Paper - What is cyberinfrastructure?

Cyberinfrastructure is a word now commonly used but lacking a single, precise definition. Everyone recognizes intuitively the analogy with infrastructure, and the use of *cyber* to refer to thinking or computing – but what exactly is cyberinfrastructure as opposed to information technology infrastructure? Indiana University has developed one of the more widely cited definitions of cyberinfrastructure:

**Cyberinfrastructure** consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and enable breakthroughs not otherwise possible.

This definition has been used as-is or with slight modifications in several publications, in reports from national workshops on cyberinfrastructure, and is prominent in the Wikipedia definition of cyberinfrastructure.

Since the word cyberinfrastructure is based on extension of the word infrastructure, we present information on the history of the latter and then detail the much shorter history of the former. In particular, we focus on cyberinfrastructure as distributed systems that uniquely enable science and engineering discovery and learning. We differentiate this from what one might otherwise call “information technology systems” and compare it to the related European usage of the term eScience. Most importantly, we describe several cyberinfrastructure projects developed by Indiana University and others, and show how these meet and validate the above definition of the word cyberinfrastructure. In a field such as computing, so rife with buzzwords, a clear definition of cyberinfrastructure will be useful to practitioners and students alike.