

Global Studies Over Time: An Online Database Collection

PI: Patrick Manning, Mellon Professor of World History

Co-PI: Siddarth Chandra, Associate Prof., GSPIA; Director of Asian Studies

PROJECT DESCRIPTION

This collection and web-based display of databases in social sciences (including links to natural sciences) will lay the groundwork for a systematic program of research in Global Studies. It begins with a general statement on the evolution of Global Studies from its initial short-term response to crises in economics and international relations to a more systematic analysis of global patterns in human affairs. The disciplinary vision and promise of Global Studies are such that this is the time to invest substantially in the expansion of this approach. The proposal then outlines the database project and its place in the new generation of Global Studies, comparing this second-generation database project with first-generation efforts to construct large-scale databases. Third, the core description of the project describes the planned project activities in four major phases: data retrieval, data definition, website construction, and review and updating of the website. Finally, summaries of the experience of the PI and Co-PI show their preparation for this project. As a small-scale initiative, this database collection will help show the way to clarifying and advancing Global Studies

The shifting trajectory of Global Studies

Global Studies is distinctive both as a topic and as a trans-disciplinary framework for study. As a topic it encompasses interactions in the human experience and the world surrounding it. As a trans-disciplinary framework it combines work of scholars from social-science and natural-science disciplines as they interpret the major patterns and dilemmas of human society. The early days of Global Studies can be traced to studies of international relations in the early Cold War era, when atomic war was a daily threat. Then, in the 1960s, fears of the population bomb and the emergence of ecological movements caused global approaches to expand into further disciplines.¹ Global Studies gathered steam in the 1990s as the term “globalization” came to represent the palpable expansion in global economic and cultural interconnections. Together, these stages comprised the first generation of Global Studies—policy-oriented, eclectic, short-term concerns about large-scale economic, political, and sometimes ecological issues.

In the twenty-first century a shift in trajectories is setting such pragmatic and immediate concerns in the context of theoretically grounded, data-based analyses ranging across disciplinary frontiers, addressing the whole of the earth, and considering long time-frames as well as immediate concerns. One may now begin to speak of a second generation of Global Studies. Two great discoveries of the 1980s underlay this more systematic approach to global thinking: the verification of patterns of continental drift (based on analysis of changes in the ocean floors) and the verification of the unity and recent evolution of the human species (based on genetic analysis). More recently, the documentation of cyclical swings in the earth’s surface temperature (through analysis of ice cores and other strata) sets the recent pattern of global warming into long-term perspective. In each case, the natural sciences have provided genuinely

¹ More distant ancestry of Global Studies can of course be traced. See David Long and Brian Schmidt, ed., *Imperialism and Internationalism in the Discipline of International Studies* (Albany: SUNY Press, 2005); Patrick Manning, *Navigating World History: Historians Create a Global Past* (New York: Palgrave Macmillan, 2003), 163-180.

new information that affects our understanding of society today as well as its underlying processes.

One of the most important perspectives emerging in the second generation of Global Studies identifies the differing temporal dynamics of the processes in human experience. For instance, technological change is accelerating as never before, but human biology remains much the same. We still give birth to most children in our twenties and thirties; for each generation, it takes just as long for individuals to mature, to acquire knowledge, skill, and culture for enjoyment or for survival. Warfare, a third area of experience, is at once an old habit and one that changes with every generation: can war be stopped completely from now on, and make way for harmonious global society? In exploring the interaction of these and other processes, Global Studies addresses a complicated mix of various social and temporal patterns. In addition to this expanded set of analytical problems, three other factors reinforce the second generation of global interpretation: new historical and scientific data, new methods of analysis, and an expanded global audience. These developments, though they arise from multiple causes, combine to give Global Studies the possibility of becoming a productive source of important new knowledge.

The Proliferation of Information on Global Studies. Sometimes the facts speak for themselves – empirical results cause excitement and questioning that lead to waves of new research. The new data developed out of various types of research open major questions in Global Studies. Most obviously, these include the world economy: data on production, trade, money, levels of wealth, growth rates, and connections among local economies. Also included are social and cultural data on all world areas: recent scholarship is collecting and combining information on all regions, their patterns of culture and society, and their interactions with each other. Data on earth sciences and geography include those on changes in temperature, sea level, and volcanism. In genetics, analysts are collecting details on human genetic unity, but showing regional subgroups. In language, we are learning more on the relations of the various language groups to each other over time. Archaeology is developing data on early remains of human bodies and more recent remains of material culture in every area of the world. Data in these areas, previously collected for study within disciplinary limits, are now recognized to have relevance across disciplinary lines, so that accessibility and accessible documentation of the datasets are of growing importance.

Methods and Theory. To analyze on a global scale, researchers are developing appropriate methods: new ways of asking questions, new frameworks and techniques, and new uses of theory. In frameworks and scales of analysis, researchers must think in fresh and imaginative ways about the scale and boundaries of their analysis: beyond analyzing nations, researchers need to define studies of the past at every relevant level. In disciplinary frameworks, Global Studies is giving increasing attention to the relevance of the various social-science, humanities, and natural-science disciplines for each other. This requires not only study of theory within each discipline, but the additional theoretical work of linking disciplinary theories to each other. Thus, long-term study of migration requires linking theory from demography, sociology, economics, genetics, and anthropology. Historians working in Global Studies, in this and other instances, have found that they must move beyond their usual pragmatic concerns to address theory; on the other hand, they have found that their experience with multiple types of data prepares them well for global analysis. Global Studies might even be said to be developing an encompassing method of analysis. That is, it includes steps from the selection of topics to the presentation of results: the steps most specific to the study of Global Studies are those of modeling, identifying parameters, comparing, connecting systems, and shifting perspectives in analysis. The collection and display of a wide range of globally-oriented databases, as proposed in this project, can facilitate the interdisciplinary analysis that can strengthen Global Studies.

The Concerns of Global audiences. Those who are expressing interest in Global Studies include researchers, policy makers, general audiences, teachers, and students at college, secondary and elementary levels. While these audiences in the U.S. have shown

particular interest, audiences in every part of the world have an equivalent need to know about Global Studies. Of all these audiences, however, researchers are of great importance because of their centrality to building a core of knowledge about Global Studies. The growth in international connections among researchers provides an opportunity for discussion of history at the planetary level.

As an example of the benefits both empirical and conceptual that can come from a global approach, current work of the PI is showing that African populations have grown at quite different rates than widely believed. In research linking demographic studies of Africa for the late twentieth century for those of the nineteenth century and earlier, he has discovered that African population estimates for earlier times, though widely agreed upon, were impossibly low in comparison to documented populations for the past fifty years. As a result, population figures for Africa before 1900 must be increased by about 50%, and analysts of population today must now understand that populations in earlier times grew much slower than has been commonly thought.² More generally, global and interdisciplinary analysis is likely to change significantly our understanding of the patterns of the present and the past, and is likely to change our expectations for the future.

“Global Studies Over Time” in the Context of Other Database Projects

The first generation of research in Global Studies created a number of major institutions and important databases. In political and social research, the Interuniversity Consortium on Political and Social Research (ICPSR), based at the University of Michigan, began collecting and distributing datasets in the era of mainframe computers.³ In anthropology, the Human Relations Area Files began even earlier, and established itself as the single greatest database for comparative ethnographic information, with data organized by ethnic units.⁴ In economics, several well-funded organizations have collected and displayed data within national units. The U.S. National Bureau of Economic Research,⁵ the Paris-based Organization of Economic Cooperation and Development,⁶ and the World Bank⁷ have each developed and displayed ample quantities of data, and have transferred them to websites. These data, however, focus primarily on current and recent data. For economic historical data over the longer run, two individual scholars (both based in England) have carried out monumental compilations. B. R. Mitchell began with European historical statistics, and gradually expanded both his geographic and temporal scope.⁸ Angus Maddison developed global estimates of population and gross domestic product that are less specific but even more extensive, addressing much of the world for the last millennium.⁹

For all of these databases, however, the data structures were conceptualized several decades ago—in isolated disciplines, within national or ethnic limits, and set up in the technology of the time. The development of Global Studies faces persistent problems of the myopia and competitiveness of national outlooks; the shortage of professionally trained researchers with global outlooks; inadequate institutions facilitating Global Studies; absence of consensus on global research priorities; and scarce funding for research. Even in fields as obviously global as economics, sponsorship of research in Global Studies has been slow to develop. The field of history generally receives very little in research funding, and the

² Patrick Manning, “African Population: Projections, 1850-1950” (forthcoming).

³ <http://www.icpsr.umich.edu/>.

⁴ <http://www.yale.edu/hraf/>.

⁵ <http://www.nber.org/data/>.

⁶ http://www.oecd.org/statsportal/0,2639,en_2825_293564_1_1_1_1_1,00.html.

⁷ <http://www.worldbank.org/> - link to “data.”

⁸ B. R. Mitchell, *International Historical Statistics: Africa, Asia and Oceania, 1750-2000*, 4th ed. (New York: Palgrave Macmillan, 2003).

⁹ These print volumes have now been made accessible online: <http://www.ggd.net/maddison/>.

established foundations in social sciences generally restrict their funding either to national-level or micro-level studies of the past or to contemporary globalization without historical depth. To a surprising degree, research remains mired within the boundaries of separate disciplines, regions, and time frames.

Datasets of the second generation of Global Studies, designed beginning in the late 1990s, are more consciously transnational and transdisciplinary, and they are structured in the more flexible and relational technology made available in recent years. In some cases, large institutions have taken on research and data display in this updated framework. Among these institutions are the Integrated Public Use Microdata Series (IPUMS) at the Minnesota Population Center,¹⁰ the dataset collection at Hitotsubashi University (Tokyo),¹¹ the Earth Institute at Columbia University,¹² the United Nations Population Division,¹³ and a few others. Much of the original research in Global Studies, however, has been carried out by individuals and small groups scholars, whose data sets and analyses therefore risk being neglected rather than integrated into the larger task of global analysis.¹⁴

In current Global Studies programs in U.S. universities, most operate within the eclectic and short-term approach of the first generation, while a few are undertaking the research and analysis of the second generation. The outstanding program of the second generation is the Earth Institute at Columbia University, which has undertaken major research in natural and social sciences as well as its policy-oriented support of the United Nations Millennium Program of addressing global inequality. Most of the National Resource Centers in International Studies, in contrast, focus primarily on curriculum in international studies and international business in the short term, without significant emphasis on research or long-term analysis.

This second generation of Global Studies promises a stimulating, interdisciplinary ethos, in which new and larger-scale programs and data structures are developed and linked to the existing programs of analysis within area-studies and national institutions. The improved nature of relational databases and dynamic program provide technical support for the intellectual realization that social-science phenomena take place at global as well as local levels, and that scholars must develop, share, and analyze social-science data at the global level.

In summary, the changing trajectory of world events opens the door to developing new levels of knowledge about humanity. Along this path, researchers should gather new information on the distant past and the recent past, integrating it and linking it to new concerns about the future, and finding ways of alternating analytical perspectives to encompass the expanding body of information into an expanded structure of knowledge, which we identify as Global Studies.

“Global Studies Over Time”: Stages of Project Activity

The overall emphases of this project are that it is to be interdisciplinary, geographically comprehensive, and address long-term as well as short-term change. While it will draw significantly on the lessons learned in creation of the various large databases of earlier work, it will also give substantial attention to retrieval and linkage of isolated databases prepared by individual scholars.¹⁵ In addition, a parallel research proposal, recently funded by the Faculty

¹⁰ <http://www.ipums.umn.edu/>

¹¹ <http://www.hit-u.ac.jp/laboratories/index-e.html>.

¹² <http://www.earthinstitute.columbia.edu/research/index.html>.

¹³ <http://www.un.org/esa/population/unpop.htm>; for the 2004 population revision, see <http://esa.un.org/unpp/>.

¹⁴ For instance, the Global Economic History Network, based at the London School of Economics, was able to conduct three years of conferences with support from the Leverhulme Foundation. See <http://www.lse.ac.uk/collections/economicHistory/GEHN.htm>.

¹⁵ As an example, the data on Caribbean economies and societies for the nineteenth and twentieth centuries, assembled by Prof. Selwyn Carrington of Howard University, could be put into electronic form and displayed online.

Research and Scholarship Program, may reinforce the efforts of this database project. That proposal, for a “global database design seminar,” will enable members of the Pitt faculty to hear presentations by the directors of such major database projects as the China Historical GIS project, and begin setting design criteria for a more general global historical database.¹⁶ The two projects, if both are funded, should have a useful synergy: meeting with designers of large-scale databases will provide us with more information on datasets to collect; practical experience with collecting and displaying databases in this project will sharpen questions in the design seminar.

Data retrieval. The PI and Co-PI will solicit databases for inclusion on the website through contact with journal editors and scholars. Our acquaintances with scholars and journal editors are sufficiently broad that we expect to be able to contact the authors of dozens of significant databases within a few months. It is our experience that authors and compilers are happy to have their data put online or mirrored on additional sites as long as they are properly attributed. To this end, we will develop a standard form for attribution, to be placed prominently and repeatedly in each dataset. From our discussion of the project with colleagues, we expect to find new ideas for additional datasets and for improved techniques of display.

Data definition. As datasets are acquired, the PI and Co-PI will implement a process of definition for each database. First, we will retain and reproduce the statement of data definition supplied by the compiler of each dataset. In addition, we will add our own comprehensive and consistent description of the dataset, including its variables, time frame, geographic extent, dimensions (weights and measures), and sources. In addition, our presentation of the data will distinguish several levels of data transformation:

- Raw documents, unsearchable (e.g., pdf files)
- Raw documents, searchable
- Original data, structured
- Transformed data
- Aggregated data

While the project in its first year will focus on documenting and displaying data in the form given to us by authors and compilers, a more fully developed version of the website in later years might aspire to systematic transformation, so that the various datasets could become maximally comparable to each other.¹⁷

Website construction. (a) Setup. The project will establish a website on the UCIS server which will be displayed within the Global Studies Section of the UCIS site.¹⁸ Two student employees will each be engaged to work for 10 hours per week for two semesters, one on Level 1 of the project and one on Level 2. A free-lance designer will create the design of the site and update it when Level 2 is released. **(b) Level 1 of website.** The first stage of the website will include introductory text and a simple menu in HTML enabling users to access the initial display of databases. This portion of the website will be released and announced in November 2007; further work at this level will continue throughout the project year. **(c) Level 2 of website.** A small-scale, pilot version of a dynamic interface for retrieving and querying databases will be created by the Level 2 graduate student employee. Work will begin at the start of the project, and this more advanced version of the website will be released in April of 2008. The pilot program will enable users to query several datasets, combine their results, and analyze the resulting aggregate files. This pilot project version of a dynamic database will begin with a series

¹⁶ “World-Historical Database Design Group,” submitted by Patrick Manning to the Faculty Research and Scholarship Program, School of Arts and Sciences, University of Pittsburgh; letter of award Feb. 9, 2007.

¹⁷ At the same time, it will be important to maintain full documentation on the original data, as well as documentation of each stage of data transformation.

¹⁸ Discussions with Dr. Wolfgang Schlör, Interim Director of UCIS and of Global Studies, and Mark Weixel, UCIS Director of Informatics, have confirmed that UCIS will be able to host the website. Mark Weixel was especially helpful in suggesting improvements to our initial plan for the project website.

of datasets well known to the PI and Co-PI, notably the Co-PI's datasets on the Dutch East Indies in the early twentieth century, the PI's data on British and French African colonies, and other comparable datasets. A focus on trade data at the initial stage will simplify the normalization of data so that the complexity of the queries is minimized. This pilot program will also include an accounting of the various levels of transformation of data, working papers on various technical and interpretive aspects of the website, and a "comments" function enabling viewers to pose questions and suggestions to the website directors and also to the authors and compilers of the databases.

Website review and updating. The PI and Co-PI will review and revise website materials. During the project year they will develop a routine and a regularized procedure for periodic review of each segment of the website, checking for broken links, errors in entry, and other technical bugs, but also seeking ways to improve the website design and to facilitate productive linkages among the datasets. Such review should facilitate the overall purpose of the project: to collect and ultimately link data from all social sciences – economy, sociology, politics, demography, anthropology – plus environmental data. As noted above, this work may also yield insights for the parallel work of designing a comprehensive global historical database.

Project time frame. The project, to begin in August 2007, will complete its work in July 2008. The PI and Co-PI expect to submit proposals for external funding to continue and expand the scale of this project.

Possible long-term implications. While this project is intended to be of modest scale within a one-year time frame, it can also explore the possibilities of substantial expansion at later stages. Two characteristics that frame the approach of the project staff may make this database project distinctive. First, it is intended to be broadly trans-disciplinary, emphasizing the gathering and linking of data not only in economics but in demography, social conditions, geography, government, and health. Second is the experience of the project staff in colonial areas of the world and also in precolonial times: this draws their attention to the importance of locating and constructing comparable data sets for all areas of the world, as a precondition for conducting properly global analysis of human society over time.

We conclude by noting four areas of possible expansion of this project after the conclusion of this initial stage. First, the effort of communicating with journal editors and the holders of databases may result in the communication of a large quantity of data to the project. The same communication may create possibilities for collaboration with other research groups developing databases in various fields. Second, the review and editing of the incoming data may facilitate the development of a comprehensive cataloguing system for describing and linking the datasets. Third, the system of cataloguing and querying databases might ultimately create a situation in which researchers could access datasets held throughout the world, query them, and combine the results to create new datasets of unparalleled breadth and flexibility. Fourth, the effort of documenting and assessing the datasets and their linkage might create an international network of scholars who would create procedures for critiquing the resulting data and analyses. Similarly, at this scale, it is perhaps appropriate to compare the direction of this project on social science data with the progress of the Google Books project on text documents in print.