Transactions in GIS

Research Article

Transactions in GIS, 2015, 19(4): 499–515

GI Expertise

Matt Duckham

School of Mathematical and Geospatial Sciences, RMIT University, Australia

Abstract

This article concerns the identification of the expertise that GI scientists have in common. Rather than attempt to be exhaustive, the approach is to identify a minimal set of expert topics shared across GI science. In this article, a set of five expert topics is proposed for this purpose: structure, uncertainty, dynamism, language and cognition, and design with geographic information. The article then examines the extent to which the literature lends support for these five topics, using two analyses: first, a qualitative analysis of the papers found in the 2006 *Classics from IJGIS*; and second a citation analysis investigating the occurrences of keywords related to the proposed expert topics. The results suggest that these five areas of expertise do indeed tend to distinguish research in GI science from that in other fields, and in particular where the expert topics occur in combination. The establishment of an acknowledged core of shared expertise should assist in interdisciplinary research and teaching collaboration, usually founded on a shared understanding of what expertise different partners can contribute.

1 Introduction

What expertise do GI scientists possess? If GI science is a coherent discipline, then it ought to be possible to identify a common core of expertise shared by all those within the discipline. Further, for GI science to be a distinct field of research worthy of its own journals and conferences, such a core of expertise would be expected to be unique and distinct from the expertise found in other, related disciplines.

This article examines the question of whether it is possible to identify a unique and shared core of GI science expertise, such as would satisfy the expectations above. The objective of doing so is primarily pragmatic. In an environment of increasing multi- and inter-disciplinarity, GI scientists routinely collaborate with experts in many other fields, such as geography, computer science, measurement science, mathematics, statistics, psychology and design, to name but a few. Broad agreement on a core of shared expertise supports increased collaboration with other disciplines, since it provides a priori expectations about what each field might be expected to contribute. Further, an inventory of the expertise common to GI scientists provides a description of the field, complementing pre-existing definitions of "GI science."

This article briefly reviews the relationship to existing definitions, competencies, and common knowledge in GI science (Section 2). Five areas of expertise are then proposed (Section 3)

Address for correspondence: Matt Duckham, School of Mathematical and Geospatial Sciences, GPO Box 2476V, RMIT University, Victoria 3001, Australia. E-mail: matt.duckham@rmit.edu.au
Acknowledgments: The author would like to thank Mike Goodchild for proposing the idea of analyzing the expertise found the *Classics from IJGIS* and for helpful comments on a draft. Tighearnach Dunne assisted greatly with performing the Scopus data analysis. The author would also like to thank Waldo Tobler and the three anonymous reviewers for their constructive comments. Finally, the author would like to than the *Classic* author respondents for their assistance: Martin Charlton, Robert Weibel, Gerard Heuvelink, Andrew Skidmore, Harvey Miller, Mike Goodchild, Jonathan Raper, Piotr Jankowski, Stuart Fotheringham, Keith Clarke, Gennady and Natalia Andrienko, David Mark, and Marcus Llobera.

 doi: 10.1111/tgis.12166

that together are argued to be the unique and common core of GI science expertise. Section 4 attempts to test this proposition first qualitatively, by examining classic papers in GI science, and then quantitatively, through citation analysis. Finally, Section 5 concludes the article, including a discussion of the weaknesses of the approach.

2 Definitions, Competencies, and Common Knowledge

There are two distinct approaches to defining a term or concept. The first is by *intension*: specifying the exact meaning of the term or concept. The second is by *extension*: with reference to the set of things that the term or concept refers to.

In defining "GI science," the field has already developed a small number of widely agreed upon *intensional* definitions. The term was originally coined in 1992 by Michael Goodchild (Goodchild 1992) to better distinguish between the technologies associated with geographic information (primarily geographic information systems, GIS) and the science behind these technologies (i.e. geographic information science, GI science). Subsequent work has elaborated on the definition of GI science as "the science behind the systems," emphasizing different aspects, including:

> the discipline that uses geographic information systems as tools to understand the world. Clarke (1997)

and:

> the development and use of theories, methods, technology, and data for understanding geographic processes, relationships, and patterns. UCGIS bylaws, quoted in Mark (2003)

An attempt at a fuller, more detailed definition was the focus of a 1999 National Science Foundation workshop (Mark 2000), while others have considered in detail specific aspects of the definition, content, history, focus, and contribution of GI science (see, for example, Goodchild 2010; Goodchild 2011; Longley 2000; Schuurman 2000).

However, even in the context of such broadly agreed intensional definitions, a definition by extension may still be helpful. Intensional definitions tend to come "after the fact." A clear and agreed definition of GI science does not necessarily entail that those who consider themselves GI scientists will adhere to the definition (in other words: "Sayin' it don't make it so"). By contrast, it can be argued that expertise is more tangible. Expertise is a primary focus of training. People are usually able to identify, to a greater or lesser degree, what expertise their peers and they themselves possess. Indeed, one common definition of experts is as "those who have been recognized within their profession as having the necessary skills and abilities to perform at the highest level" (Shanteau 1992). Thus, an extensional definition should have the characteristic of being based on "how things are" rather than "how things should be." Section 4 attempts to test this proposition.

Another existing source of information about the expertise connected with GI science is the *Geographic Information Science and Technology Body of Knowledge* (GIS&T BoK) (DiBiase et al. 2006). The GIS&T BoK identifies 10 "knowledge areas" and more than 330 topics from a top-down perspective as a resource for curriculum design. Ongoing community-led efforts to develop a second edition GIS&T BoK (http://gistbok.org) have 11 top-level "concepts" (closely related to the first edition knowledge areas) and identify more than 1,600 concepts in total. The GIS&T BoK is all-encompassing, aiming to include any topic that might reasonably be part of the GI science "educational infrastructure." Many units in the GIS&T BoK are acknowledged to overlap substantially with expertise in other disciplines (e.g. GD7 "Land surveying and GPS" with overlaps to surveying and geodesy; GD10 "Aerial imaging and photogrammetry "with overlaps to photogrammetry; GD11 "Satellite and shipboard remote sensing" with overlaps to remote sensing; and DM2 "Database management systems," with overlaps to computer science).

Rather than be all encompassing, the objective in this article is to identify a *minimal* set of areas of expertise that together uniquely define the shared knowledge of GI scientists. There are GI scientists who certainly possess expertise on high-level GIS&T BoK units such as "Cellular automata" (GC3) and "Genetic algorithms" (GC5), "Legal aspects" of GIST (GS1), or "Managing the GI system operations and infrastructure" (OI2). However, one is likely to encounter many GI scientists (including the author) who would not regard themselves as experts in these and other significant expertise areas. Thus, it is not surprising that while the core areas of expertise identified in this article may all be found within the GIS&T BoK, many more topics found in the GIS&T BoK do not feature as core areas of expertise in Section 3 below.

Similarly, the core competencies listed in the US Department of Labor (DOL) Core Geospatial Abilities and Knowledge (http://www.careeronestop.org/competencymodel/competency-models/geospatial-technology.aspx) clearly intersect with GI science expertise. But as a result of the broader scope of the core competencies model (geospatial abilities and knowledge generally, rather than GI science expertise) it is again to be expected that the listed competencies would seem neither individually to uniquely define GI science, nor in aggregate to be shared across all GI scientists.

Other related disciplines, including geography, have also grappled with expertise. Adopting a developmental perspective has provided insights into the characteristics of geography education necessary to train experts (e.g. Downs 1994, Huynh and Sharpe 2013). However, here too it is acknowledged that there is "as yet no generally accepted definition of geographic expertise" (Downs 2014).

Thus, even in the context of existing definitions, competencies, and common knowledge, the key question addressed in this article remains unanswered: "What unique expertise is common across GI science?"

3 Core Expertise

As foreshadowed in the discussion above, the desirable characteristics of a set of core expertise areas for GI science include:

Shared: The areas of core expertise should be shared across all those who identify themselves as GI scientists. In short, we would expect the selected topics to be ones that *any* GI scientist could be expected to discuss knowledgeably with another acknowledged expert in that area.

Minimal: Rather than attempt to be exhaustive, we aim for a minimal list of core expertise. In such a list, the removal of any single expertise would be acknowledged as failing to adequately capture the essential nature of the field.

Despite these clear criteria, the precise choice of what topic areas should be included as "core" expertise in GI science is necessarily somewhat subjective. Nevertheless, in this section five leading candidates are proposed and briefly introduced: expertise in uncertainty, in dynamism, in spatial language and cognition, in design, and in the structure of geographic information. These five candidates arose out of the 2014 Waldo Tobler Distinguished Lecture in GIScience at the AAG annual meeting in Florida, and the "Geographic Information Science" entry for the forthcoming Wiley/AAG International Encyclopaedia of Geography (Duckham 2015) by the author. The expertise areas elaborated upon in the current article are substantially the same as those earlier proposals, with the exception that a sixth area of expertise, "scale dependence," has been subsumed under the "structure of geographic information".

The subsections below introduce and justify the inclusion of each of the five core areas of expertise, as well as briefly explaining the exclusion of other likely candidates (under the heading of "Auxiliary expertise"). Section 4 then attempts to test this selection against the literature.

3.1 Structure

Ask any GI scientist what he or she regards as the most important principle that underpins his or her expertise, and the answer will almost always be: "Tobler's First Law" (TFL). TFL posits that near things in space are more similar than distal things, and is measured statistically as "spatial autocorrelation": the degree of interdependency of observations in geographic space.

Autocorrelation is not the only inherent structure in geographic information. Spatial heterogeneity (also referred to as the statistical concept of "non-stationarity") is an inherent feature of geographic information sometimes cited as a companion fundamental "law" to TFL (e.g. Goodchild 2011). Spatial heterogeneity arises because the systematic similarity of nearby places (TFL) itself varies spatially, leading to an expectation of "clumps" of things in geographic space. Spatial heterogeneity is closely related to the modifiable areal unit problem (MAUP), another topic familiar to any GI scientist. MAUP arises because the results of any analysis derived from information about arbitrary (i.e. "modifiable") geographic regions (i.e. "areal units") may not be independent of those regions. Arguably, the topic of spatial analysis is chiefly concerned with this structure (cf. Openshaw 1990; Ding and Fotheringham 1992).

Scale-dependence can be regarded as another aspect to the inherent structure of geographic information. Many geographic phenomena and processes look different depending on whether they are being observed in more or less detail. For example, a topographic feature that looks like a "peak" (higher than its neighbors) at a fine level of detail may instead look like a "pit" (lower than its neighbors) as we zoom out to a coarser, more granular level of detail (as neatly illustrated in Wood 1996).

Regardless of how this structure is described or manifests itself, a sophisticated understanding of the inherent and non-random structure of geographic information is a clear candidate for inclusion in the domain of expertise of any GI scientist.

3.2 Uncertainty

Uncertainty is acknowledged as an endemic feature of geographic information (cf. Chrisman 1984; Fisher 1989; Goodchild 1995). Uncertainty arises from the inherent imperfection of information. Imperfection is often decomposed into imprecision (a lack of detail), inaccuracy (a lack of correctness), and vagueness (the existence of borderline cases) (Worboys and Clementini 2001). Any measurement of location, for example, is necessarily of limited accuracy and precision. Many topographic features are vague (e.g. "Mount Snowdon"), and arguably most toponyms are too (at least in how they are used, even if not in official gazetteers or maps). Common spatial relations, such as "near" or "left of," are similarly vague. Attempts to describe and quantify the unavoidable imperfection of geographic data sets also have a long history, for example in the spatial data quality standards literature (e.g. NCDCDS 1988).

One might argue that despite decades of work, today's spatial databases and GIS still perform poorly in managing and querying information about imperfection, and supporting decision-making under uncertainty. However, the lack of a solution to this problem does not preclude expertise. Indeed, experts in solved problems are typically in low demand. Thus, a sophisticated understanding of the causes and consequences of uncertainty in geographic information is a reasonable contender for a core expertise in GI science.

3.3 Dynamism

Like uncertainty, dynamism is another endemic feature of geographic information. There exist no geographic phenomena that are truly static, and so changes to our information about the geographic world are always to be expected. Attempts to represent and reason with change have consistently been amongst the most active research topics in GI science over its history. GI science frequently incorporates influential ideas about time from disciplines such as geography (e.g. time geography, cf. Hägerstrand 1970; Miller 2005), philosophy (e.g. SNAP/SPAN ontology, cf. Grenon and Smith 2004; Goodchild et al. 2007), and design (e.g. Minard's famous figurative map, cf. Kraak 2014).

As for uncertainty, it is true that today's GIS and spatial database technologies are not especially well adapted to dealing with dynamic information. Nevertheless, a sophisticated understanding of the importance of events, processes, change, and movement are clearly evident across the breadth of GI science.

3.4 Language and Cognition

Geography is fundamentally an experienced space (cf. Zubin's C-space or Montello's geographical space; Montello 1993). Thus, GI science has always been intimately involved with the unavoidable mismatch between human understanding and automated reasoning about geographic space. Systematic distortions are not simply unwanted "bugs" but are acknowledged to be integral features of human spatial cognition (Tversky 1992). Hierarchies, for example, may provide efficient structures for humans to remember and reason about space. But the same structures are implicated in well-studied discrepancies between objective measurements and human understanding of geographic space (e.g. Stevens and Coupe 1978; Hirtle and Jonides 1985). On the other hand, the shared meaning humans ascribe to space is notoriously difficult to capture using automated systems. Toponyms, for example, are frequently used by humans to share and communicate information about geographic places. However, toponyms continue to present significant challenges to automated information retrieval (cf. Hollenstein and Purves 2011).

It seems reasonable to argue, then, that the expertise in the differences between, on the one hand, human language and cognition about space and, on the other, automated computing with information about space, is common to all GI scientists.

3.5 Design

The history of GIS is intimately connected with cartographic design and automated mapping. A substantial body of early work in the field, for example, was concerned with algorithms for cartographic generalization and map labeling (e.g. Brassel and Weibel 1988). Basic cartographic design principles remain at the foundation of geovisualization, concerned not only with static map-based representations, but also with dynamic, interactive, and exploratory interfaces to geographic information (e.g. Maceachren 1994). Today, topics such as cartography, geovisualization, and geovisual analytics are well represented across the breadth of the conferences, journals, and taught curricula closely connected with GI science. Thus, GI science continues to be concerned with designing interfaces to geographic information in a way that promotes human understanding of that information.

Expertise in geographic information design, whether traditional cartographic products or more interactive computer-based interfaces, seems then another strong contender as a core expertise for GI science.

3.6 Auxiliary Expertise

While most GI scientists would hopefully agree with the inclusion of these five core areas of expertise, many, if not most, GI scientists may also possess expertise *outside* of our five core topics. Thus, some justification is required for those areas of expertise that are excluded, including:

- databases and algorithms, for which, of course, considerable expertise exists within computer science;
- qualitative spatial reasoning, cellular automata, agent-based modeling, and other topics frequently studied in connection with artificial intelligence;
- the ontology of spatial and spatiotemporal information, topics strongly associated with philosophy and computer science;
- Earth measurement and sensing, as the primary focus of study, for example, in geodesy, the measurement sciences, and remote sensing; and
- the social context of GIS, a focus, for example, in critical geography.

In these and other examples, it is not hard to identify acknowledged experts in those areas who would also be acknowledged as expert GI scientists. However, for each of these topics our argument is the same: that these topics are not sufficiently common across the breadth of GI science to count as *shared* expertise. Note that this is not to say that those with expertise in such areas are not also expert GI scientists. Rather, the argument here is that expertise in these auxiliary topics is not required to be considered an expert GI scientist.

4 Evaluation

Having proposed five areas of core expertise in GI science, what evidence exists to support this proposal? This section presents two separate evaluations, examining the extent to which the five areas of GI expertise capture aspects of the GI science research literature. Although the evaluations generate numerical data, both are also unavoidably qualitative and subjective in key aspects. The first evaluation looks specifically at the expertise encapsulated in Fisher 2006, "Classics from IJGIS" published by Taylor and Francis (Section 4.1). The second compares a citation analysis from a selection of quality journals inside and outside the field (Section 4.2).

4.1 Classics from IJGIS

IJGIS (the *International Journal of Geographical Information Science*) is widely regarded as one of the leading journals in the field of GI science. To mark the 20th anniversary of the journal, Peter Fisher edited a retrospective collection of 19 of the leading IJGIS articles from the previous two decades, called "Classics from IJGIS" (Fisher 2006). A natural question to ask, then, is how are the five areas of expertise represented across these classic IJGIS papers?

In choosing the classics, Peter Fisher was guided primarily by numbers of citations to articles, but with the constraints that: the final selection was spread relatively evenly over the 20 years; no two articles were written by the same first author; application or literature review articles were excluded; and the final selection contains a representative spread of topics (see Fisher 2006, Chapter 1). Clearly, the selection is subject to a range of biases, not only in Peter Fisher's selection process, but also the biases inherent in citation counts and the scientific review process. However, none of these expected biases would seem to rule out this corpus as a reasonable candidate for further investigation of GI expertise.

To make the assessment process more concrete, five questions were asked of each paper. These five questions were:

"To what extent does each paper concern:

1. the underlying structure of geographic information, including autocorrelation and scale dependence (G);
2. uncertainty (accuracy, imprecision, and/or vagueness) in spatial information (U);
3. dynamism (changes, events, processes, movement) in spatial information (D);
4. human spatial language or human spatial cognition (C); or
5. cartographic design or geovisualization (V)."

For consistency, three answers were possible: (1) "primary" (P) if the topic is the main focus of an article; (2) "secondary"(S) if the topic is covered, but is not a main focus of an article; or (3) "unrelated" (.) if the topic is not connected with the article. The further constraint was made that each article was allowed at most one primary topic. However, any other combinations of "P," "S," and "." were permitted.

Table 1 summarizes the results derived from this qualitative analysis. The top row from each article shows the responses of the author for this article. Where available, the second row for each article provides the responses to the same questions from one of the authors of the classic article itself (this author is indicated in brackets as the "responding" author). To ensure independence, the responses of the author of this article were made before seeking responses from the classic article authors; and the classic article authors were not aware of these nor any other responses when formulating their own answers. Responses were received for 13 out of the 18 papers where a response was sought (sadly Peter Fisher himself, author of the 19^{th} paper, chapter 9, died in 2014).

4.1.1 Discussion

Comparing the two sets of responses, the results do indicate a relatively high degree of agreement. For example, in 10 out of 12 cases, the primary focus "P" identified by the author of this article exactly matched that identified by the responding author of the article itself (83% agreement). Overall, 43 out of 65 answers agreed exactly (66% agreement). Neither these results nor the methodology will be argued to be statistically robust. However, the relatively high level of independent agreement for the primary and secondary areas of expertise does lend confidence in these otherwise subjective judgments. Sceptical readers are encouraged to repeat the exercise for themselves, and compare the results in Table 1 with their own answers or those of a colleague.

If we accept the argument that the results in Table 1 are indeed a relatively robust and unbiased assessment of the expertise evident in these classic papers, what can these tell us about our chosen areas of expertise? Overall the topic of the "structure of geographic information" (G) is, as might be expected, the most frequently cited, especially as a secondary focus.

Table 1 Relationship between the five expert topic areas (G: structure of geographic information, U: uncertainty, D: dynamism, C: language and cognition, V cartographic design and geovisualization) and 19 *Classics from IJGIS* (Fisher 2006). "P" indicates the primary focus of a paper; "S" indicates a secondary focus; "." indicates the topic is not a focus of the article. The top line in each row indicates the responses of the author for this article; the second line, where present, indicates the responses of one of the authors of the classic article itself

Chap	First author (Responding)	Title	G	U	D	C	V
2	Openshaw (Charlton)	A Mark 1 Geographical Analysis Machine for the Automated Analysis of Point Data Sets	P (P)	S (S)	S .	. .	. .
3	Brassel (Weibel)	A Review and Conceptual Framework of Automated Map Generalization	. (S)	. (S)	. .	S .	P (P)
4	Heuvelink (Heuvelink)	Propagation of Errors in Spatial Modelling with GIS	S (S)	P (P)	. .	. .	. .
5	Skidmore (Skidmore)	A Comparison of Techniques for Calculating Gradient and Aspect from a Gridded DEM	S (S)	P (P)	. .	. .	. (S)
6	Worboys	Object-Oriented Data Modelling for Spatial Databases	S	.	.	.	.
7	Egenhofer	Point-Set Topological Spatial Relations	P	.	.	.	.
8	Miller (Miller)	Modelling Accessibility Using Space-Time Prism Concepts within GIS	. (S)	. .	P (P)	. .	. .
9	Goodchild (Goodchild)	Geographical Information Science	S (S)	S (S)	S (S)	S .	S (S)
10	Fisher	Algorithm and Implementation Uncertainty in Viewshed Analysis	S	P	.	.	.
11	Raper (Raper)	Development of a Geomorphological Spatial Model Using Object-Oriented Design	. (S)	. (S)	P (P)	. (S)	. .
12	Jankowski (Jankowski)	Integrating Geographical Information Systems and Multiple Criteria	S .	. (S)	. .	S (P)	. (S)

(Continued, next page)

Table 1 *Continued*

Chap	First author (Responding)	Title	G	U	D	C	V
		Decision-Making Methods					
13	Fotheringham (Fotheringham)	The Geography of Parameter Space	P (P)	. (S)	. (S)	. .	S (S)
14	Frank	Qualitative Spatial Reasoning	S	S	.	P	.
15	Kiiveri	Assessing, Representing, and Transmitting Positional Uncertainty in Maps	S	P	.	.	S
16	Gaydos (Clarke)	Loose-Coupling a Cellular Automaton Model and GIS	S (S)	. (S)	S (P)	. .	. (S)
17	Bishr	Overcoming the Semantic and Other Barriers to GIS Interoperability	S	.	.	S	.
18	Andrienko (Andrienko)	Interactive Maps for Visual Data Exploration	. .	. .	. .	. .	P (P)
19	Smith (Mark)	Geographical Categories: An Ontological Investigation	. .	. .	. .	P (P)	. .
20	Llobera (Llobera)	Extending GIS-Based Visual Analysis	. (S)	. .	. (S)	P (P)	S .

However, it is notable that the expert topics have a reasonably even level of representation across the results. For example, looking at the responses from the classic paper authors, each of the expert topic areas has either two or three papers with that topic area as a primary focus. Recall that these responses from classic paper authors were made without knowledge of other classic authors' responses or of the responses by the author for this article.

From the other perspective, three articles were scored as having *no* primary focus within the five areas of expertise (chapters 6, 9, and 17; first authors Worboys, Goodchild, and Bishr). In the cases of chapters 6 and 17 ("Object-Oriented Data Modelling for Spatial Databases" and "Overcoming the Semantic and Other Barriers to GIS Interoperability"), these classic papers represent a "bridge" to GI science from related areas of expertise (specifically from artificial intelligence and computer science). Arguably, these papers are classics in the GI science field because they were amongst the first to place fundamental concepts from AI (object-orientation and semantic interoperability) within a GI science context. The underlying topics, though, were at the time already long-established areas of expertise in their home disciplines. (Both object-oriented concepts and computing with semantic networks first emerged in AI in the 1960s.) The third exception, (chapter 9, "Geographical Information Science," Goodchild) is not strictly an article *containing* new research in the domain of GI science; rather it is an article *about* GI science research. As such, it is more of a meta-article about the characteristics and state of research in the field, much as this current article is.

In summary, while somewhat subjective, these results do appear to be relatively robust, with moderately good agreement between independent assessments. The different areas of expertise all seem to be well represented. Further, even in cases where classic papers appear to exhibit low levels of expertise in any of the core areas, there exists a plausible rationale to explain these exceptions. Thus, although qualitative and necessarily somewhat subjective, the spread of responses to the five questions of the *Classics from IJGIS* papers does seem to support our key contention: that the five areas of expertise provide a reasonable characterization of classic research in GI science.

4.2 Citation Analysis

While the analysis in the previous section yields initial evidence to support the selection of the five areas of expertise, it does not easily scale to larger bodies of published work or facilitate comparison with other fields. This section presents a comparative citation analysis upon a much larger body of literature and across multiple fields.

In brief, our analysis looks for occurrences of keywords related to our five areas of expertise in articles published in a range of journals associated with GI science and other related fields. In the following subsections we examine first the selection of journals, and then the selection of keywords, before presenting the analysis results.

4.2.1 Journal selection

Table 2 provides a list of journals, drawn from across a range of disciplines, that forms the corpus for the citation analysis. Given the sheer number of journals published today, combined with the increasingly blurred lines between disciplines, the selection of this corpus was inevitably the most subjective part of this citation analysis. Although somewhat arbitrary, the rationale and process of selection for inclusion in the list is as follows.

Our selection began with four of the leading GI science journals listed on Scopus: *IJGIS*, *Transactions in GIS*, *Cartography and Geographic Information Science*, and *Geoinformatica*. From this core we then added more than 20 comparable journals from fields close to GI science, including leading journals that publish a significant quantity of GI science research (even if GI science is not the primary focus), such as *Computers, Environment and Urban Systems*, and *the Annals of the AAG*. Specifically, the list includes journals from the fields of computer science, geography, remote sensing, planning, bioinformatics and ecology, health and economics informatics. These topics were chosen because, like GI science, they address long-standing research questions that have been transformed by new information technologies. A further constraint on the wider selection that was that: (1) only journals listed in Scopus were considered (to facilitate our subsequent analysis); and (2) as far as possible, selected journals had comparable levels of impact to the core GI science journals (as measured by SCImago Journal Rank indicator, SJR). While the core selected journals varied in SJR score from 0.564 to 0.996, the entire set of selected journals varied from 0.462 to 4.118. The highest SJR score for any journal in SJR is 45.894, with the median score 0.214.

Although no purely objective procedure for the journal selection could be formulated, it is argued that the list in Table 1 is at least *representative* of leading journals in GI science and similar journals in closely related fields. It can also be argued that the list is not strongly biased, beyond the constraints identified above (informatics-related fields, citation levels, and inclusion

Table 2 Selected journals including highlighted core selection of leading journals with a significant focus on GI science contributions

Journal	SCImago Journal Rank indicator (SJR)
ACM Computing Surveys	3.113
ACM Transactions on Computer Systems	0.526
Annals of the Association of American Geographers	1.635
Behavioral Ecology	1.749
BMC Bioinformatics	1.929
Cartography and Geographic Information Science	0.848
Computers, Environment and Urban Systems	1.079
Ecological informatics	0.836
Environmental Science and Policy	1.674
Environmental Science and Technology	2.979
Geoinformatica	0.564
Global Ecology and Biogeography	4.118
Health Informatics Journal	0.462
IEEE Transactions on Industrial Informatics	2.666
IEEE Transactions on Visualization and Computer Graphics	1.402
IEEE/ACM Transactions on Computational Biology and Bioinformatics	0.72
Information Economics and Policy	0.537
International Journal of Geographical Information Science (/Systems)	0.996
International Journal of Remote Sensing	0.776
ISPRS Journal of Photogrammetry and Remote Sensing	2.055
Journal of Biomedical Informatics	1.097
Journal of Environmental Management	1.211
Journal of Geographical Sciences	0.58
Journal of Geotechnical and Geoenvironmental Engineering	2.824
Landscape and Urban Planning	1.327
PLoS Computational Biology	3.274
Transactions In GIS	0.689

in Scopus). Again, the sceptical reader is encouraged to select different journals and repeat our simple analysis below to compare further results.

4.2.2 Keyword selection

The second partially subjective part of our analysis is to define a set of keywords that are unambiguously associated with our five areas of expertise. To reduce the set of possible keywords to terms that are clear candidates for GI science expertise, only words that appeared in entry titles in the 2008 "Encyclopedia of Geographic Information Science" (Kemp 2008) were considered. Then from that set for each expertise area, the two keywords were selected that are most closely associated with that area of expertise. The selected terms are given in Table 3.

Terms that were likely to be excessively ambiguous were, however, discarded. As a result, the expertise in "dynamism" had only one keyword describing it ("temporal") as the other two

Table 3 Keywords selected to describe the five areas of expertise, drawn from the set of entry title words in the Encyclopedia of Geographic Information Science (Kemp 2008)

Label	Expertise	Selected keywords
G	Structure of geographic information	autocorrelation, scale
U	Uncertainty	uncertainty, accuracy
D	Dynamism	temporal
C	Language and cognition	cognition (or cognitive), reasoning
V	Design	visualization (or visualisation), cartography (or cartographic)

options ("change" and "event") were deemed too frequently used in senses quite different to the technical sense of "dynamism in geographic information."

Clearly, the selected keywords do not capture the complexity of the expertise area (as indeed no set of keywords could). However, they are expected to occur frequently in connection with any expert discussion of that area. Consequently, these keywords are arguably adequate terms for searching for literature in that area.

4.2.3 Occurrence of keywords

The analysis then proceeded by searching through the articles published in the selected journals for occurrences of the keywords associated with the expertise areas (in Table 3).

Rather than search through the full text or titles of each article, it was decided to search only the titles of cited references in each article. This approach was chosen in an attempt to provide a degree of standardization against idiosyncratic or ambiguous uses of keywords. In short, the titles of references for an article are expected to provide a more robust indication of the expertise upon which that article draws, than the title of the article itself. However, a similar analysis of the terms in the titles of articles showed this decision had only a marginal effect on the results, as the occurrence of terms in titles and cited references is highly correlated (Pearson's correlation coefficient of 0.84 to 0.99 across the five expertise categories).

Using the Scopus bibliographic engine, data was generated about the number of articles that cited references with titles which contained terms from our keyword list (Table 3). The data was generated for each journal (for all issues in Scopus prior to 2015), and for each area of expertise (G, U, D, C, V). For example, Figure 1 shows the percentage of articles in each journal where the titles of cited references contained keywords relating to at least one of the identified expertise areas (blue bars) as a proportion of the total number of articles published in the journal (prior to 2015).

It is immediately noticeable from Figure 1 that those journals most strongly associated with GI science also tend to be associated with higher proportions of articles that cite references with expertise-related keywords in their titles. *IJGIS*, *Cartography and Geographic Information Science*, *Transactions in GIS*, and *Geoinformatica* all appear in the top seven journals.

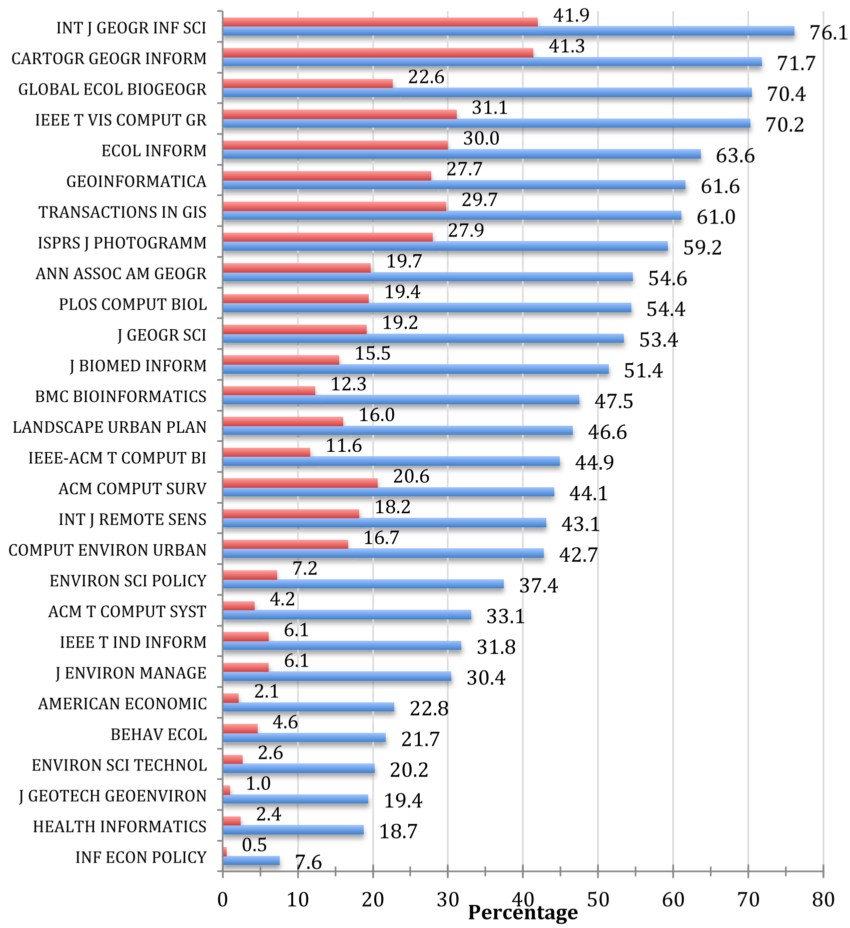

Figure 1 Proportion of articles in selected journals where the titles of cited references contain keywords from: (1) at least one of the different expertise areas (blue bars); or (2) at least two of the different expertise areas (red bars). Full titles of selected journals are listed in Table 3. Data: Scopus

4.2.4 *Co-occurrence of keywords*

In addition to the occurrences of keywords in cited references, Figure 1 also shows the proportion of articles containing co-occurrences in cited reference titles of keywords related to at least two expertise areas (shorter, red bars). These co-occurrences might occur in different *cited* references, but must occur in references cited in the same published article (in the specified journal). (For example, this article contains citations with the terms "cognitive" and "scale" in the titles, even though it does not contain a reference with both terms "cognitive" and "scale" in a single title.) The intuition behind this refinement is to begin to examine the *combination* of expertise found in GI science. Indeed, a random selection of 100 journals from the entire Scopus listing (more than 20,000 active English language journals) exhibited a much lower frequency of such co-occurrences in general. The median proportion of papers with such co-occurences across the randomly selected journals was 0.3%, with a maximum proportion of 19% out of any tested (found in *Mental Health Services Research*).

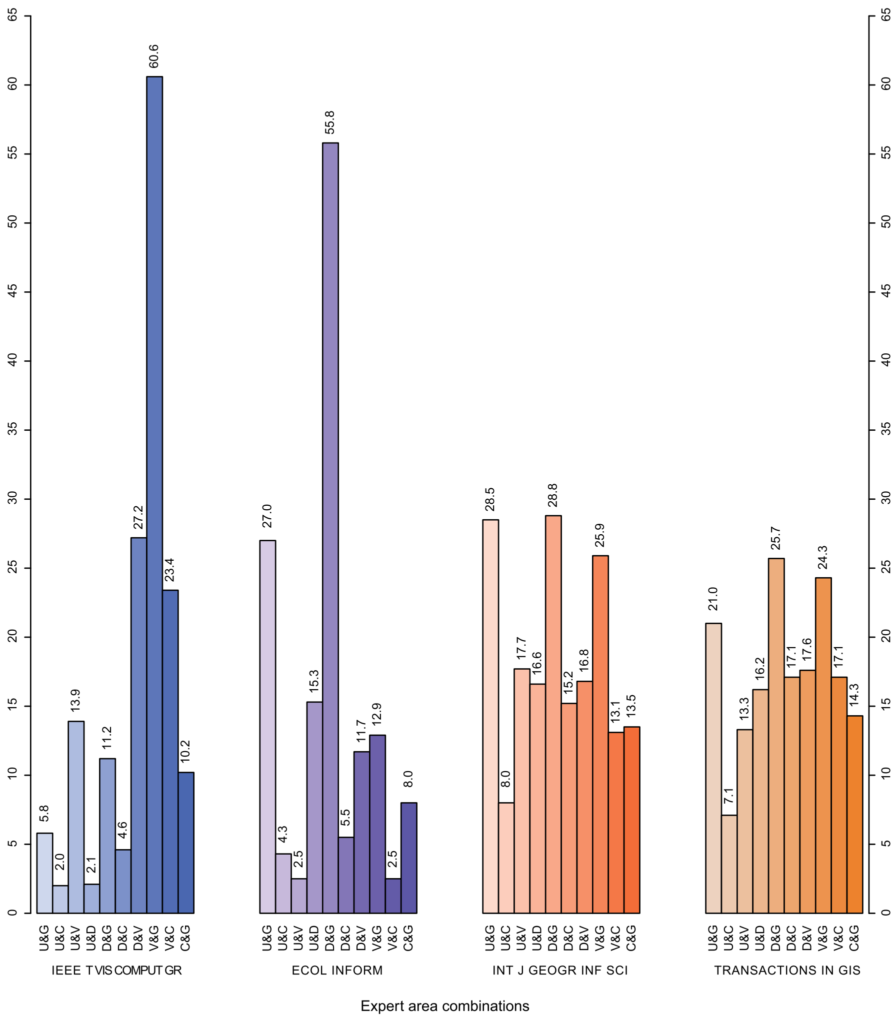

Figure 2 Percentage of articles where specific pairs of expert-topic keywords co-occur in the titles of cited references, from four selected journals. The pairs of expert-topic keywords relate to those specified in Table 3, e.g. "G&U" concerns the pairs of keywords for G, structure of geographic information ("autocorrelation" or "scale"), and U, uncertainty ("uncertainty" or "accuracy"). Data: Scopus

Thus, the co-occurrence of our keywords appears to provide a partial "fingerprint" for GI science journals. GI science journals are frequently associated with higher proportions of articles citing references with keywords in the titles relating to two or more different expertise areas. However, it is noticeble that even using co-occurence of terms, in our sample of 28 a number of journals less closely associated with GI science remain highly ranked: most notably *IEEE Transactions on Visualization and Computer Graphics* (rank 3rd) and *Ecological Informatics* (rank 4th).

Drilling down, Figure 2 shows the co-occurences of keywords from specific pairs of expertise areas (which together make up the overall co-occurrence data shown in Figure 1b) for these two anomalous journals, *IEEE Transactions on Visualization and Computer Graphics* and *Ecological Informatics* as well as for two acknowledged GI science journals, *IJGIS* and *Transactions in GIS*. The results highlight that keywords related to certain specific pairs of expertise areas occur frequently in the cited references of *IEEE Transactions on Visualization and Computer Graphics* (for example, those related to the G and V keyword pair) and *Ecological Informatics* (for example, the G and D keyword pair). Conversely, the keywords related to other specific pairs of expertise areas occur much less frequently in those two journals. Turning to the journals *IJGIS* and *Transactions in GIS*, a different pattern is evident. These GI science journals exhibit a much more even distribution across the different pairs of expertise areas. An F-test confirms this pattern, indicating that the differences in variance of co-occurrence frequency between any of the GI science journals and the two anomalous journals is signficant at the 5% level (e.g. comparing *IJGIS* and *IEEE Transactions on Visualization and Computer Graphics*, the differences are significant, $f(9) = 0.1559$, $p = 0.01073$; comparing *IJGIS* and *Transactions in GIS* the differences are not significant, $f(9) = 1.6971$, $p = 0.4429$).

4.2.5 Discussion

The results presented above appear to indicate some clear patterns in the cited references in leading GI science journals, when compared with other journals.

First, it appears that GI science journals tend to have a higher proportion of articles that cite papers with titles related to at least one of the proposed five areas of expertise. This pattern seems to strengthen further when considering the proportion of articles that cite papers with titles related to two or more different areas of expertise.

Second, even in those cases where journals that might be considered outside the core of GI science exhibit frequent co-occurrences of pairs of keywords, the pattern of these citations is significantly different from the GI science journals tested. Whereas the GI science journals examined tended to have an even coverage across all pairs of keywords, those outside the core of GI science tended to have co-occurrences of only specific pairs of keywords, with other pairs poorly represented. In other words, the GI science journals can be expected to have higher proportions of co-occurences of keywords of *any* pair out of the five expertise areas than journals related to other fields.

This basic analysis has been repeated a number of times, varying the precise journals and keywords, and appears to be relatively robust. However, sceptical readers are again encouraged to try this analysis for themselves on any other journals they feel have been missed here.

5 Conclusions

This article has proposed five core areas of expertise in GI science: structure, uncertainty, dynamism, language and cognition, and design of geographic information. The article has further argued that these topics constitute a minimal set of expertise shared across GI scientists. While many GI scientists may also possess expertise in some other related topics, from algorithms to user-generated content, this article argues that these topics are not sufficiently common across the field to count as shared expertise.

To test this contention, the article presented evidence from classic papers in the field and from wider citation analysis. The results do indeed lend support for this thesis, in particular indicating that the five areas of expertise are well-represented across GI science publications, more so than across other fields of study. Further, the results indicated that even in closely related fields, such as visualization or informatics, it is the *combination* of this expertise that makes GI science unique. While other fields may possess expertise in visualization or uncertainty or cognition and language, say, none of those tested here can claim to have the unique combination of expertise across those topics.

These results have been carefully compiled, and constructed to avoid bias as far as possible. However, the analysis contained in this article inevitably relies on a small but significant number of partially subjective decisions, for example, in selecting the journals under study and the keywords used in each expertise area. Although the results appear robust to slight differences in choices (of journals and keywords), the author would welcome further suggestions and analysis of other alternatives.

This article includes data from analysis of recent and past articles in GI science journals and books. However, the analysis still represents a snapshot of the expertise in the field. Like any snapshot, this analysis may incorporate some appreciable distortions. Recent trends in expertise are undoubtedly under-represented in the corpus from the *Classics from IJGIS*; conversely the increasing rate of publication means that more recent expertise is most likely over-represented is the analysis of journal articles. More generally, it seems certain that expertise in any field is not immutable. A degree of flux in research focus and evolution in expertise is to be expected over time. Adaptations of the quantitative citation-based methods in this article applied to publications over restricted ranges of time might help to track this evolution in expertise. Indeed, an appreciation of these changes in expertise over time provides us with a historical context for our field.

A better understanding of our shared expertise as GI scientists may help us in a range of other endeavors. Education is one domain where a picture of the desired expertise of students is of undoubted value. For example, the method used here might contribute to structuring or populating learning artifacts, such as the latest edition of the GIS&T BoK. Understanding overlaps in expertise and topics could help in the management of scientific activities, for example assisting academics in identifying new conferences, or assisting conferences to better target specific topics and domains.

Most importantly, the article set out with the aim of aiding interdisciplinary collaboration. The analysis in this paper supports the contention that GI science does indeed have a unique expertise "fingerprint," distinct from other related fields. However, although that fingerprint is unique, it is not exclusive. Many GI scientists possess other, additional expertise; and many non-GI scientists may also possess some parts of the fingerprint. Ultimately, identifying such a concise set of expert topics shared across GI science makes the task of explaining "what we do" easier, assisting in identifying the contribution of GI scientists to collaborators with complementary expertise.

References

Brassel K E and Weibel R 1988 A review and framework of automated map generalization. *International Journal of Geographical Information Systems* 2: 229–44

Chrisman N 1984 The role of quality information in the long-term functioning of a geographic information system. *Cartographica* 21: 79–87

Clarke K C 1997 *Getting Started with Geographic Information Systems*. Upper Saddle River, NJ, Prentice Hall

DiBiase D, DeMers M, Johnson A, Kemp K, Luck A T, Plewe B, and Wentz E (eds) 2006 *Geographic Information Science and Technology Body of Knowledge*. Washington, DC, Association of American Geographers and University Consortium for Geographic Information Science

Ding Y and Fotheringham A S 1992 The integration of spatial analysis and GIS. *Computers, Environment and Urban Systems* 16: 3–19

Downs R M 1994 Being and becoming a geographer: An agenda for geography education. *Annals of the Association of American Geographers* 84: 175–91

Downs R M 2014 Time and practice: Learning to become a geographer. *Journal of Geography* 113: 189–97

Duckham M 2015 Geographic Information Science. In Richardson D, Castree N, Goodchild, M F, Kobayashi A L, Liu W, and Marston R (eds) *International Encyclopedia of Geography*. New York, John Wiley and Sons: in press

Fisher P 1989 Knowledge-based approaches to determining and correcting areas of unreliability in geographic databases. In Goodchild M F and Gopal S (eds) *Accuracy of Spatial Databases*. London, Taylor and Francis: 45–54

Fisher P (ed.) 2006 *Classics from IJGIS: Twenty Years of the International Journal of Geographical Information Science and Systems*. Boca Raton, FL, CRC Press

Goodchild M F 1992 Geographical information science. *International Journal of Geographical Information Systems* 6: 31–45

Goodchild M F 1995 Sharing imperfect data. In Onsrud H and Rushton G (eds) *Sharing Geographic Information*. New Brunswick, NJ, CUPR Press: 413–25

Goodchild M F 2010 Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science* 1: 3–20

Goodchild M F 2011 Challenges in Geographic Information Science. *Proceedings of the Royal Society A* 467: 2431–43

Goodchild M F, Yuan M, and Cova T J 2007 Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science* 21: 239–60

Grenon P and Smith B 2004 Snap and span: Towards dynamic spatial ontology. *Spatial Cognition and Computation* 4: 69–104

Hägerstrand T 1970 What about people in regional science? *Papers in Regional Science* 24: 6–21

Hirtle S and Jonides J 1985 Evidence of hierarchies in cognitive maps. *Memory and Cognition* 13: 208–17

Hollenstein L and Purves R 2010 Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science* 1: 21–48

Huynh N T and Sharpe B 2013 An assessment instrument to measure geospatial thinking expertise. *Journal of Geography* 112: 3–17

Kemp K (ed) 2008 *Encyclopedia of Geographic Information Science*. London, Sage Publications

Kraak M-J 2014 *Mapping Time*. Redlands CA, Esri Press

Longley P A 2000 The academic success of GIS in geography: Problems and prospects. *Journal of Geographic Systems* 1: 37–42

MacEachren A M 1994 *How Maps Work: Issues in Representation and Design*. New York, Guilford Press

Mark D M 2000 Geographic information science: Critical issues in an emerging cross-disciplinary research domain. *Journal of the Urban and Regional Information Systems Association* 12(1): 45–54

Mark D M 2003 Geographic information science: Defining the field. In Duckham M, Goodchild M F, and Worboys M F (eds) *Foundations of Geographic Information Science*. New York, Taylor and Francis: 3–18

Miller H J 2005 A measurement theory for time geography. *Geographical Analysis* 37: 17–45

Montello D 1993 Scale and multiple psychologies of space. In Frank A U and Campari I (eds) *Spatial Information Theory: A Theoretical Basis for GIS*. Berlin, Springer: 312–21

NCDCS (National Committee for Digital Cartographic Data Standards) 1988 The proposed standards for digital cartographic data. *American Cartographer* 15: 131–35

Openshaw S 1990 Spatial analysis and geographical information systems: A review of progress and possibilities. In Scholten H J and Stillwell J C H (eds) *Geographical Information Systems for Urban and Regional Planning*. Berlin, Springer: 152–64

Schuurman N 2000 Trouble in the Heartland: GIS and its critics in the 1990s. *Progress in Human Geography* 24: 569–90

Shanteau J 1992 Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes* 53: 252–62

Stevens A and Coupe P 1978 Distortions in judged spatial relations. *Cognitive Psychology* 13: 422−37

Tversky B 1992 Distortions in cognitive maps. *Geoforum* 23: 131–38

Wood J 1996 Morphometric Characterisation in the Geomorphological Characterisation of Digital Elevation Models. Unpublished PhD Dissertation, University of Leicester

Worboys M F and Clementini E 2001 Integration of imperfect spatial information. *Journal of Visual Languages and Computing* 12: 61–80

