

## **Report of the Ad-hoc Review of the Commission for the Management and Application of Geoscience Information, CGI, October 2019**

1. On 29 May 2019, the IUGS Executive Committee established an Ad-hoc Committee comprising Edmund Nickless (Chair, IUGS Councillor), Claudia Mora (Secretary, IUGS Councillor), Marko Komac (European Federation of Geologists) and Helen M<sup>c</sup>Kenzie (Sellafield Limited) to review the Commission for the Management and Application of Geoscience Information (CGI) with the main goals of:
  - i. Analysing performance since the establishment of the Commission in 2004;
  - ii. Assessing the alignment of the work of CGI with the objectives and aims of IUGS;
  - iii. Identifying ways to increase the impact and significance of the Commission, including concrete objectives, for the near future.
  
2. We convened in London at the Geological Society on 14/15 October 2019. On the first day we met with François Robida, CGI Chair, and by Skype with Oliver Raymond, CGI Treasurer, to receive a presentation and discuss background information, including annual reports, newsletters and financial statements that had been provided prior to the meeting. The following day, the Review Committee worked together to draft a report, which has been subsequently edited, reviewed and agreed upon by all Committee members.

### **Background on the Commission for the Management and Application of Geoscience Information (CGI)**

3. As a science that requires the observation, storage and interpretation of spatially and physically explicit information, the geological sciences face the challenge of defining, establishing, writing and adopting digital standards. There has been an exponential rise in data collection and archiving, resulting in many, and large, data sets. The public funding sources of most scientific research now require that data collected are open to access by the scientific community and the general public. The ability to use and examine “big data” may advance our understanding of geological processes, and may reveal previously unknown geological phenomena and associations.
  
4. Several challenges pose a barrier to the development of information systems to enable the broad application of geological data, including:
  - i. Past geological data have been collected using many different, non-digital standards;
  - ii. Different approaches to digital data have been taken, using different languages and semantics;
  - iii. The geological disciplines use data for different purposes which may require different approaches to data management and manipulation.
  
5. Despite these challenges, there is great interest in improving the quality and range of geological information and the applications used to acquire, analyse, process and disseminate geoscience data. Numerous user groups, national, regional, and commercial interests are working to develop geoscience databases and software to

enable broad application and interpretation of geological information. An important key to the success of these efforts is the development of global geoscience information standards that enable the exchange of knowledge, support harmonisation of multiple databases both within and across geoscience disciplines, and result in more rapid advancement of geoscience information efforts, with greater impact and global usage.

6. Recognizing both the challenges and the great scientific potential that would be unleashed by the development of robust, interoperable geological information systems, IUGS established the Commission for the Management and Application of Geoscience Information, or CGI. The Commission was launched in 2002 and formally ratified by the IUGS in 2004, with the mission *to foster the interoperability and exchange of geoscience information by active community leadership, collaboration and education, and the development and promotion of geoscience information standards, and best practice*. The Commission brings together the best experts in the field of geoinformatics and IT from around the globe, many with more than two decades of experience in the fields of management and application of geoscience information. The vision of the Commission is:
  - i. that geoscience information can be exchanged, understood, and used without limitation,
  - ii. that geoscience information can be readily integrated with standards-based information from other knowledge domains;
  - iii. that geoscience information is semantically rich and structured to enable seamless interaction in all environments;
  - iv. that global education about the management, modelling, exchange, and use of geoscience information enables its best possible application;

for the benefit of all society.

### **Accomplishments of CGI**

7. Since 2004, CGI has been extremely successful in meeting its original intent and goals. Its principal achievements are the development of several leading data model and data transfer standards for geological information captured in geological maps, studies, boreholes, databases, etc. In particular, it has developed flagship standards: GeoSciML and EarthResourcesML (ERML) that have been successfully tested and broadly adopted by the geological informatics community, including geological surveys around the world. An important partnership with One Geology provided a testbed for the GeoSciML standard and also an international vehicle by which the useful nature of GeoSciML could be established. The GeoSciML standard has been brought into and adopted by other large research infrastructure initiatives such as EPOS (European Research Infrastructure on Solid Earth), AuScope (Australian research infrastructure for geoscience), EarthCube (US Geoscience cyberinfrastructure), and the Open Geospatial Consortium (OGC). The USGS identified GeoSciML in its spatial data infrastructure roadmap. CGI experts invested significant efforts to work with software developers such as QGIS, ESRI, and Dassault systems to implement their standards. Because of this successful record of contributions to the development of digital spatial standards, CGI is identified as a geoscience information pioneer by the Commission on Data of the International Council for Science (ICS).

8. In addition to the great success of their GeoSciML and ERML standards, CGI has also made extensive progress in developing Vocabularies and multi-lingual Thesaurus that help to establish global consistency in geological terminology. Over the past few years, CGI has focused attention to extending these standards to other geoscience domains, such as borehole data and 3D geological models. To assist the global community in understanding the use of geoinformation standards, *i.e.*, to “train the trainers”, CGI has prepared tools and guides for the implementation of the standards, and the membership has organised training classes for geoscientists.
9. CGI leadership is drawn primarily from geoscientists employed by national geological surveys, which hold the majority of the geological data needed to populate geological models and databases. The surveys will ultimately benefit from the development and implementation of geoscience information systems that are standardized and have extensive interoperability. Such systems support the distribution and use of public data and can be used for both academic research and applied problems of economic and public interest. CGI conducts its efforts through Working Groups focused on GeoSciML, ERML, Geoscience Terminology, and Geoscience Domains. Other Working Groups have concluded their activities, including those focused on Interoperability and Multi-lingual Thesaurus. While a large amount of the standard development effort has been completed by a core group of members, representing <10% of the Commission membership, training activities have been organized and led by regional membership, who are based primarily within national geological surveys. CGI has a strong network of continental working groups representing Africa, Asia, Central and Latin America, Europe, North America, and Oceania, whose membership has coordinated very visible regional activities that have been reported through CGI annual reports and newsletters. CGI was a partner on the Geoscience Information in Africa (GIRAF) network initiative to stimulate sharing of geological data and good data practices across Africa and to support African nations in becoming more active partners in the global geoscience information community. Despite the many activities sponsored by the GIRAF initiative, Africa is still under-represented in global geoscience information efforts.
10. One of the most important results of the most recent 4 years has been the development of a mutually-beneficial partnership between CGI and the Open Geospatial Consortium (OGC). The OGC is an international industry consortium of over 538 companies, government agencies and universities, which is driven to make geospatial information and services findable, accessible, interoperable and reusable. CGI, with its emphasis on data structure and semantics, fits very well within the much broader umbrella of OGC, which focuses on data systems, syntax and data structure. CGI participates in OGC as a non-voting, non-paying member, and has helped to create a geoscience working group within the OGC structure, leading to the adoption of GeoSciML v4.1 as a formal OGC standard. The geoscience working group takes advantage of regular OGC meetings to expand knowledge of the standard to other working groups and to identify how the standard might be improved. Other activities of CGI within OGC are focused towards the Hydrology Domain Working Group of OGC, with the aim of coordinating the development of GroundWaterML.

## **Observations and Recommendations of the ARC Review Committee**

11. Over fifteen years of dedicated effort, CGI members have not only stimulated interest in the need for geoinformation standards, but led the actual structuring and development of the GeoSciML and ERML standards, and standard Vocabularies. The rapid expansion of global efforts to develop geological information systems, and international recognition of CGI standards, the work of the Commission remains extremely timely and, indeed, CGI appears to be on the cusp of even more significant recognition of their hard work.
12. In its presentation to the ARC Committee, CGI suggested that one of its upcoming goals is to broaden the application of geological information standards to support the development of integrated physical and social science models. That would require simplification of GeoSciML for general cases. While ultimately a good service, the Review Committee recommends that CGI keeps its geological focus and that, the Commission turn its focus to activities leading to broader knowledge about and adoption of their standards. The focused promotion of their standards to consortia and other, potentially well-funded, groups may not only result in further adoption of the standards, but their further development. There are multiple, integrative geological information systems currently in planning or development (EPOS, One Geology, EarthCube) and large, international consortia aimed at organizing multiple geoinformation systems to be accessible and interoperable, such as the OGC. Ultimately, the development and evolution of geological data standards should be recognised as a cooperative and collaborative, international effort, which no one national geological survey can deliver. The most successful outcome for CGI is to be adopted and embedded into large, international geoscience information infrastructures as they are developed. By partnering and embedding CGI-developed standards into these systems, some of which are supported by substantial government and/or industry funding, the work of refining and evolving the standards is transferred to another generation of geoscientists, and the work of the Commission can be considered to be very successfully completed, indeed, a model of an effective IUGS Commission. Thus, direct external sponsorship should not be a primary focus of CGI's efforts in the next 4 years. Instead, its impact will be solidified by adoption of its standards, or at least an understanding of the factors necessary for acceptable standards.
13. To retain the intent and structure of the developed standards and platforms, and to ensure the standard and tools remain up to date as new information and technology becomes available, CGI should be careful to note changes and refinements to the standard (version control and archiving of old versions) and to maintain a robust audit trail. It is also recommended that a record of useful lessons and mistakes learned in the development of GeoSciML be developed to enable more efficient development of future tools and standards.
14. CGI leadership and contributors have been supported in these efforts by their host institutions, predominantly national geological surveys, which hold the majority of the geological data needed to populate geological models and databases. The surveys will ultimately benefit from the development and implementation of geoscience information systems that are standardized and have extensive interoperability. Such systems

support the distribution and use of public data and can be used for both academic research and applied problems of economic and public interest.

15. Several factors limit the impact and effectiveness of CGI -factors that are common to other many IUGS Commissions. Specifically, member time and effort must be institutionally supported, and therefore the work that the members commit to Commission goals must have high value to the members' institutions. With small IUGS budgets, many of the additional resources needed to support training and to increase collaborations must be funded by member institutions, as well. Finally, the work of the Commission must be led by appropriate subject matter experts. The geoscience information challenge requires multiple expertise, including geoscientists with significant domain knowledge and ability to conceptualize and capture the essential descriptors of geological features, data scientists with the ability to extract and create effective ontologies, and IT experts.
16. The CGI leadership team has remained largely the same over the past 20 years. There is a core of about 30 very active members, with more occasional contributions by other members, who may face employer or economic constraints. The CGI leadership team needs to be "refreshed" by next generation scientists who can learn from the existing team so that important knowledge is not lost. It is recommended that CGI members develop a robust succession plan to mitigate against leadership changes driven by retirement, a change in institution or redeployment of valued members to other job priorities. The succession plan should consider the skills and experience required to deliver the remaining elements of the CGI's work/mission/objectives. It may also consider the training and development needs for those in the early stages of their career to ensure that, whether or not the Commission itself remains active, the next generation of geoscientists understands the value of data management standards. As part of the succession planning and capability building, gaps in global and gender representation should be considered. The core CGI team is drawn primarily from Europe, Australia, China and South America; additional focus on gaps in membership from African and Central American countries should be addressed. Membership and training initiatives should continue to be locally driven, possibly through national surveys, with support from CGI through the provision of training for trainers. The developing African Mining Vision initiative presents an opportunity for CGI to have profound effect.

### **IUGS Support for CGI**

17. The great success and significant contributions of CGI to the global community have been supported by the member employers, predominantly national geological surveys, which provide salary, expenses and project contributions of very significant economic value, probably in excess of \$100k annually. The value of that institutional support towards completion of their goals cannot be overstated, and is orders of magnitude greater than the IUGS support which helps to leverage it. Historically, support from IUGS to the Commission has been ~ \$8.5k annually. However, the 2018 budget provided by IUGS was \$2.5k. The Review Committee believes the present, nominal budget from IUGS is misaligned with the rapidly rising international interest in the management and application of geoscience information, particularly given the

designation of Deep-time Digital Earth as an IUGS-recognized big programme. Support at such a nominal level calls into question IUGS commitment to this important area at a point in time when understanding of geoinformation standards is especially needed. Given their earlier success and the pressing need for geoinformation standards, the Review Committee recommends that the Commission receive at least one more 4-year cycle of funding, at a level commensurate with the tasks it must complete to expand knowledge and wider adoption of the standards developed by the Commission with a focus on GeoSciML and followed by ERML and GroundWaterML.

18. Even at the larger level of annual support previously provided by IUGS, the Commission has been unable to complete important tasks that required additional assistance, such as the completion of dictionaries and thesauruses. To manage those requirements, the Commission elected to save IUGS support over multiple years, in order to build sufficient funds to commission completion of dictionaries by the British Geological Survey. The Review Committee encourages a return to, or increase over, previous funding levels. To enable the Commission to plan projects extending over more than one year, the IUGS should consider providing an *indicative* level of future financial support over a 2-4 year period, recognising that such support may not be achievable in practice, being subject to affordability, the financial health of the Union, and a continued high level of achievement by the Commission.
  
19. Other ways that the IUGS can support CGI in completing its mission include:
  - a. directing all of its relevant Commissions and Task Groups to work towards understanding of and adoption of the geological data standards in their current and future efforts;
  - b. encouraging the use of both the IUGS and CGI brand logos in all relevant communications to enhance the profile of any communications and promotion of GeoSciML, ERML and other products. Similarly, organizations adopting GeoSciML and ERML should be encouraged to display the CGI and IUGS logos, together with their own logo, on their products;
  - c. by promoting CGI activities and success in the IUGS e-newsletter. This should be relatively easy to do, as CGI conscientiously publishes its own annual reports, newsletters and highlights. To date, there has been very limited feedback or acknowledgement of these reports by IUGS, and the Review Committee observes that annual feedback is a professional courtesy to reporting commissions and can help Commissions like CGI remain focused on their objectives;
  - d. CGI members contributing significant work should be formally recognized and commended by IUGS by a letter acknowledging their contributions to the project, as well as a statement highlighting the importance of the project's aims and objectives. This communication should be shared with the contributors' management.

### **Summary Recommendations of the ARC**

20. In summary, international interest in the development and application of geoscience information platforms remains very high and such systems are likely to touch every

aspect of our discipline in the future. The CGI has successfully developed several data standards and tools that will enable geoinformation systems to meet their potential to deliver new scientific findings and understanding. Therefore, the Review Committee recommends:

- i. that the life of the Commission is extended for four years;
- ii. that emphasis is placed on promoting interest and understanding of GeoSciML and ERML, expanding the adoption of the standards, and embedding CGI into much larger, international consortia and programs dedicated to the evolution of geoinformation systems;
- iii. that CGI create a new program of work aligned with that emphasis, developing key performance indicators, and deliverables due at the end of the second and fourth years;
- iv. that the Commission is funded at a level commensurate with its agreed program, and that IUGS indicates a provisional level of funding in future years;
- v. that there should be a further review at the end of three years to consider the case for further extending, or wrapping up the work of, the Commission;
- vi. that the Executive Committee provides feedback at least annually on progress achieved.

21. The Review Committee is grateful to the Geological Society of London for hosting this review and to François and Ollie for their good humour and patience in answering our many questions.

Agreed by the Review Committee 11 November 2019