Séminaire / Seminar **AMAP**

**Edward SPANUOLO** is currently a PhD Student at Pennsylvania State University. He is interested in the evolution of Southeast Asian rainforests and their paleo-heritage and has worked on the use of computer vision to provide new characters to identify extant and fossil leaves.

**Email**: ejs5902@psu.edu

**Website**: https://www.researchgate.net/profile/Edward-Spagnuolo

**Thursday Dec. 8, 2022**
3 pm (Paris time, UCT +2)

https://umontpellier-fr.zoom.us/meeting/register/tJEqd-mprTMtE90EkWhx6b3lyRSPs-4G9UDA

# Decoding family-level features for modern and fossil leaves from computer-vision heat maps

*presented by*

## Edward Spagnuolo

**Pennsylvania State University**

***ABSTRACT:*** Angiosperm leaves present a classic identification problem due to their morphological complexity. Computer-vision algorithms can identify diagnostic regions in images, and heat map outputs illustrate those regions for identification, providing novel insights through visual feedback.
This talk will focus on a recent publication in which we investigate the potential of analyzing leaf heat maps to reveal novel, human-friendly botanical information with applications for extant- and fossil-leaf identification. We developed a manual scoring system for hotspot locations on published computer-vision heat maps of cleared leaves that showed diagnostic regions for family identification. Heat maps of 3114 cleared leaves of 930 genera in 14 angiosperm families were analyzed. The most important ("hotspot") regions of highest diagnostic value were scored for 21 leaf locations. The resulting data were viewed using box plots and analyzed using cluster and principal component analyses. We manually identified similar features in fossil leaves to informally demonstrate potential fossil applications. The method successfully mapped machine strategy using standard botanical language, and distinctive patterns emerged for each family. Hotspots were concentrated on secondary veins (Salicaceae, Myrtaceae, Anacardiaceae), tooth apices (Betulaceae, Rosaceae), and on the little-studied margins of untoothed leaves (Rubiaceae, Annonaceae, Ericaceae). Similar features drove the results from multivariate analyses. The results echo many traditional observations, while also showing that most diagnostic leaf features remain undescribed.
Machine-derived heat maps that initially appear to be dominated by noise can be translated into human-interpretable knowledge, highlighting paths forward for botanists and paleobotanists to discover new diagnostic botanical characters.

***KEY WORDS:*** machine learning, computer vision, leaf architecture, leaf identification, fossil identification

| | |
|---|---|
| **Invited and animated by:** | Dr. Anne-Laure DECOMBEIX (UMR AMAP) |
| **Type:** | Research questions & results |
| **Oral language:** | English |
| **Language of PPT:** | English |