

A New Digital Divide? The Deep Politics of Digital Humanities in Asia and the Non-Western World

Thomas S. Mullaney

Associate Professor of Chinese History Stanford University

E-mail: tsmullaney@stanford.edu

Over the past decade, a powerful new suite of spatial, textual, and social network analysis tools – broadly understood as the Digital Humanities – has begun to reshape the methods that we as Humanists and Social Scientists bring to bear on our questions, and indeed the very questions we ask. Looking out over the terrain of Digital Humanities (DH) initiatives, the vista is a marvelous and dynamically changing one. At Stanford University alone, one can point to award-winning programs such as the Mapping the Republic of Letters project, myriad initiatives based at the Stanford Literary Lab, the Kindred Britain project, and the ORBIS Geospatial Network Model of the Roman World, to cite only a handful of examples. When we extend our view across the United States and worldwide, the roster of DH initiatives becomes ever more compelling and exciting.

At the same time, an impartial view of Digital Humanities scholarship in the present day reveals a stark divide between “the West and the rest.” With notable exceptions, such as the Markus platform co-developed by Hilde De Weerd at Leiden University; the China Biographical Database Project at Harvard University, the Japanese and Chinese text analysis work of Hoyt Long and Richard Jean So; and the Digital Islamic Humanities Project at Brown University, far fewer large-scale DH initiatives have focused on Asia and the Non-Western world than on Western Europe and the Americas.

This divide runs very deep, and is not primarily a question of scholarly interest or orientation. The “Asia deficit” within Digital Humanities is in no small part the outcome of more entrenched divides within the platforms and digital tools that form the foundation of DH itself. Digital databases and text corpora – the “raw material” of text mining and computational text analysis – are far more abundant for English and other Latin alphabetic scripts than they are for Chinese, Japanese, Korean, Sanskrit, Hindi, Arabic, and other Non-Latin orthographies. This deficit, in turn, derives in large part from the widespread unavailability of Optical Character Recognition (OCR) platforms, text parsers, and tokenizers capable of handling and processing Non-Latin scripts – not in any way due to the paucity of primary source materials. Even for historic maps of Asia – primary source materials which face none of the same challenges as Non-Latin textual materials – the number of such maps that have been metadata-tagged, digitized, and georectified (three essential processes that must be carried out before digital spatial analysis can be undertaken) are consistently far less than their Western European and U.S. counterparts. As a result, when we look at DH in Western Europe and the Americas, we find a vibrant intellectual environment in which even college and university undergraduates – let alone more advanced researchers – can download off-the-shelf analytical platforms and data corpora, and venture into new and cutting-edge research questions; while, in the context of Asian Studies, we find an environment in which many of the most basic elements of DH research remain underdeveloped or non-existent.

Drawing upon his work on the DHAsia program at Stanford, as well as his research on the 19th, 20th, and 21st-century history of Chinese information technology, historian Thomas S. Mullaney explores the deep politics of Digital Humanities in Asia and the Non-Western world.