[CP-01-013] Cyberinfrastructure

Abstract

Cyberinfrastructure (sometimes referred to as e-infrastructure and e-science) integrates cutting-edge digital environments to support collaborative research and education for computation- and/or data-intensive problem solving and decision making (Wang 2010).

Keywords: computing infrastructures, cyberGIS, High Performance Computing

Author & citation

Wang, S. (2019). Cyberinfrastructure. The Geographic Information Science & Technology Body of Knowledge (2nd Quarter 2019 Edition), John P. Wilson (Ed.).
DOI: 10.22224/gistbok/2019.2.4.

Explanation

1. Background
2. CyberGIS
3. Synergistically Advancing CI and CyberGIS

1. Background

The term "cyberinfrastructure" (CI) was coined by a National Science Foundation (NSF) Blue-Ribbon Committee to reflect how the traditional modes of all fields of research are being enhanced and revolutionized by the integrative capabilities of high-performance computers, advanced collaboration tools and facilities, data storage and visualization tools, and sensor webs in the digital age (Atkins et al. 2003; Wang and Armstrong 2009).

Like the physical infrastructure of roads, bridges, power grids, telephone lines, and water systems that support modern society, CI stands for distributed computers, information and communication technologies combined with personnel and integrating components that provide a long-term digital ecosystem to empower research and education practices (Wang 2013). E-Science is a similar term coined in the Europe with a main emphasis on how science and engineering practices are transformed by CI. Though coming from slightly different perspectives, both CI and e-science represent the powerful paradigm in which high-performance and distributed computers, data-intensive knowledge systems, and information and communication technologies are integrated to provide services to empower research, education, and scholarship, ranging from engineering and science to the humanities (Figure 1).

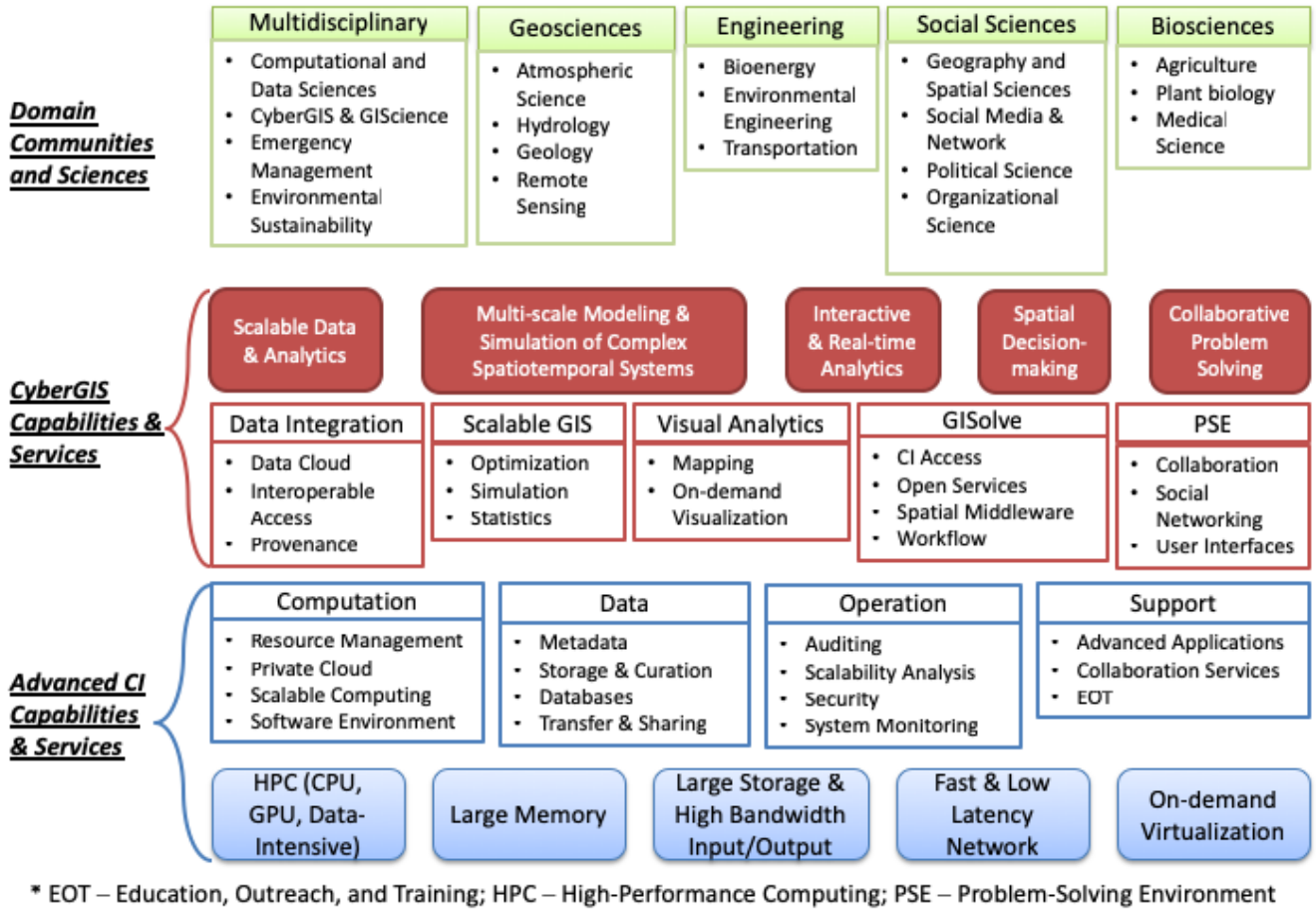


Figure 1. Advanced cyberinfrastructure and cyberGIS for enabling new knowledge environments for research and education. Source: author.

Many research domains have developed and used CI to transform their state of the art and have achieved significant results (Anselin and Rey 2012; Li 2018; Wang and Goodchild 2018). However, significant challenges remain to innovate and exploit CI in research and education practices mainly because 1) CI technologies are sophisticated and continue to be advanced rapidly; and 2) CI workforce in research and education requires significant efforts to develop (Shook et al. 2019; Yang et al. 2010).

From a technological point of view, CI is composed of high-performance computing systems, data, information resources, networking, digitally enabled-sensors, instruments, virtual organizations, and observatories, along with an interoperable suite of software services and tools (Wang and Zhu 2008). These capabilities collectively offer great opportunities to transform the way geographic information science and systems (GIS) are developed and used to achieve high-performance, distributed, and collaborative GIS and to support large-scale and multi-scale geospatial problem solving and discovery (Wang et al. 2013).

2. CyberGIS

CyberGIS is defined as GIS based on advanced computing and cyberinfrastructure (Wang 2017). It has emerged during the past decade as a vibrant interdisciplinary field (Wang 2010; Wang and Goodchild 2018; Wright and Wang 2011). CyberGIS synergizes advanced cyberinfrastructure and e-science, GIS, spatial analysis and modeling, and a number of geospatial domains to improve research productivity and enable scientific advances (Wang et al. 2013). It represents a fundamentally new GIS modality based on holistic integration of high-performance and distributed computing, data-driven knowledge discovery, visualization and visual analytics, and collaborative problem-solving and decision-making capabilities (Wang 2010).

CyberGIS has provided a solid foundation for breakthroughs in diverse science, technology and application domains, and contributed to the innovation of CI overall (Wang and Goodchild 2018). For example, NSF-funded a $4.8 million multi-institution project: CyberGIS Software Integration for Sustained Geospatial Innovation involved a number of academic institutions, industrial partners (e.g. Esri), U.S. government agency partners (e.g. US Geological Survey), and U.S. federally funded research and development laboratory (e.g., Oak Ridge National Laboratory), and multiple international partners (Wang et al. 2013). With an international scope, the project has established a sustainable cyberGIS software framework (e.g., the CyberGIS Science Gateway and the CyberGIS Toolkit) while achieving major research advances in tackling multi-scale environmental and geospatial challenges.

3. Synergistically Advancing CI and CyberGIS

Tremendous challenges and opportunities exist to synergistically advance CI and cyberGIS. The recent computation- and data-intensive transformation of broad research and education practices has posed significant challenges revolving around cyberGIS and geospatial data science (Wang 2016). The proliferation of geospatial data sources (e.g., remote sensing, sensor networks, and social media) have brought up new challenges such as poor or unknown data quality, missing metadata, and lack of well-defined sampling schemes (Miller and Goodchild 2016). The proliferation of these data streams represents a pressing big data problem that is not likely to slow down in the foreseeable future, thus innovation of advanced CI-enabled geospatial data and software capabilities is required to address it. This innovation opportunity is significant and unique because geospatial data science often simultaneously involves computing-, data-, and visualization-intensive capabilities with responsive user interactivity expected and, thus, cannot be effectively supported by conventional CI. Major scientific and technological breakthroughs need to be pursued through holistic approaches to geospatial integration with advanced CI and cyberGIS.

It is worth noting one challenging question. How to systematically integrate geospatial data, analysis and models of varying assumptions and quality, different spatial and temporal resolutions, and diverse data models and semantic representations for understanding the complexity of coupled digital, environmental, and human systems as they relate to a number of grand science challenges? This important question needs to be addressed for adapting to a number of science domains while scaling to big data and computing to benefit many researchers. Such holistic approaches will likely produce a suite of innovative,

community-driven, and extensible and open tools for data-rich geospatial integration, and associated new knowledge and science on synthesizing multi-source and multi-provenance geospatial data with advanced CI and cyberGIS approaches. Specific approaches of geospatial data science include but not limited to combining: (1) qualitative and quantitative inferences, (2) compatible representations of place and space with uncertainty explicitly addressed, (3) geospatial and spatiotemporal interactions, (4) geospatial patterns and related processes and their feedbacks, and (5) data-intensive computational, geospatial, and statistical methods.

References

Anselin, L. and Rey, S. (2012). Spatial Econometrics in an Age of CyberGIScience. International Journal of Geographical Information Science. 26(12): 2211–2226.

Atkins, D. E., Droegemeier, K. K., Feldman, S. I., et al. (2003). Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure.

Li, W. (2018). Lowering the Barriers for Accessing Distributed Geospatial Big Data to Advance Spatial Data Science: The PolarHub Solution. Annals of the American Association of Geographers, 108(3), 773-793.

Miller, H. J. and Goodchild, M. F. (2016) Data-Driven Geography. GeoJournal, 80: 449-461.

Shook, E., Bowlick, F. J., Kemp, K. K., Ahlqvist, O., Carbajeles-Dale, P., DiBiase, D., ... Wang, S. (2019). Cyber Literacy for GIScience: Toward Formalizing Geospatial Computing Education. The Professional Geographer, 71(2), 221–238.

Wang, S. (2010). A CyberGIS Framework for the Synthesis of Cyberinfrastructure, GIS, and Spatial Analysis. Annals of the Association of American Geographers, 100(3): 535–557.

Wang, S. (2013). CyberGIS: Blueprint for Integrated and Scalable Geospatial Software Ecosystems. International Journal of Geographical Information Science, 27(11): 2119–2121.

Wang, S. (2016). CyberGIS and Spatial Data Science. GeoJournal, 81(6): 965-968.

Wang, S. (2017). CyberGIS. In International Encyclopedia of Geography: People, the Earth, Environment and Technology (eds D. Richardson, N. Castree, M.F. Goodchild, A. Kobayashi, W. Liu and R.A. Marston). John Wiley & Sons.

Wang, S. and Armstrong, M. P. (2009). A Theoretical Approach to the Use of Cyberinfrastructure in Geographical Analysis. International Journal of Geographical Information Science, 23(2): 169–193.

Wang, S. and Goodchild, M. F. (2018). CyberGIS for Geospatial Innovation and Discovery. Springer, Dordrecht, Netherlands.

Wang, S. and Zhu, X-G. (2008). Coupling Cyberinfrastructure and Geographic Information Systems to Empower Ecological and Environmental Research. BioScience, 58(2): 94–95.

Wang, S., Anselin, L., Bhaduri, B., Crosby, C., Goodchild, M. F., Liu, Y., and Nyerges, T. L., (2013). CyberGIS software: a synthetic review and integration roadmap. International Journal of Geographical Information Science, 27(11), 2122-2145.

Wright, D. J. and Wang, S. (2011). The Emergence of Spatial Cyberinfrastructure. Proceedings of the National Academy of Sciences USA, 108(14): 5488–5491.

Yang, C., Raskin, R., Goodchild, M. F., and Gahegan, M. (2010). Geospatial Cyberinfrastructure: Past, Present and Future. Computers, Environment and Urban Systems, 34(4): 264-277.