
PROFILES IN PRESERVATION

The Core Historical Literature of Agriculture: Agricultural History Rejuvenated for Today's Users

Linda Stewart
Joy Paulson

The
CORE HISTORICAL LITERATURE
Of **AGRICULTURE**
Albert R. Mann Library, Cornell University

ABOUT
SEARCH
BROWSE
CONTACT
HELP
HOME

CORNELL

HISTORY CELEBRATES the battlefields whereon we meet our death, but scorns to speak of the plowed fields whereby we thrive. It knows the names of the king's bastards but cannot tell us the origin of wheat. This is the way of human folly." -- *Jean-Henri Fabre*

The Core Historical Literature of Agriculture (CHLA) is a core electronic collection of agricultural texts published between the early nineteenth century and the middle to late twentieth century. Full-text materials cover agricultural economics, agricultural engineering, animal science, crops and their protection, food science, forestry, human nutrition, rural sociology, and soil science. Scholars have selected the titles in this collection for their historical importance. Their evaluations and 4,500 core titles are detailed in the seven volume series *The Literature of the Agricultural Sciences*, Wallace C. Olsen, series editor.

Current online holdings: Pages: 840,344 **Books:** 1,834 (1,893 Volumes) **Journals:** 6 (288 Volumes)

For a related collection of core texts in the disciplines of home economics, see **Home Economics Archive: Research, Tradition and History (HEARTH)** at <http://hearth.library.cornell.edu>

Albert R. Mann Library, 2005. Core Historical Literature of Agriculture (CHLA). Ithaca, NY: Albert R. Mann Library, Cornell University. <http://chla.library.cornell.edu> (Version January 2005).

© 2005 Cornell University Library. Questions? Comments? Please [contact us](#).

[\(Back to top of page\)](#)

The homepage for the Core Historical Literature of Agriculture database is located at <http://chla.library.cornell.edu/>.

The Core Historical Literature of Agriculture (CHLA) is a database of books and journals, mostly from 1850-1950, related to North American agriculture. Still growing, it currently contains over 1,800 English-language monographs and six journal titles. Including both scholarly and popular works, CHLA is a rich source of descriptive

Linda Stewart (E-mail: lgs1@cornell.edu) is Life Sciences Bibliographer and Special Collections Librarian and Joy Paulson (E-mail: jp243@cornell.edu) is Preservation Librarian, both at the Albert R. Mann Library, Cornell University, Ithaca, NY 14853.

information about rural culture and activities during an important historical period.

Authors as diverse as Horace Greeley, Julian Huxley, Oscar Mayer, Cyrus McCormick, and John Muir are represented.

This older information is not useful just to historians. Today, many people pursue farming and crafts as economic sidelines, even hobbies. Practitioners of alternative agriculture, in particular, can mine this lode of pre-World War II techniques, developed before the widespread use of chemical pesticides and fertilizers.

GENESIS OF THE DATABASE

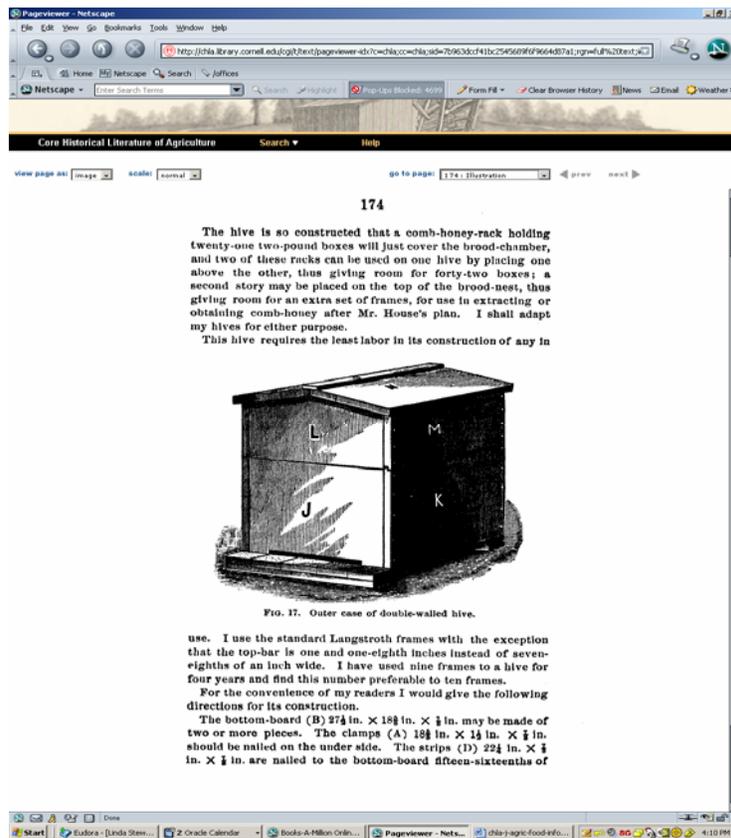
CHLA, which began as a selection project at Cornell University's Albert R. Mann Library, evolved into a preservation project and expanded into a full-text database. During the late 1980s, Wallace Olsen, then at Mann Library, began a Rockefeller Foundation funded project to identify "the most valuable journal and monographic literature for academic teaching and agricultural research." His aim was to create bibliographies in the areas of agricultural economics, agricultural engineering, animal science, crops and crop protection, food science, forestry, human nutrition, rural sociology, and soil science. The monograph titles, gleaned from standard bibliographies and overview publications, were ranked by scholarly reviewers who were familiar with their respective fields and had a historical perspective.

Preservation staff at Mann used the final lists, published as the series *The Literature of the Agricultural Sciences*, edited by Wallace Olsen, as a basis for a microfilming preservation project funded by a New York State Cooperative Preservation

Grant. By 1993, the National Endowment for the Humanities was sufficiently interested in digitization as a prospective preservation medium to fund conversion of many of the volumes to electronic format. Title II-C funding for digitization followed in 1994.

LATEST ADDITIONS TO CHLA

Since 1994, the database has been available in a variety of interfaces and new materials periodically added. The summer of 2004 saw the addition of six journals. Building on CHLA, *The Hive and the Honeybee*, consisting of historic beekeeping books ranked by scholars, was also added in 2004. Digitization of these materials was funded by beekeepers from around the United States.

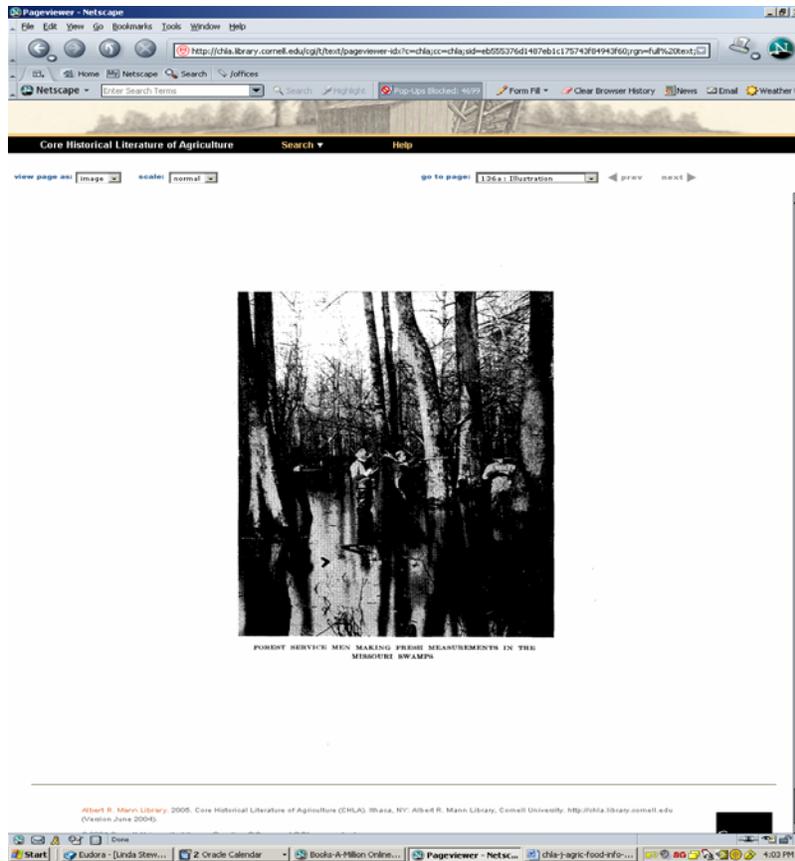


Much space in beekeeping books is devoted to descriptions and drawings of hives, such as this one recommended in Henry Alley's 1883 *Beekeeper's Handy Book*.

Most of the materials in CHLA are in the public domain and were published prior to 1923; however, in a number of cases, permission was obtained from copyright holders to mount later works. Several of the journal titles, such as *Rural Sociology* and *American Journal of Agricultural Economics* are as current as 1989 and 1995, respectively.

DATABASE USE

Since the database is freely available on the Internet, it is receiving considerable use. The most heavily used monographic title in recent months, Andrew Winton's 1932 *Structure and Composition of Foods*, received over 10,000 page views in the last six months of 2004. The runner-up, Selvidge's 1925 *Blacksmithing: a Manual for Use in School and Shop*, received over 8,000 page views. In general, the most heavily used titles are serial runs, large reference sets, such as L.H. Bailey's *Cyclopedia of American Agriculture*, and instructional manuals like the blacksmithing title. The transaction logs show searches for do-it-yourself information, specific facts, general agricultural topics, and historical subjects. All titles in CHLA can be retrieved by Google searches as well as through the database front page and Cornell's regular online catalog, and it appears that a variety of users, from scholars to the general public, have located them.



Gifford Pinchot was America's first professionally trained forester. This photograph from his 1914 *Training of a Forester* shows a forest examiner, on skis, running a compass line.

THE DIGITIZATION PROCESS

Preservation-Quality Scanning

The preservation and digital library communities have agreed that it is important to scan materials at much higher levels of resolution than can currently be displayed on monitors or easily transmitted via the Internet. Scanning materials at high resolutions creates master files rich in detail that can continue to be used as Internet bandwidth expands and high-resolution monitors are developed. This will allow us to avoid having to re-scan rare or fragile items or to have to settle for images of lower quality when the

paper original is no longer available. While there is some room for latitude, at Cornell we have adopted the following standards:

- Master files are in TIFF format, a standardized format that can be migrated over time.
- TIFF files must either be uncompressed or compressed using a “lossless” compression scheme.
- Materials are scanned at the following resolutions:
 - Bitonal: 600 dpi, one-bit
 - Gray-scale: 400 dpi, 8 bit
 - Color: 300 to 400 dpi, 24 bit
- Bound materials are scanned with an overhead digital camera in conjunction with a book cradle.

For the CHLA project, the vast majority of materials were scanned as 600 dpi, one-bit images because they were text. We are in the process of adding a few color images for titles that have high quality color plates. To date, none of the materials in the collection have needed gray-scale scanning.

Enhancing Access: OCR and Structural Metadata

Digital still images, in and of themselves, only enhance access to the information contained in them in one way—they can be accessed at any time via the Internet. However, students and scholars, who are our primary audience, often desire enhanced levels of access to make the collection more useful to them. We achieve this in a number of ways.

First, we use Optical Character Recognition software (OCR) to create full-text that is searchable. In order to keep costs at a reasonable level, we use uncorrected OCR. Correcting OCR or alternatively rekeying the text adds a great deal of cost to a project. We have found that OCR of monographs in English is highly accurate. Serials with a relatively simple page layout (most academic serials fall into this category) also OCR well. In fact, the OCR is good enough for us to make it directly available to the user, so he or she can manipulate the text for research purposes. However, serials with complex page layouts, such as popular journals, with multiple columns and fonts as well as a mix of text, illustrations, and advertisement on a single page, do not OCR as well. While the OCR is accurate enough for searching purposes, the OCR text itself is problematic.

We also create structural metadata in order to facilitate navigation within a title. We use a system developed in-house, which allows us to do the document structuring and quality control of the images simultaneously. For monographs, we mark important structures within the volume, such as title pages, tables of contents, prefaces and introductions, lists of illustrations and maps, bibliographies, and indexes. Unless a monograph is heavily illustrated, we can structure and inspect approximately four volumes per hour, and student assistants are usually able to perform this work.

THE APPLE-TREE

I

WHERE THERE IS NO APPLE-TREE

The wind is snapping in the bamboos, knocking together the resonant canes and weaving the myriad flexible wreaths above them. The palm heads rustle with a brisk crinkling music. Great ferns stand in the edge of the forest, and giant arums cling their arms about the trunks of trees and rear their dim jacks-in-the-pulpit far in the branches; and in the greater distance I know that green parrots are flying in twos from tree to tree. The plant forms are strange and various, making mosaic of contrasting range of leaf-size and leaf-shape, palm and grass and fern, epiphyte and liana and clumpy mistletoe, of grace and clumsiness and even misproportion, a tall thick landscape all mingled into a symmetry of disorder that charms the attention and fascinates the eye.

It is a soft and delicious air wherein I sit. A torrid drowse is in the receding landscape. The people move leisurely, as befits the world where there is no preparation for frost and no urgent need of laborious apparel. There are tardy bullock-carts, unconscious donkeys, and men pushing vehicles. There are odd products and unaccustomed cakes and cookies on little stands by the roadside, where the turbaned vendor sits on the ground unconcernedly.

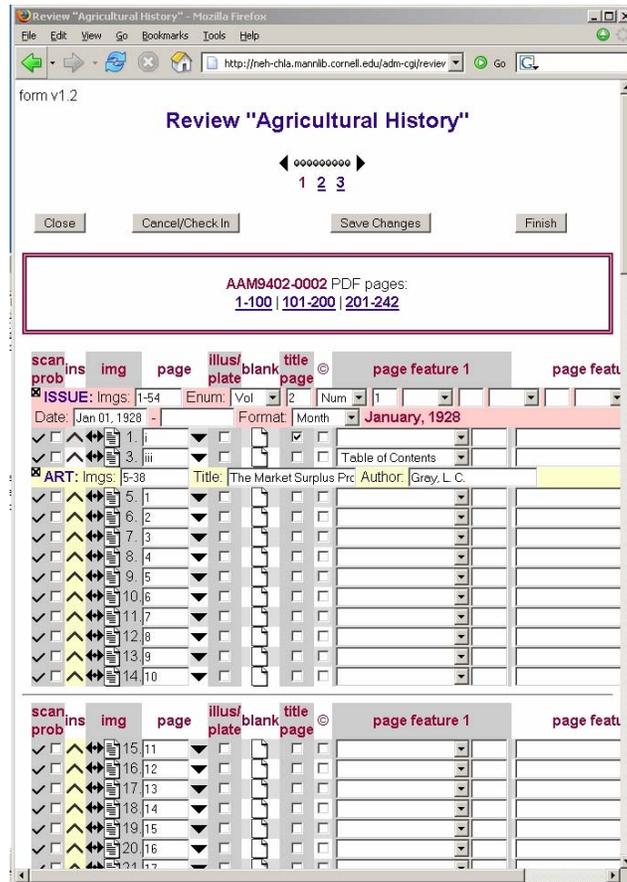
7

Sample of OCR'd text from a monograph.

Enhancing Access: Serial Literature

Creating access to serial literature at the article level, which is desired by scholars and students, is a much more complex and time-consuming process. Our structuring tool was designed to work with serials as well as monographs. For each issue we identify the volumes and/or issue number, the date of publication, and the authors, titles, and page ranges of every article in the volume, as well as those structures identified in

monographs. The structuring process for serials requires substantially more time than for monographs.



Example from serials structuring tool.

Depending on the number of pages and articles in a volume, it takes an average of one to four hours to structure a volume. For volumes with very large number of pages and many articles and complex page layouts, such as popular journals, it can take even longer. We are rethinking this issue with the popular journals, due to the high cost and length of time needed for completion.

The structural metadata created for serials has other uses besides navigation. It allows us to create lists of authors of articles, which can be browsed. The keying-in of article level information also allows users to be able to search author names, article titles, and keywords from the titles.

Enhancing Access: Browsing Features and Google

The OCR and structural metadata we create also allows us to create several browse features that enhance access to the materials in the collection. Monographs can be browsed by author or title. Monographs can also be browsed by date range. Journals can be browsed by journal titles and then volume or year.

The Core Historical Literature of Agriculture
Albert R. Mann Library, Cornell University

Agricultural history
Agricultural History Society, Fargo, N.D.

Title from journal information screen (viewed June 2004).

- [1927](#) (v.1 n.1 - v.1 n.2)
- [1928](#) (v.2 n.1 - v.2 n.4)
- [1929](#) (v.3 n.1 - v.3 n.4)
- [1930](#) (v.4 n.1 - v.4 ind.)
- [1931](#) (v.5 n.1 - v.5 ind.)
- [1932](#) (v.6 n.1 - v.6 ind.)
- [1933](#) (v.7 n.1 - v.7 ind.)
- [1934](#) (v.8 n.1 - v.8 ind.)
- [1935](#) (v.9 n.1 - v.9 ind.)
- [1936](#) (v.10 n.1 - v.10 ind.)
- [1937](#) (v.11 n.1 - v.11 ind.)
- [1938](#) (v.12 n.1 - v.12 ind.)
- [1939](#) (v.13 n.1 - v.13 ind.)
- [1940](#) (v.14 n.1 - v.14 ind.)
- [1941](#) (v.15 n.1 - v.15 ind.)
- [1942](#) (v.16 n.1 - v.16 ind.)
- [1943](#) (v.17 n.1 - v.17 ind.)
- [1944](#) (v.18 n.1 - v.18 ind.)
- [1945](#) (v.19 n.1 - v.19 ind.)
- [1946](#) (v.20 n.1 - v.20 ind.)
- [1947](#) (v.21 n.1 - v.21 ind.)
- [1948](#) (v.22 n.1 - v.22 ind.)
- [1949](#) (v.23 n.1 - v.23 ind.)
- [1950](#) (v.24 n.1 - v.24 suppl.)
- [1951](#) (v.25 n.1 - v.25 ind.)
- [1952](#) (v.26 n.1 - v.26 ind.)
- [1953](#) (v.27 n.1 - v.27 ind.)
- [1954](#) (v.28 n.1 - v.28 ind.)
- [1955](#) (v.29 n.1 - v.29 ind.)
- [1956](#) (v.30 n.1 - v.30 ind.)
- [1957](#) (v.31 n.1 - v.31 ind.)
- [1958](#) (v.32 n.1 - v.32 ind.)
- [1959](#) (v.33 n.1 - v.33 ind.)
- [1960](#) (v.34 n.1 - v.34 ind.)
- [1961](#) (v.35 n.1 - v.35 ind.)
- [1962](#) (v.36 n.1 - v.36 ind.)
- [1963](#) (v.37 n.1 - v.37 ind.)
- [1964](#) (v.38 n.1 - v.38 ind.)
- [1965](#) (v.39 n.1 - v.39 ind.)

1927

- [January, 1927](#) Volume 1, Number 1 (20 pages)
- [July, 1927](#) Volume 1, Number 2 (31 pages)

1928

- [January, 1928](#) Volume 2, Number 1 (51 pages)
- [April, 1928](#) Volume 2, Number 2 (65 pages)
- [July, 1928](#) Volume 2, Number 3 (50 pages)
- [October, 1928](#) Volume 2, Number 4 (59 pages)

1929

- [January, 1929](#) Volume 3, Number 1 (53 pages)
- [April, 1929](#) Volume 3, Number 2 (68 pages)
- [July, 1929](#) Volume 3, Number 3 (32 pages)
- [October, 1929](#) Volume 3, Number 4 (62 pages)

Access page for individual issues of the journal *Agricultural History*.

It is also possible to browse authors of articles.

V

<ul style="list-style-type: none"> • Vaidya, Kirit G. (1) • Vaile, Roland S. (1) • Vaile, Ronald S. (1) • Valdes, Alberto (2) • Valdes E., Alberto (1) • Valentine, J. M. (1) • Valpey, John (1) • Van Arsdall, Roy N. (1) • Van Duyn, Carl (1) • Van Duzee, Edward P. (6) • Van Duzee, M. C. (6) • Van Duzee, E. P. (1) • Van Dyke, Edwin C. (2) • van Kooten, G. C. (3) • Van Schilfgaarde, Jan (1) • Van Vliet, Hadley (1) • VanArsdall, R. Thomas (1) • Vance, A. M. (1) • Vance, Arlo M. (2) • Vance, Lawrence L. (1) • Vance, Rupert B. (1) • Vandeman, Ann (1) • Vandenborre, R. J. (1) • Vandenborre, Roger J. (2) • Vandeputte, J. M. (1) • Vanderleck, J. (1) • Vandeveer, Monte L. (1) • Vandiver, Joseph S. (1) • VanLandingham, Eugenia P. (1) • Vanneman, Reeve D. (1) • VanSickle, John J. (2) • Vanthana, C. (1) • Vantreese, Valerie L. (1) • Vanvig, Andrew (2) • Vanzetti, David (1) • Variyam, Jayachandran N. (2) • Varley, A. P. (1) • Varner, D. B. (1) • Vary, Karl A. (1) • Vasavada, Uptal (1) • Vasavada, Utpal (4) • Vasey, Tom (4) • Vasievich, Josphe M. (1) • Vass, A. F. (2) • Vatter, Harold G. (1) • Vaughan, L. M. (3) • Vaughan, Theo L. (1) • Vaughn, Gerald F. (1) • Vaux, H. J., Jr. (1) • Vaux, Henry J. (1) • Vaux, Henry J., Jr. (2) • Vazquez-Calcerrada, P. B. (1) • Vazquez-Calcerrada, Pablo B. (1) • Veeman, Michele M. (2) • Veeman, Terence S. (1) • Veevers, J. E. (2) • Vellianitis-Fidas, Amalia (1) 	<ul style="list-style-type: none"> • VenKataraman, Ravi (1) • Vera, Heman (1) • Veracammen, James (1) • Vercammen, James (2) • Vergara, Jose (1) • Verleger, Philip K. (1) • Vermeer, James (2) • Verstegen, Jos A. A. M. (1) • Vertrees, James G. (1) • Vickery, Edward (1) • Vickery, Raymond E. (2) • Vickery, Robert K. (1) • Viereck, H. L. (2) • Viereck, Henry L. (8) • Vilet, H. Van (1) • Villa-Issa, Manuel R. (2) • Villarreal, Fernando J. Gonzalez (2) • Vincent, Clark E. (1) • Vine, Margaret Wilson (1) • Viner, Jacob (1) • Vining, Rutledge (1) • Viscusi, W. Kip (1) • Vissa, Anuradha (1) • Vitaliano, Peter (1) • Vith, Albert (1) • Vivino, A. Earl (1) • Vlasin, Raymond D. (2) • Vliet, H. Van (1) • Vo, Trang T. (1) • Vocke, Gary (1) • Voelker, Stanley W. (3) • Voet, Hillary (1) • Vogel, Harold A. (1) • Vogel, Robert C. (1) • Vogelsang, Donald L. (1) • Vogt, Evon Z., Jr. (1) • Vogt, George B. (2) • Vogt, Paul L. (1) • Voland, Maurice E. (3) • Volin, Lazar (7) • von Ciriacy-Wantrup, Siegfried (2) • Von Frauedorfer, Sigmund (1) • Von Frauedorfer, Sigmund (1) • Von Pischke, J. D. (1) • Von Tungeln, George H. (1) • vonOppen, Matthias (1) • Voon, Jan P. (1) • Voon, Thomas J. (1) • Voris, Ralph (2) • Voss, Paul R. (2) • Vosti, Stephen A. (1) • Voth, Donald E. (5) • Vries, Egbert D. (1) • Vrooman, C. W. (2) • Vroomen, Harry (1) • Vucinich, Alexander (1) • Vukina, Tomislav (1)
---	--

Access page for article authors whose names begin with V.

While this feature allows a user to browse through an alphabetical list of all the authors of the articles in the collection, it does not allow one to browse just the article authors of a specific journal title.

FUTURE OF THE PROJECT

Only about half the titles we have identified as important for preservation have been digitized. We will continue to add titles, as funding is available. With internal funding, we will slowly add further serials and monographs. If outside funding becomes available, we will undertake a larger project. Our hope is to create a definitive agricultural archive, preserving and providing access to the most significant agricultural texts of the 1850-1950 period.