal from which a non-technical user can access the application, and having imported an ASCII file with the data, can select the number of processors on which to operate. The portal then schedules the operation via the Globus Middleware based on the schema showed in Figure 6.

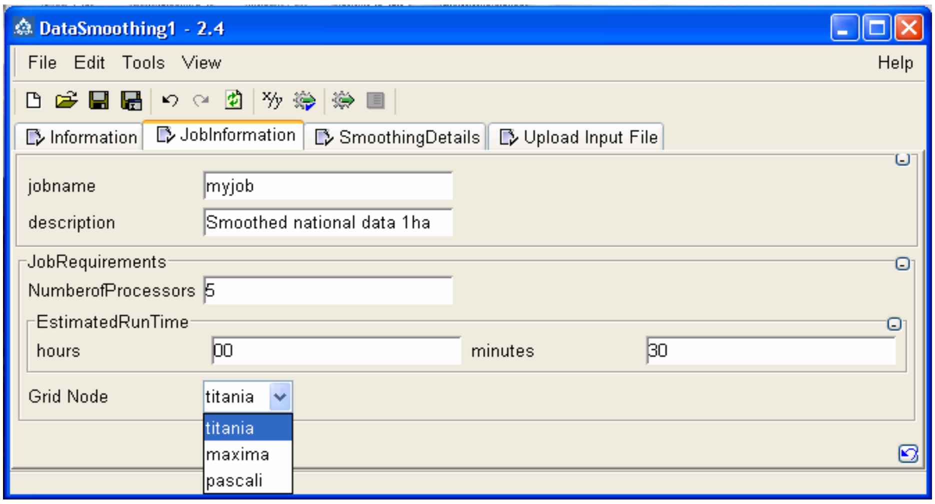


**Figure 5: Scheduling the job from the White Rose Grid Portal**

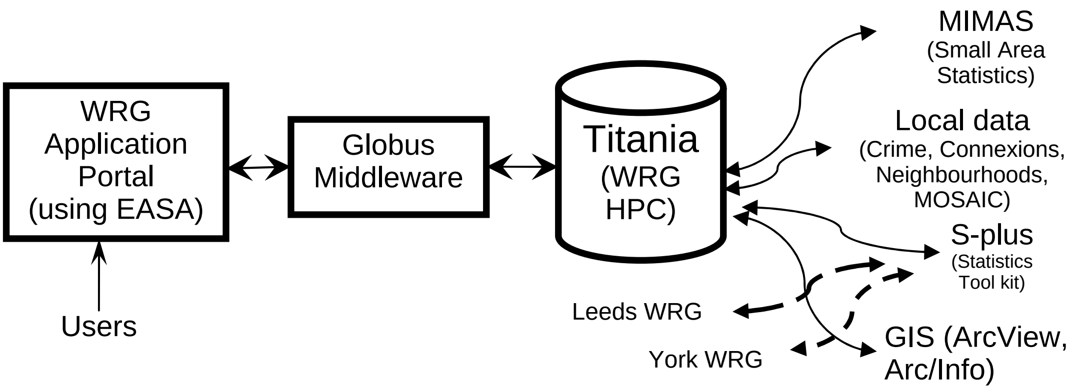


**Figure 6: Overall Architecture of the Application**

The advantages of using multiple high performance computers in the smoothing application are indicated by an analysis of its performance. As an example, the smoothing at 10 kilometers, required 6 and quarter hours on a single Pentium 4 processor (3 GHz) which was reduced to 1 hour when performed over 15 parallel processors on the White Rose Grid. It must be noted however that this performance refers to a bespoke smoothing algorithm because the same smoothing operation done utilizing the internal routines of ArcView Spatial Analysis or ArcGIS is too onerous in computing terms. So the key lesson here is that Grid computing is very useful provided: i) the user is able to access the Grid infrastructure and parallelize the process, and ii) the user has access, or is able to write, the appropriate algorithm since off-the-shelf software he or she might be familiar on the desk top cannot harness the Grid (see also Clematis et al. 2003; and Shi et al. 2002).

The solution found of writing the appropriate algorithm and publishing it on a portal for easy access and use by non-expert users is potentially extremely valuable in opening up the Grid to wide range of users from different disciplines. On the other hand, this would imply that potentially hundreds or thousands of such algorithms currently embedded in desk-top software would need to rewritten and published in this way. This plug-and-play approach would fit with the service-oriented Open Grid Service Architecture towards which the Grid is moving, and the endorsement of large industrial players such as IBM and Sun of the Grid concept and architecture is clearly a welcome step (Hey and Trefethen, 2002). However, this approach would also require the development of interoperable catalogues of services, and appropriate metadata declaring the service capabilities, provenance, conditions and so on. Such catalogues and services would also need to be easily searchable and accessible across multiple portals of the type experimented with in this pilot project, and be able to interoperate regardless of the specific configurations of the Grid as deployed across different projects and communities. Here we see a clear convergence with current efforts underway in the spatial information community aimed at developing Spatial Data Infrastructures (SDIs).

Whilst it is not the purpose of this paper to discuss SDIs in any great detail, it is worth noting that SDIs, defined as "frameworks of policies, institutional arrangements, technologies, data, and people that enable the effective sharing and using of geographic information" share many of the research and operational issues relevant to the Grid, and the Data Grids in particular. Many governments across Europe, and beyond, are currently developing SDIs often in conjunction to their e-government activities (Craglia et al. 2003; Crompvoets and Bregt, 2002; Williamson et al. 2003). At the European level, the European Commission's initiative to develop an Infrastructure for Spatial Information in Europe (http://INSPIRE.jrc.it) is providing a focus for agreements between international standards organisations (ISO, CEN), industry (e.g. Open Geospatial Consortium) and government organisations on many of the technical issues of relevance also to the Grid community. For example, the requirements on service interoperability are providing focus and urgency to research on geospatial semantics (e.g. Kuhn, 2005; Bishr 1998) that would help address some of the challenges highlighted above. Conversely, further development of the Grid across multiple communities of users is of increasing interest to SDIs as they move from the current emphasis on distributed access to data towards also distributed processing. There are therefore many complementarities and synergies to be exploited in joining the efforts between the Grid and the geospatial community (see also Craglia et al. 2005). Where both communities have also a potential shared interest is in engaging with the social dimension of technology, and what Kling (1999) labels as the field of social informatics studies. This is a rich field of research with proven credentials largely within organizational boundaries, and we have yet to see an equivalent body of research addressing either SDI or Grid design and implementation to the same extent. As indicated in the next section, this pilot project was explicitly design to address some of these issues.

# Technological development within a social science research project

One of the distinctive features of the pilot project is that it is both a substantive social science research project and a development of technology project. To this end the participation of an ethnographer helped to explore the development process of e-social science through a co-construction approach to Grid-enabled social science. Through participation, observation and reflection with the ethnographer, a reflexive dimension was encouraged, informing the development process. Collaboration, through the demonstrator, prompted each respective social scientist to reflect on the meaning of their own and colleagues' research practices, methodologies and underlying theoretical frameworks. Discussions in the early part of the development process involved a significant amount of time to create understanding of each others' data so that the multi-dimensional aspect of social exclusion could be explored. Furthermore, the networks of policy-makers and service providers also had to be incorporated in the project. This aspect was dependent on the existing networks that the urban planners, criminologists and sociologists already had developed over time. Computer scientists, technological developers and the social scientists had to learn to understand each other's discourses, facilitated by workshops led by technologists to get ‘behind the tools’. This helped to develop trust and enough knowledge to enable users to shape the Grid. In relation to the technical development, the applied technologist from the University’s Corporate Computing Service participated in discussions among social scientists to gain an understanding of what ‘doing social science’ meant in the context of the pilot. Through this participation the applied technologist could start to undertake some technical development, often using off-the-shelf type solutions (the EASA portal described earlier) in reconfiguring the Grid for the pilot project. This was not always possible, for example the smoothing of the model output required ad hoc technical development due to inability of off-the-shelf GIS software to access the Grid. Although time-consuming, the development has been user-led and demonstrated a relevance to the user-community, alongside robust and secure technical development. Additionally, the research was neither socially nor technically determined at the situated level of the Grid, but data-driven, where the negotiation on the suitability of the datasets and related analyses were revealed as important dimensions in this project, in turn impacting on approaches to training.

One of the key characteristics of this project was that the research team wanted to explore the ways in which participants in the pilot developed e-science through ‘learning by doing’. This meant that the team developed an approach to technological development, design and use that differs from many of the existing approaches by emphasizing the user as a producer of new technology. The team needed to be able to understand the ways in which actors imagined, developed new tools and methods, tried them out, reflected and evaluated on their action, and then developed things further from those reflections. The concern with the formation of agents’ intentions and meaning that create change, and in a process being changed, means that the project needed to address the constitution of ‘meaning’ in the innovation process. To this end, the ethnographer used the concept of performance to establish the link between culture and technological forms, by addressing the relationship between spontaneous action and reflexivity, in which artefacts, meanings and values become meaningful in various contexts of usage. When this is put in the context of the processes of technological innovation it creates an understanding of technological change that is embedded within culture, which gives that change its distinctive characteristic(s) and influences ongoing technological development and research practice.

The connection between theory and practice is significant within an interpretive approach. As computer technology is more malleable, design, in the form of interpretation and adaptation, generates more user-driven development (Dittrich et al., 2002). The interplay of work practices and organisation on the one hand, and ICT on the other, is highly situational. This interplay requires design methodologies that take this situational character into account. The pilot showed how multi-practice design constituencies, interlaced design and use and reflective development of evolving practices and technology shape the definition of e-social science. Research to further such understanding, and to develop means to support such situated design processes, involves understanding the situation at hand, by both contributing to and learning from it. What we call situated innovation, situational adaptation, rejection and provocation of theoretical knowledge, contribute to involved practice and is necessary for the academic production of knowledge that furthers the development and deployment of ICT, such as e-social science using Grid technologies. And from this situated innovation approach, and by the team going through a ‘learning by doing’ approach to developing both method and practice a version of e-social science was conducted. The demonstrator also contributed to informed user-focused training opportunities and facilities that make Grid computing easier for non-experts and beginners, as well as providing some of the technical and social infrastructure for the new Informatics Collaboratory for the Social Sciences (ICoSS) at the

University of Sheffield. For example, it is now possible to generalise operational data on crime and youths, and publish it through a Web Map Server (OGC, 2001) in formats that protect confidentiality on the one hand, but allow overlay and further analysis with distributed policy information from other stakeholders in the region.

# Conclusions

The paper presented and discussed the results of one of the one-year pilot projects funded by the ESRC in the context of its e-social Science Programme. The outcomes of the project have contributed to the theoretical and policy debates in environmental criminology by showing the extent to which it is possible to account for 80% of the spatial distribution of offenders through socio-economic and neighborhood variables Moreover, the strong relationship between the spatial distribution of offenders and that of victims (over 7-0%) adds value to the results. From a technical perspective, the project demonstrated the value of Grid computing to social science. Without it, it would have not been possible to smooth the model results at different scales and identify patterns in their geographical distribution. At the same time, the limitations of the Grid became also apparent. The gap between the application software with which users are familiar and the middleware of the Grid is still very wide, and without the strong input of technologists able to bridge such gap via ad-hoc applications published on a portal, users would have not been able to access the Grid facilities. The portal-based solution is enormously promising for the further use for the Grid by non-specialists, but it emphasizes the need for catalogues of such applications, and interoperable services transparent to the users, and able to chain at the machine level to provide information products to users. The complexity of the research agenda in this area is highlighted further by considering the requirements of multi-disciplinary collaborations. As indicated in this paper, the team had to spend a significant amount of time and resources to learn to understand each other, and develop a shared understanding of each other scientific paradigm, mode of researching, and theoretical perspective, in addition to understanding the meaning of the data available, and of the research questions that could be addressed. Learning about the technology came last. This means that whilst training on technology can address some of the barriers faced, it would take a major effort to try and explicit disciplinary perspectives and unstated assumptions into a formal language that also machines can read. Semantic interoperability is therefore not just about data, but about a much broader social context that provides meaning to data, its analysis, and its interpretation. The research agenda in this field is very rich and needs o be addressed across multiple communities, including the Grid and geospatial communities.

**Acknowledgements**
The pilot project described in this paper was undertaken at the University of Sheffield with funding from the Economic and Social Research Council’s grant number RES-149-25-0027. The views expressed are those of the authors alone and do not necessarily represent those of the organizations involved.

# References

Bishr, Y., Overcoming the Semantic and Other Barriers to GIS Interoperability. *IJGIS*, 1998. 12(4): pp. 299-314.

Clematis A., Mineter M. and Marciano R. 2003. High Performance Computing with geographical data. *Parallel Computing*, 29, 1275-1279.

Craglia M. et al. 2003. *Geographic Information in the Wider Europe*. (www.ec-gis.org/ginie)

Craglia M., Grifiths M. and Costello A. 2005. From SDIs towards the Spatial Grid. *Proceedings of the GISPlanet Conference*, Estoril, Portugal, 30thMay-2nd June 2005. www.gisplanet.org

Crompvoets J. and Bregt A. 2002. World Status of National Spatial Data Clearinghouses. *URISA Journa*l http://www.urisa.org/Journal/APANo1/crompvoets.pdf

Dittrich, Y., Floyd, C. and Klischewski, R. (2002) *Social thinking, software practice*. MIT; London.

Friedrichs, J. and Blasius, J. (2003) Social Norms in Distressed Neighbourhoods: Testing the Wilson Hypothesis. *Housing Studies* 18(6): 807-826.

Haining R. (2003) *Spatial Analysis: Theory and Practice*. Cambridge University Press.

Hey T and Trefethen A. (2002). The UK e-ScienceCore Programme and the Grid. *Future Generation Computer Systems* 18, 1017-1031.

Kling R. 1999. What is Social Informatics, and why does it matter? *D-Lib Magazine*, January, Vol5(1). (http://www.dlib.org/dlib/january99/kling/01kling.html

Kuhn, W. Geospatial Semantics: Why, of What, and How? *Journal on Data Semantics*, Spring 2005 (Special Issue on Semantic-based Geographical Information Systems), Springer LNCS 3534

Office of the Deputy Prime Minister 2004. Indices of Multiple Deprivation, http://www.odpm.gov.uk/stellent/groups/odpm_control/documents/contentservertemplate/odpm_index.hcst?n=4610&l=3

Open GIS Consortium, Inc. *Web Map Service Specification, OpenGIS Implementation Specification*, OGC 01-068r2, January 2001. www.opengeospatial.org

Sampson, R.J., Raudenbush, S.W. and Earls, F. (1997), Neighborhoods and Violent Crime: A Multi-Level Study of Collective Efficacy. *Science* 277, 918-924.

Shi Y. Shortridge, and Bartholic J. 2002. Grid Computing for Real Time Distribute Collaborative Geoprocessing. Paper at the Symposium on Geospatial Theory, Processing, and Applications, Ottawa, 2002 (http://www.isprs.org/commission4/proceedings/pdfpapers/128.pdf)

Shaw, C.R. and McKay, H.D. (1942) *Juvenile Delinquency and Urban Areas*. University of Chicago Press.

Wessels B., Craglia M. and Smith R. 2005. Sharing Techniques and Experience in a Collaborative e-Social Science Pilot Project. *Proceedings of the First International Conference on e-Social Science*, Manchester 22-24 June.

Wikström, P-O.H. and Loeber, R. (2000) Do Disadvantaged Neighborhoods Cause Well-adjusted Children to Become Adolescent Delinquents? A Study of Male Serious Juvenile Offending, Individual Risk and Protective Factors, and Neighborhood Context. *Criminology* 38(4): 1109-1142.

Wiles, P. and Costello, A. (2000) *The Road to Nowhere: The Evidence for Travelling Criminals* Home Office Research Study 207 London: Home Office

Williamson I., Rajabifard A., Feeney M.E. 2003. *Geospatial data infrastructures: from Concepts to Realit*y. London: Taylor and Francis

Wilson, W.J. (1987) *The Truly Disadvantaged*. Chicago: University Press.