The 2014-2015 University Research Committee (URC) studied the topic of Academic Computing at USC, and in particular sought to answer the overarching question of "how academic computing at USC needs to evolve to take advantage of, and adapt to, technologies becoming available to support communication, computation, data storage, and data sharing." The URC's focus areas for this report were:

- Identifying software that are most desired and most used by faculty in their research, representing strong candidates for university level site licenses.
- Determining needs for the effective use of software, including staff support, training, a centralized and simple database of such resource, building user communities.

To prepare to answer this question, the URC consulted several information technology (IT) experts across campus, who provided valuable information and insight into the USC software, computing, and IT resources. A university-wide faculty survey was also conducted, in which faculty were asked questions regarding their software use and needs. The survey did not include questions regarding computing resources. In general, the current URC study focused more on software resources and data storage than computing resources and infrastructure.

The following individuals were invited to meetings with the URC, during which they provided an overview of resources at their respective units, and engaged in discussions with the Committee about strengths, weaknesses, and plans related to their programs:

- Yibu Chen (Bioinformatics program coordinator)
- Candace Borland and Joe Cevetello (ITS Assistant CIOs)
- Sam Gustman (USC Libraries)

The Committee findings resulting from these meetings as well as the faculty survey are summarized below. First, individual findings from each unit that participated in meetings with the URC are summarized as well as general findings that may apply across units and campus (Findings 1-8), followed by faculty survey results about software needs (Finding 9).
Finding 1:
The Bioinformatics Program provides access to many commercial software and data resources free of charge to research users, with financial support provided by the Provost Office and the Norris Medical Library. They also provide in-house workshops, training, and vendor tutorials. There are approximately 10 major packages provided to users, with more than 1,200 registered users. The number of seats for each package is determined based on demand and available budget, and varies from package to package. Given the large number of users, it is estimated that the cost saving in having the multi-seat site license as opposed to individual purchase is at least one order of magnitude. It is also estimated that for most packages, as long as there are about 10 or more users, the costs of acquiring a site license are less than purchasing individual copies.

Finding 2:
The Bioinformatics Program model for providing software support to research users appears to be a successful one: they provide a wide range of software tools to a large number of users; no major gaps in coverage have been identified by them or by their users. At present, this is the only university level resource integrating software, training and support from a single university unit. However, looking ahead, the Program will need enhanced computing resources (for example by enhancing their computing condo at the High Performance Computing Center, HPCC), additional seats for certain packages, tools for upgrading data security, and data storage capabilities in the next 3 years. These needs have not yet been quantified, but should be in the near term.

Finding 3:
ITS provides software support across campus units for several common computing and statistics packages such as Matlab, Mathematica, Qualtrics, and SAS. The ITS website has information on all supported software. It was, however, noted by the Committee that the information contained in the ITS website is not always easy to find. Some of the available packages, notably SAS, are difficult to install, but currently there is no standard user help or training for any of the “supported” software. Support is available at the ITS UPC location for walk-in users.

Finding 4:
ITS has been engaged in discussions within campus and with other universities to investigate methods for acquiring Adobe Creative Cloud site license. This particular package is mentioned because it has very high demand across USC, but is extremely expensive. Adobe site license costs are also very high because Adobe charges not per seat but per head. Therefore, even though the cost per head is one-tenth of retail cost, the total will be unaffordable by this model.
Finding 5: Funding for ITS provided software is in part contributed by various Schools and Colleges across campus and by ITS discretionary funds. The ITS procedures for acquiring software site licenses vary, and appear to be based on individual (or groups of) faculty putting requests to their Deans. The procedure does not appear to be well advertised or understood by many faculty members. ITS has an advisory committee to help with software acquisition decisions, but it was not clear to the URC how the advisory committee interfaces with the faculty or the Deans.

Finding 6: USC Libraries constitutes the University’s core expertise in archiving extremely large data sets. The USC Digital Repository, operated jointly with ITS, currently holds 94 petabytes of data with a goal of 2 exabytes of data. They support digital data for organizations such as Warner Brothers and the Shoah Foundation, among many others. USC Libraries refresh their media once every 3 years to ensure no data are lost, as well as keeping a mirror site at Clemson University. The USC Libraries expertise is in archiving image data, photos, sound, and music; they specialize in archiving collections and content, not software or analytics. They are also working on getting HIPPA-B certification and approval for health-related data and imaging archives.

Finding 7: USC Libraries carries out its work primarily through contracts (for example with Warner Brothers or Shoah Foundation). Any data archiving, hosting, metadata generation, web page maintenance, etc., needs to be supported by individual projects (i.e., faculty funding). Also, the repository is not currently conceived as a resource to enable broad access to USC generated data, but is instead conceived as a place for long terms storage of data for individual projects.

Finding 8: The URC noted that in many cases, it is difficult to find information about many of the above resources. These include supported and existing software, which could cause unnecessary individual purchases. One example is Endnote, a frequently used program that is individually purchased by many faculty members despite the free availability of this software through the USC library. This is partially because of incomplete information in websites and other outlets, and partially because the computing, data handling, and software resources are scattered among several different entities, making an exhaustive search a difficult task.

Finding 9: The faculty software and database access survey had 445 respondents. A significant majority indicated that their primary software interest, whether already at-hand or desired, was under the “Statistics / Data Analysis / Data Capture / Data Visualization” category. Among the top 10 most popular software and database packages, only 4 are currently being supported by ITS. None of the next 10 most favored are being supported. Results are summarized in the appendix.
Recommendation 1:
It appears that software, database, data archiving, and computational resources are scattered among several different entities on campus. Furthermore, it appears that there is a large degree of variability/inconsistency among various schools/college/units in availability of such resources. The URC recommends that investment and effort be made to harmonize and/or unify these resources.

Recommendation 2:
Based on the survey results, currently there does not appear to be a sufficient correlation between the faculty research software/database needs and available packages through ITS. A more direct input mechanism from faculty research groups is needed in the ITS software acquisition decision-making process. In its Mission statement, ITS states that “The mission of ITS is to ensure the effective application and integration of advanced information technology in support of the teaching, learning, research, and administrative missions of the university. Clear leadership and guidance within ITS enable the organization to fulfill its mission in order to serve the entire USC community with distinction.” To stay true to its mission statement and its key goals (http://itservices.usc.edu/about/), ITS needs to institute effective measures for seeking and implementing inputs from the entire USC community including faculty, staff, and students. To help ITS achieve its goals and vision, the URC recommends that an advisory committee composed of these key stakeholder groups advise ITS and evaluate its performance on an annual basis.

Recommendation 3:
Currently, there appears to be no baseline allocation for faculty research group data archives. The URC recommends that a reasonable amount of secure data storage and backup be provided to all faculty.

Recommendation 4:
Better information dissemination is needed by ITS and other computing units. Although several different USC units and resources are involved, an individual faculty member may not be aware of such distinctions, so that a one-stop webpage to orient and link to all university resources should be established to help the users parse through the multitudes of available resources. This will also make the gaps in available resources more clear. Such unified portal could enable efficient searching for available site-licensed software, accessing tutorials, and requesting staff help. It could also provide a forum for suggesting new software purchases. A community of users/experts among the faculty could also be linked together through this page. An option could exist to receive email notifications of updates.