SAMPLING CULTURAL UNIVERSES OF NOTABLE INDIVIDUALS FROM WIKIPEDIA

Luis A. Miccio [1,2], Paschalis Agapitos [1], Juan Luis Suarez [3] and Gustavo A. Schwartz[1,4]

[1] *Donostia International Physics Center, P. M. de Lardizábal 4, 20018 San Sebastián, Spain.*

[2] *Institute of Materials Science and Technology (INTEMA), National Research Council (CONICET), Colón 10850, 7600 Mar del Plata, Buenos Aires, Argentina.*

[3] *CulturePlex Lab, Western University, London, Ontario N6A 3K7, Canada.*

[4]*Centro de Física de Materiales (CSIC-UPV/EHU) - Material Physics Centre (MPC), P. M. de Lardizábal 5, 20018 San Sebastián, Spain.*

ABSTRACT

The traditional study of disciplines such as philosophy, science, and art has often focused on exceptional individuals, disregarding the broader social context, chance, and collective culture. However, to grasp significant paradigm shifts throughout history, it is also essential to understand the complex network of relationships that surround these persons. Digital humanities and cultural analytics are interdisciplinary fields that utilize computational methods and data analysis to gain insights into these complex cultural phenomena and its relationship with general human behaviour. Among these, the application of tools such as complex networks have been extensively utilized, proving to be valuable in uncovering and quantifying context interactions. In this study, we employ complex network analysis to reveal interdisciplinary knowledge within Wikipedia. As shown in the scheme in Figure 1, we construct graphs representing the cultural universes of influential figures from different knowledge domains. By statistically sampling the resulting complex cultural network we gain valuable insights into the historical context that shaped these exceptional individuals and their contributions. Thus, we explore the possibility of sampling the underlying cultural network, which is inherently diffuse and inaccessible. Our proposed approach involves randomly sampling triads of individuals, representing a diverse range of disciplines. We found that this method enables a quantitative comparison of different fields, and of notable individuals' characteristic network in a given period.

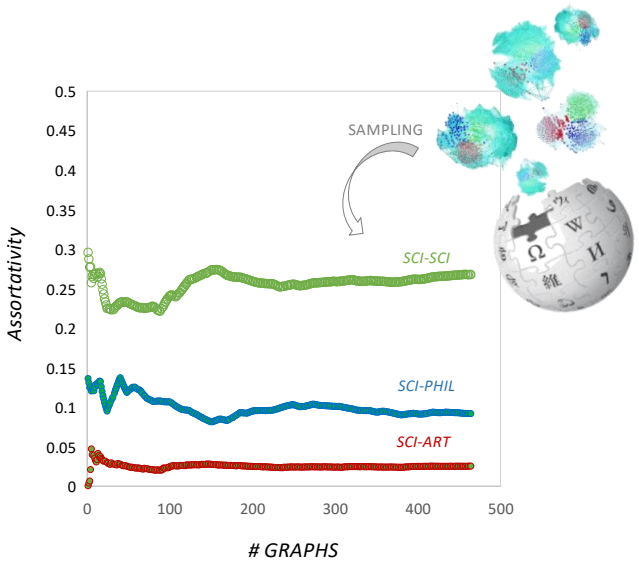

Figure 1. Schematic picture of network sampling from Wikipedia and the resulting inter-cluster relationships through assortativity coefficients A_{12}, A_{22} and A_{23} (as the number of samples increases the observed values stabilize and the comparison becomes more reliable).