East Asian Digital Humanities

Dr. Molly Des Jardin

reading/video assignments revised March 21, 2019

Course Description

This course covers a wide range of current and emerging digital projects and topics in East Asian studies. Students will engage with digital projects focused on East Asia (encompassing Japanese, Chinese, and Korean languages) as well as research being done on digital methodologies for the humanities in those areas. Coursework consists of project and research analysis, active discussion, and learning about the implementation of various digital projects. No technical expertise is required but students must have reading knowledge of Chinese, Japanese, or Korean at the high-intermediate or advanced level. Class presentations, papers, discussions, and all course readings are in English, but final projects involve reading articles and/or critiquing projects in the language and geographic area of students' expertise.

Assignments and Grading:

1. Final Project (40%)

Undergraduate Final Project: Literature Review

Students will conduct a literature review on a topic of interest. It should be based on one or a combination of the class units so students have a foundation for their review based in the articles or projects introduced in the syllabus. The review should include the current state of the field, such as implemented projects or research papers from the past 10 years. Some background reading older than 10 years may be used. **At least one** project or research paper reviewed must be in the student's language of expertise, but the paper should be in English and any quotations from the foreign-language source translated into English. Students may use any citation style they are comfortable with. **Maximum 15 pages, double-spaced (12-point Arial or Times New Roman font), not including bibliography.**

Graduate Final Project: Project Pitch and Analysis

Students will outline and pitch a project or research proposal of their devising, and contextualize it in the current state of the field and technological developments. The assignment may combine multiple units of class topics, but should be a coherent project with a clear research aim. This assignment consists of:

   a. Project pitch to a funding organization or student's department/advisor. (1 single-spaced page ONLY)
   b. Literature review and critical analysis of current research related to the student's idea. This can include digital projects and research papers. **At least one** project or research paper reviewed must be in the student's language of expertise, but the paper should be in English and any quotations from the foreign-language source translated into English.
   c. Introduction to and outline of the student's project idea, including research question and why it is novel or important in context of the existing field.
   d. Outline of technology, methodologies, and tools (if available) needed to complete the project.

e. Bibliography of literature review and critical analysis/contextualization sections

The project pitch and analysis should consist of **no more than 20 double-spaced pages in 12-point Arial or Times New Roman font**, excluding part (a) and (e). Students should take the space they need -- no more and no less -- to explain, contextualize, and justify their projects, even if they don't reach the 20-page maximum.

2. Class Participation & Discussion (40%)
Attendance and participation in class discussions (as well as thorough preparation by reading materials and previewing projects, software, and tutorials) is mandatory. Generally, the first half of the seminar meeting will cover the specifics of research and projects (including Q&A among students and guest speakers), and the second will be more general critical discussion of the week's topic.

3. Response Papers (20%)
Students will write to the week's readings and projects, due the day after class each week. The grade is pass/fail and is an additional participation assessment. The papers are not to be final, polished work, but rather an initial reaction and thoughts on the week's topic and issues, and should be no longer than two pages.

Week 1-2. What is Digital Humanities?
Introductions and student perspectives
Readings:
- Vierthaler, Paul. "Imperial Chinese studies and trends in the digital humanities." *China Policy Institute: Analysis* (2016). https://cpianalysis.org/2016/06/07/imperial-chinese-studies-and-trends-in-the-digital-humanities/
- Cha, Javier. "Digital / humanities: new media and old ways in South Korea." *Asiascape: Digital Asia* 2, no. 1-2 (2015): 127–48. doi:10.1163/22142312-12340022.
- Cha, Javier. "Digital Korean studies: recent advances and new frontiers." *Digital Library Perspectives* 34, no. 3 (2018): 227-244.
- Agarwal, Kritika. "A history dissertation goes digital." *AHA Today* (2017). http://blog.historians.org/2017/08/a-history-dissertation-goes-digital/
- De Weerdt, Hilde. "Digital interpretations." http://chinese-empires.eu/blog/digital-interpretations/
- Stockmann, Daniela. "Towards area-smart data science: critical questions for working with big data from China." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2718120

Further resources:
- De Weerdt, Hilde. "Digital perspectives on imperial Chinese political history." https://vimeo.com/168242706

Week 3. Mapping/GIS
Readings:

- D'Ignazio, Catherine. "What would feminist data visualization look like?" https://civic.mit.edu/feminist-data-visualization
- Murrieta-Flores, Patricia, Christopher Elliott Donaldson, and Ian Norman Gregory. "GIS and literary history: advancing digital humanities research through the spatial analysis of historical travel writing and topographical literature." *Digital Humanities Quarterly* 11, no. 1 (2017). http://www.digitalhumanities.org/dhq/vol/11/1/000283/000283.html
- Bauch, Nicholas. "Digital geohumanities: visualizing geographic thought." *International Journal of Humanities and Arts Computing* 11, no. 1 (2017): 1-15.

Projects to explore:
- Hidenori Watanave projects:
    - http://n.mapping.jp/index_en.html
    - http://1941.mapping.jp/
    - http://en.wasurenai.mapping.jp/
- JapanMap: https://worldmap.harvard.edu/japanmap/
- Digital Archive of Japan's 2011 Disasters: http://jdarchive.org
- China Historical GIS: http://maps.cga.harvard.edu/chgis/
- 江戸後期 武蔵・相模国 村名マップ: https://fudoki.midoriit.com/ (Japanese only)
- 近代書物流通マップ: https://the-artifacts.firebaseapp.com/ (Japanese only)

Further resources:
- *International Journal of Humanities and Arts Computing* 11, no. 1 http://www.euppublishing.com/toc/ijhac/11/1

Week 4. Metadata

Readings:
- Vierthaler, Paul. "Analyzing printing trends in late imperial China using large bibliometric datasets." *HJAS* 76 (2016): 87-133.
- Drabinski, Emily. "Queering the catalog: queer theory and the politics of correction." *The Library Quarterly: Information, Community, Policy* 83, no. 2 (2013): 94–111.
- Wilson, Emma Annette and Mary Alexander. "When metadata becomes outreach: indexing, describing, and encoding for DH." *dh+lib* 2016 Special Issue (2016). https://acrl.ala.org/dh/2016/07/29/when-metadata-becomes-outreach/

Projects to explore:
- Buddhist Studies bibliographic visualization http://www.acmuller.net/zotero-visualization/output/index.html (see also: https://networks.h-net.org/node/6060/discussions/190645/resource-visualization-zotero-buddhist-studies-bibliography)

Further resources:
- Buckner, Cameron, Mathias Niepert, and Colin Allen. "From Encyclopedia to Ontology: Toward Dynamic Representation of the Discipline of Philosophy." *Synthese* 182 (2011): 205-233.

ASSIGNMENT FOR WEEK 5: Email the professor 30 documents (*plain unicode text*) in contemporary Chinese, Japanese, or Korean (pick one) from online news site(s). These will be compiled into a sample corpus. All filenames should end in .txt. Feel free to use our subscription databases at Penn for news sources; ask your librarian for a list in your language. See courseware page for detailed explanation and naming conventions.

Week 5. Digital Literary Studies

Readings:
- Long, Hoyt, and Richard Jean So. "Literary pattern recognition: modernism between close reading and machine learning." *Critical Inquiry* 42, no. 2 (2016): 235-267.
- Heuser, Ryan, and Long Le-Khac. "Learning to read data: bringing out the humanistic in the digital humanities." *Victorian Studies* 54, no. 1 (2011): 79–86. doi:10.2979/victorianstudies.54.1.79. (PDF)
- "little magazine, worldwide network." In Eric Bulson, *little magazine, world form* (New York: Columbia University Press, 2017): 33-73 (notes on 280-286).
- Vierthaler, Paul. "Fiction and history: polarity and stylistic gradience in late imperial Chinese literature." *Journal of Cultural Analytics* 23 (2016). https://culturalanalytics.org/2016/05/fiction-and-history-polarity-and-stylistic-gradience-in-late-imperial-chinese-literature/

Videos:
- Vierthaler, Paul. "Digital research into the authorship of the Jin Ping Mei" https://vimeo.com/163570659
- Vierthaler, Paul. "Digital approaches to late imperial Chinese literature: exploring quasi-historical texts" https://youtu.be/adRlz2HbDD0

Week 6. Network Analysis
Readings:
- Clifford, Timothy. "Ming anthologies of ancient-style prose." In "In the Eye of the Selector: Ancient-Style Prose Anthologies in Ming Dynasty (1368-1644) China" (University of Pennsylvania, 2017).
- Lee, Jae-Yon. "Before and after the 'age of literary coteries': a diachronic analysis of writers' networks in Korea, 1917-1927." *Korea Journal* 57, no. 2 (Summer 2017): 35-68.
- So, Richard Jean, and Hoyt Long. "Network analysis and the sociology of modernism." *boundary 2* 40, no. 2 (2013): 147-182.
- Lee, Kihwang, and Jae Yun Lee. "Quantitative analysis of the 17th-19th century Korean culinary manuscripts." In *New Eyes for Discovery: Foundations and Imaginations of Digital Humanities* 數位人文研究的新視野 : 基礎與想像 (2011): 189–204. Taipei: Taiwan da xue chu bang zhong xin.

Videos:
- Cha, Javier. "Intellectual history and computing: modeling and simulating the world of Korean yangban" https://vimeo.com/169044260

Projects to explore:
- Network Maps (Tom Mazanec; exchange poetry in late medieval China): http://tommazanec.com/projects/network-maps/

Further resources:
- Gephi tutorial http://www.briansarnacki.com/gephi-tutorial/
- De Weerdt, Hilde. "Digital perspectives on imperial Chinese history" https://vimeo.com/168242706

Week 7. Corpus building and why it's hard

Readings:
- Sturgeon, Donald. "Unsupervised extraction of training data for pre-modern Chinese OCR." *Proceedings of the Thirtieth International Florida Artificial Intelligence Research Society Conference* (2017): 613-618.
- Bode, Katherine. "The equivalence of 'close' and 'distant' reading; or, toward a new object for data-rich literary history." *Modern Language Quarterly* 78.1 (2017): 77-106.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K Gray, The Google Books Team, Joseph P Pickett, et al. "Quantitative analysis of culture using millions of digitized books." Science 331, no. 6014 (2011): 176–82.
- Pechenick, Eitan Adam, Christopher M. Danforth, and Peter Sheridan Dodds. 2015. "Characterizing the Google Books corpus: strong limits to inferences of socio-cultural and linguistic evolution." *PLOS One* 10, no. 10 (2015): 1–24.

Reference: Word segmentation tools/software
- MeCab for Japanese http://taku910.github.io/mecab/
- Rakuten MA for contemporary Japanese and Chinese https://github.com/rakuten-nlp/rakutenma and demo at http://rakuten-nlp.github.io/rakutenma/
- Stanford Word Segmenter for contemporary Chinese https://nlp.stanford.edu/software/segmenter.shtml - and it can be used with NLTK https://www.quora.com/How-does-Python-NLTK-process-Chinese
- Jieba for Chinese https://github.com/fxsjy/jieba

Further Resources:
- Bode, Katherine. "A Response to some responses." https://katherinebode.wordpress.com/a-response-to-some-responses/

*** We will use your sample contemporary news corpora from your previous assignment (for Chinese, Japanese, and/or Korean, segmented by the professor) for weeks 8-10 ***

Week 8. Text Analysis basics

Readings:
- Catalinac, Amy. "Quantitative text analysis with Asian languages: some problems and solutions." *Polimetrics* 1, no. 1 (2014): 14-16.

- *Digital Humanities at Berkeley* blog. "Ghost in the machine." http://digitalhumanities.berkeley.edu/blog/16/02/29/ghost-machine
- Deng, Ke, Peter K. Bol, Kate J. Li, and Jun S. Liu. "On the unsupervised analysis of domain-specific Chinese texts." *PNAS* 113, no. 22 (2016): 6154-6159.
- Slingerland, Edward, Ryan Nichols, Kristoffer Neilbo, and Carson Logan. "The distant reading of religious texts: a 'big data' approach to mind-body concepts in early China." *Journal of the American Academy of Religion* 85, no. 4 (2017): 985-1016.

Videos:
- So, Richard Jean and Hoyt Long. "A Macroanalysis of modern Japanese and Chinese texts" https://vimeo.com/164368902

Tools available:
- Voyant Tools: http://voyant-tools.org
  - Download your own: https://github.com/sgsinclair/VoyantServer
- AntConc: http://www.laurenceanthony.net/software/antconc/
  - Intro videos: https://www.youtube.com/playlist?list=PLiRIDpYmiC0Ta0-Hdvc1D7hG6dmiS_TZj

Further resources:
- Froehlich, Heather. "An introductory bibliography to corpus linguistics." https://hfroehli.ch/2014/05/11/intro-bibliography-corpus-linguistics/
- Froehlich, Heather. "Corpus analysis with Antconc." https://programminghistorian.org/lessons/corpus-analysis-with-antconc

Week 9. Text Analysis advanced topics:
Text reuse and comparison

Readings:
- Bingenheimer, Marcus. "Collation strategies for the Buddhist canon -- as seen in the frequency and impact of character variance in canonical editions of the Song Gaoseng Zhuan 宋高僧傳 (T.2061)." *East Asian Publishing and Society* 4, no. 2 (2014): 155–74.

- Sturgeon, Donald. "Unsupervised identification of text reuse in early Chinese literature." *Digital Scholarship in the Humanities* (2017): 1-15. Accessed November 11, 2017.
- Vierthaler, Paul and Mees Gelein. "A BLAST-based, language-agnostic text reuse algorithm with a MARKUS implementation and sequence alignment optimized for large Chinese corpora." *Journal of Cultural Analytics* (3/22/2019). http://culturalanalytics.org/2019/03/a-blast-based-language-agnostic-text-reuse-algorithm-with-a-markus-implementation-and-sequence-alignment-optimized-for-large-chinese-corpora/
- Text Tools introduction. https://digitalsinology.org/text-tools/
- Text Tools tutorial. https://dsturgeon.net/texttools/

Projects to explore:
- Text Tools: https://digitalsinology.org/text-tools/
- Viral Texts: http://viraltexts.org

- CText Parallel Passages: https://ctext.org/tools/parallel-passages
- Text Encoding Initiative (Wikipedia): https://en.wikipedia.org/wiki/Text_Encoding_Initiative

Week 10. Text Analysis advanced topics:
Topic modeling

Readings:
- Underwood, Ted. "Topic modeling made just simple enough." https://tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough/
- Roland, Teddy. "Topic modeling: what humanists actually do with it." http://digitalhumanities.berkeley.edu/blog/16/07/14/topic-modeling-what-humanists-actually-do-it-guest-post-teddy-roland-university
- London, Daniel. "Of nuance and algorithms: what conceptual history can learn from topic modeling." https://jhiblog.org/2016/04/13/of-nuance-and-algorithms-what-conceptual-history-can-learn-from-topic-modeling/
- Nichols, Ryan, Edward Slingerland, Kristoffer Nielbo, Uffe Bergeton, Carson Logan, and Scott Kleinman. "Modeling the contested relationship between *Analects, Mencius,* and *Xunzi*: preliminary evidence from a machine-learning approach." *Journal of Asian Studies* 77, no. 1 (2018): 19-7.
- Allen, Colin, Hongliang Luo, Jaimie Murdock, Jianghuai Pu, Xiaohong Wang, Yanjie Zhai, and Kun Zhao. "Topic modeling the Hàn diǎn Ancient Classics (汉典古籍)." *Journal of Cultural Analytics* (10/12/2017). http://culturalanalytics.org/2017/10/topic-modeling-the-han-dian-ancient-classics-%e6%b1%89%e5%85%b8%e5%8f%a4%e7%b1%8d/

Tools to explore:
- Topic Modeling Tool: https://github.com/senderle/topic-modeling-tool

Further resources:
- Weingart, Scott B. "Topic modeling for humanists: a guided tour." http://scottbot.net/topic-modeling-for-humanists-a-guided-tour/
- *Journal of Digital Humanities* 2, no. 1 (Winter 2012) special issue on topic modeling: http://journalofdigitalhumanities.org/2-1/
- Blei, David M. "Probabilistic topic models." *Communications of the ACM* 55, no. 4 (April 2012): 77-84.
- Seroussi, Yanir, Ingrid Zukerman, and Fabian Bohnert. 2014. "Authorship attribution with topic models." *Computational Linguistics* 40 (2): 269–310.
- Mimno, David. "Topic modeling bibliography." https://mimno.infosci.cornell.edu/topics.html

Week 11. APIs and how to communicate with a database
Professor presentation:
Quick intro to APIs (preview slides before class)

Readings:

- Sturgeon, Donald. "Crowdsourcing, APIs, and a digital library of Chinese." *China Policy Institute: Analysis* (2016). https://cpianalysis.org/2016/06/08/crowdsourcing-apis-and-a-digital-library-of-chinese/
- De Weerdt, Hilde. "Collaborative innovation and the Chinese digital humanities." *China Policy Institute: Analysis* (2016). https://cpianalysis.org/2016/06/09/collaborative-innovation-and-the-chinese-digital-humanities/
- "Database" (Wikipedia). https://en.wikipedia.org/wiki/Database (skim)
- "Database: design and modeling" (Wikipedia). https://en.wikipedia.org/wiki/Database#Design_and_modeling (read this part carefully)

Projects to explore:

- China Biographical Database: https://projects.iq.harvard.edu/cbdb
- CTP Digital Humanities Tools: http://ctext.org/tools/digital-humanities
- MARKUS: http://dh.chinese-empires.eu/markus/beta (Google Chrome only)
- Hiroshima Cat Street View: http://hiroshima-welcome.jp/kanpai/catstreetview/index_2_off.html
- 近代書物流通マップ: https://the-artifacts.firebaseapp.com/ (Japanese only)

Further resources:

- CTP API tutorial: https://digitalsinology.org/classical-chinese-dh-python-ctext-api/
- CTP API documentation: http://ctext.org/tools/api
- MARCUS tutorial videos http://dh.chinese-empires.eu/markus/beta/video.html

Week 12. Crowdsourcing and public engagement

Readings:

- Drouin, Jeffrey. "Close- and distant-reading modernism: network analysis, text mining, and teaching *The Little Review*." *The Journal of Modern Periodical Studies* 5, no. 1 (2014): 110–35. doi:10.1353/jmp.2014.0001.
- Brabham, Daren C. "The Myth of amateur crowds." *Information, Communication & Society* 15, no. 3 (2012): 394–410.
- Mostern, Ruth and Marieka Arksey. "Don't just build it, they probably won't come: data sharing and the social life of data in the historical quantitative social sciences." *International Journal of Humanities and Arts Computing* 10, no. 2 (2016): 205-224.
- Owens, Trevor. "Crowdsourcing digital heritage: the objectives are upside down." http://www.trevorowens.org/2012/03/crowdsourcing-cultural-heritage-the-objectives-are-upside-down/
- Hashimoto, Yuta, et al. "The kuzushiji project: developing a mobile learning application for reading early modern Japanese texts." *Digital Humanities Quarterly* 11, no. 1 (2017). http://www.digitalhumanities.org/dhq/vol/11/1/000281/000281.html

Videos:

- Sturgeon, Donald. "Towards a sustainable infrastructure for historical Chinese texts." https://youtu.be/D6x-d5RqkTs

Projects to explore:
- CText: http://ctext.org
- Digital Archive of Japan's 2011 Disasters: http://jdarchive.org
- Minna de Honkoku (みんなで翻刻): https://honkoku.org (including intro video in English)

Week 13. Putting projects and portfolios online

Projects to explore:
- MLA Humanities Commons: https://hcommons.org
- Creative Commons: https://creativecommons.org

To prepare:
Look at sample projects and also brainstorm what you might do. Be prepared to share!
- Wordpress: https://wordpress.com (compare with https://wordpress.org)
- Omeka: https://omeka.org/classic/showcase/
- Scalar: https://scalar.usc.edu/scalar/
- Mukurtu: http://mukurtu.org
- Knight Lab tools: https://knightlab.northwestern.edu/projects/
- Jekyll + Github.io: https://programminghistorian.org/lessons/building-static-sites-with-jekyll-github-pages

Videos:
- "History Respawned: Assassin's Creed Origins and Digital Giza." https://youtu.be/ANQL5z_JEpQ

Week 14. Project presentations

Student lightning talks (5 minutes):
- Final project lightning round
- 5-minute following Q&A each talk
- Get feedback before turning in the final project assignment

Week 15. Wrap up

Readings:
- Piper, Andrew. "There will be numbers." *Journal of Cultural Analytics* (5/23/2016). http://culturalanalytics.org/2016/05/there-will-be-numbers/
- Da, Nan Z. "The computational case against computational literary studies." *Critical Inquiry* 45 (Spring 2019): 601-639.
- Treat, John Whittier. "Japan is interesting: modern Japanese literary studies today" *Japan Forum* 30, no. 3 (2018): 421-440.
- Pick two interviews with DH figures from *LA Review of Books:* https://lareviewofbooks.org/feature/the-digital-in-the-humanities/
- De Weerdt, Hilde. "Isn't the Siku quanshu enough? Reflections on the impact of new

digital tools for classical Chinese." http://chinese-empires.eu/blog/isnt-the-siku-quanshu-enough-reflections-on-the-impact-of-new-digital-tools-for-classical-chinese/

- Gibbs, Frederick W. "New forms of history: critiquing data and its representations." http://tah.oah.org/february-2016/new-forms-of-history-critiquing-data-and-its-representations/

Resources of note:

Penn Libraries Professionals

Molly Des Jardin, Japanese Studies Librarian, Korean Studies Liaison (mollydes@upenn.edu)

Brian Vivier, Chinese Studies Librarian (vivier@pobox.upenn.edu)

Scott Enderle, Digital Humanities Specialist (enderlej@upenn.edu)

Sasha Renninger, Digital Humanities Specialist (sashafr@upenn.edu)

Girmay Misgna, Mapping & Geospatial Data Librarian (gmisgna@upenn.edu)

Laurie Allen, Director for Digital Scholarship (laallen@upenn.edu)

Dot Porter, Curator of Digital Research Services (dorp@upenn.edu)

General DH

http://dhdebates.gc.cuny.edu/ (*Debates in the Digital Humanities*)

http://culturalanalytics.org/ (*Journal of Cultural Analytics*)

http://www.digitalhumanities.org/dhq/ (*Digital Humanities Quarterly*)

http://digitalhumanities.org/answers/ (Digital Humanities Questions & Answers)

http://ach.org/ (The Association for Computers and the Humanities)

http://www.themacroscope.org/2.0/ (*Exploring Big Historical Data: The Historian's Macroscope* - book and companion site)

A New Companion to Digital Humanities, ed. Susan Schreibman, Ray Siemens, and John Unsworth. (Reserves)

Lynda.com training videos (https://lynda.upenn.edu)

Asia

Asiascape: Digital Asia journal https://franklin.library.upenn.edu/catalog/FRANKLIN_9977082724103681

https://vimeo.com/dhasia (DHAsia @ Stanford talk videos)

Shu wei ren wen yan jiu cong shu series (some English content) (https://franklin.library.upenn.edu/catalog?utf8=%E2%9C%93&search_field=title_search&q=%22Shu+wei+ren+wen+yan+jiu+cong+shu%22)

Japan

http://codh.rois.ac.jp/ (Center for Open Data in the Humanities)

https://www.jadh.org/ (Japanese Association for Digital Humanities)

http://www.dhii.jp/nagasaki/blog/ (Digital Humanities Notes in Japan blog)

https://www.jstage.jst.go.jp/browse/jjadh (*Journal of the Japanese Association for Digital Humanities*) (appears defunct)

http://21dzk.l.u-tokyo.ac.jp/DHI/ (University of Tokyo Digital Humanities Initiative)
http://guides.library.upenn.edu/japanesetext (research guide on Japanese text analysis resources)
http://guides.library.upenn.edu/japanesedigitalresources (research guide on online resources for Japanese studies)
Shirīzu Nihon bunka dejitaru hyūmanitīzu (Japanese content only) (https://franklin.library.upenn.edu/catalog?utf8=%E2%9C%93&search_field=title_search&q=Shiri%CC%84zu+Nihon+bunka+dejitaru+hyu%CC%84maniti%CC%84zu)
Digital humanities JP (Japanese content only) (https://franklin.library.upenn.edu/catalog/FRANKLIN_9963041733503681)

China
https://digitalsinology.org/ (Digital Sinology blog)
http://maps.cga.harvard.edu/chgis/ (China Historical GIS)
Big Data and the Study of Chinese History (Chinese content only) https://mp.weixin.qq.com/s/ALkqC4x_VsgTTlFHJL2ZOw
@digitalsinology on Twitter
https://digitalsinology.org/classical-chinese-dh-getting-started/ (Classical Chinese DH: Getting Started)
http://asbc.iis.sinica.edu.tw/ (Academia Sinica Balanced Corpus of Modern Chinese)
http://lwc.daanvanesch.nl/ (Leiden Weibo Corpus)

Korea
http://guides.library.upenn.edu/koreanstudies