COLLEGE OF AGRICULTURE & LIFE SCIENCES CYBERINFRASTRUCTURE NEEDS REPORT

JULY 2018

Table of Contents

EXECUTIVE SUMMARY

The purpose of this report is to identify cyberinfrastructure needs required for the College to effectively meet its mission goals. For this survey, we purposely chose a broad definition of cyberinfrastructure and assessed its application across all college mission areas. Briefly, the committee's recommendations are that:

- Cyberinfrastructure Advisory Group (CAG) & Communications and Cyber Technologies Team (CCT) coordinate to disseminate and promote existing CALS and UA technology resources;
- CAG & CCT coordinate and provide information to ensure all employees have equal access and opportunity to utilize new technologies;
- CAG work with data science faculty and the Office of Career and Academic Services to ensure that CALS students are cyber-savvy;
- CAG do biennial surveys of CALS faculty, staff, and students to identify unmet and emerging cyberinfrastructure needs;
- CAG & CCT coordinate to implement and report progress on EC approved recommendations arising from this report; and
- progress on implementing approved recommendations is widely disseminated to CALS administrators, faculty and staff for comment and feedback.

A. SCOPE AND PURPOSE OF REPORT
Underpinning many of the specific goals designed around achieving the College of Agriculture & Life Sciences (CALS) Strategic Plan is a College-wide commitment to developing cyberinfrastructure to support research, education and extension missions. For the purposes of this report, cyberinfrastructure is broadly defined as "information technology systems and expertise which enables innovation and discovery". Technology systems may include (but are not limited to) data storage, data management, data analysis, data visualization, data sharing and reuse, and expertise in using these technologies. The *scope* of this report is to (i) broadly assess how individual units within CALS are utilizing technology; (ii) identify barriers to deploying new technologies across CALS mission areas; and (iii) identify pathways to help CALS employees more effectively apply new technologies in their research, extension and teaching.

The rapidly changing nature and application of these technologies mean that this report is necessarily not comprehensive and detailed; rather it is intended to serve as a starting point for developing cyberinfrastructure and raising awareness of the value of new technologies within CALS. We recommend periodic surveys (every 2-3 years) for CALS employees and students to identify emerging and unmet needs within the College.

B. CURRENT SUPPORT FOR CYBERINFRASTRUCTURE WITHIN CALS
Currently, the College's cyberinfrastructure and technology needs are supported by a concerted effort consisting of Communications and Cyber Technologies Team (CCT) and IT professionals embedded directly in the Academic Units. CALS' Federated IT model allows for particular services tailored for a specific academic area to be served by the Unit's IT professional(s), while overlapping Division needs provided by CCT. CCT also provides expertise in four priority areas: (i) Web and Mobile Development; (ii) Data Systems and Infrastructure Solutions; (iii) Video Production; and (iv) Multimedia Technologies. Together, CCT and Academic Unit IT meet regularly to discuss IT governance, vision and strategy, emerging technologies, IT security policies and best practices, and cross-training of software tools.

The Cyberinfrastructure Advisory Group (CAG) was appointed by the Dean to advise on how CALS could leverage existing and emerging technologies to ensure that CALS is able to effectively meet its mission goals. CAG is a faculty-lead group that advises the CCT Director and the Dean. Unless otherwise directed by the CALS Executive Council (EC), we recommend that the CAG work closely with CCT, Executive Leadership and Units to implement recommendations approved by the EC, monitor progress of implementation and assess future cyberinfrastructure needs.

C. ASSESSMENT OF COLLEGE NEEDS AND RECOMMENDATIONS
Assessment Method
To develop a broad view of College-wide cyberinfrastructure needs, members of the CALS Cyberinfrastructure Advisory Group (CAG) met with CALS unit leaders and IT personnel. The purpose of these meetings was twofold: first, to make these personnel aware of the existence and goals of the faculty-led CAG, and second, to get an overview of how cyberinfrastructure is

utilized across each unit. Discussion was deliberately broad ranging across research, extension and education activities. In many cases HODs identified faculty working in areas utilizing new technologies and/or surveyed their faculty to solicit additional input. The CAG thanks HODs, IT personnel and faculty for their time and contributions to this report. We expect that, as we address specific needs identified by this initial survey that we will continue to rely on the knowledge and time of these CALS faculty and IT professionals.

Summary of Findings
A complete list of CALS Cyberinfrastructure needs identified by each unit is available at https://goo.gl/RwE8fs. This initial list was summarized and categorized by mission (Research, Extension, Teaching). Needs that crossed multiple missions or were identified by multiple Units are prioritized in the recommendations from this report.

In general, much discussion focused around making existing CALS and UA technology resources more accessible, and easier to identify and disseminate. We note that tech-savvy users in their fields often reported no bottlenecks to employing this technology, while others who were newer to the field or less experienced often were enthusiastic about applying these technologies to their work but unsure how to achieve this. The committee considers that a key goal to making CALS more cyber-savvy is ensuring that all employers have equal access and opportunity to utilize new technologies.

Another broad concern was the need for training and education in managing large or complex data sets. A data science training white paper has already been produced (https://goo.gl/fPvsa5); the CAG supports this effort and is actively working to support the development of data science curriculum within CALS. We note that in addition to education, information about campus-wide resources and training opportunities for both students and employees is also seen as a priority.

Identified Needs and Recommendations
Priority cyberinfrastructure needs & recommendations (based upon needs that were identified across multiple CALS units or mission areas) are:

1. College-level support to access and assess compute, storage and archival options.
While some units requested assistance with data storage and archiving, others are specifically interested in compute capacity (e.g., cores, RAM) and software support. We note that in each case, often individuals are unaware of University-level initiatives providing compute capacity.
Recommendation: That the CCT and CAG serve as source points to get information about college and University compute systems and opportunities (e.g., resources provided by UA HPC, CyVerse, UA Libraries). Where multiple, similar resources exist CAG and CCT should provide recommendations (e.g., pros and cons) to help IT personnel best advise Faculty. These options should be distributed to IT personnel and HODs for distribution to Faculty. To ensure that this information is both accessible and periodically updated, it should also be provided via the CCT website.

2. Educational courses on programming, data management and data analytics. Multiple CALS units are either developing data science courses, while others recognize the need for these skills in graduating students. Related needs specifically mentioned the development of data analytics courses and the ability to incorporate large data sets into teaching.

Recommendations: That CAG collaborate with the Office of Career and Academic Services to identify faculty interested in data science undergraduate curriculum/course development. This collaboration should focus on development of data science competencies which can be applied to both new and existing courses, and ensure that any new courses complement existing courses with data science content. It is strongly advised that course development and curriculum be distributed across multiple CALS units to avoid siloing. We recommend a focus on developing tech-savvy students within CALS, rather than training data scientists.

Given the range and diversity of data used across CALS, it is difficult to address all data analytics requirements. We recommend providing faculty with data science competencies (including data analytics) and developing lower-level courses for students to develop generalized data science skills. Individual majors can then focus on data analytics cases across their senior and graduate level courses, as suitable for their specific domain areas. We note that development of new data science themed courses will also require concomitant development of computer classroom facilities (addressed in point 15, below).

3. Develop a single point resource for students and employees to access for data thinking, training & education opportunities. Several HODs and faculty pointed out that – while there may be education and training opportunities on campus – it is difficult to identify these opportunities, particularly for those not already expert in at least some areas of data science.

Recommendation: The Data Science Institute's Data Science Resources & Training (DSRT) group is collecting and collating data science education and training opportunities and making these available online (datascience.arizona.edu/dsrt). Members of CAG should continue to support the DSRT initiative and collaborate with the Offices of Research, Academic Services and Cooperative Extension to ensure that information about this resource is disseminated to all CALS students and employees. CAG and CCT should also co-ordinate to identify and organize training based upon UA resources (e.g., HPC, CyVerse, UA Library resources).

4. College-level support to move courses online. While some Units have been very successful at providing online courses, other Units have less access to supporting resources to make this transition.

Recommendation: The CAG should work with the Office of Academic Services to identify, summarize and make accessible information about developing online courses provided by the Office of Instruction & Assessment (OIA) and Office of Digital Learning (ODL). This effort should also include discussion about creating a CALS-level approach to developing and providing online courses. We also note that the Office of Research should be included in any discussion of graduate courses.

Additional cyberinfrastructure needs & recommendations that were identified across multiple CALS units or mission areas are:

5. Ability to identify and connect with CALS personnel who have cyberinfrastructure/data science expertise. A specific example suggested was formation of College-level research computing group to advise faculty.

Recommendation: CCT and CAG should serve as starting points to identify experts, both within the College and across campus. The CAG should arrange for information to be circulated to CALS personnel that these groups are available to assist with identifying experts in these areas.

6. Improved network access at sites off main campus.

Recommendation: While we note that this is an ongoing need, a recently completed three-year effort addressed the network issues at many off site locations. Employees should inform CCT if a site needs better network capacity so that they can enable this in future upgrades.

7. Information/directions from CALS & UITS on new technologies and their development and/or deployment time scale. This concern focused primarily on communication flow between UITS, CCT and Unit IT personnel.

Recommendation: It is strongly recommended that all Unit IT personnel attend the CCT network managers' meetings (held monthly) where these issues can be raised and addressed. This forum provides a conduit for Unit IT to be informed about campus wide initiatives.

8. Development of a College calendar as a single, definitive source for all College events, seminars, etc.

Recommendation: This has been accomplished with the launch of the new website: https://cals.arizona.edu/calendar. The next iteration of this includes the ability for content managers to input calendar events on department websites and have it auto-populate to the CALS website's calendar. This implementation will require moving all department-based websites to CALS QuickStart template (a Drupal-based starterkit available to everyone, http://quickstart.cals.arizona.edu). Five of the ten CALS units already use CALS QuickStart (July 2018). We recommend that CCT work with unit IT and HODs to inform about this opportunity and update department-based websites, as requested.

9. College-level support for development and showcasing novel applications of technology. While some units specifically interested in College-level support for app development, others requested examples of technology application to promote innovation.

Recommendation: College-level support for web and mobile development is available through CCT's Web and Mobile Development Team (https://cct.cals.arizona.edu). We recommend that CCT work with department-level IT and HODs to publicize this service. We further recommend that CAG identify opportunities for showcasing novel application through College websites, newsletters and campus-wide initiatives.

10. Develop more responsive systems for Unit web presences.

Recommendation: CCT has developed website systems that can be rolled out and then managed by Departmental personnel. We recommend that CAG & CCT inform HODs and faculty of this service.

11. CALS support for databases.

Recommendation: CCT already provides trained personnel and technical support to assist those who wish to develop, deploy, maintain databases. This information needs to be communicated with CALS personnel. CAG should investigate providing examples for inclusion with Research and Extension announcements/updates to highlight these services.

Additional needs identified by individual CALS Units are listed below. We note that although these issues may have only been brought up in discussions with a single unit, they are likely to apply to more than one CALS unit.

12. White-glove service in CALS for HPC, CyVerse, Amazon. This white-glove service would consist of an in-house group of data scientists and systems administrators versed in on-campus cyberinfrastructure offerings to provide faculty in-person consulting services. These service members would have the time and expertise necessary to learn the requirements, build-out the cyberinfrastructure, provide data pre- and post-processing scripts, and assist in delivering finalized reports.

Recommendation: CCT currently provides limited version of this service. Capabilities beyond this (including training opportunities) are provided on campus by the High Performance Computing, Amazon and CyVerse services. UITS has also hired a new Director and Assistant Director of Research to help facilitate a white-glove service at the campus level. We recommend that CCT assess the College need for these services and collaborative links and capacity to support College needs.

13. Support to anonymize & de-personalize own data sets rather than paying third party entities.

Recommendation: That CAG investigate processes and resources for anonymize and de-personalize data (medical school, IRB, libraries) and recommendation to CCT how to offer this service. We note that given the diversity of data types and data sets used by CAL personnel, this is likely to be a complex procedure with no single method that can cover all eventualities.

14. Mechanisms for promoting/marketing new educational programs, etc.

Recommendation: We recommend that individual Units work closely with the Office of Career and Academic Services to develop promotional materials and make potential students and employers aware of these programs. CAG should provide information about updating Unit websites to reflect new programs (see point 10, above). Although not our purview, we suggest that the Office of Career and Academic Services develop an online information bulletin (similar to those used by the Research and Extension Offices).

15. Increased support and access for computers in classrooms. UA provides specialized IT support for class rooms listed in the central scheduling catalog, however there are many additional classrooms that are used and/or maintained by CALS Units.

Recommendation: CAG should work with Unit heads to identify these classrooms and make information about the classroom capacity, technology and IT support readily accessible. This information, along with usage, can serve as a way to improve classroom services in an integrated and coordinated manner. CAG should also work with Office of Instructional Assessment to identify and provide faculty with information about campus-wide computer classrooms.

16. Increased support for data science badge development. Badging or micro-certifications are a useful way for potential employers to identify skills not explicit on transcripts.

Recommendation: CAG to work with the Office of Career and Academic Services and Office of Instructional Assessment to provide information about micro-certification to HODs and Unit curriculum committees. The faculty-led Data Science Curriculum group should link micro-certifications to competencies as an example of how this process works.

17. Include IT training & information in student orientation.

Recommendation: CAG to coordinate with the Office of Career and Academic Services and Library Services to identify aspects of IT training that can be incorporated into student orientation. This information will be provided to HODs and personnel involved in orientation at the Unit level.

18. Subscription to commonly used/useful software.

Recommendation: CAG should periodically (annually) survey HODs and Faculty to identify commonly used software for research, education and extension; collated information should be shared with HODs and CCT to increase buying power and make recommendations to UA Libraries and Bookstores. Information about CALS classrooms (identified in point 15) should also include information about installed software.

19. Improved surveillance at sites off main campus.

Recommendation: Personnel responsible for security at sites off main campus should work with CCT to identify problem areas and develop solutions. We recommend that HODs be provided with a CCT contact to address any security concerns.

D. TRACKING PROGRESS AND IDENTIFYING FUTURE NEEDS.
As already noted, the technologies we are trying to promote and utilize within CALS are rapidly changing, as is their use and application. This report is not intended to be definitive, but rather a document that is reviewed, assessed and edited to meet the changing requirements of the College. With this in mind we recommend that:

- CAG and CCT provide CALS administrators with annual reports on progress towards implementing report recommendations, and that these reports are made available to CALS Faculty;

- CAG organize a system for collecting additional feedback on needs from CALS faculty; and
- CAG and CCT actively assess cyberinfrastructure needs periodically, and develop new recommendations to address any novel needs.

This report is available online for comment at: https://goo.gl/LVidhH

Prepared July 2018
By the CALS Cyberinfrastructure Advisory Group
 Fiona McCarthy, Matt Rahr, Jeremy Weiss, Katy Prudic, Mike Crimmins, Eric Lyons