AgNIC Pre-conference 2009

"If It's Digital and in Google –
Then They Will Come"

Presented at the National Agricultural Library

By

Dr. Rob McGeachin

Texas A&M University Libraries

April 21, 2009

Digitization Basics

Overview

- What is digital preservation
- Scanners
- Digitization standards
- Planning a project
- Mechanics / Demo of scanning

Why Digitize ?

- Make an archival preserved copy
- Make material accessible to the world
- More usage of materials
- Make Metadata Records Harvestable to Other Indexing and Search Sites

Desire a More Accessible Preserved Format than Microfilm

- TAES Bulletin started in 1888 so early volumes are becoming very brittle
- TAES Bulletin was first preserved by 1980-82 microfilming project
- Microfilm in only a few library locations
- Customers don't like to use microfilm!

Digitization Concepts

- Digital image – representation of a two dimensional image in digital bits of ones and zeroes
- AKA Raster image -- **raster graphics** image or **bitmap**, is a data structure representing a rectangular grid of pixels

Digitization Concepts – Image Quality

- Resolution
  - Resolution in dots per inch (dpi)
  - The higher the resolution the more dpi
  - Examples of commonly used resolutions: 200, 300, 400, 600 dpi
  - Film photograph or printed page resolution is up to 30,000 dpi

Digitization Concepts -- Image Quality

- Bit depth or color depth
  - number of bits used to represent the color of a single pixel in a bitmapped image
  - 1 bit monochrome (bi-tonal black and white)
  - 8 bit greyscale
  - 24 bit color

Digitization Concepts -- Image Quality

- Uncompressed raw image files
  - Bitmap (.bmp)
  - Tagged Image File Format (TIFF) (.tif)

Digitization Concepts -- Image Quality

- Compressed derivative files
  - Used for easier transmission over the web due to smaller size
  - Joint Photographic Experts Group (JPEG) (.jpg)
  - Portable Network Graphics (.png)
  - Graphics Interchange Format (.gif)
  - Portable Document Format (.pdf)

Types of Scanners

- Flat bed (may have sheet feeder)
- Book
  - Book edge
  - Book cradle
- Microform
  - Microfilm
  - Microfiche

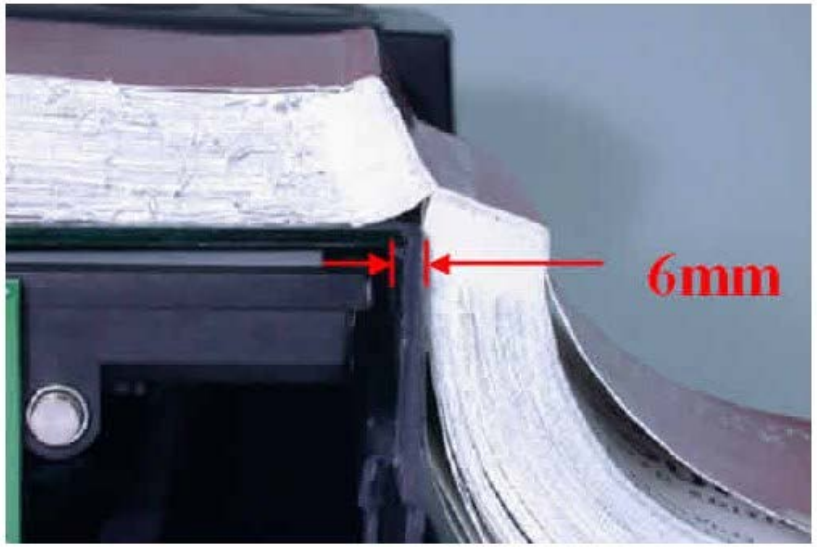
Flatbed Scanner ($500 to $10K)

Flatbed Scanner with Document Feeder

Book Edge Scanners

Book Edge Scanners ($300 to $5K)

Book Edge Scanner



Book Cradle Scanner ($6K to $25K)

Book Cradle Scanner with page turning ($35K)

Book Cradle Scanner with page turning - Kirtas' APT BookScan 1200 ($500K)

Project (Collection) Design

- Have a well thought out plan before doing anything!
- However realize you will still need to be flexible once operational

Example Project: Design of Texas A&M University Libraries' Digitization of the Bulletin of the Texas Agricultural Experiment Station

Workforce Requirements

- 1 Librarian – 20% time
- 2 Student Workers – ½ time at $8.00 per hour

- Digital Repository minimum
  - 2 Librarians (head and metadata cataloger)
  - 2 Programmers

Project Design – Equipment and Software

- Computer Workstation
- 500 GB External Hard Drive
- Plustek OpticBook 3600 Book Edge Scanner and Software
- Adobe Acrobat Ver. 7.0
- D-Space Software
- Lots of IR File Storage – about 600 GB

Project Design – TIFF Image Files

- One image file per page, including blank pages
- Scan at 400 dpi for text pages
- Scan at 600 dpi for illustration pages
- Scan as grayscale unless color illustration
- Text pages = 6 MB, illustrations = 14 MB

Project Design – File Naming

- Name folders by Publication and Publication Year
- Ex: TAES Bulletin
  - 1888
  - 1889

- Filename: Bulletin Number, Page Number
- Ex: b0001 0001.tif …. b0001 0004.tif;

- Filename: Bulletin Number
- ex: Bull0001.pdf …. Bull1400.pdf

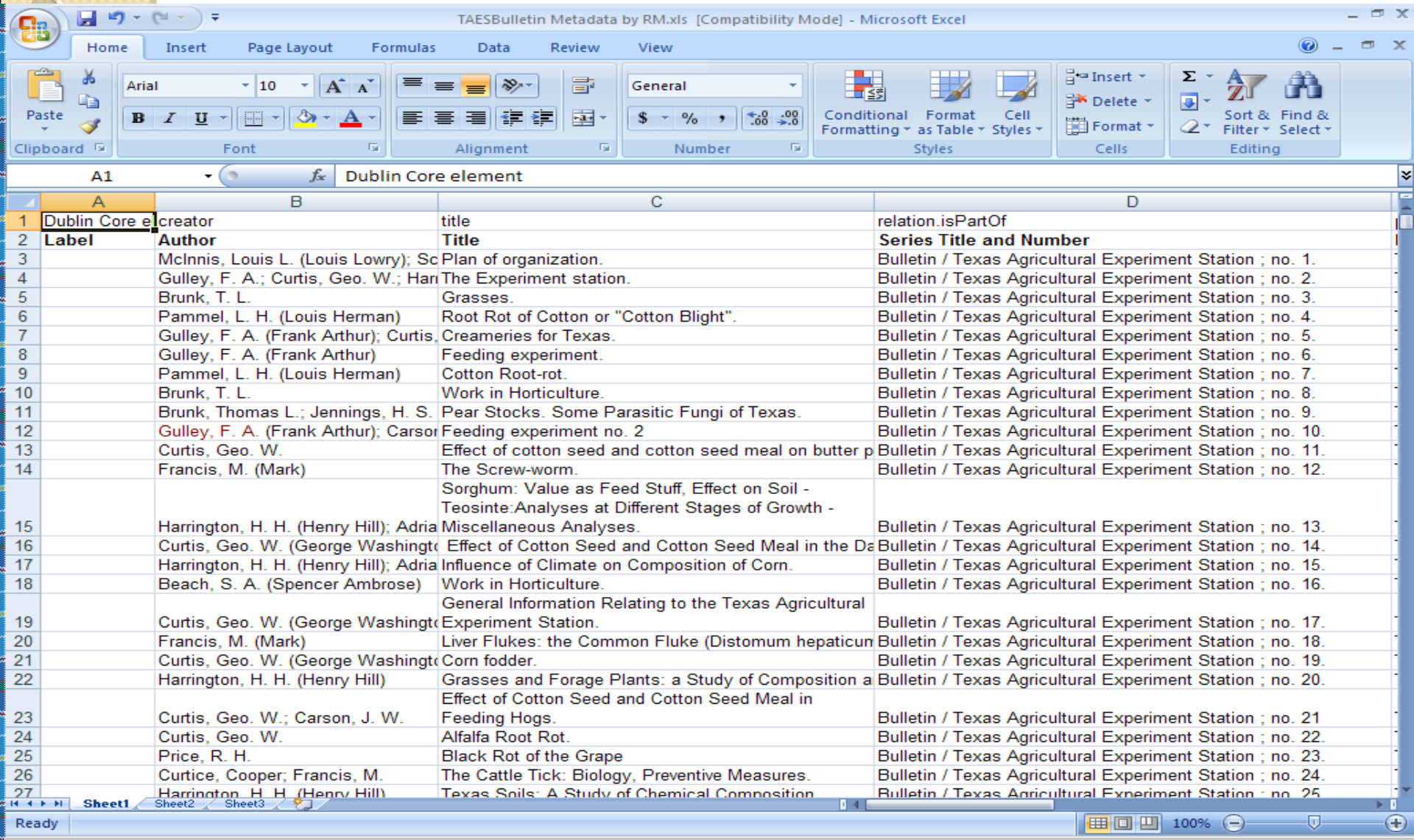
Project Design – PDF Files

- Add each page image file to document
- Perform OCR on all pages to generate text
- Rotate illustration & chart pages 90 degrees
- Save as fully searchable PDF file
- Full Bulletin text searchable in Repository

Project Design -- Metadata

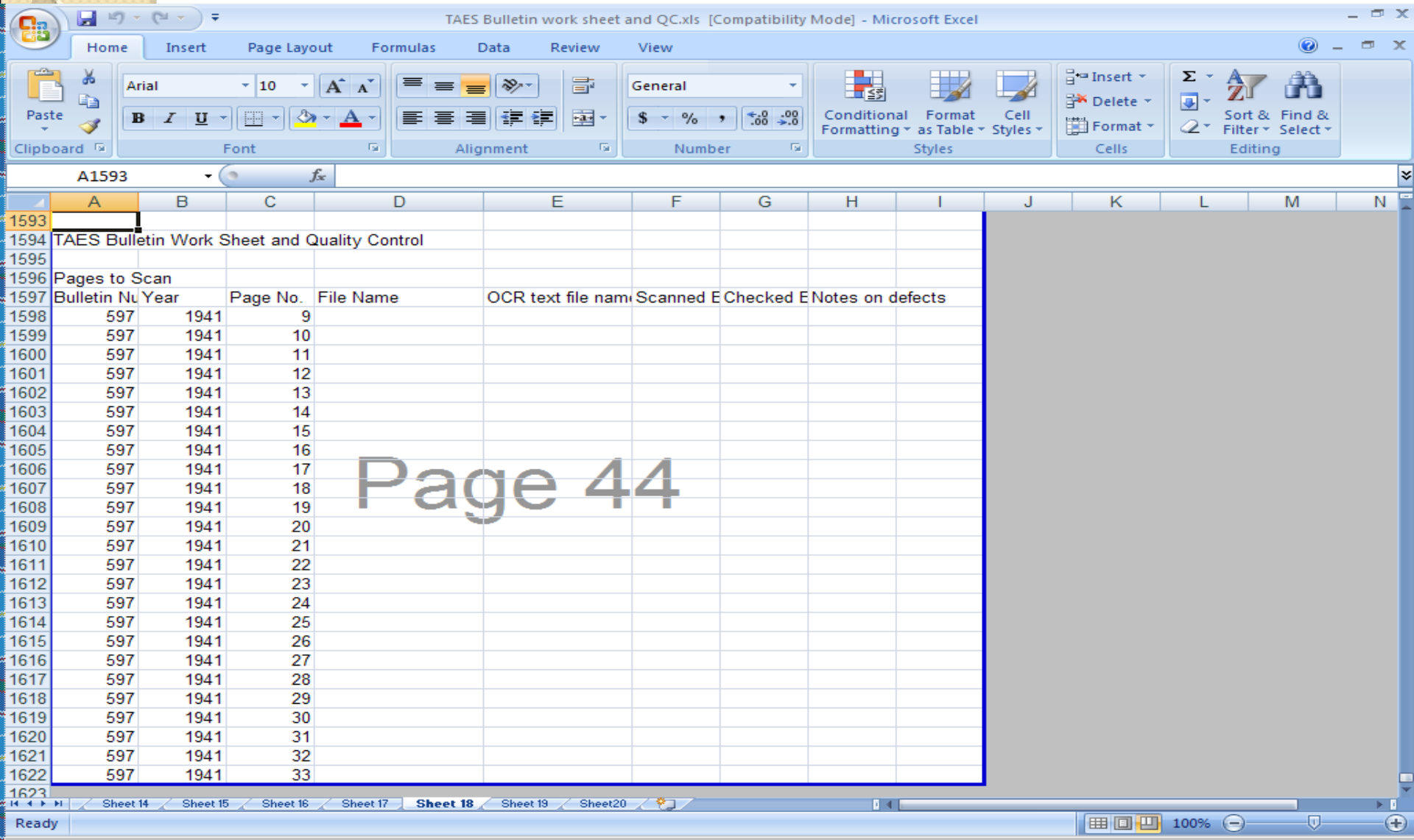
- NAL modified Dublin Core
- Include author, title, keywords, date, series
- Include NAL Thesaurus Subject Terms
- Initial Records in Excel Spreadsheet
- Entry part of DSpace workflow
- Metadata OAI harvestable

Spreadsheet of Metadata

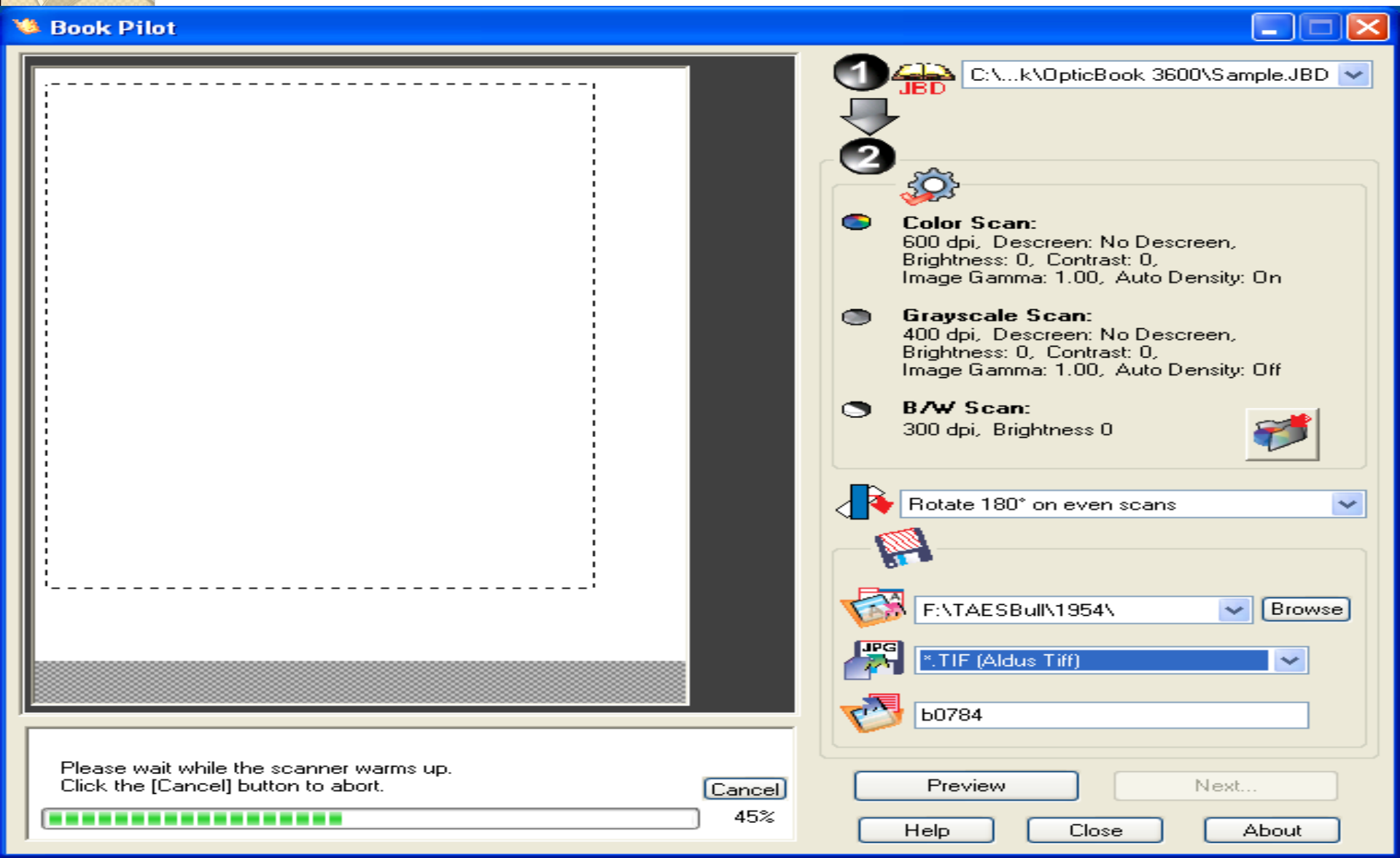


	A	B	C	D
1	Dublin Core el	creator	title	relation.isPartOf
2	**Label**	**Author**	**Title**	**Series Title and Number**
3		McInnis, Louis L. (Louis Lowry); Sc	Plan of organization.	Bulletin / Texas Agricultural Experiment Station ; no. 1.
4		Gulley, F. A.; Curtis, Geo. W.; Har	The Experiment station.	Bulletin / Texas Agricultural Experiment Station ; no. 2.
5		Brunk, T. L.	Grasses.	Bulletin / Texas Agricultural Experiment Station ; no. 3.
6		Pammel, L. H. (Louis Herman)	Root Rot of Cotton or "Cotton Blight".	Bulletin / Texas Agricultural Experiment Station ; no. 4.
7		Gulley, F. A. (Frank Arthur); Curtis,	Creameries for Texas.	Bulletin / Texas Agricultural Experiment Station ; no. 5.
8		Gulley, F. A. (Frank Arthur)	Feeding experiment.	Bulletin / Texas Agricultural Experiment Station ; no. 6.
9		Pammel, L. H. (Louis Herman)	Cotton Root-rot.	Bulletin / Texas Agricultural Experiment Station ; no. 7.
10		Brunk, T. L.	Work in Horticulture.	Bulletin / Texas Agricultural Experiment Station ; no. 8.
11		Brunk, Thomas L.; Jennings, H. S.	Pear Stocks. Some Parasitic Fungi of Texas.	Bulletin / Texas Agricultural Experiment Station ; no. 9.
12		Gulley, F. A. (Frank Arthur); Carson	Feeding experiment no. 2	Bulletin / Texas Agricultural Experiment Station ; no. 10.
13		Curtis, Geo. W.	Effect of cotton seed and cotton seed meal on butter p	Bulletin / Texas Agricultural Experiment Station ; no. 11.
14		Francis, M. (Mark)	The Screw-worm.	Bulletin / Texas Agricultural Experiment Station ; no. 12.
15		Harrington, H. H. (Henry Hill); Adria	Sorghum: Value as Feed Stuff, Effect on Soil - Teosinte:Analyses at Different Stages of Growth - Miscellaneous Analyses.	Bulletin / Texas Agricultural Experiment Station ; no. 13.
16		Curtis, Geo. W. (George Washingto	Effect of Cotton Seed and Cotton Seed Meal in the Da	Bulletin / Texas Agricultural Experiment Station ; no. 14.
17		Harrington, H. H. (Henry Hill); Adria	Influence of Climate on Composition of Corn.	Bulletin / Texas Agricultural Experiment Station ; no. 15.
18		Beach, S. A. (Spencer Ambrose)	Work in Horticulture.	Bulletin / Texas Agricultural Experiment Station ; no. 16.
19		Curtis, Geo. W. (George Washingto	General Information Relating to the Texas Agricultural Experiment Station.	Bulletin / Texas Agricultural Experiment Station ; no. 17.
20		Francis, M. (Mark)	Liver Flukes: the Common Fluke (Distomum hepaticun	Bulletin / Texas Agricultural Experiment Station ; no. 18.
21		Curtis, Geo. W. (George Washingto	Corn fodder.	Bulletin / Texas Agricultural Experiment Station ; no. 19.
22		Harrington, H. H. (Henry Hill)	Grasses and Forage Plants: a Study of Composition a	Bulletin / Texas Agricultural Experiment Station ; no. 20.
23		Curtis, Geo. W.; Carson, J. W.	Effect of Cotton Seed and Cotton Seed Meal in Feeding Hogs.	Bulletin / Texas Agricultural Experiment Station ; no. 21
24		Curtis, Geo. W.	Alfalfa Root Rot.	Bulletin / Texas Agricultural Experiment Station ; no. 22.
25		Price, R. H.	Black Rot of the Grape	Bulletin / Texas Agricultural Experiment Station ; no. 23.
26		Curtice, Cooper; Francis, M.	The Cattle Tick: Biology, Preventive Measures.	Bulletin / Texas Agricultural Experiment Station ; no. 24.
27		Harrington, H. H. (Henry Hill)	Texas Soils: A Study of Chemical Composition	Bulletin / Texas Agricultural Experiment Station ; no. 25

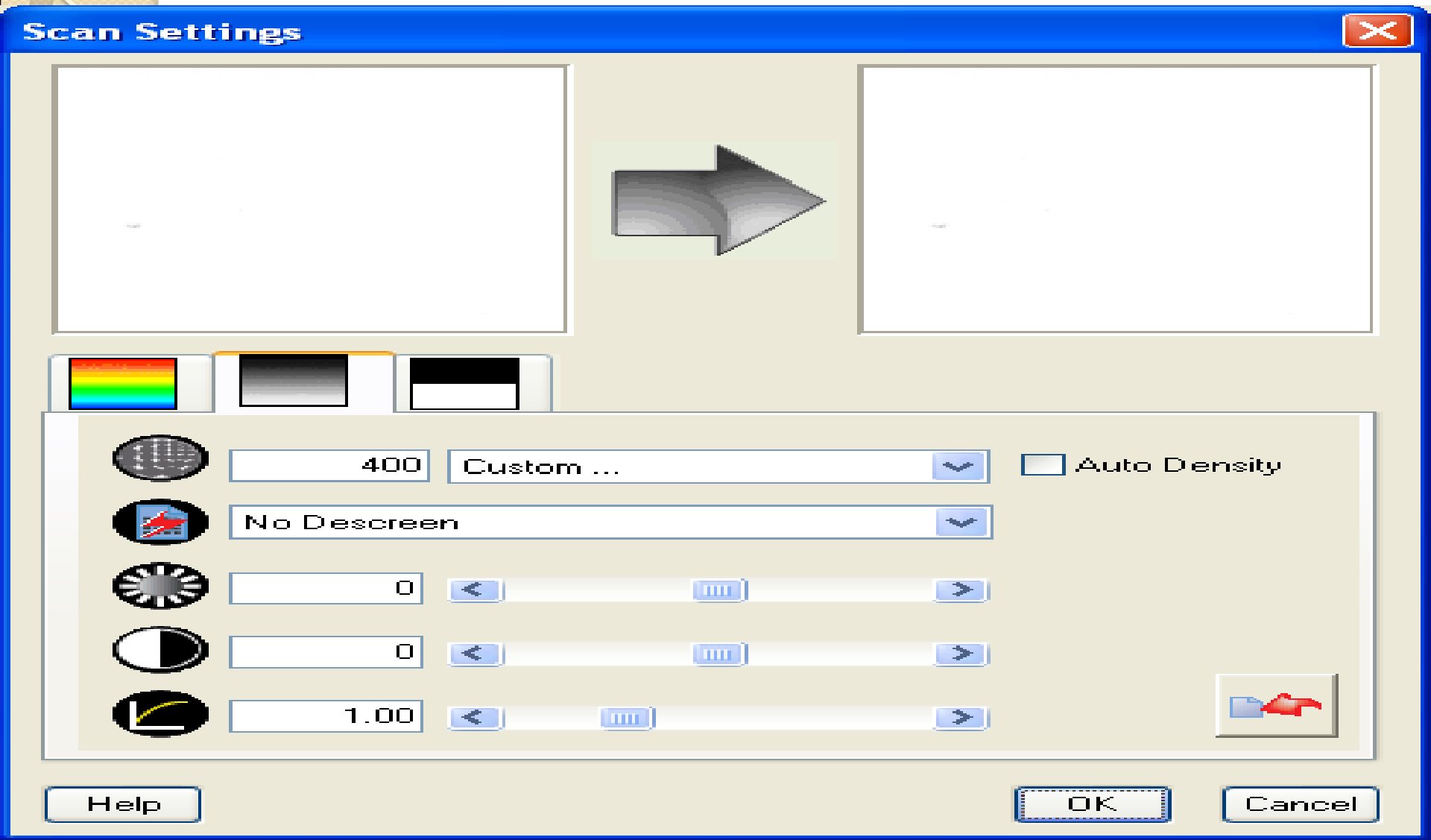
Digitization Workflow Spreadsheet



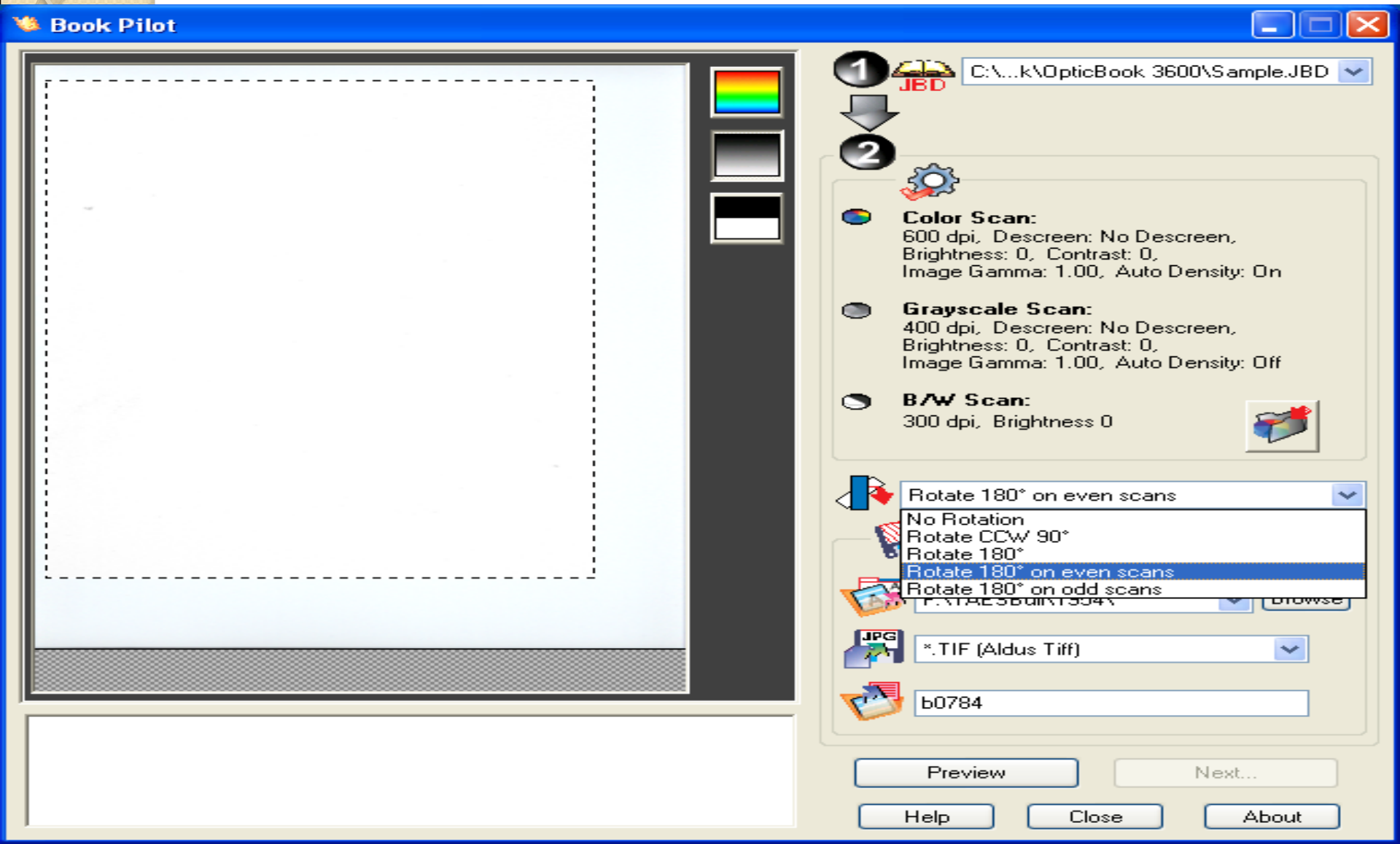
Scanner Setup



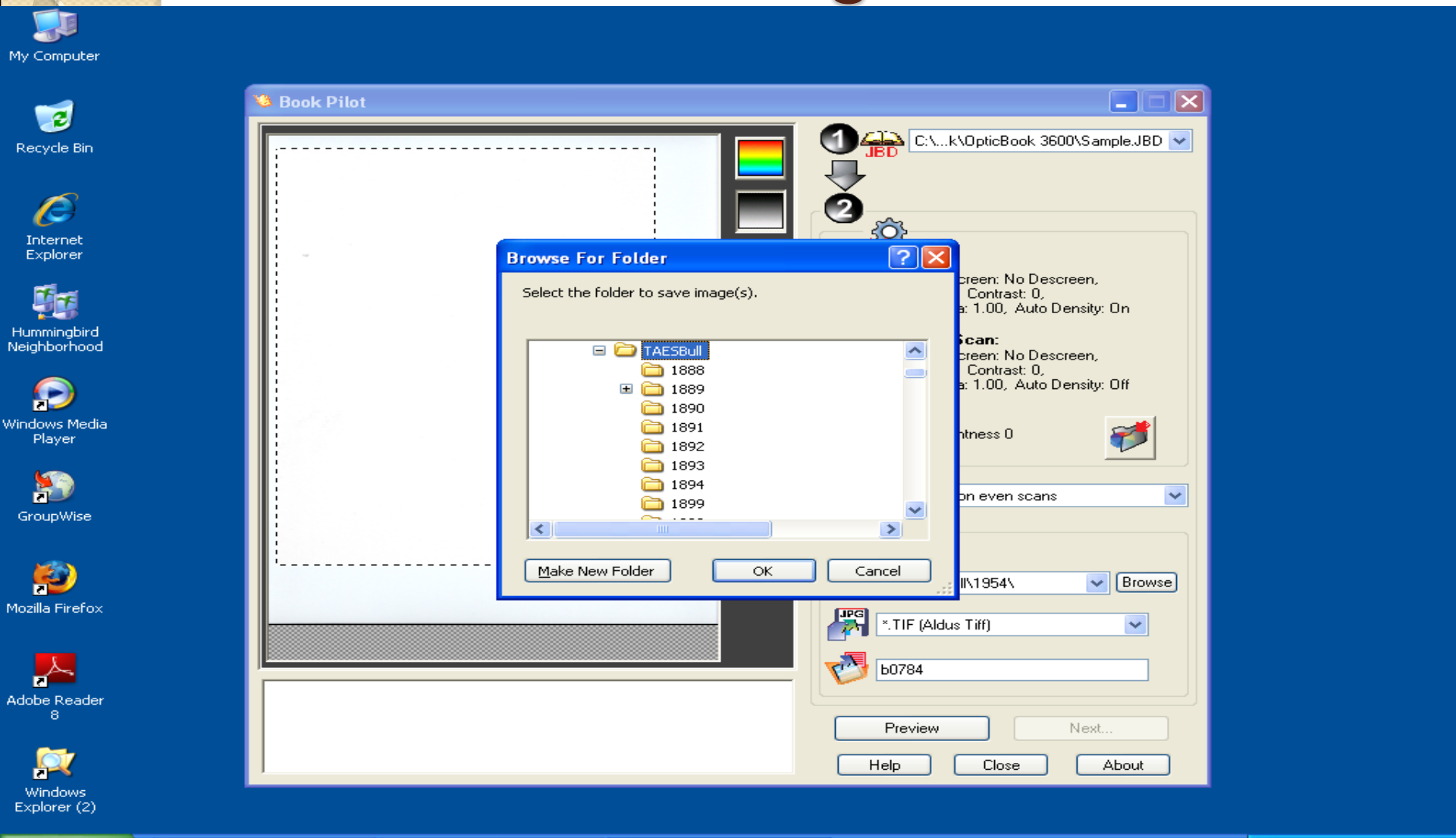
Scan Settings – Grayscale



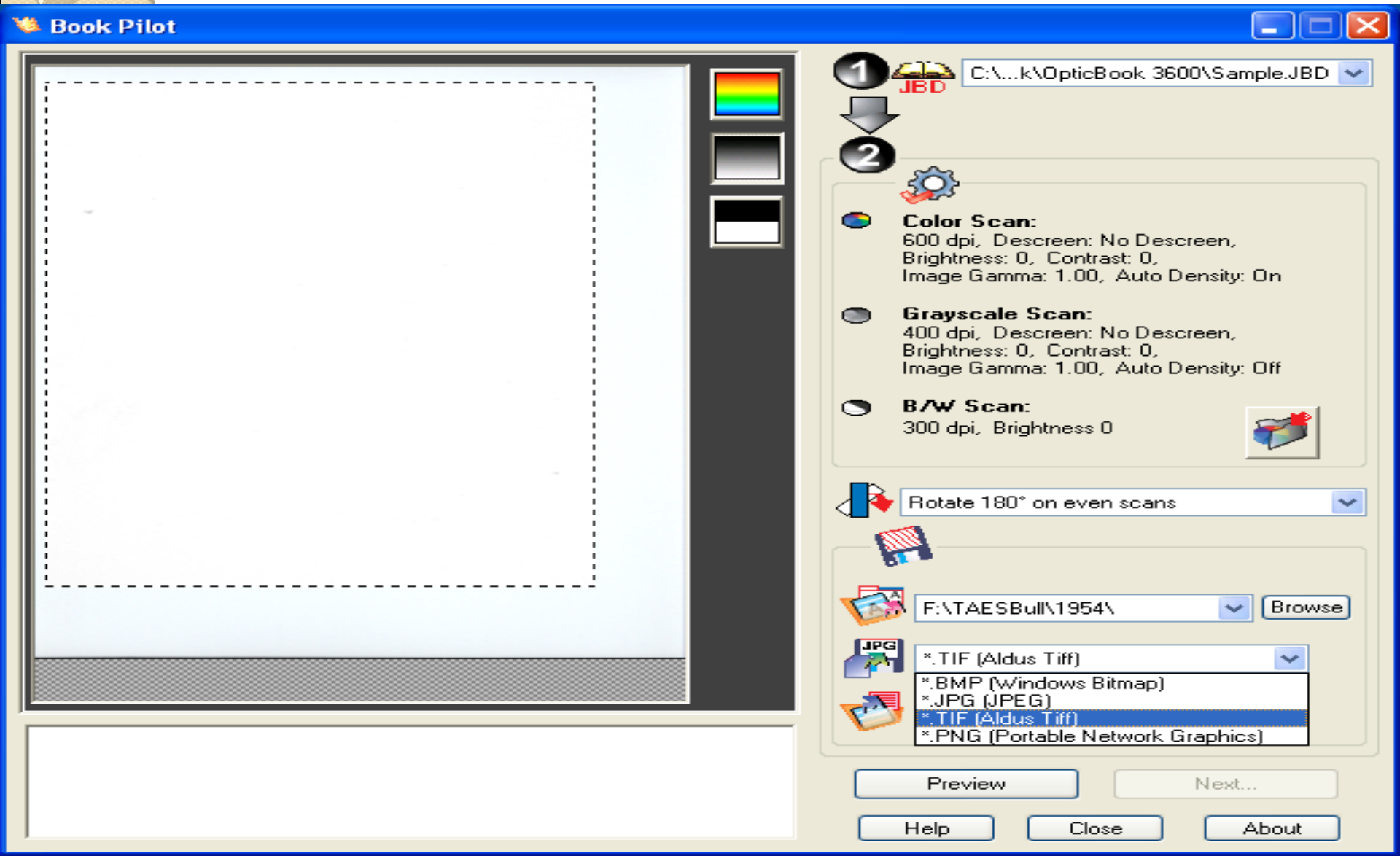
Rotate 180 on Even Scans



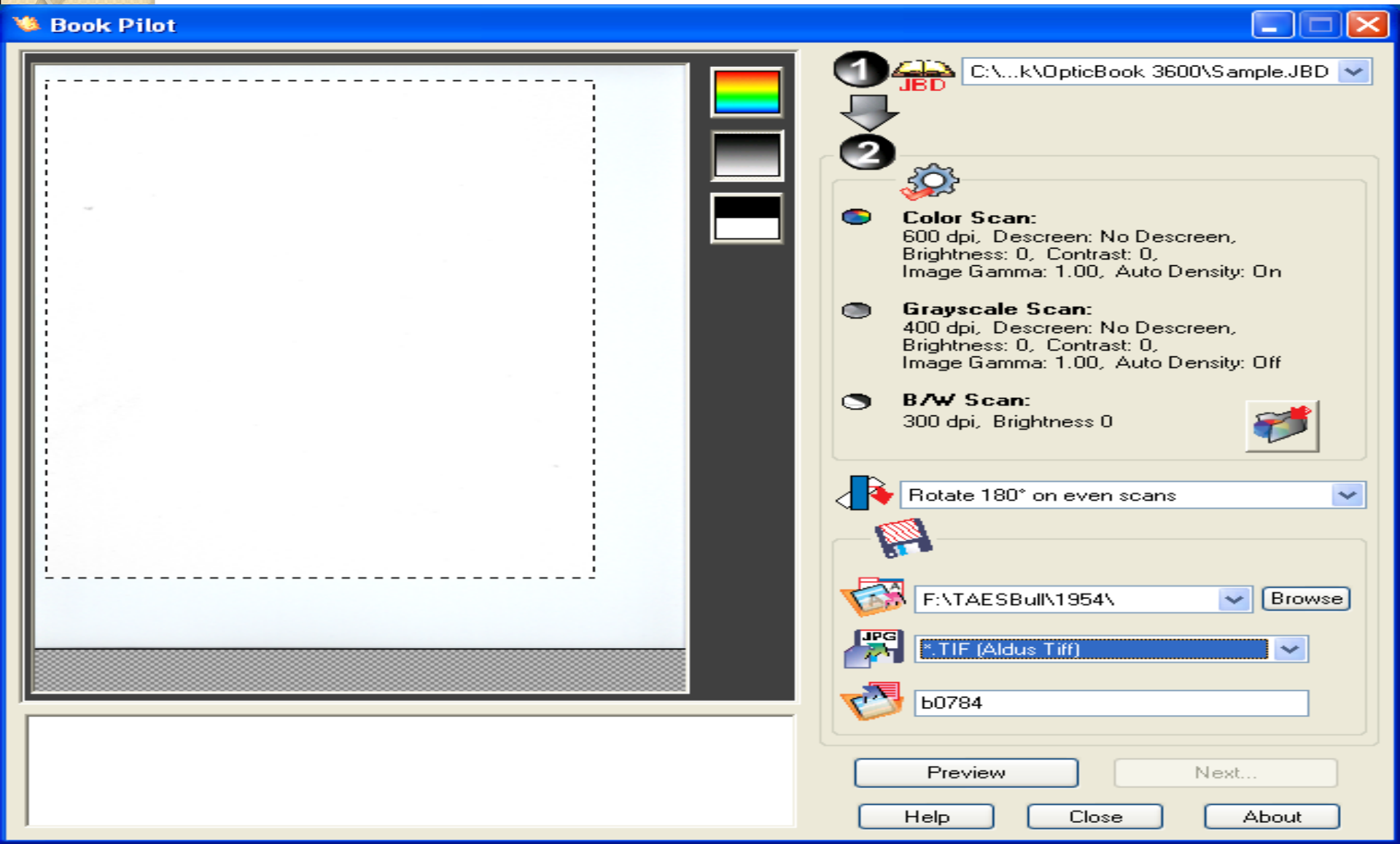
Select File Folder Path to Save Images



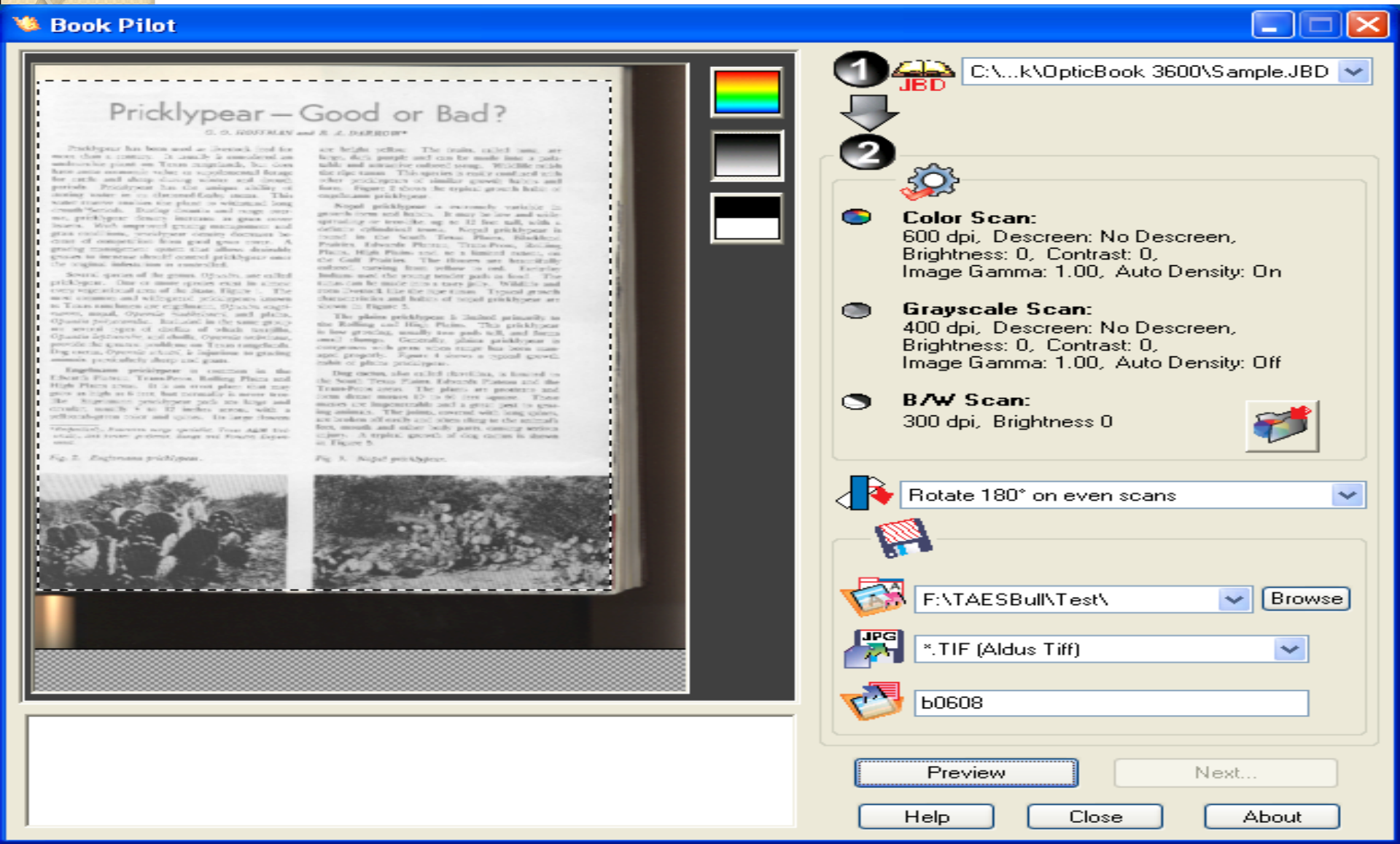
Select File Type -- TIF



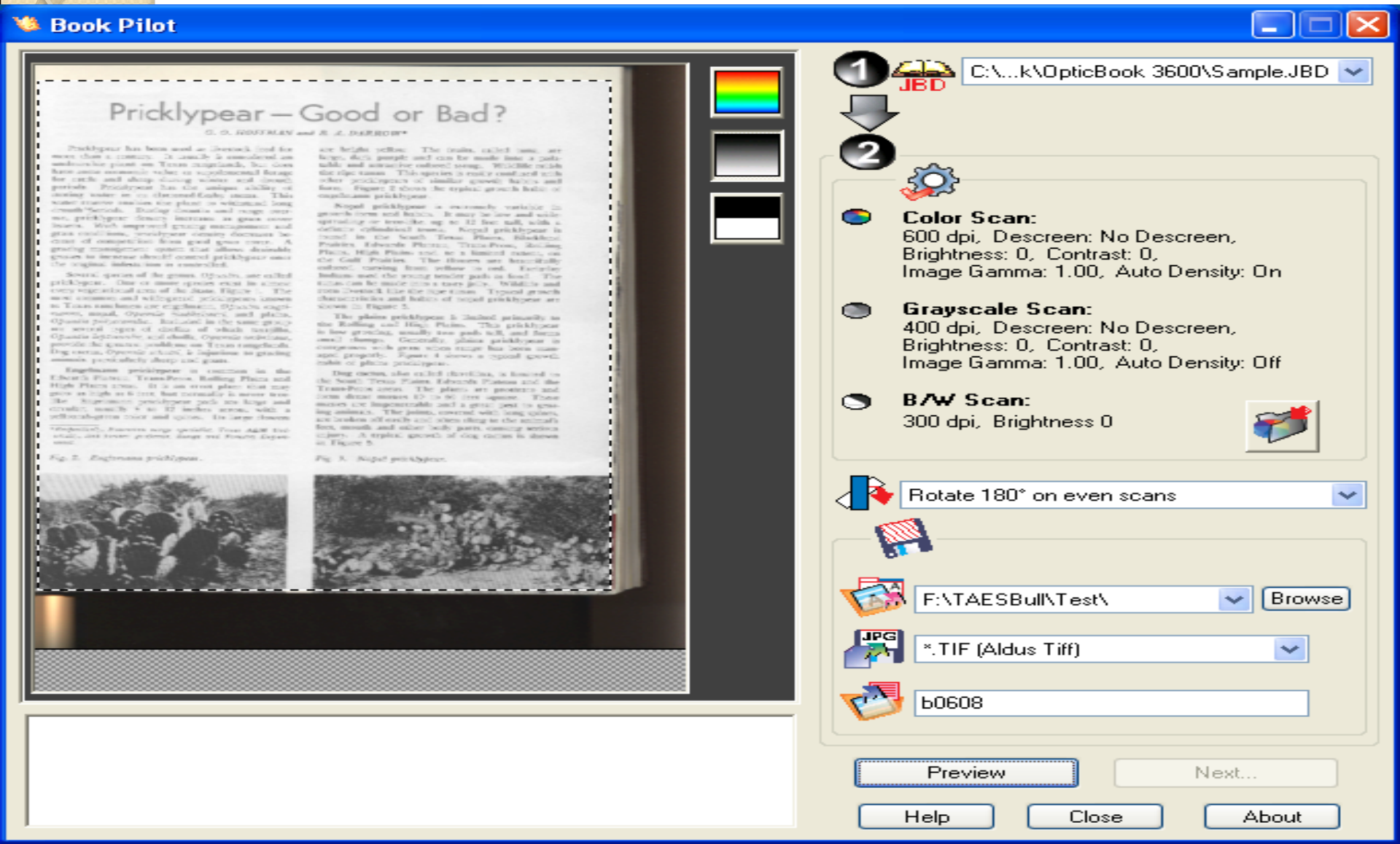
Enter Base File Name– b#####



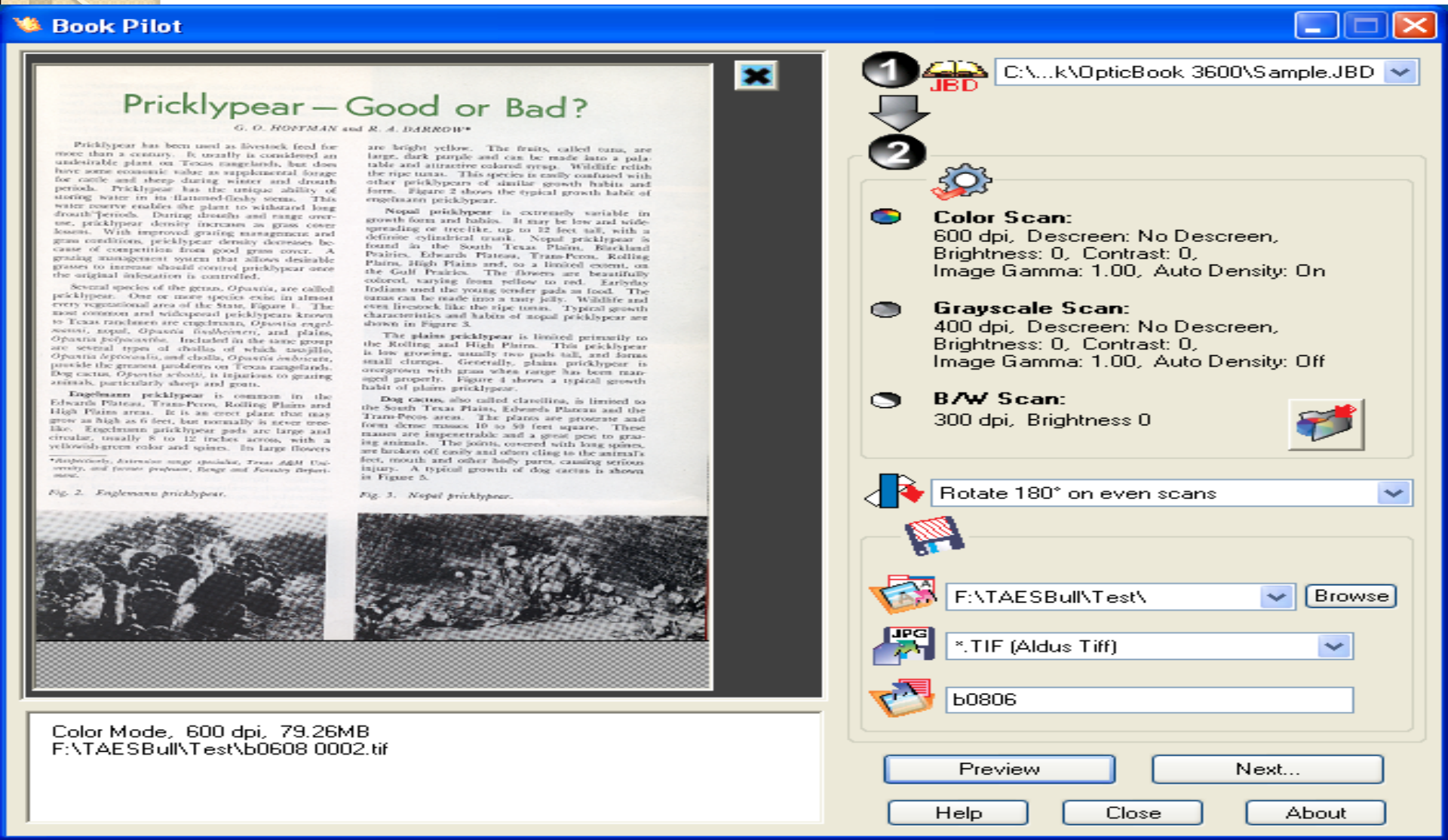
Preview Scan



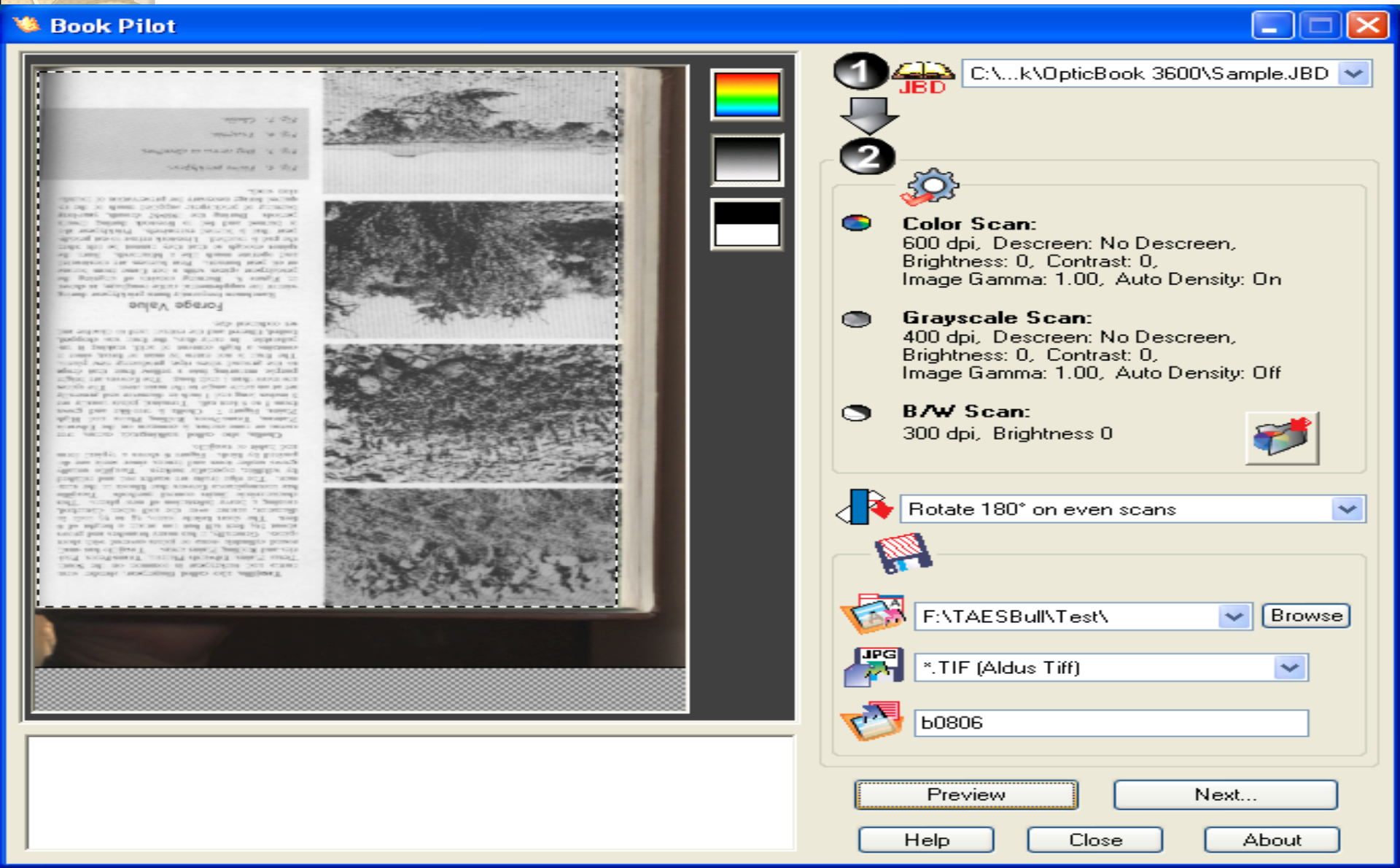
Adjust Scanning Margins



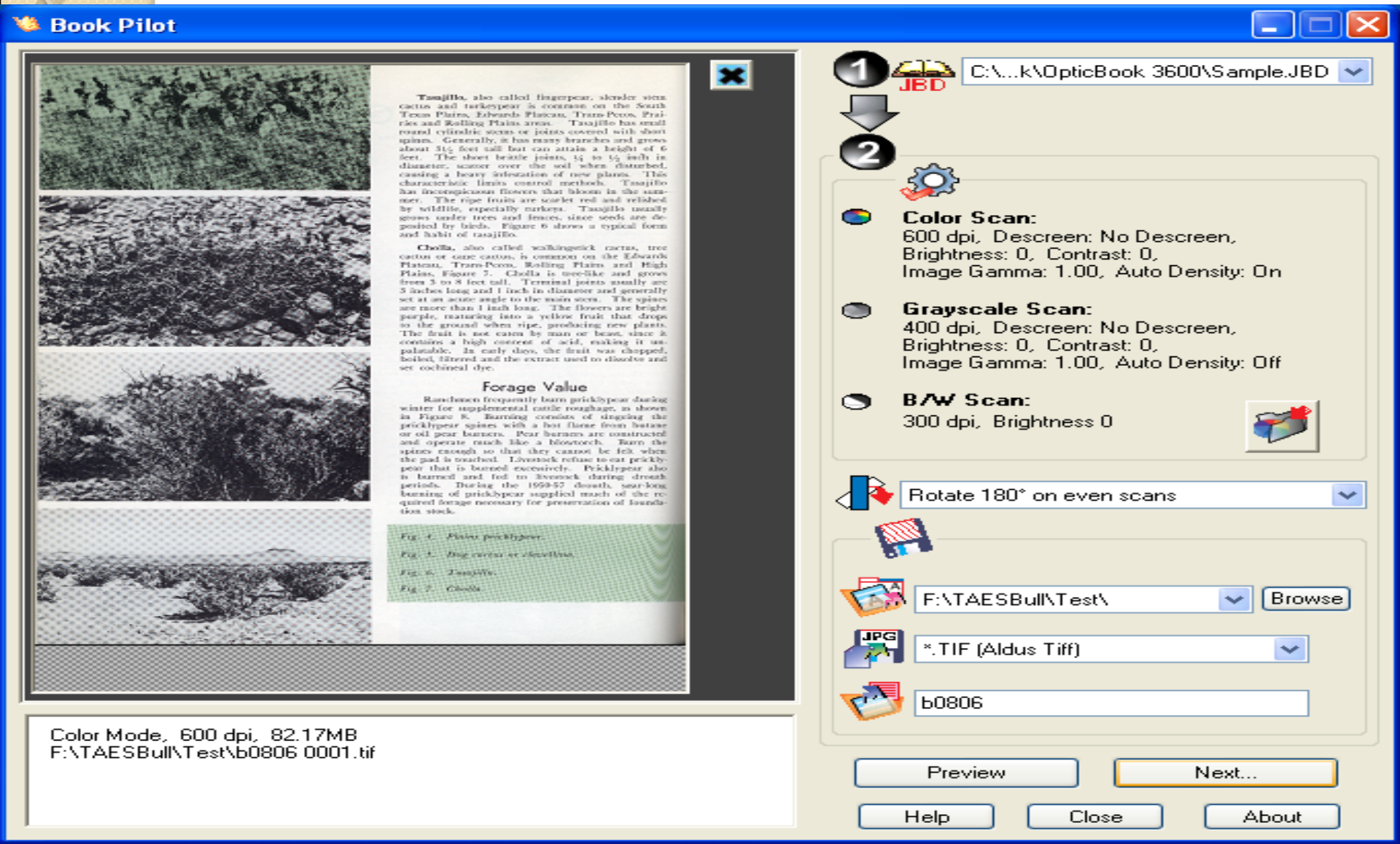
Scan the Page Image



Next Page (rotated 180) Preview



Scan Page (Auto-rotated Image)



Resulting TIF Image of Page

b0806 0002.tif - Windows Picture and Fax Viewer

Pricklypear — Good or Bad?

G. O. HOFFMAN and R. A. DARROW*

Pricklypear has been used as livestock feed for more than a century. It usually is considered an undesirable plant on Texas rangelands, but does have some economic value as supplemental forage for cattle and sheep during winter and drouth periods. Pricklypear has the unique ability of storing water in its flattened fleshy stems. This water reserve enables the plant to withstand long drouth periods. During drouths and range overuse, pricklypear density increases as grass cover lessens. With improved grazing management and grass conditions, pricklypear density decreases because of competition from good grass cover. A grazing management system that allows desirable grasses to increase should control pricklypear once the original infestation is controlled.

Several species of the genus, *Opuntia*, are called pricklypear. One or more species exist in almost every vegetational area of the State, Figure 1. The most common and widespread pricklypears known to Texas ranchmen are engelmann, *Opuntia engelmanni*, nopal, *Opuntia lindheimeri*, and plains, *Opuntia polyacantha*. Included in the same group are several types of chollas of which tasajillo, *Opuntia leptocaulis*, and cholla, *Opuntia imbricata*, provide the greatest problems on Texas rangelands. Dog cactus, *Opuntia schotti*, is injurious to grazing animals, particularly sheep and goats.

Engelmann pricklypear is common in the Edwards Plateau, Trans-Pecos, Rolling Plains and High Plains areas. It is an erect plant that may grow as high as 6 feet, but normally is never tree-like. Engelmann pricklypear pads are large and circular, usually 8 to 12 inches across, with a yellowish-green color and spines. Its large flowers are bright yellow. The fruits, called tuna, are large, dark purple and can be made into a palatable and attractive colored syrup. Wildlife relish the ripe tunas. This species is easily confused with other pricklypears of similar growth habits and form. Figure 2 shows the typical growth habit of engelmann pricklypear.

Nopal pricklypear is extremely variable in growth form and habits. It may be low and wide-spreading or tree-like, up to 12 feet tall, with a definite cylindrical trunk. Nopal pricklypear is found in the South Texas Plains, Blackland Prairies, Edwards Plateau, Trans-Pecos, Rolling Plains, High Plains and, to a limited extent, on the Gulf Prairies. The flowers are beautifully colored, varying from yellow to red. Earlyday Indians used the young tender pads as food. The tunas can be made into a tasty jelly. Wildlife and even livestock like the ripe tunas. Typical growth characteristics and habits of nopal pricklypear are shown in Figure 3.

The **plains pricklypear** is limited primarily to the Rolling and High Plains. This pricklypear is low growing, usually two pads tall, and forms small clumps. Generally, plains pricklypear is overgrown with grass when range has been managed properly. Figure 4 shows a typical growth habit of plains pricklypear.

Dog cactus, also called clavellina, is limited to the South Texas Plains, Edwards Plateau and the Trans-Pecos areas. The plants are prostrate and form dense masses 10 to 50 feet square. These masses are impenetrable and a great pest to grazing animals. The joints, covered with long spines, are broken off easily and often cling to the animal's feet, mouth and other body parts, causing serious injury. A typical growth of dog cactus is shown in Figure 5.

*Respectively, Extension range specialist, Texas A&M University, and former professor, Range and Forestry Department.

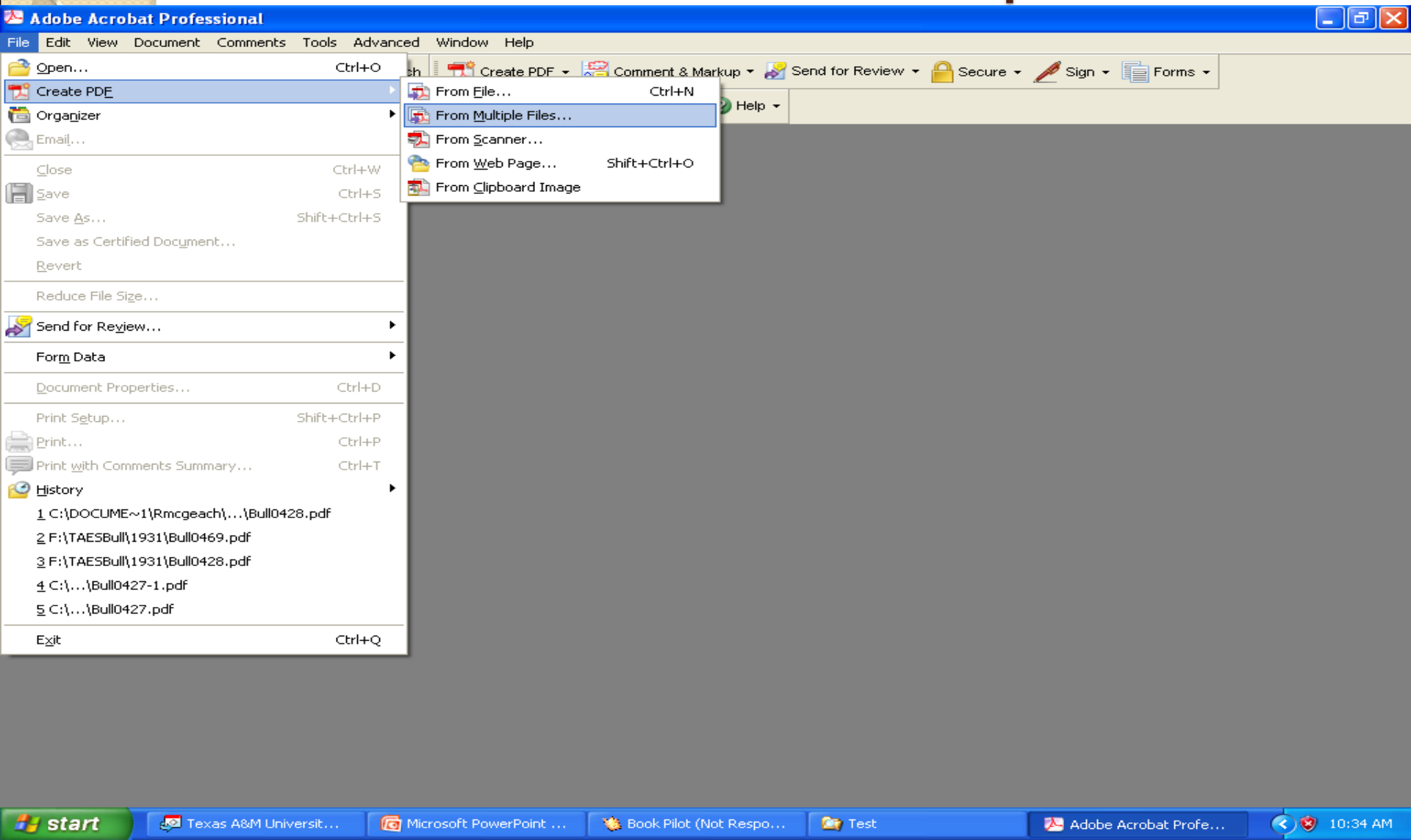
Fig. 2. Englemann pricklypear.

Fig. 3. Nopal pricklypear.

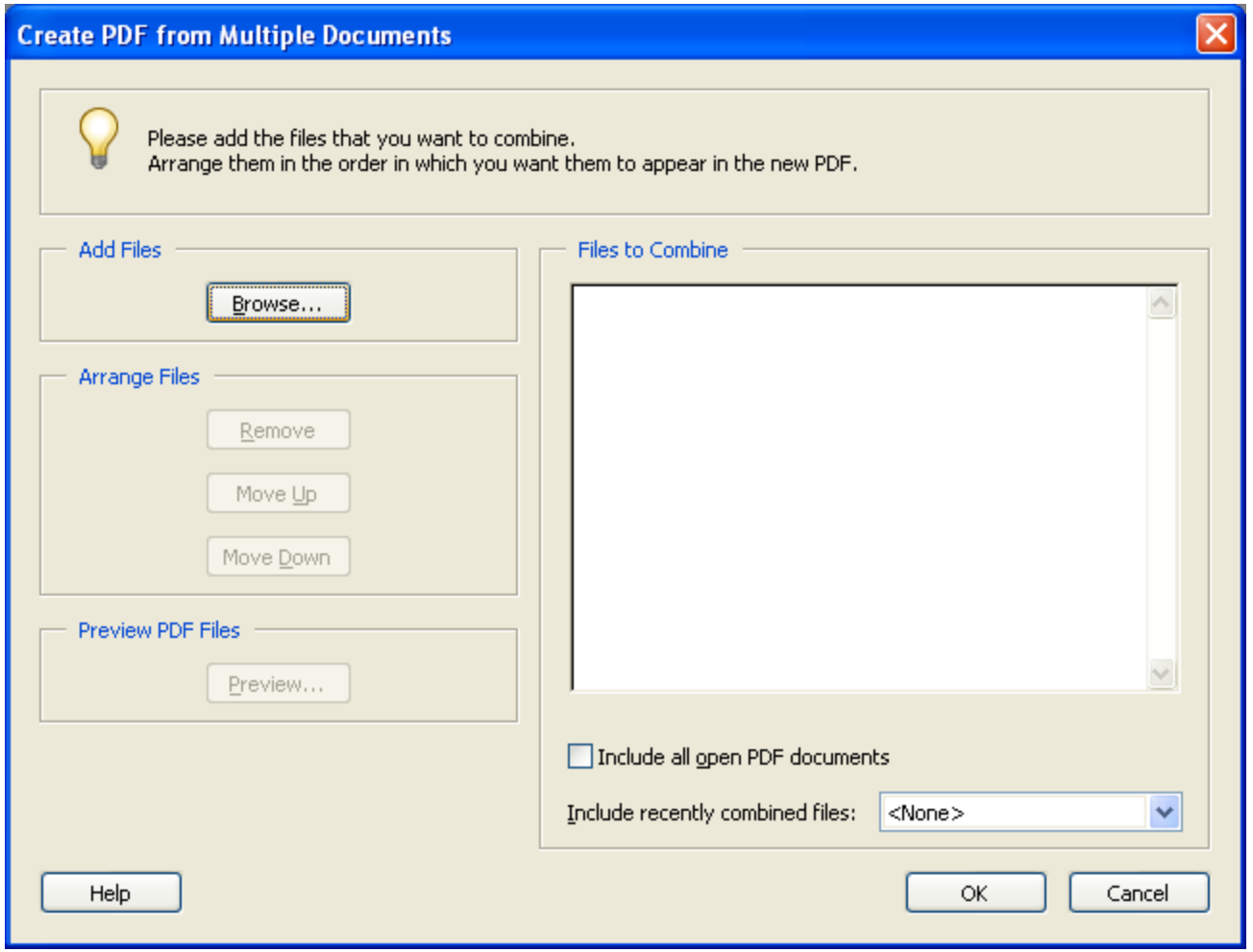
Demo of Problem Scans

- B&W scans too all or nothing
- Image too skewed, not straight
- Too much dark page gutter

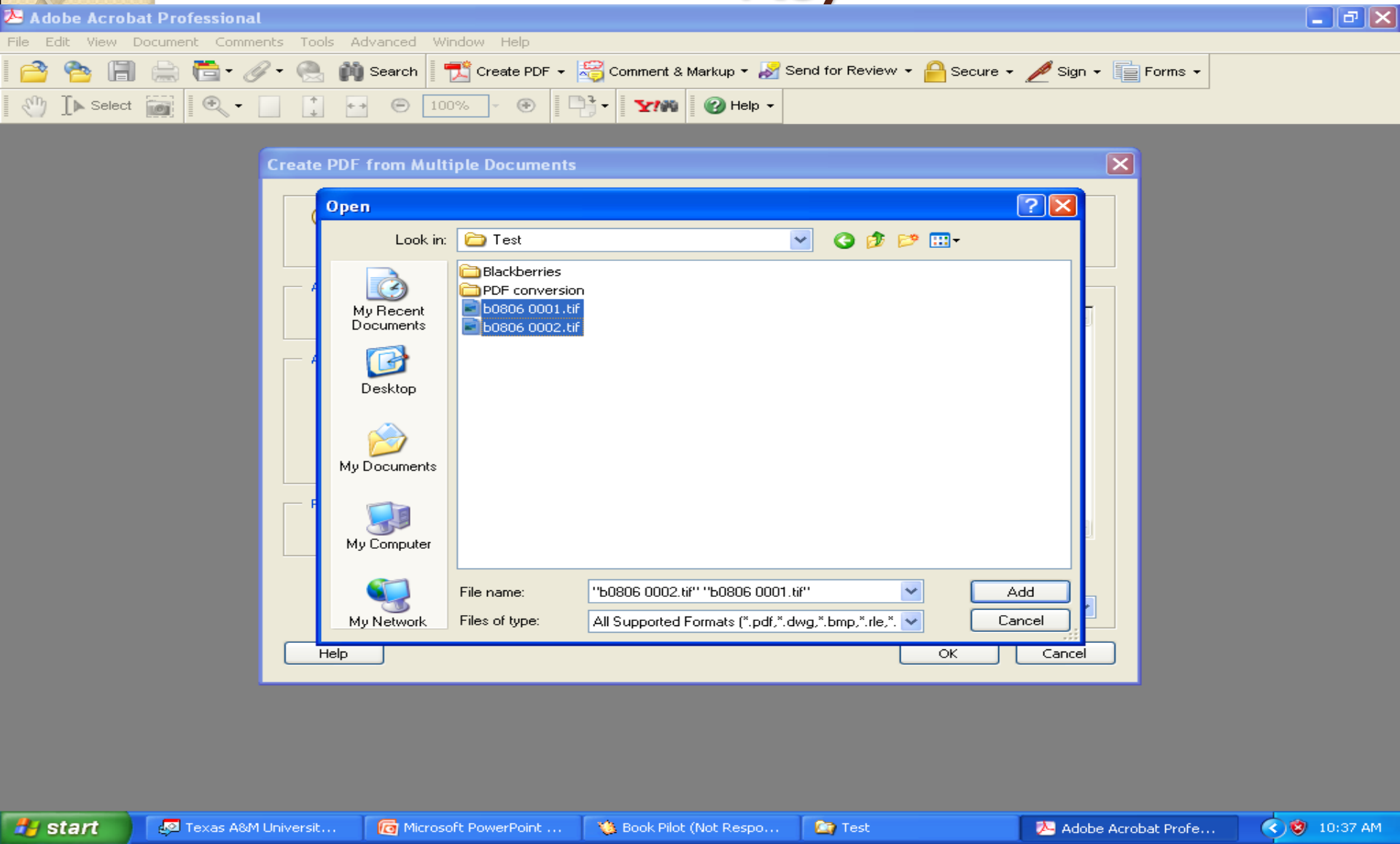
Adobe Acrobat Professional: Create New PDF from Multiple Files



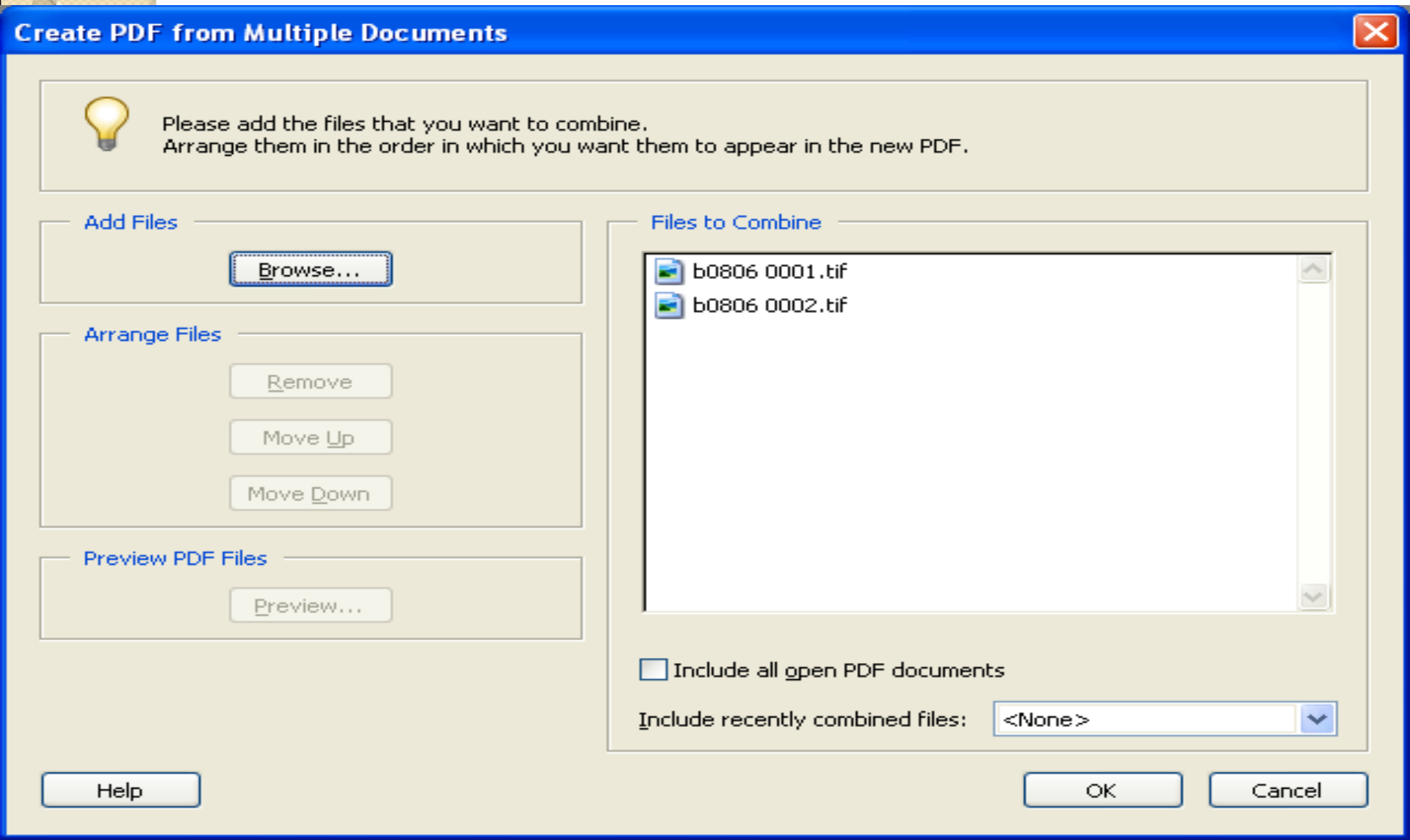
Browse to File Folder



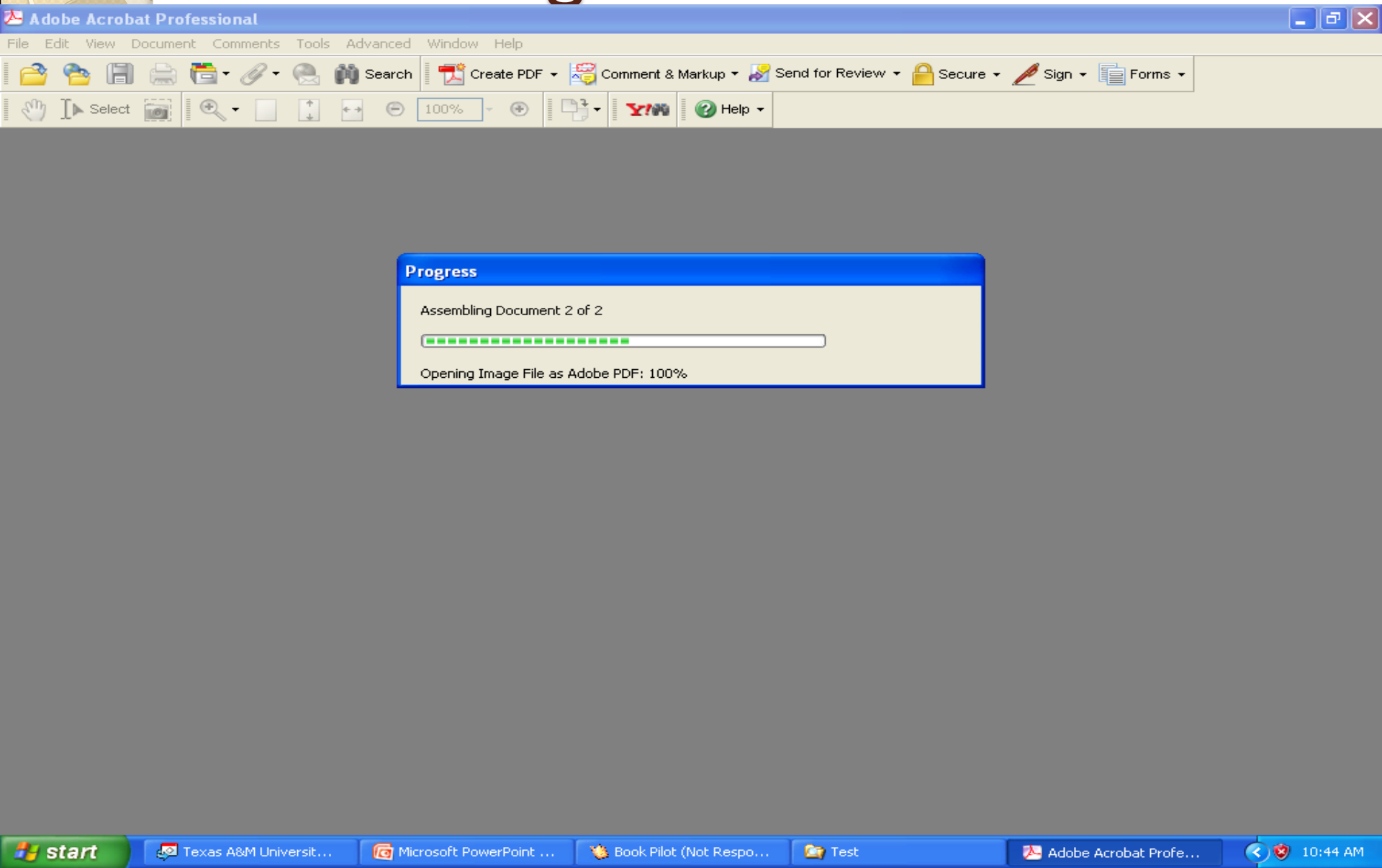
Select Multiple Files Using Control Key



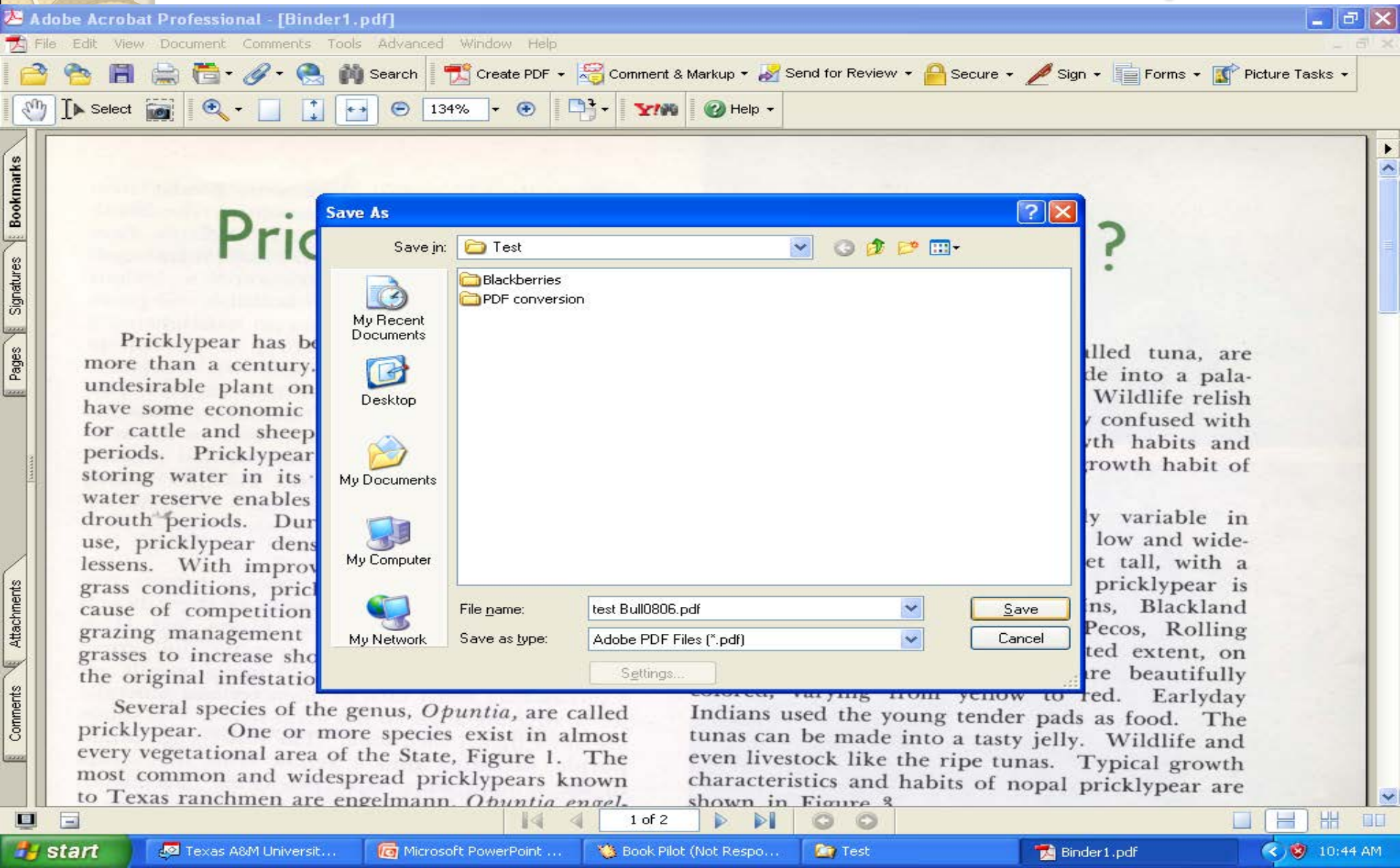
Added File List



Processing into PDF File



Save New File as Bull#####.pdf



New PDF File

Adobe Acrobat Professional - [test Bull0806.pdf]

File Edit View Document Comments Tools Advanced Window Help

Search | Create PDF | Comment & Markup | Send for Review | Secure | Sign | Forms | Picture Tasks

Select | 100% | Help

Bookmarks | Signatures | Pages | Attachments | Comments

Pricklypear — Good or Bad?

*G. O. HOFFMAN and R. A. DARROW**

Pricklypear has been used as livestock feed for more than a century. It usually is considered an undesirable plant on Texas rangelands, but does have some economic value as supplemental forage for cattle and sheep during winter and drouth periods. Pricklypear has the unique ability of storing water in its flattened-fleshy stems. This water reserve enables the plant to withstand long drouth periods. During drouths and range overuse, pricklypear density increases as grass cover lessens. With improved grazing management and grass conditions, pricklypear density decreases because of competition from good grass cover. A grazing management system that allows desirable grasses to increase should control pricklypear once the original infestation is controlled.

Several species of the genus, *Opuntia*, are called pricklypear. One or more species exist in almost every vegetational area of the State, Figure 1. The most common and widespread pricklypears known to Texas ranchmen are engelmann, *Opuntia engelmanni*, nopal, *Opuntia lindheimeri*, and plains, *Opuntia polyacantha*. Included in the same group are several types of chollas of which tasajillo, *Opuntia leptocaulis*, and cholla, *Opuntia imbricata*, provide the greatest problems on Texas rangelands. Dog cactus, *Opuntia schotti*, is injurious to grazing animals, particularly sheep and goats.

Engelmann pricklypear is common in the Edwards Plateau, Trans-Pecos, Rolling Plains and High Plains areas. It is an erect plant that may

are bright yellow. The fruits, called tuna, are large, dark purple and can be made into a palatable and attractive colored syrup. Wildlife relish the ripe tunas. This species is easily confused with other pricklypears of similar growth habits and form. Figure 2 shows the typical growth habit of engelmann pricklypear.

Nopal pricklypear is extremely variable in growth form and habits. It may be low and widespreading or tree-like, up to 12 feet tall, with a definite cylindrical trunk. Nopal pricklypear is found in the South Texas Plains, Blackland Prairies, Edwards Plateau, Trans-Pecos, Rolling Plains, High Plains and, to a limited extent, on the Gulf Prairies. The flowers are beautifully colored, varying from yellow to red. Earlyday Indians used the young tender pads as food. The tunas can be made into a tasty jelly. Wildlife and even livestock like the ripe tunas. Typical growth characteristics and habits of nopal pricklypear are shown in Figure 3.

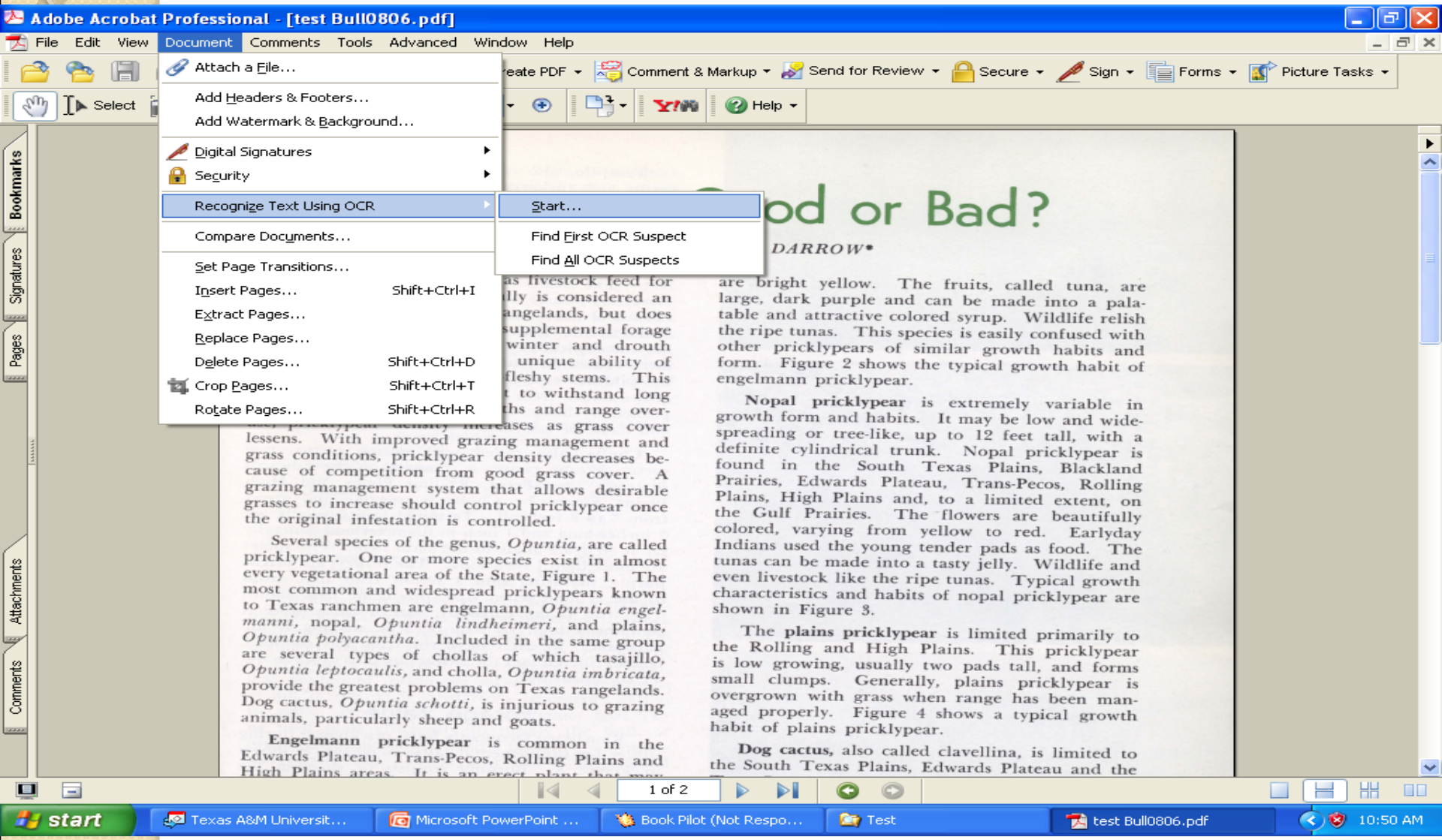
The **plains pricklypear** is limited primarily to the Rolling and High Plains. This pricklypear is low growing, usually two pads tall, and forms small clumps. Generally, plains pricklypear is overgrown with grass when range has been managed properly. Figure 4 shows a typical growth habit of plains pricklypear.

Dog cactus, also called clavellina, is limited to the South Texas Plains, Edwards Plateau and the

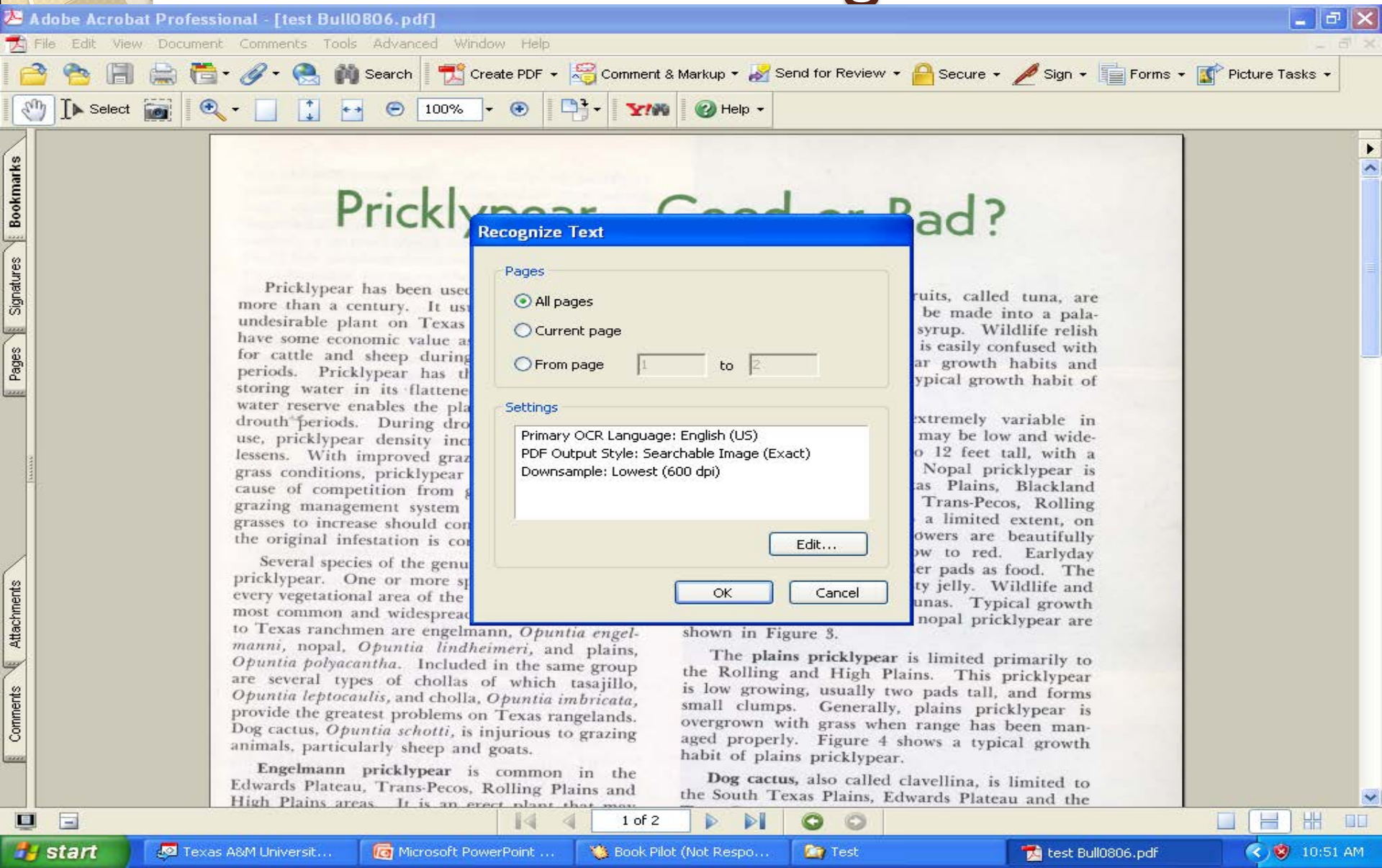
1 of 2

start | Texas A&M Universit... | Microsoft PowerPoint ... | Book Pilot (Not Respo... | Test | test Bull0806.pdf | 10:47 AM

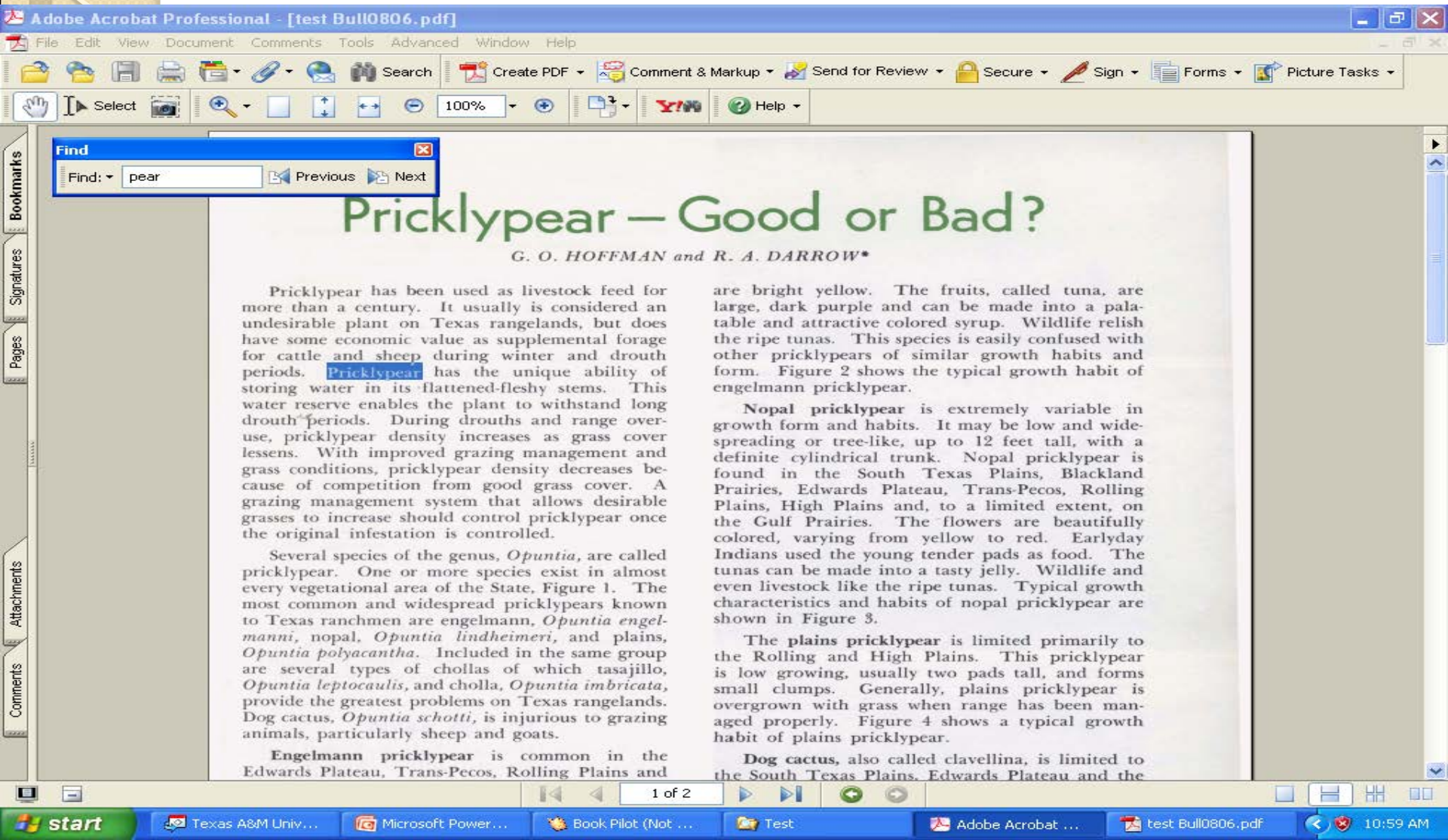
Do OCR on the File



Select -- On All Pages



Makes PDF fully text searchable



PDF File with OCRed Text

- Allows searching with in retrieved documents
- Also allows full text indexing of the item if the Repository is designed to include text of PDF, Word, TXT type files in the indexing it generates in addition to the Metadata in item records

Questions on Digitization and Scanning Basics?

This Slide Blank in Original

(Time for a break!)

Digitization – In Depth

Metadata, Repositories, and Access

Outline:

- Standards Initiatives
- Metadata Standards
- Repositories
- Access to Repositories

Standards Initiatives

- National Digital Infrastructure Preservation Program (NDIPP)
  - http://www.digitalpreservation.gov/
- Federal Agencies Digitization Guidelines Initiative
  - http://www.digitizationguidelines.gov/stillimages/documents/DigitizationStandards.pdf
  - http://www.digitizationguidelines.gov/stillimages/documents/Guidelines_Bibliography-DRAFT-2008-07-20-1.pdf

NAL Scanning Specifications

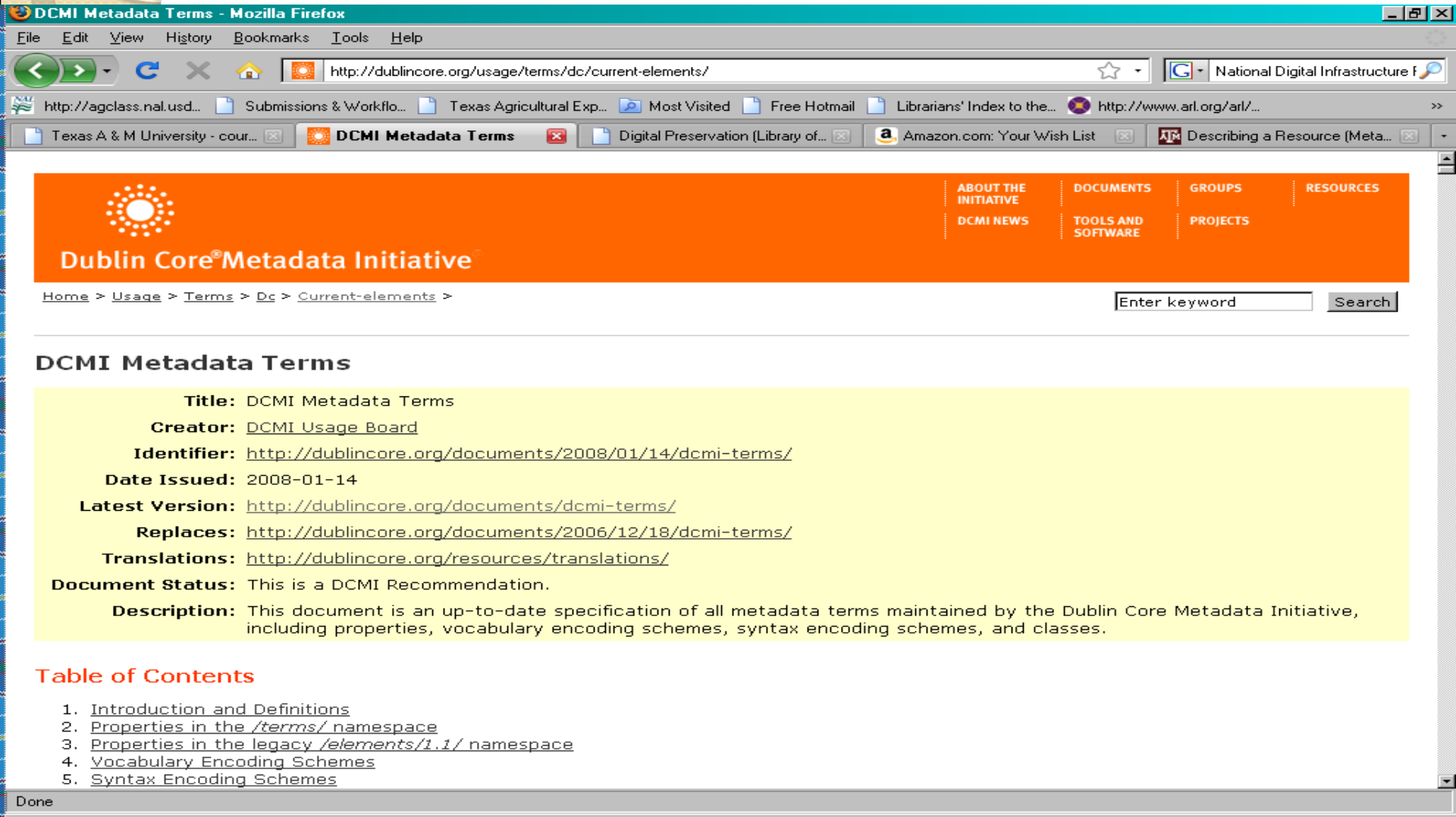
- Home / NAL Collections / Preservation/ Scanning Specifications
  - http://riley.nal.usda.gov/nal_display/index.php?info_center=8&tax_level=3&tax_subject=158&topic_id=2009&level3_id=6471

NAL Scanning Specifications

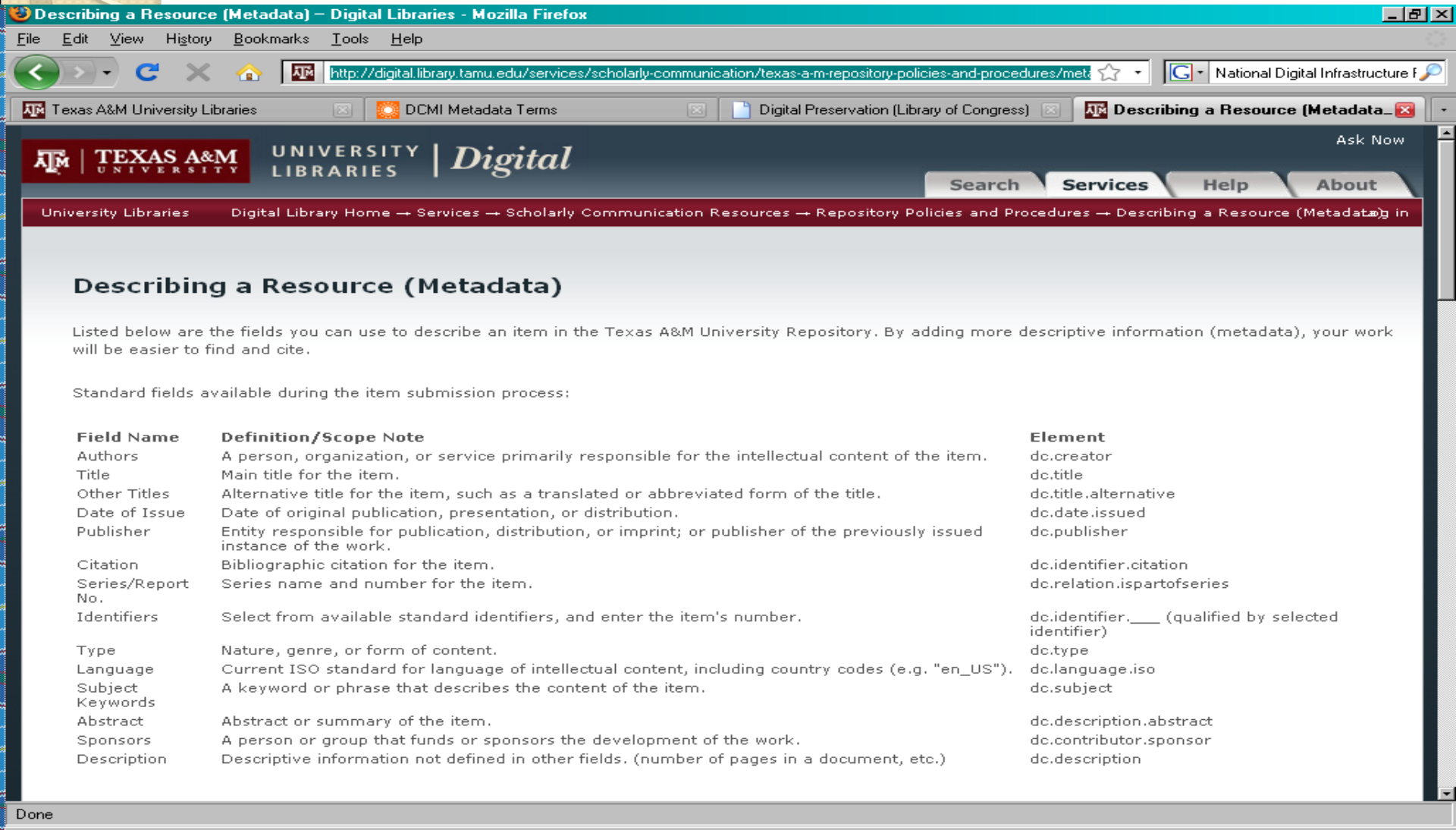
Metadata Standards

- Dublin Core Metadata Initiative
  - http://dublincore.org/usage/terms/dc/current-elements/
- Example Digital Library Metadata Policy
  - http://digital.library.tamu.edu/services/scholarly-communication/texas-a-m-repository-policies-and-procedures/metadata-policy

Dublin Core Metadata Initiative



Texas A&M University Digital Library Metadata Policy



Metadata Elements

- Authors
  - dc.creator
- Title
  - dc.title
- Date of Issue
  - dc.date
- Publisher
  - dc.publisher
- Citation
  - dc.bibliographicCitation

Metadata Elements

- Series/Report Number
  - dc. relation.ispartofseries
- Identifiers
  - dc.identifier
- Type
  - dc.type
- Language
  - dc.language

Describe Item with Metadata

- Subject Keywords
  - dc.subject
  - dc.subject.nalt
- Abstract
  - dc.abstract
- Sponsors
  - dc. contributor.sponsor
- Description
  - dc.description

Digital Repository Basics

- Software – Commercial or Open Source
- Just one "service" of a Digital Library
  - Repository
  - Peer Reviewed Open Access E-Journals
  - Blogs
  - Wikis
  - Scholarly communications website

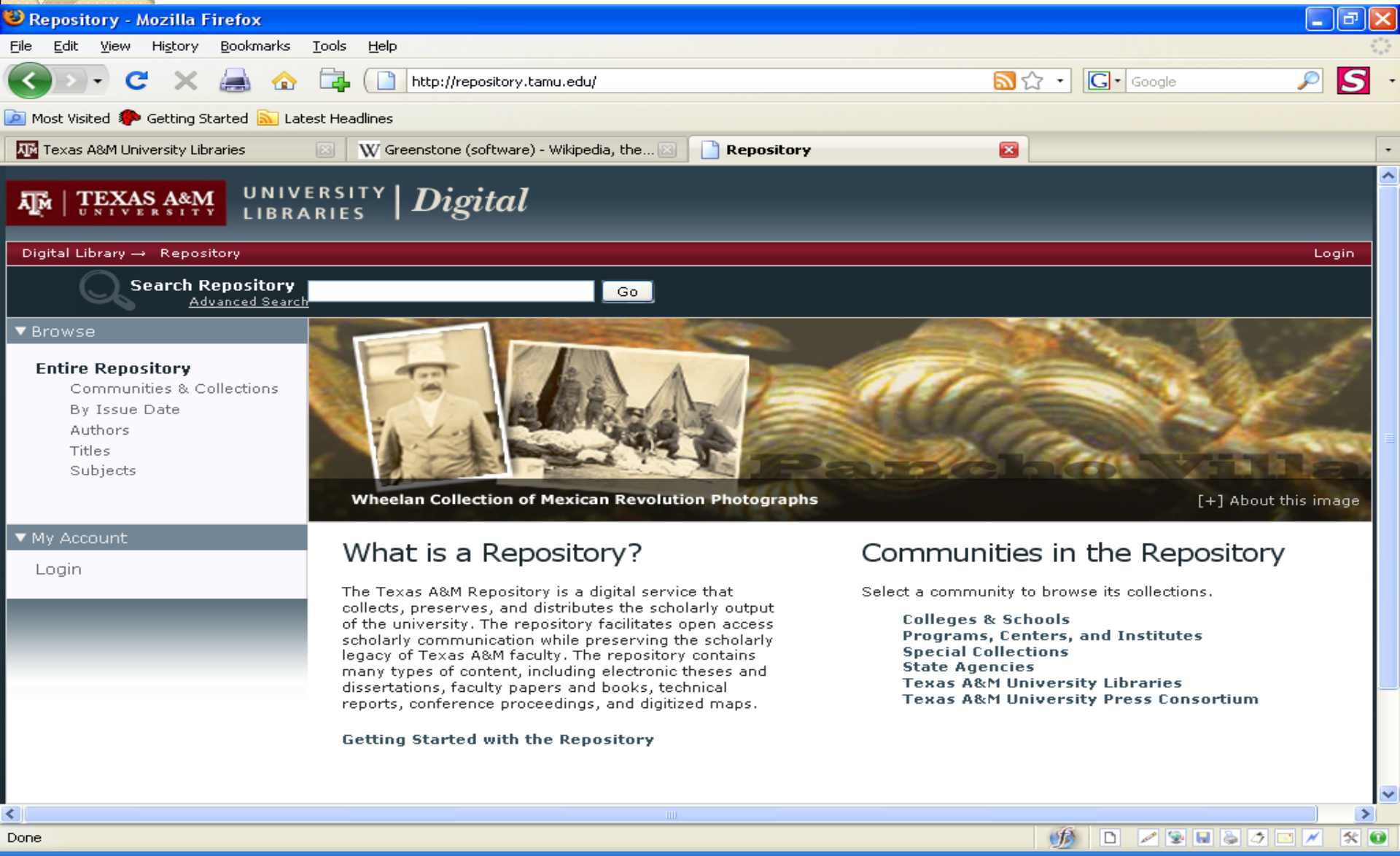
Digital Repository Software

- Open Source
  - DSpace
  - Eprints
  - Fedora
  - Greenstone
- Commercial
  - CONTENTdm
  - Digital Commons
  - DigiTool

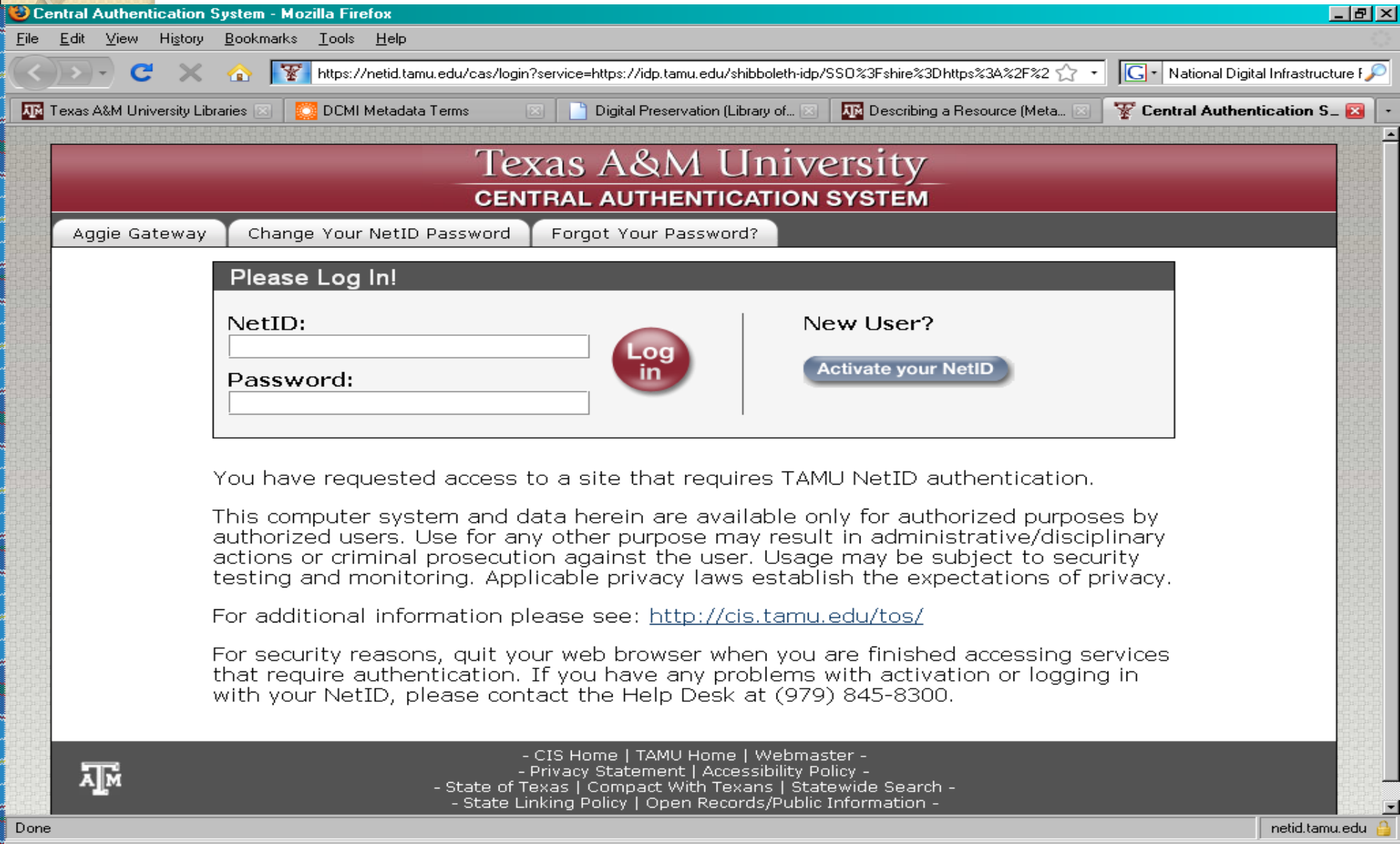
Submitting Items to the Repository

- Components of a Submission
  - Item Metadata
  - Archival Files
  - Derivative files for viewing over the web
- License
- Quality Control Check
- Approval and "Publication" to go live
- Post Publication Editing

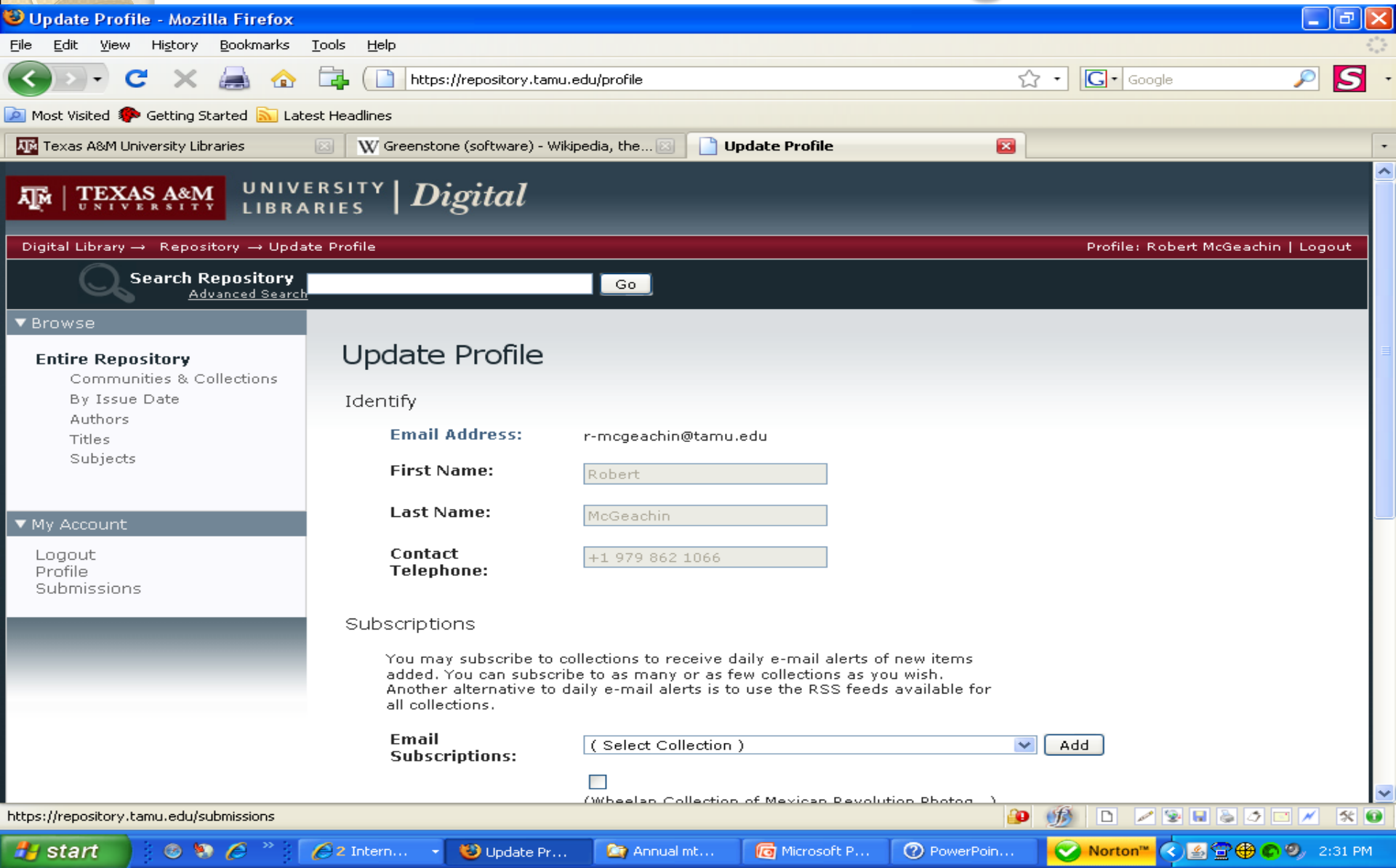
Repository Home Page (DSpace)



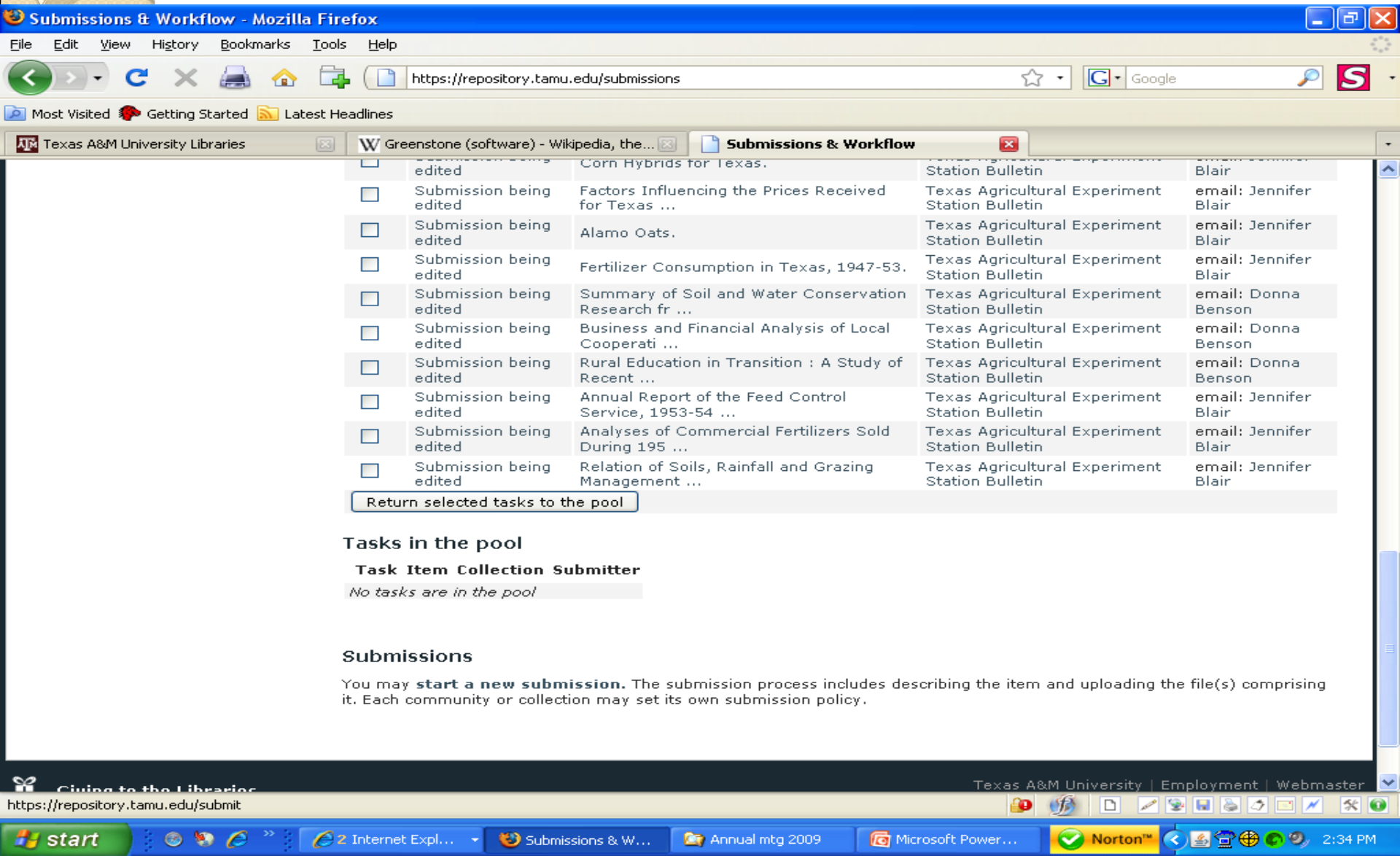
Login to Repository Account



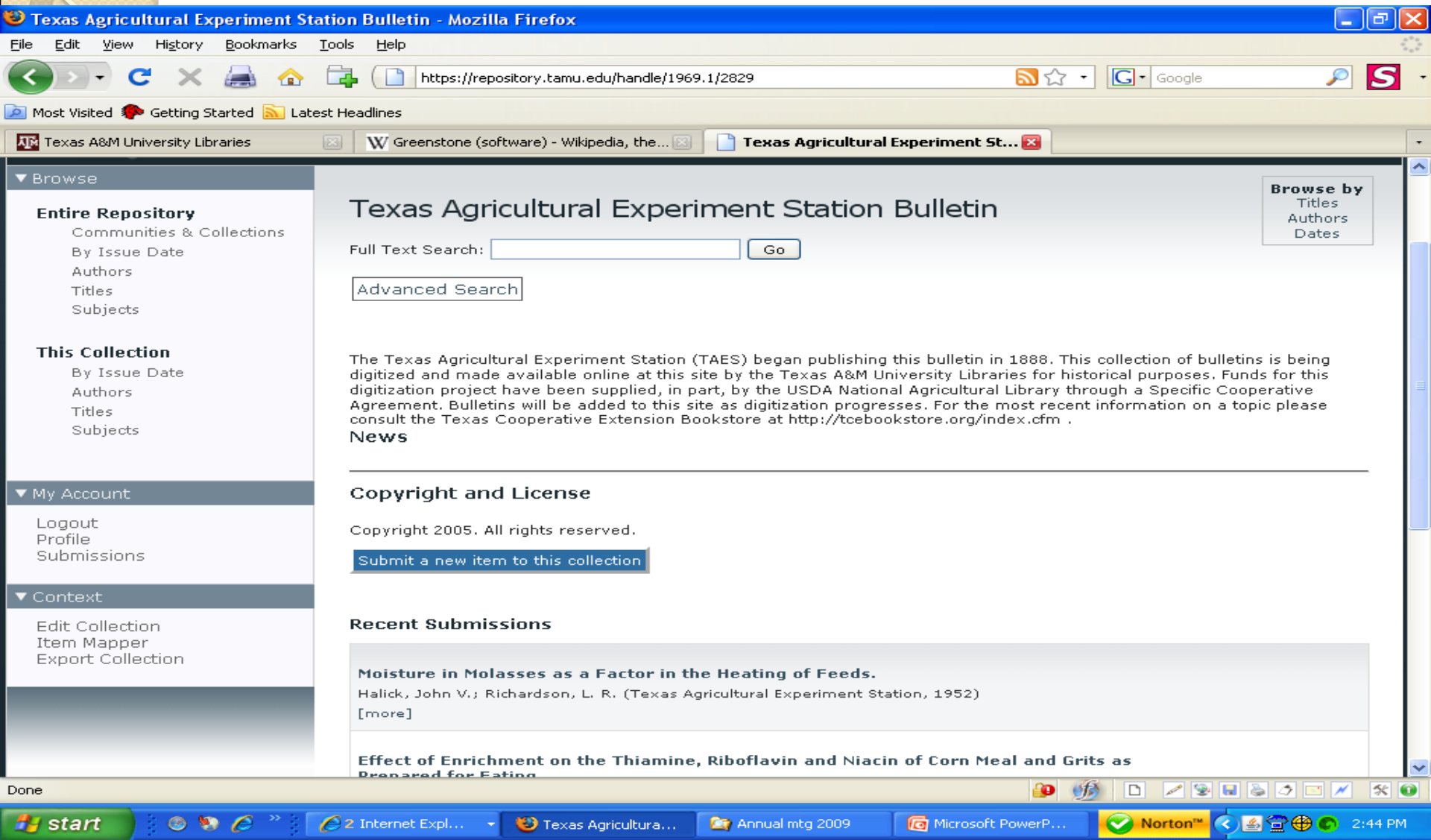
Go To Submissions Page



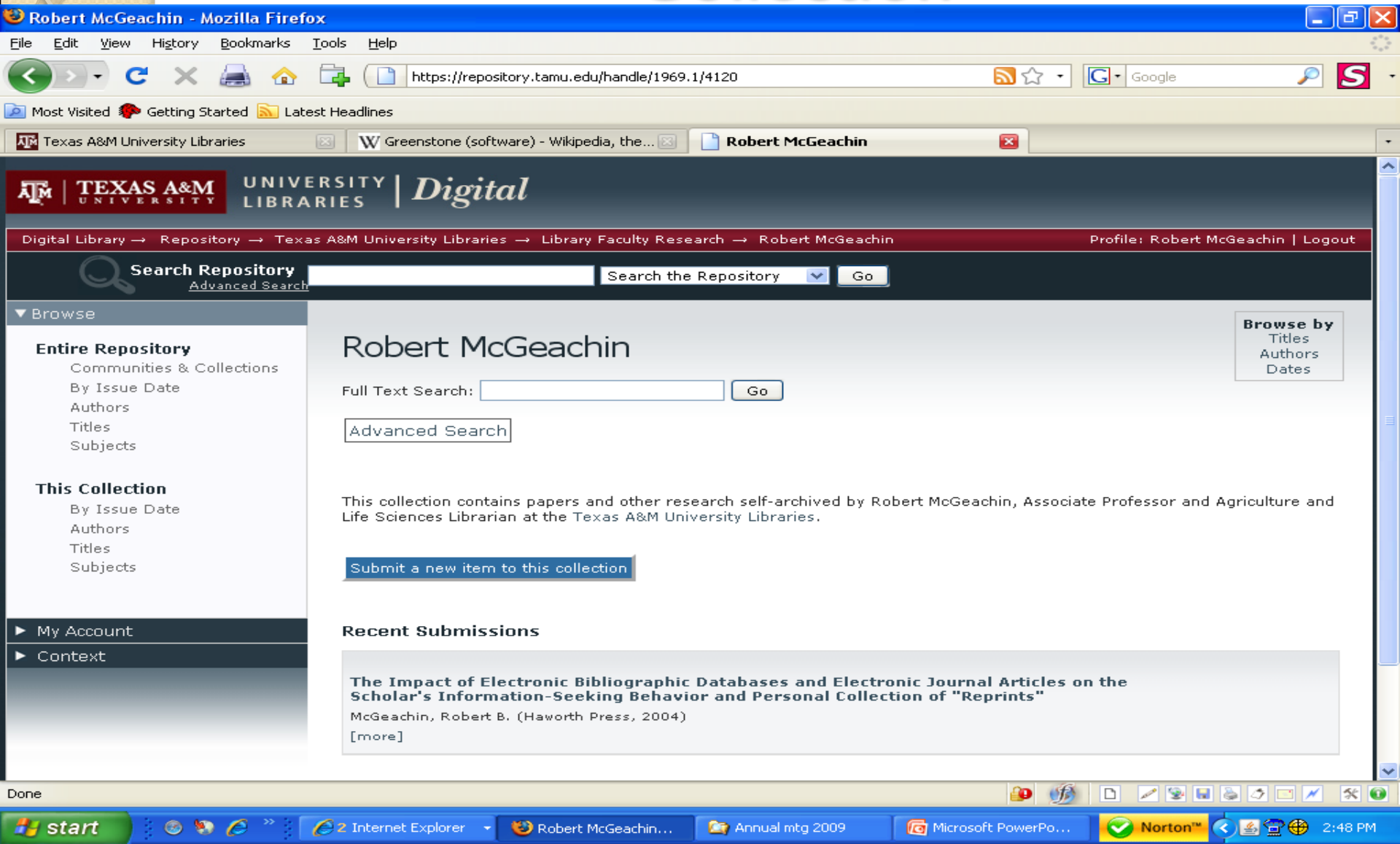
Start A New Submission



Or Start New Submission from a Collection Page



Go To Submit New Item in My Collection

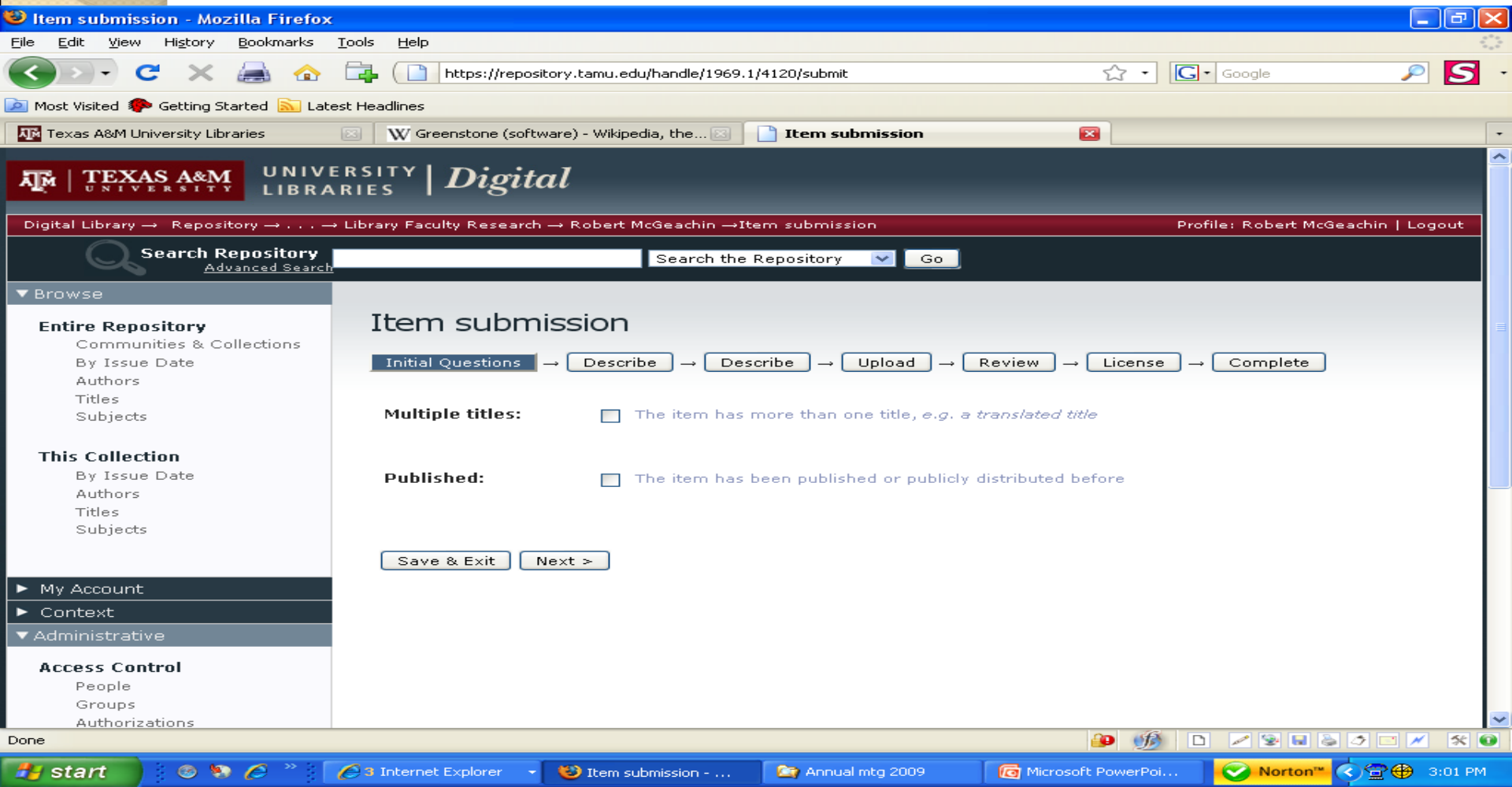


Robert McGeachin Collection

- https://repository.tamu.edu/handle/1969.1/4120
- Demo of Submitting This Presentation to the Digital Repository

Initial Questions

- Previously Published and Multiple Titles



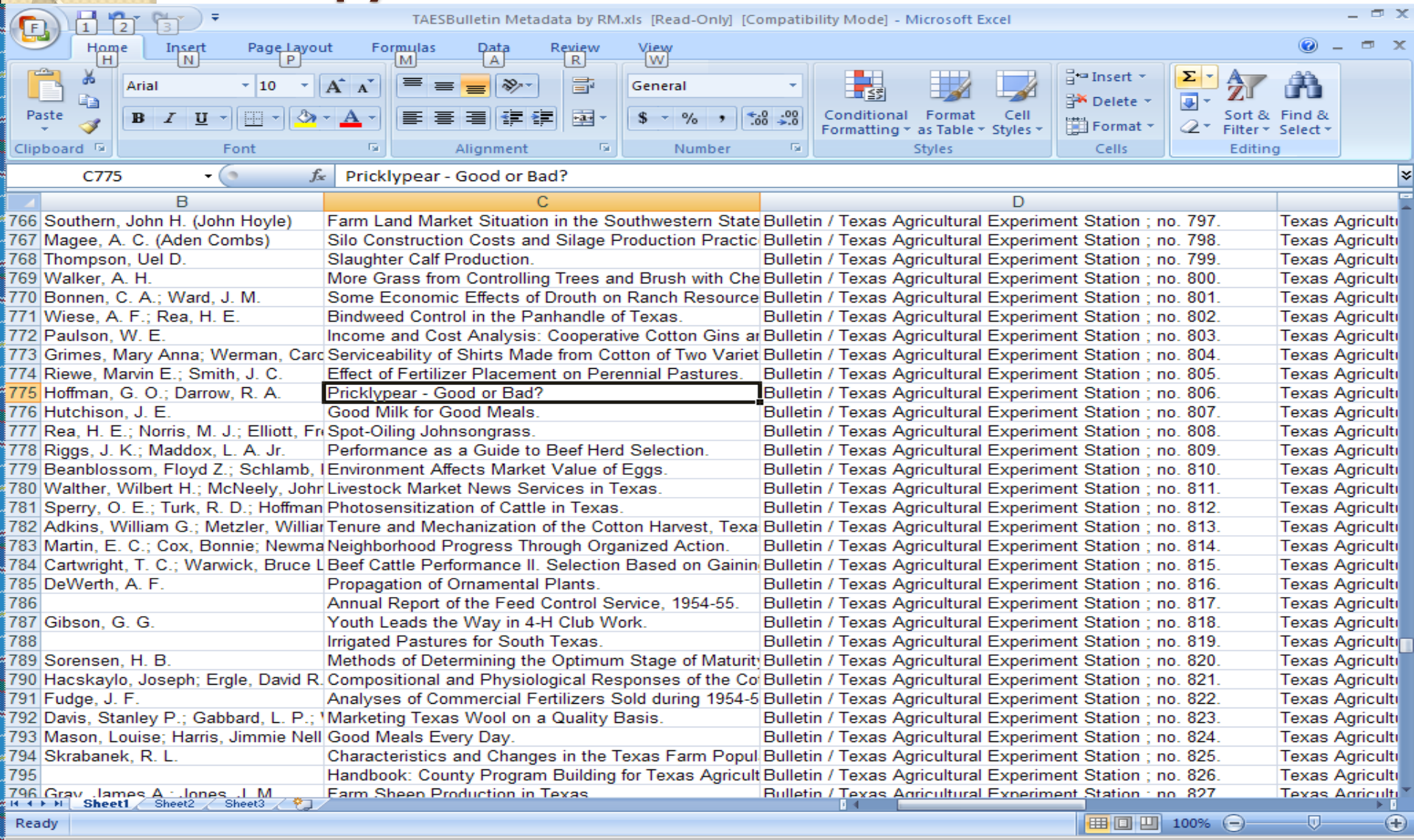
Describe Item with Metadata

- Authors
- Title
- Date of Issue
- Publisher
- Citation
- Series/Report Number
- Identifiers
- Type
- Language

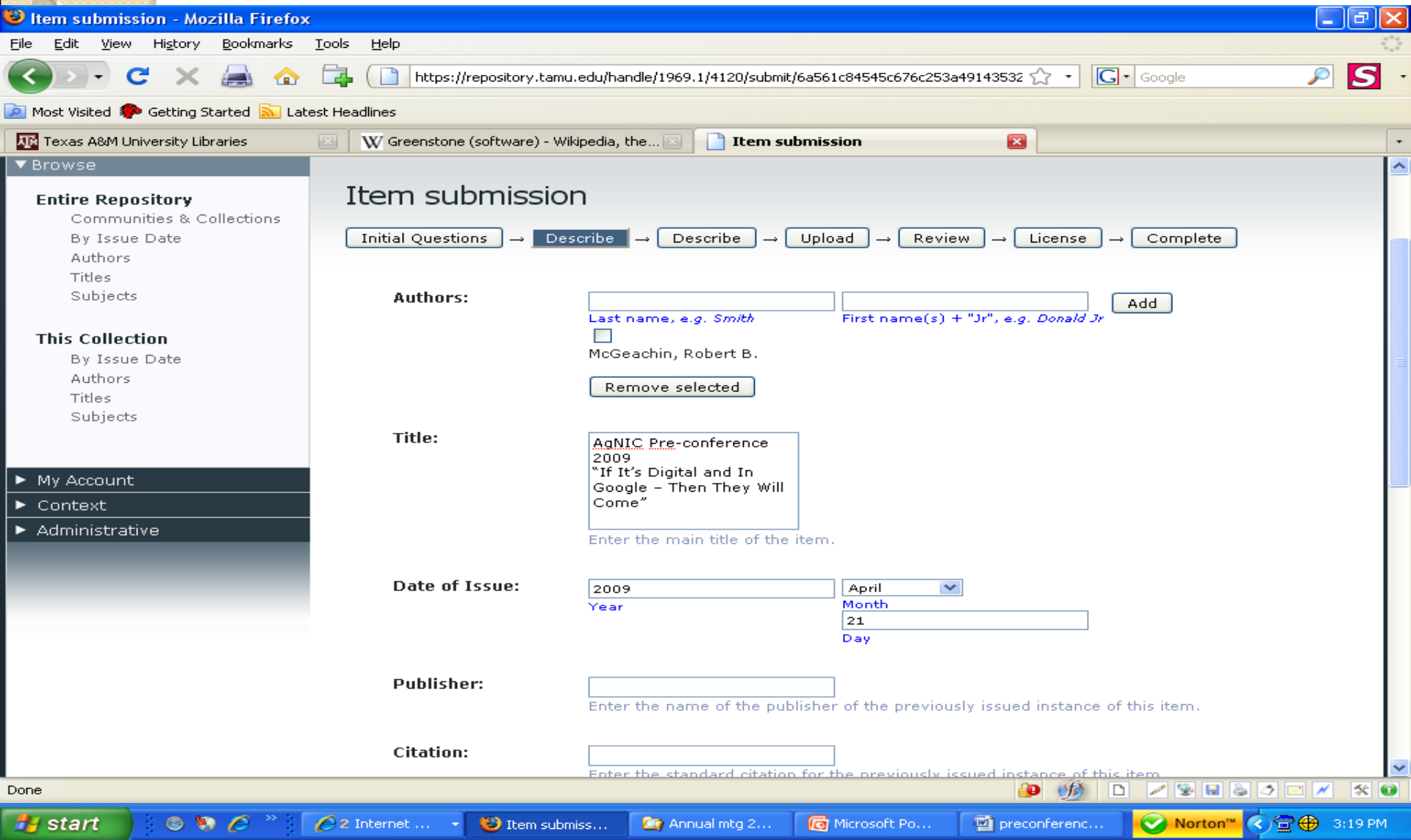
Describe Item with Metadata

- Subject Keywords
- Abstract
- Sponsors
- Description

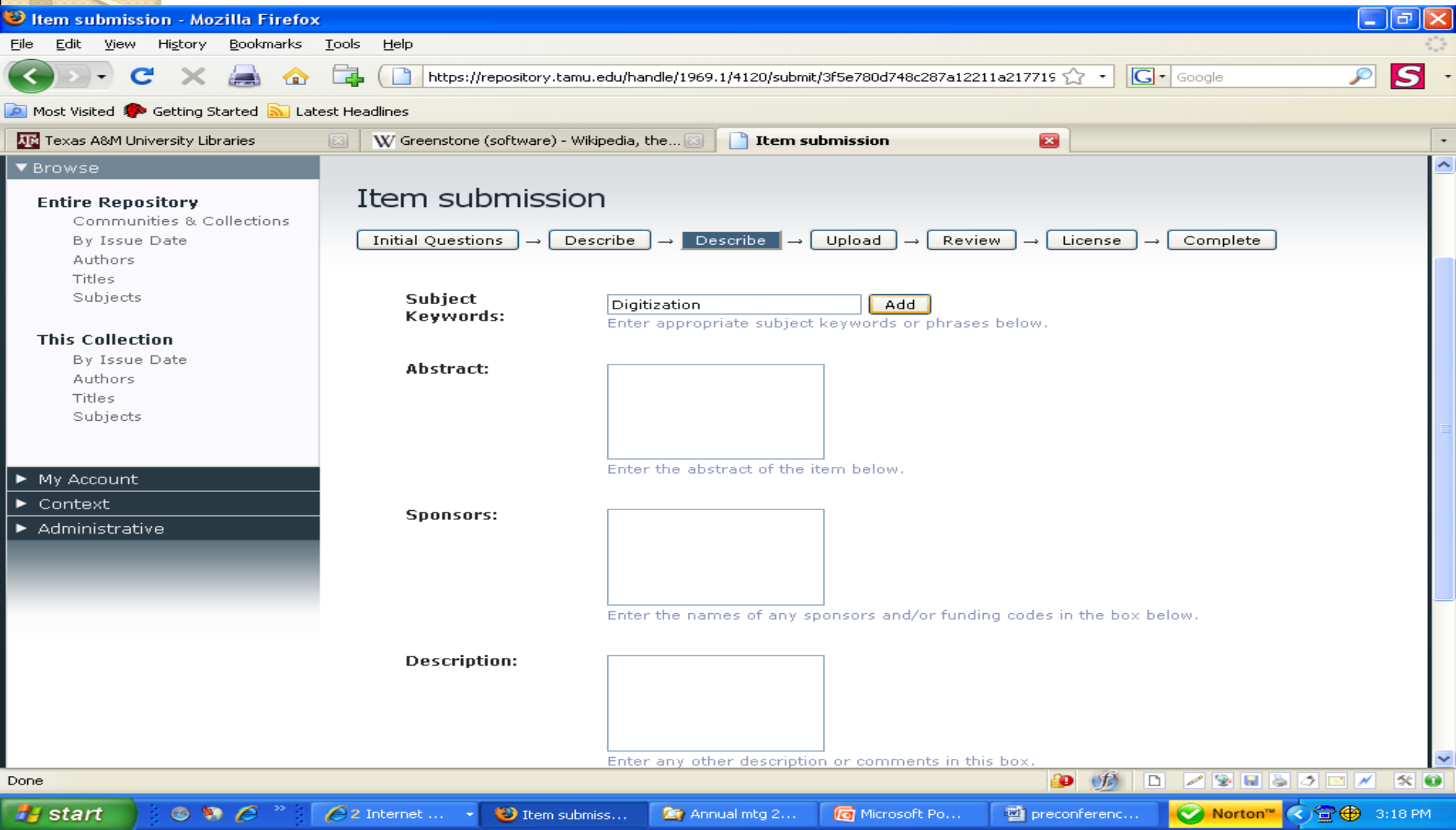
Item Metadata in Spreadsheet – Copy and Paste into Submission



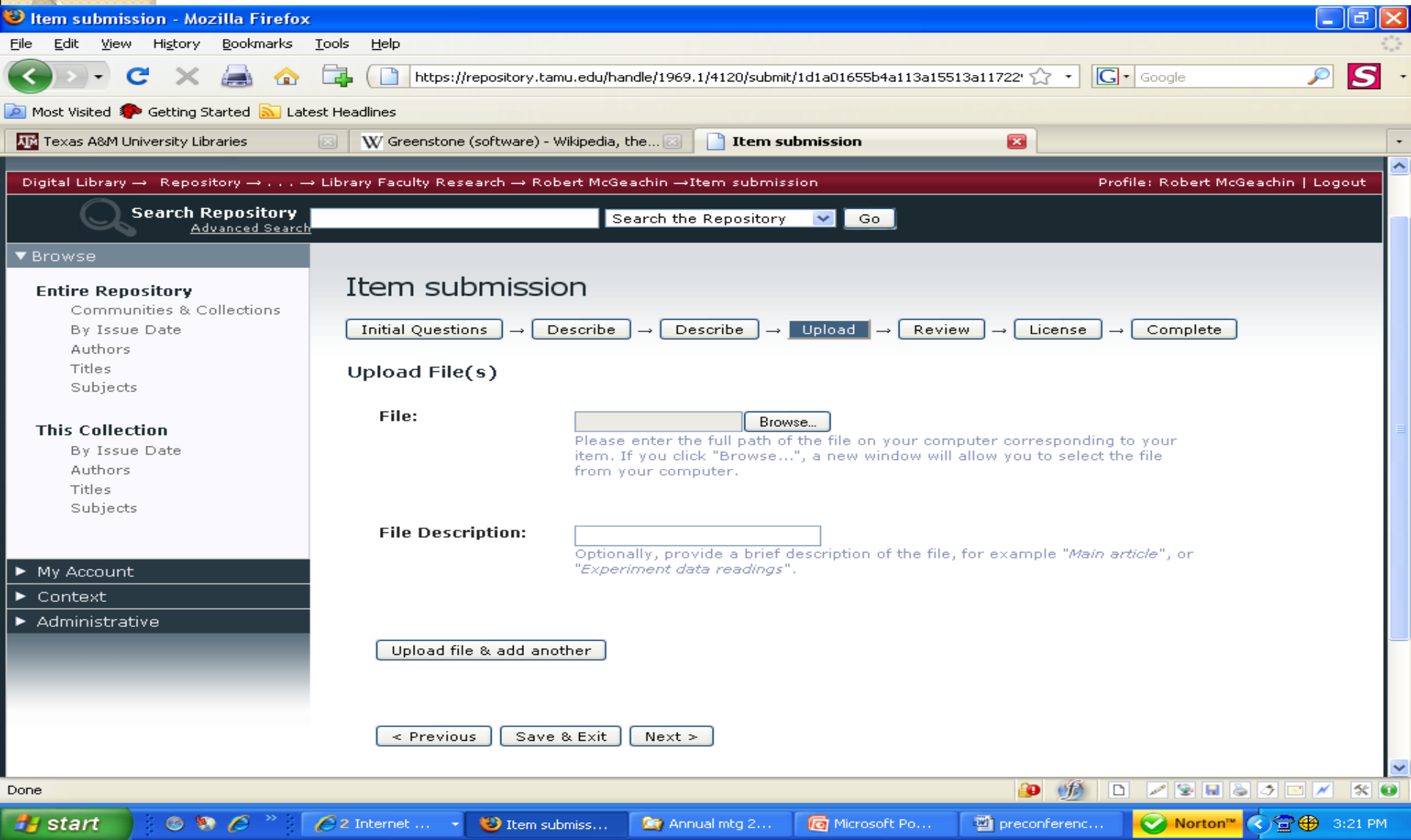
Describe Item with Metadata



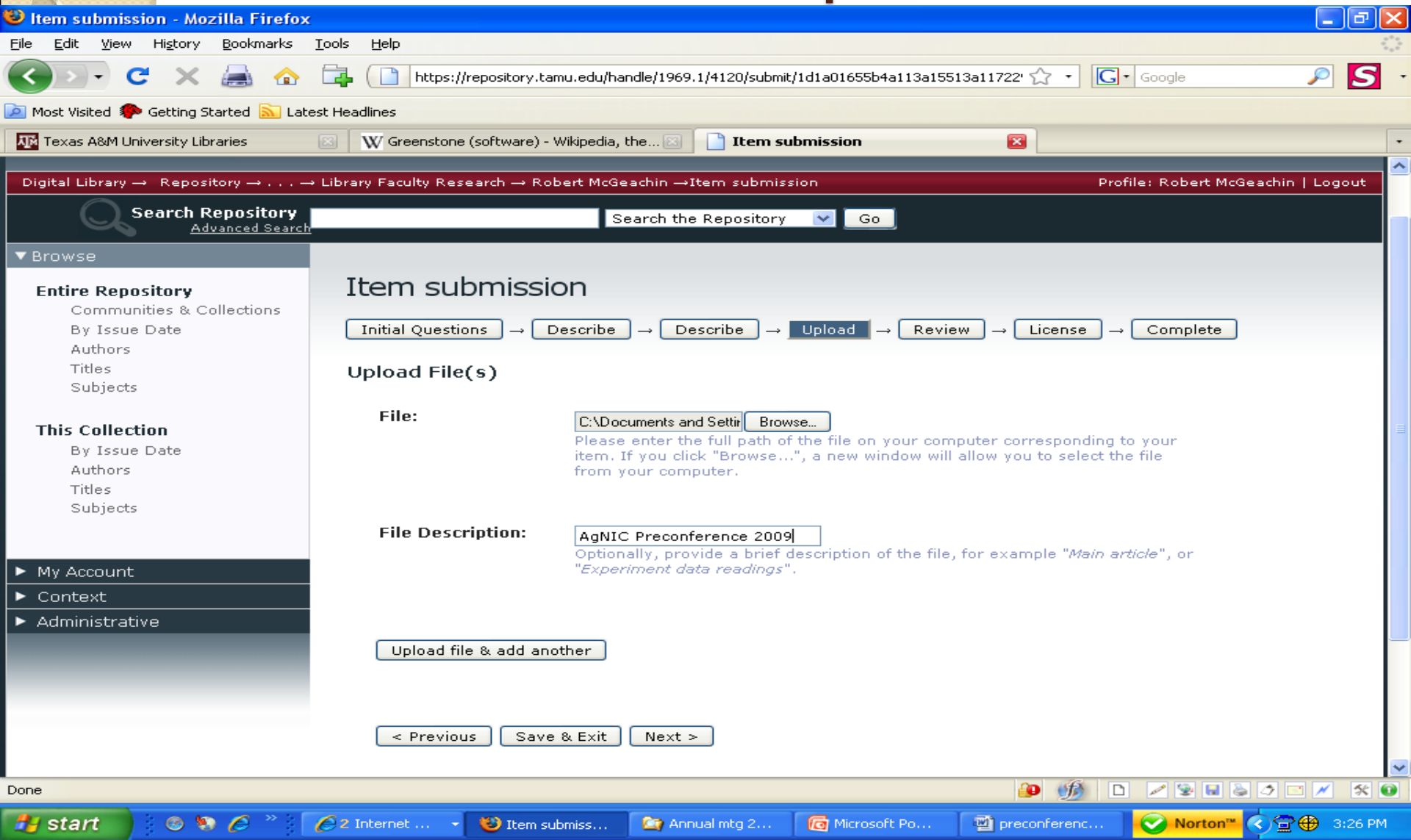
Describe Item with Metadata



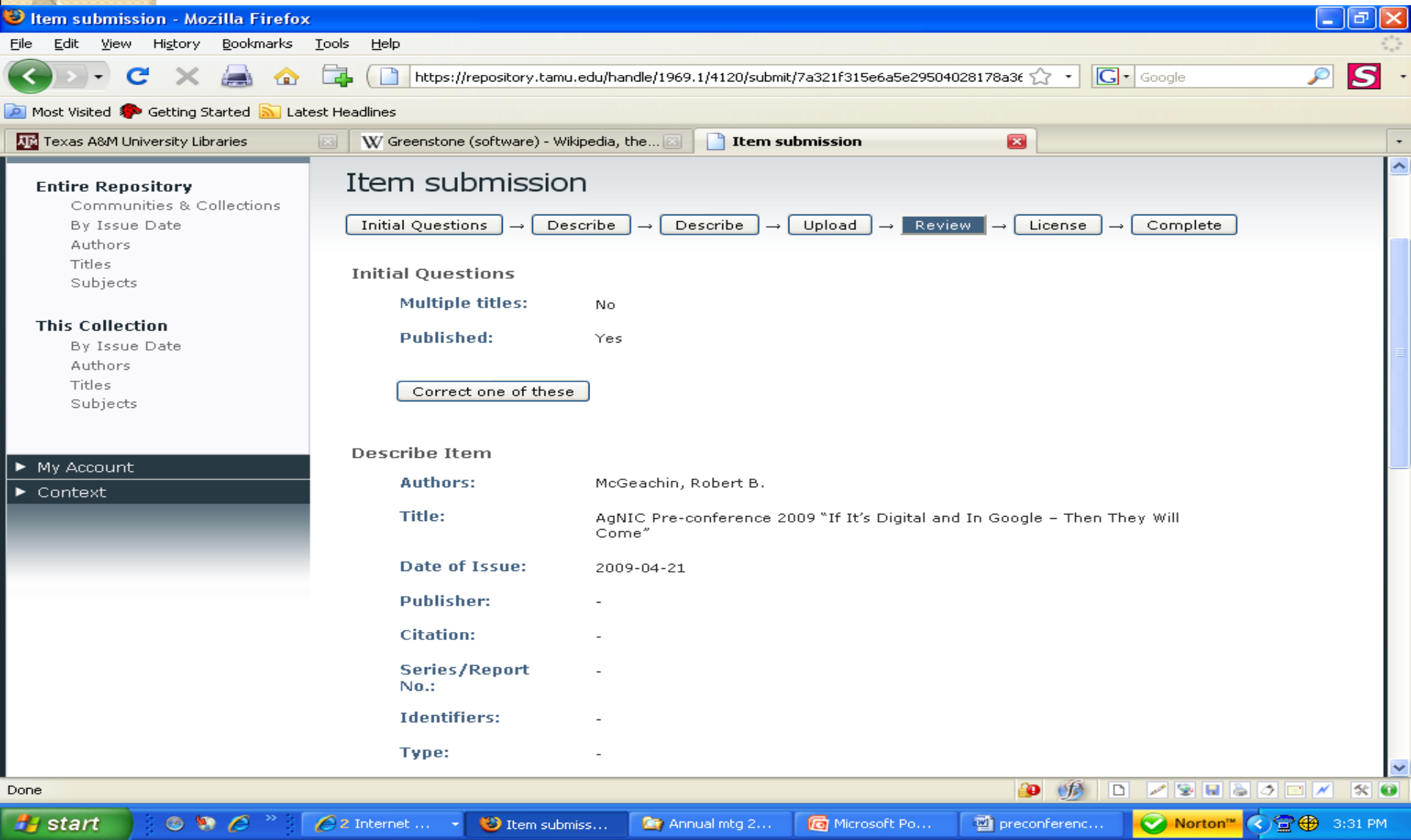
Upload Files to Repository



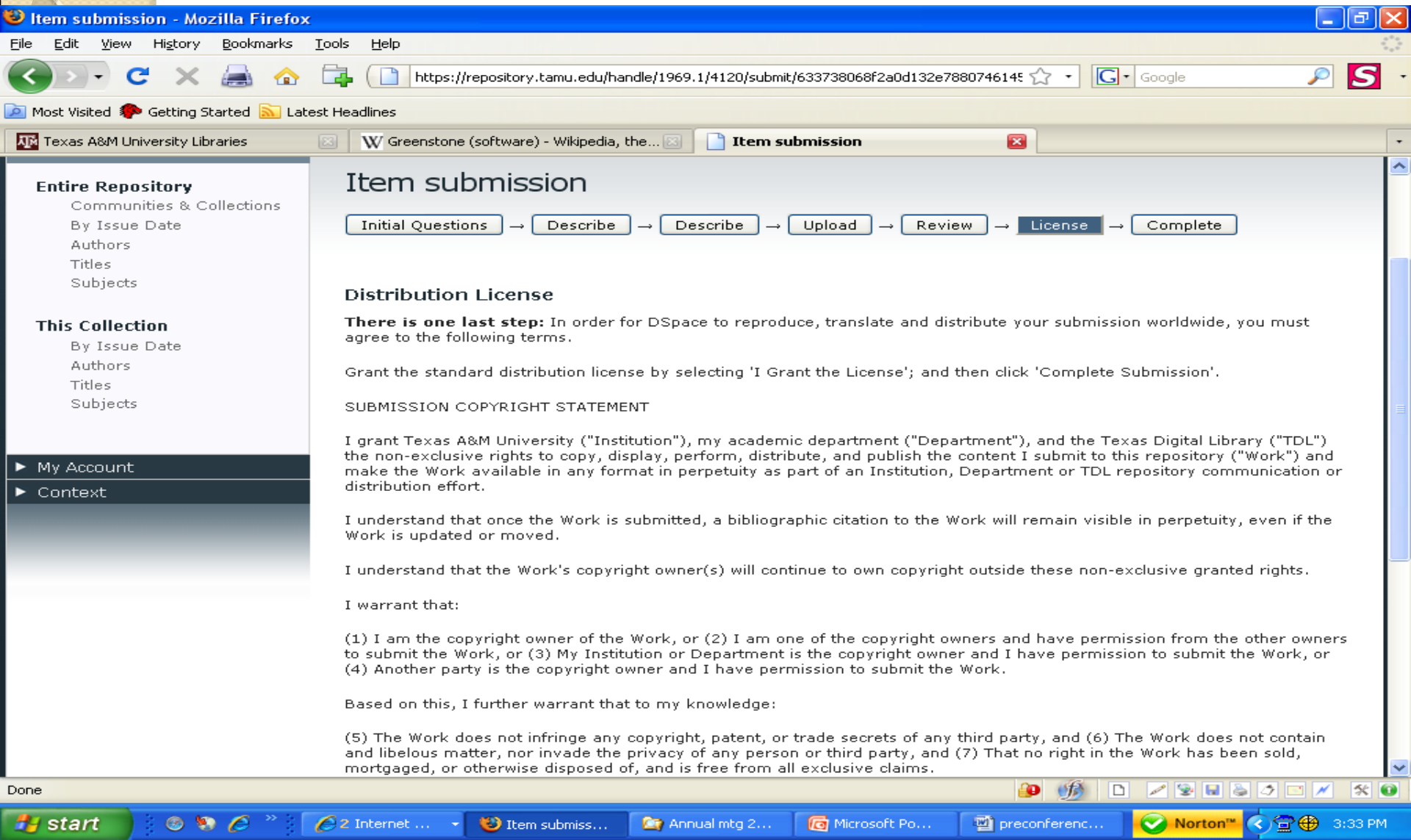
Browse to File Locations & Add File Description



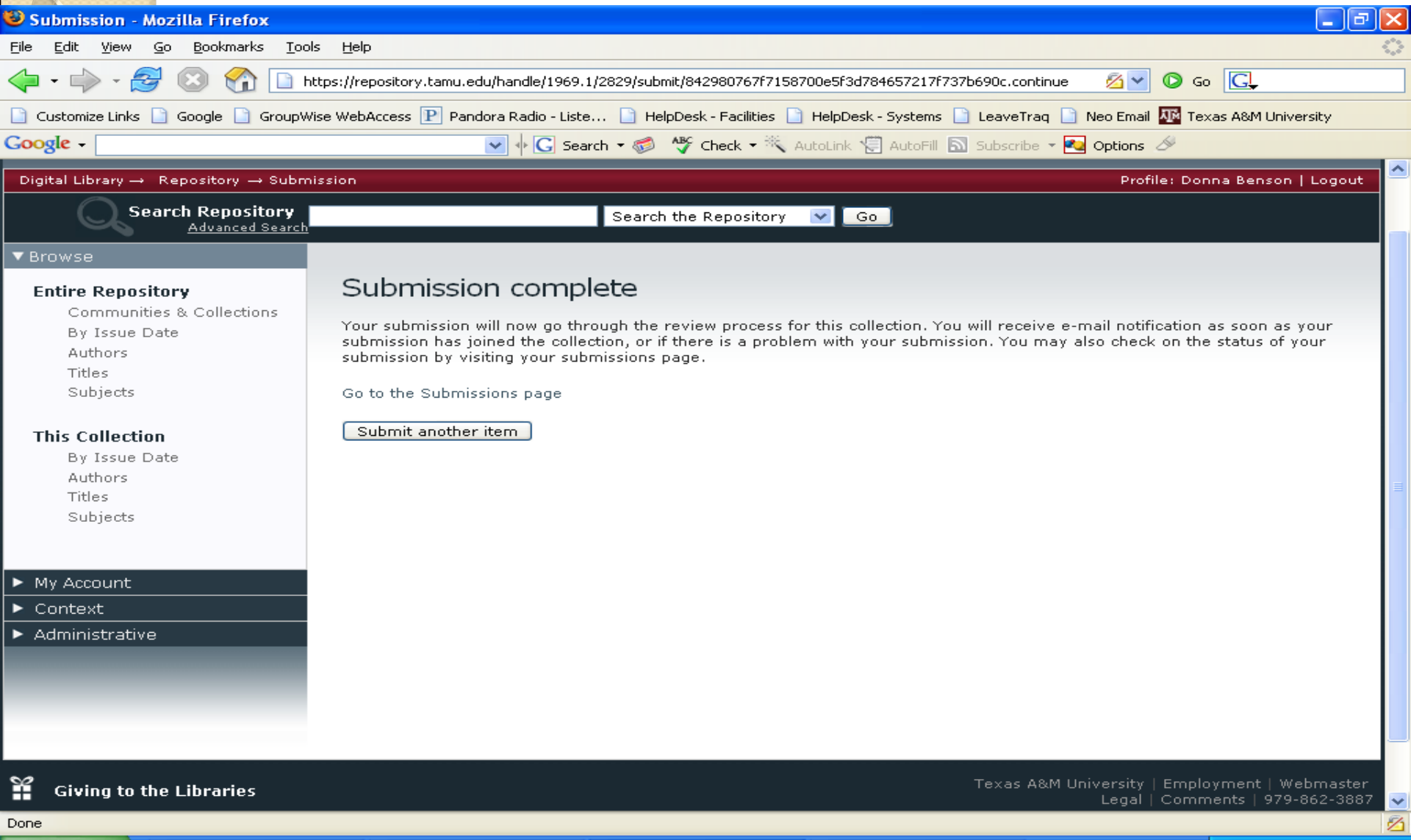
Review Item


Item submission - Mozilla Firefox
File Edit View History Bookmarks Tools Help
https://repository.tamu.edu/handle/1969.1/4120/submit/7a321f315e6a5e29504028178a3
Most Visited Getting Started Latest Headlines
Texas A&M University Libraries
Greenstone (software) - Wikipedia, the...
Item submission
Entire Repository
Communities & Collections
By Issue Date
Authors
Titles
Subjects
This Collection
By Issue Date
Authors
Titles
Subjects
My Account
Context
Item submission
Initial Questions → Describe → Describe → Upload → Review → License → Complete
Initial Questions
Multiple titles: No
Published: Yes
Correct one of these
Describe Item
Authors: McGeachin, Robert B.
Title: AgNIC Pre-conference 2009 "If It's Digital and In Google – Then They Will Come"
Date of Issue: 2009-04-21
Publisher: -
Citation: -
Series/Report No.: -
Identifiers: -
Type: -
Done
start
2 Internet ...
Item submiss...
Annual mtg 2...
Microsoft Po...
preconferenc...
Norton™
3:31 PM

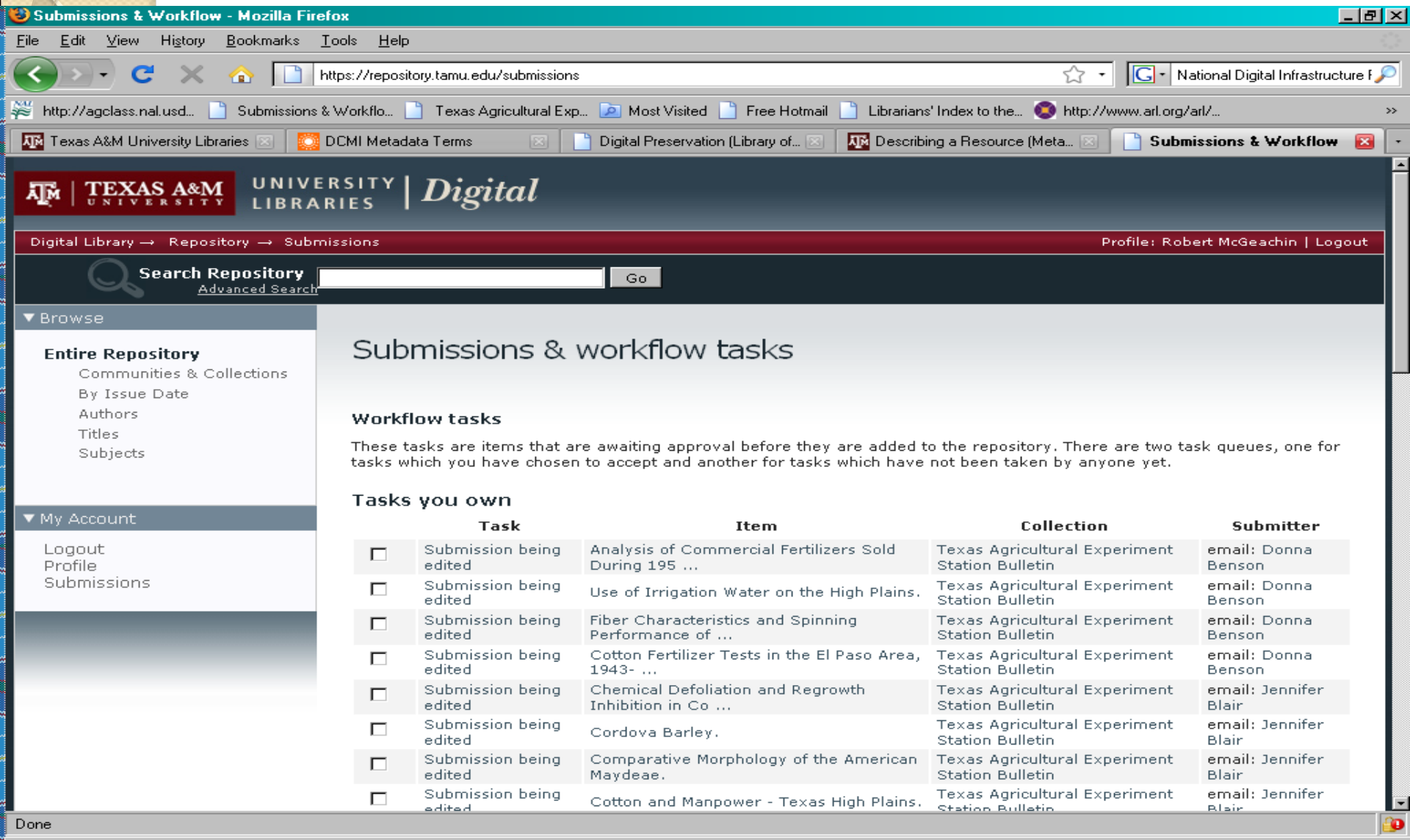
Grant Distribution License



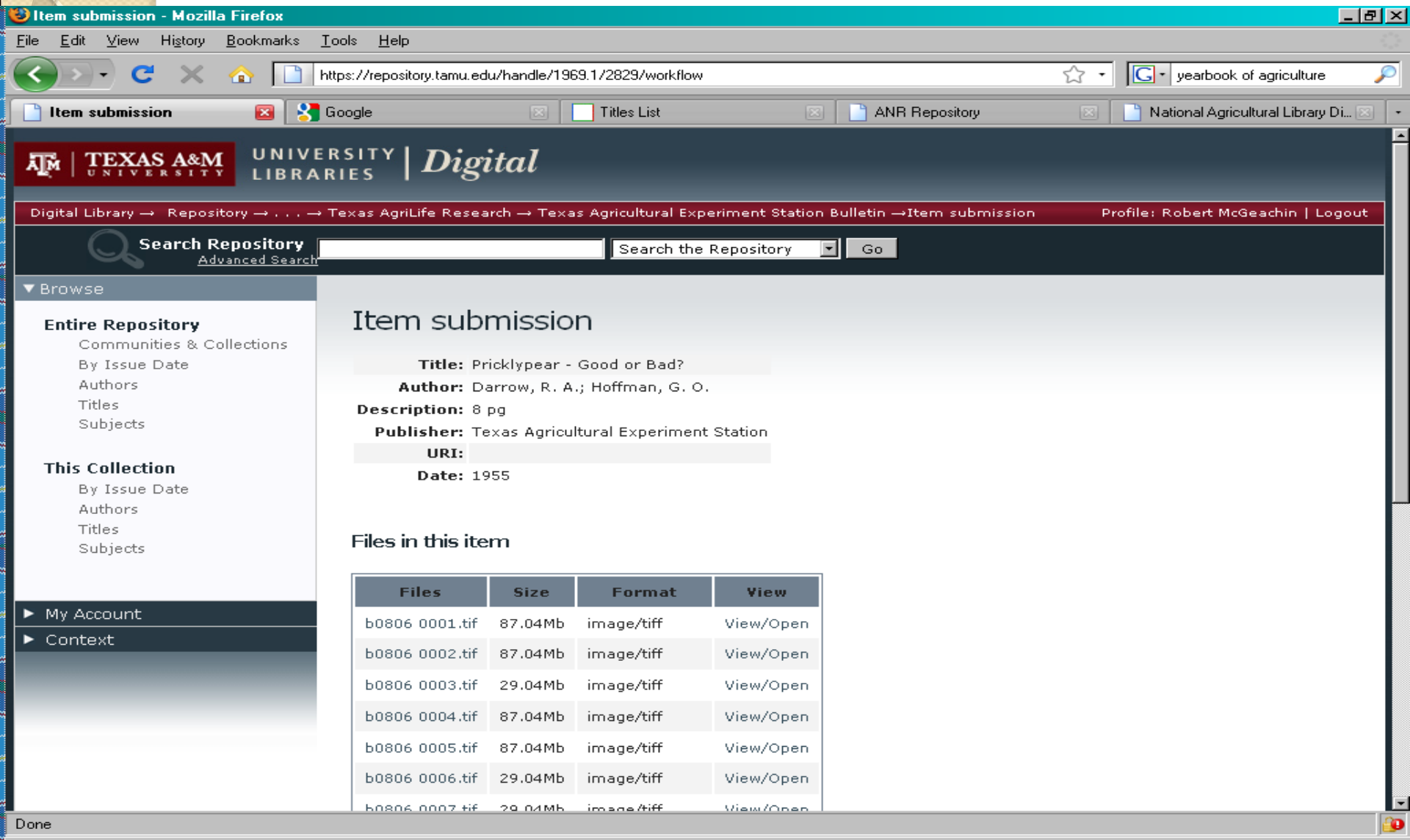
Complete Submission



Quality Control Check



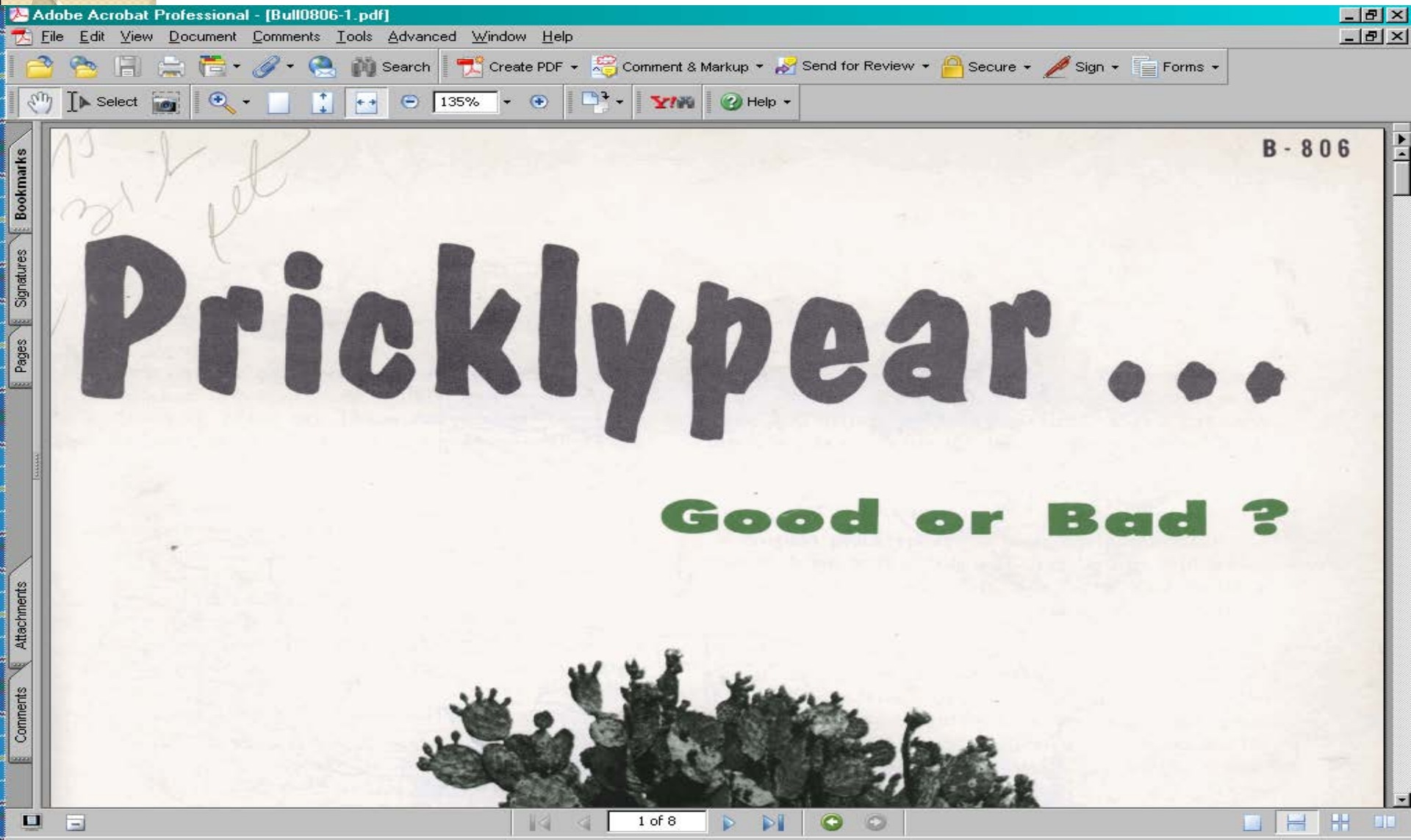
Editor Examines Item Submission



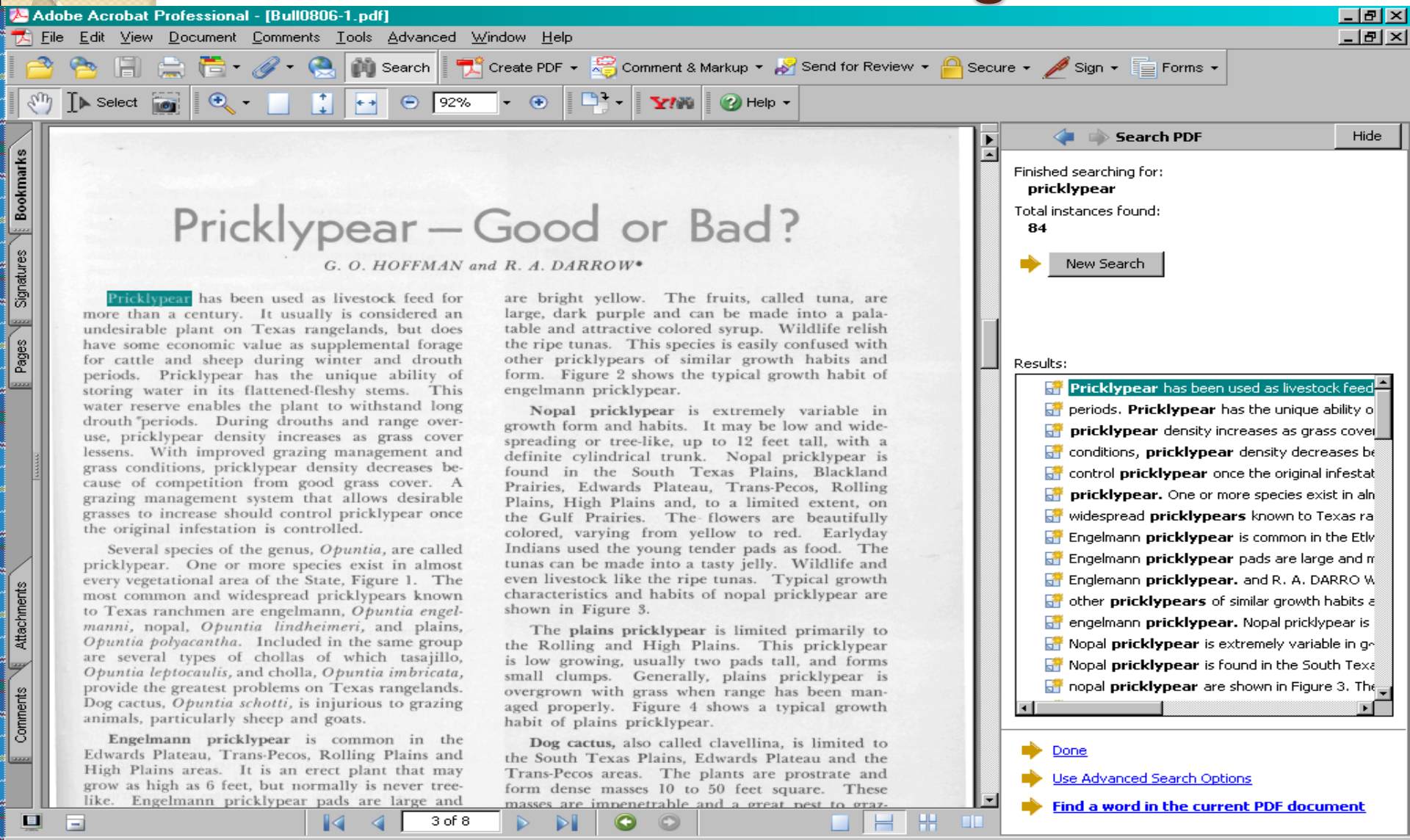
Editor Examines Item Submission

- Check Metadata
- Check for all image files
- Examine PDF version
  - Check for all pages in order and proper orientation
  - Check images cover full page of content
  - Check quality of page images
  - Check full text in "background" from OCR

Editor Examines Item Submission – PDF



Editor Examines Item Submission – PDF Indexing



Check Page Orientation

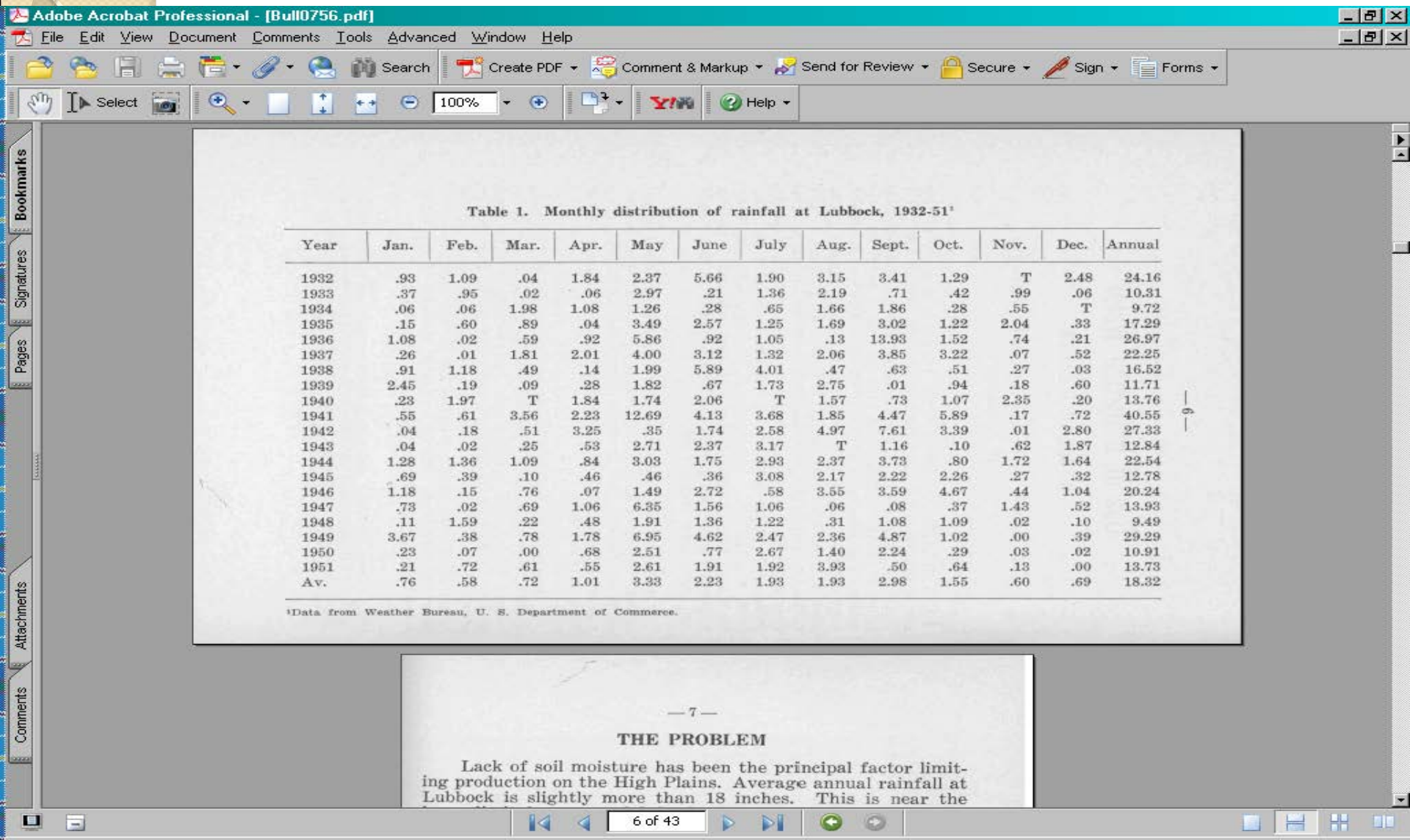


Table 1. Monthly distribution of rainfall at Lubbock, 1932-51[1]

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Annual
1932	.93	1.09	.04	1.84	2.37	5.66	1.90	3.15	3.41	1.29	T	2.48	24.16
1933	.37	.95	.02	.06	2.97	.21	1.36	2.19	.71	.42	.99	.06	10.31
1934	.06	.06	1.98	1.08	1.26	.28	.65	1.66	1.86	.28	.55	T	9.72
1935	.15	.60	.89	.04	3.49	2.57	1.25	1.69	3.02	1.22	2.04	.33	17.29
1936	1.08	.02	.59	.92	5.86	.92	1.05	.13	13.93	1.52	.74	.21	26.97
1937	.26	.01	1.81	2.01	4.00	3.12	1.32	2.06	3.85	3.22	.07	.52	22.25
1938	.91	1.18	.49	.14	1.99	5.89	4.01	.47	.63	.51	.27	.03	16.52
1939	2.45	.19	.09	.28	1.82	.67	1.73	2.75	.01	.94	.18	.60	11.71
1940	.23	1.97	T	1.84	1.74	2.06	T	1.57	.73	1.07	2.35	.20	13.76
1941	.55	.61	3.56	2.23	12.69	4.13	3.68	1.85	4.47	5.89	.17	.72	40.55
1942	.04	.18	.51	3.25	.35	1.74	2.58	4.97	7.61	3.39	.01	2.80	27.33
1943	.04	.02	.25	.53	2.71	2.37	3.17	T	1.16	.10	.62	1.87	12.84
1944	1.28	1.36	1.09	.84	3.03	1.75	2.93	2.37	3.73	.80	1.72	1.64	22.54
1945	.69	.39	.10	.46	.46	.36	3.08	2.17	2.22	2.26	.27	.32	12.78
1946	1.18	.15	.76	.07	1.49	2.72	.58	3.55	3.59	4.67	.44	1.04	20.24
1947	.73	.02	.69	1.06	6.35	1.56	1.06	.06	.08	.37	1.43	.52	13.93
1948	.11	1.59	.22	.48	1.91	1.36	1.22	.31	1.08	1.09	.02	.10	9.49
1949	3.67	.38	.78	1.78	6.95	4.62	2.47	2.36	4.87	1.02	.00	.39	29.29
1950	.23	.07	.00	.68	2.51	.77	2.67	1.40	2.24	.29	.03	.02	10.91
1951	.21	.72	.61	.55	2.61	1.91	1.92	3.93	.50	.64	.13	.00	13.73
Av.	.76	.58	.72	1.01	3.33	2.23	1.93	1.93	2.98	1.55	.60	.69	18.32

[1]Data from Weather Bureau, U. S. Department of Commerce.

— 7 —

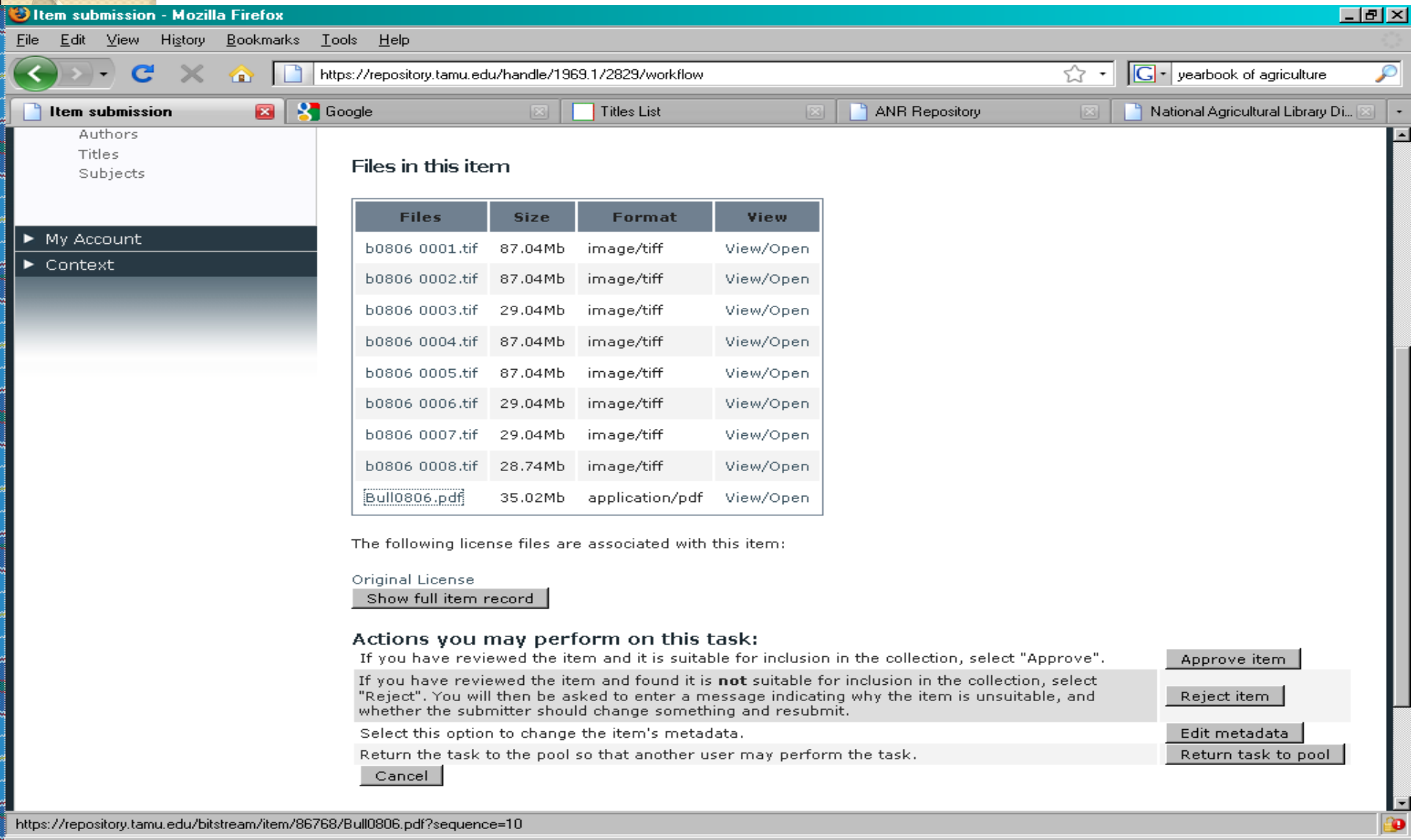
THE PROBLEM

Lack of soil moisture has been the principal factor limiting production on the High Plains. Average annual rainfall at Lubbock is slightly more than 18 inches. This is near the

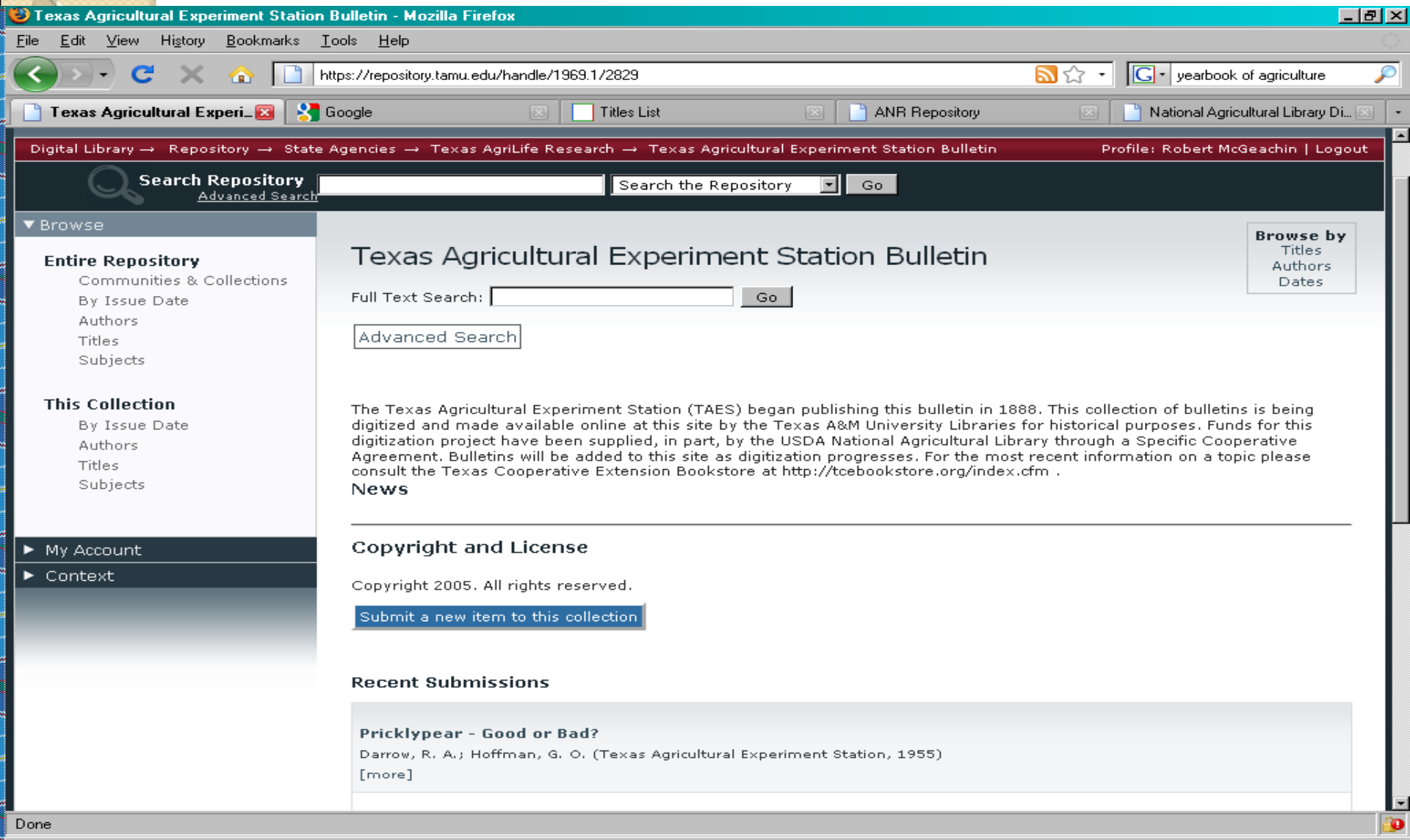
Editor Options

- Approve Item
- Reject Item (removes Item from Repository)
- Edit Metadata (to correct errors)
- Return Task to Pool of Submissions

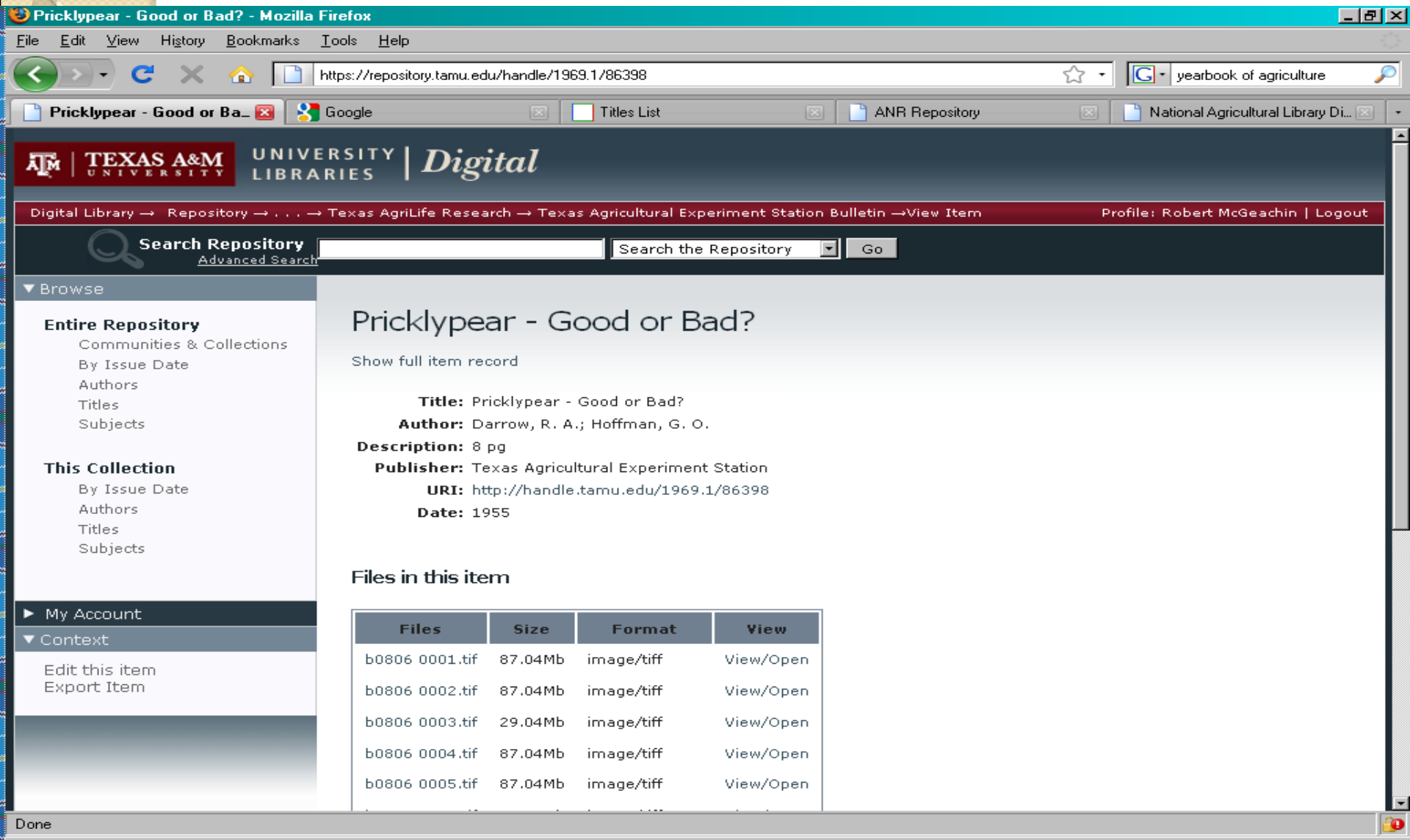
Editor Options



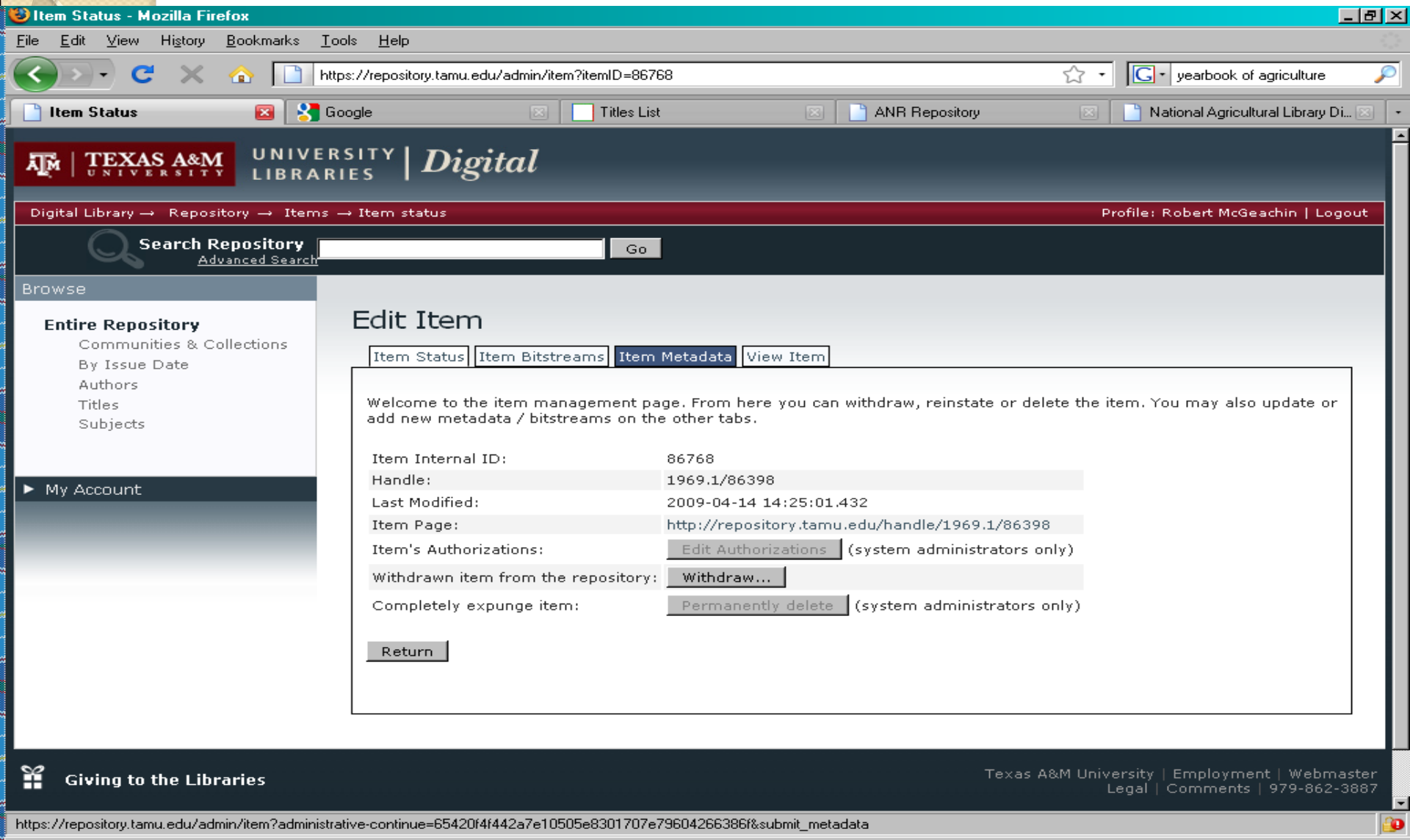
Publishes Item in Repository



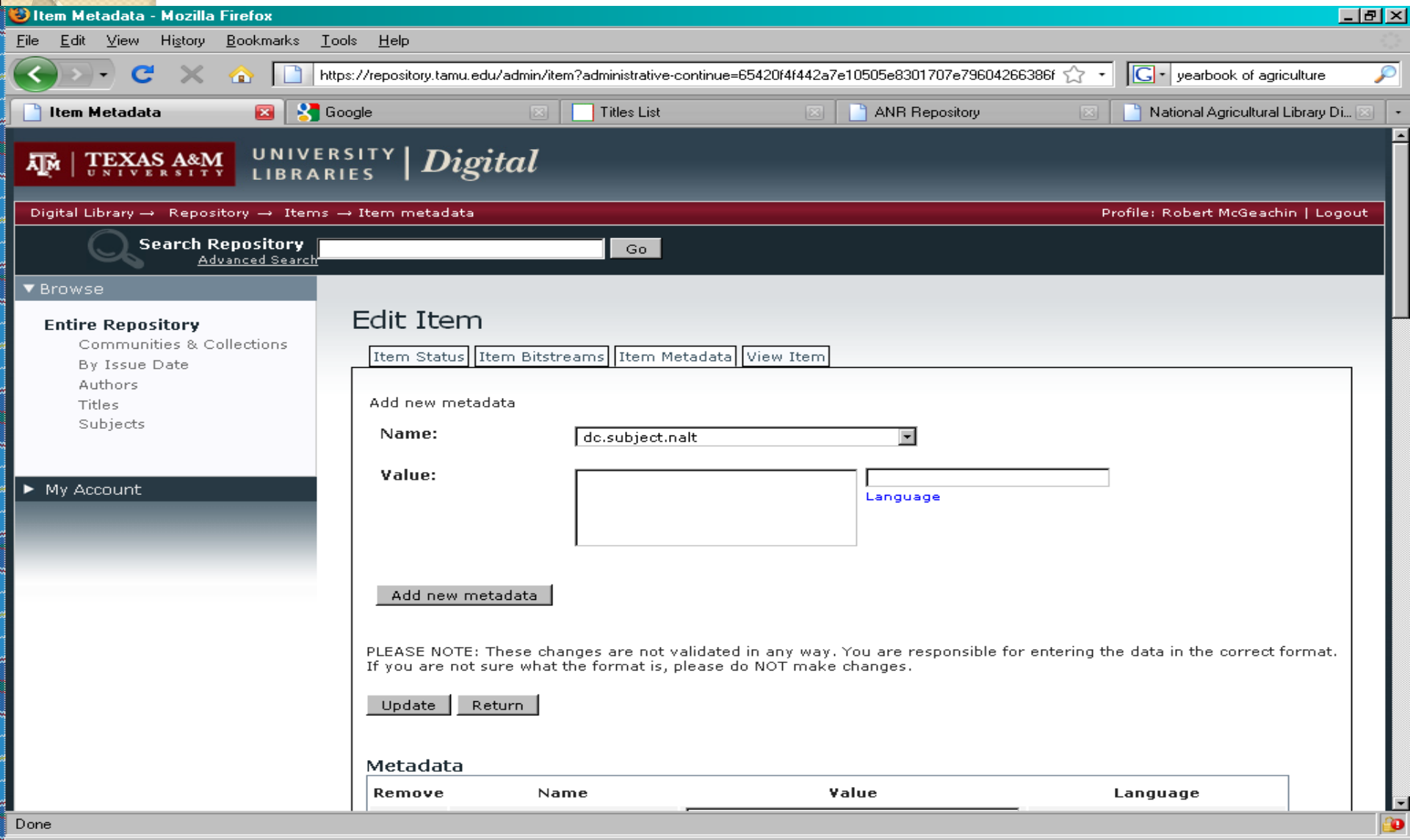
Item in Repository



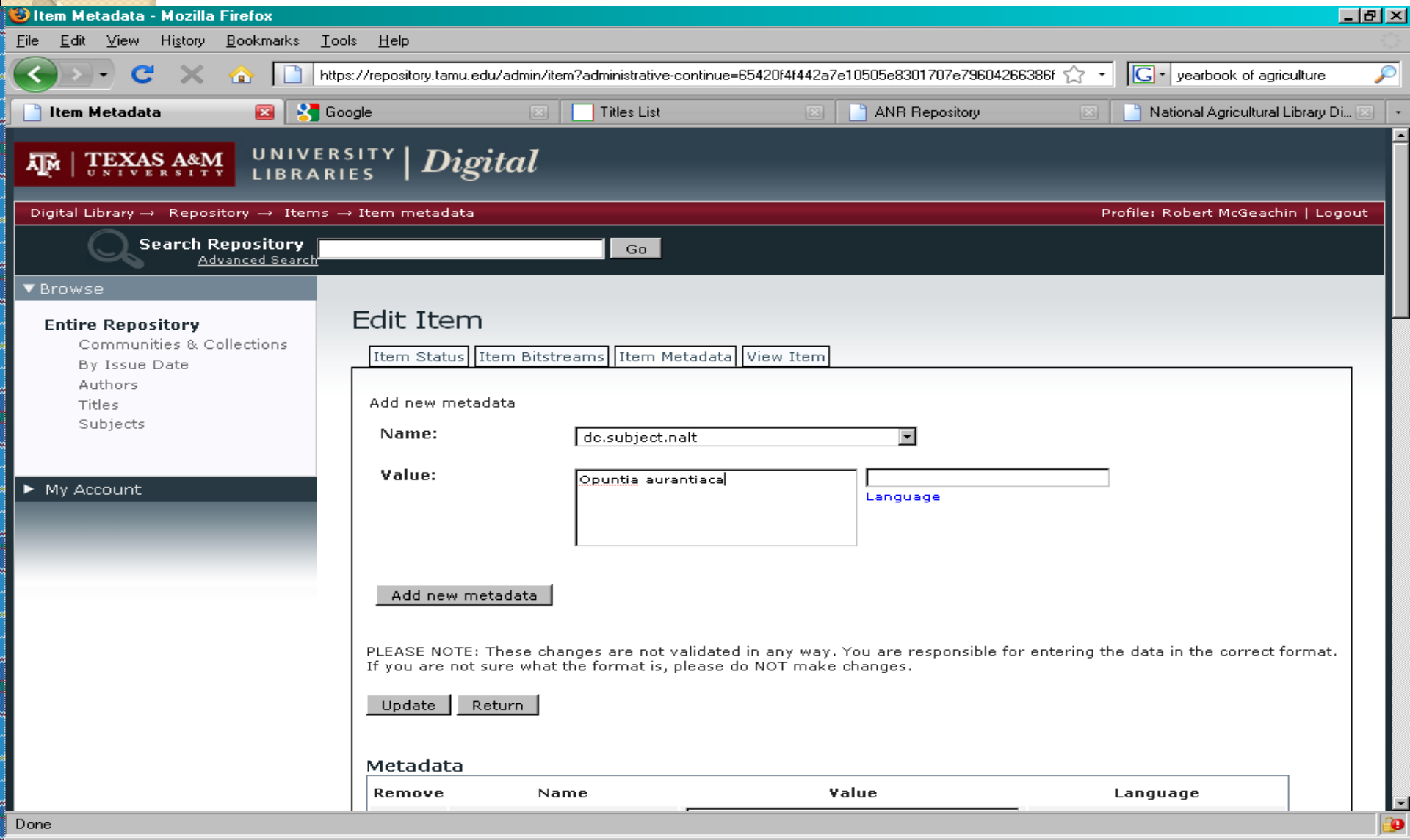
Post Publication Editing



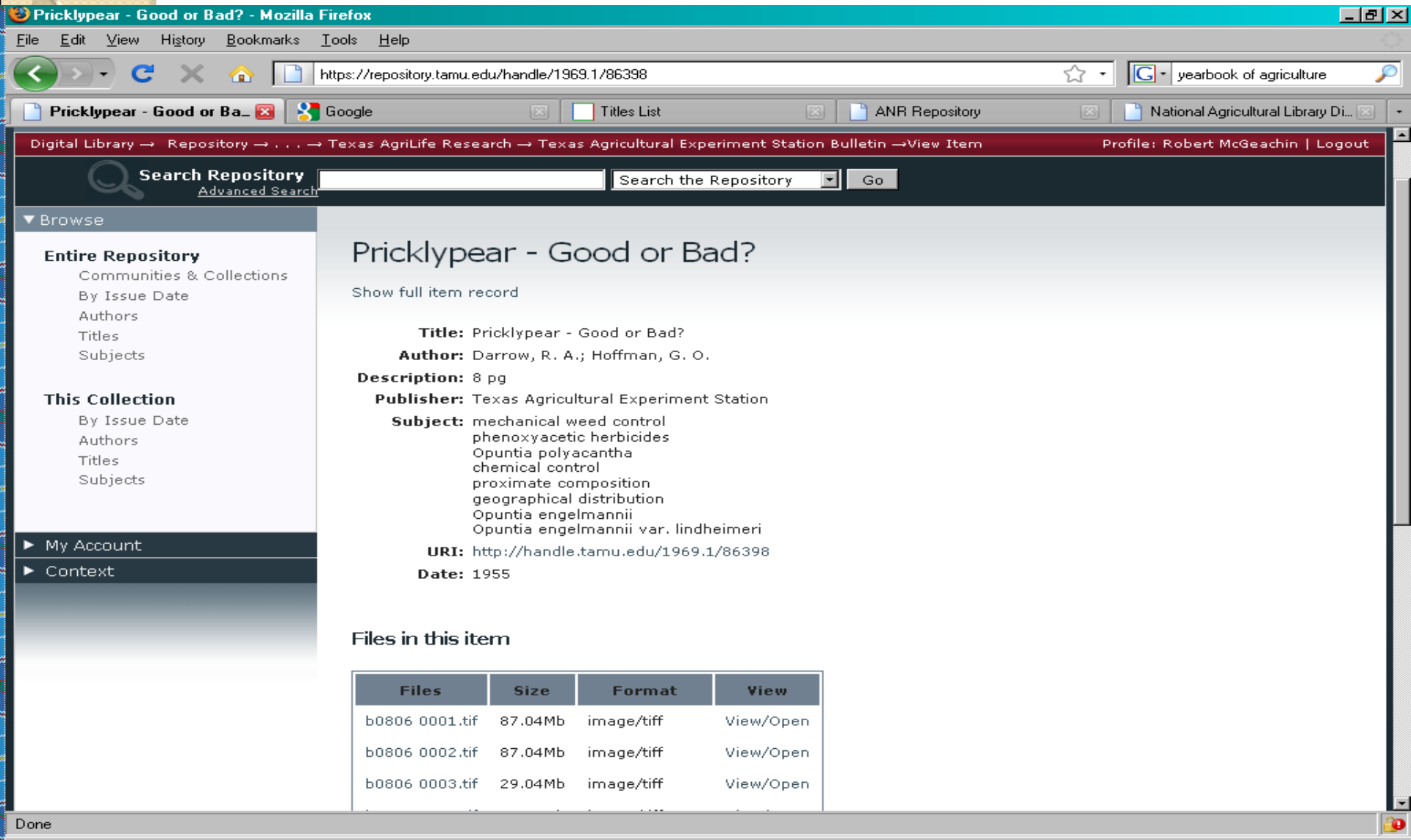
Edit Metadata



Add NALT Subjects



Final Item with Metadata



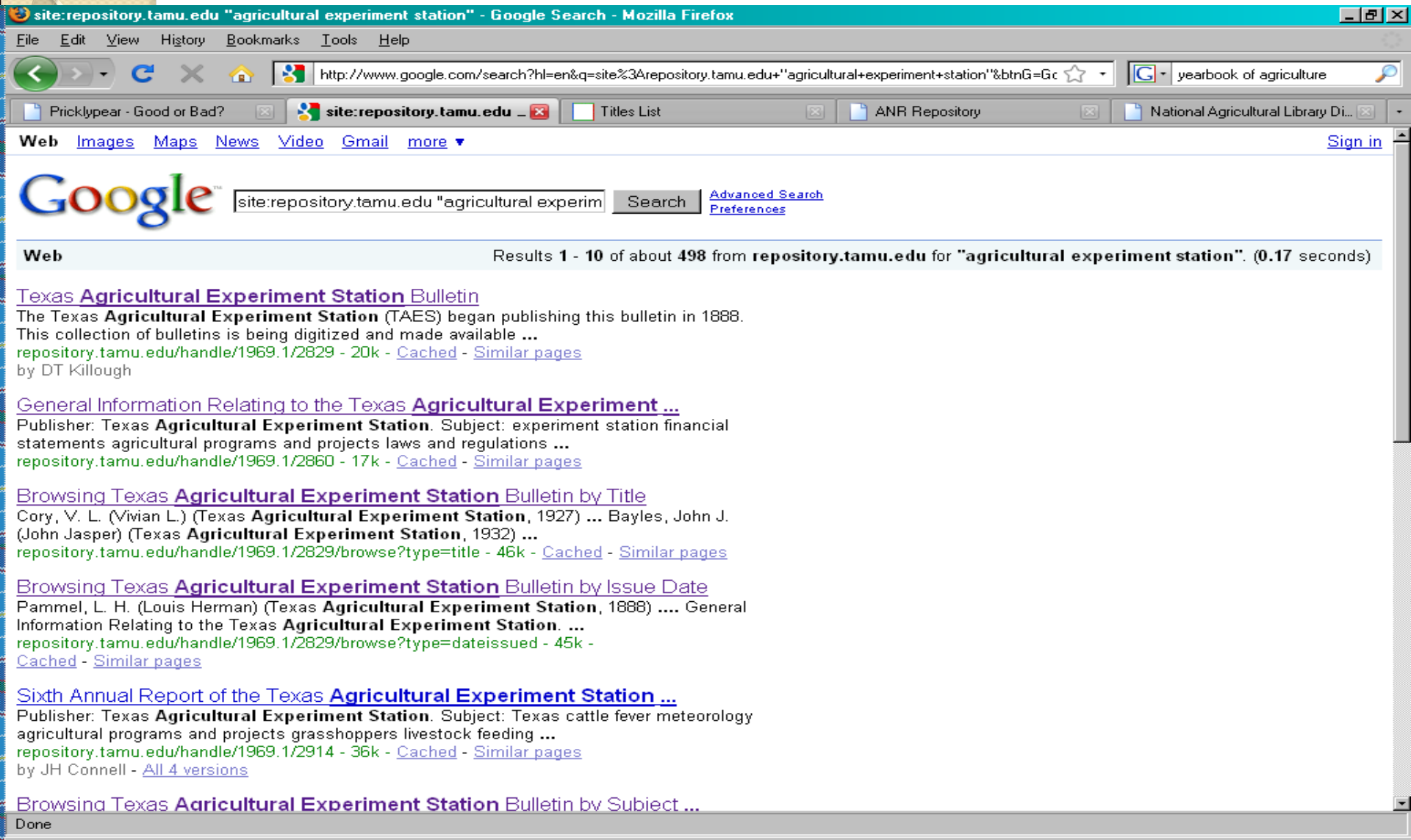
Access to Items in Repository

- Open Archives Initiative (OAI) Harvestable Metadata
- Makes item records harvestable by web crawlers
- Items indexed by many search engines
  - Commercial – Google, Yahoo, etc.
  - Specialized searches – AgOAI, NDLA, etc.

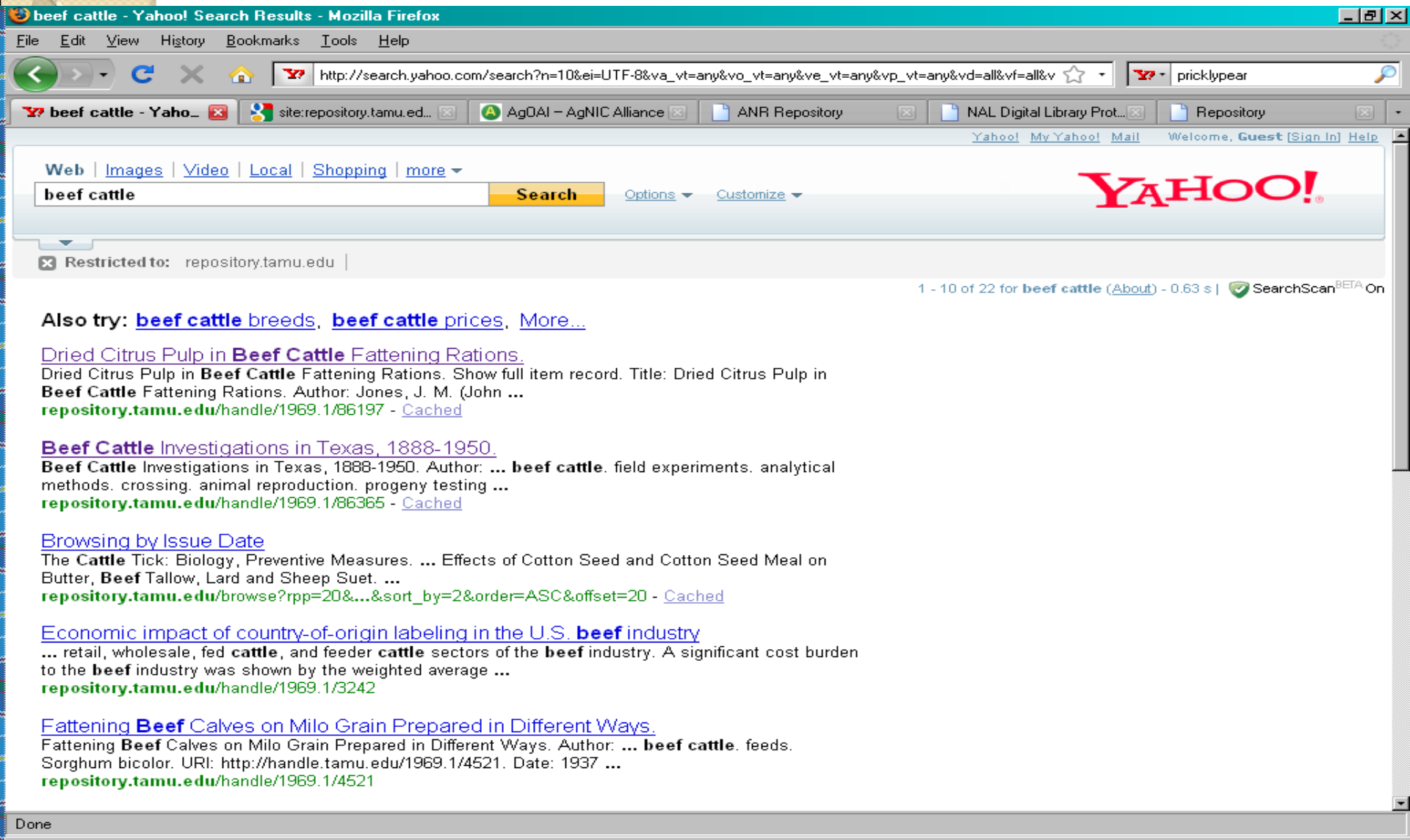
Commercial Search Engines

- Google
  - Search -- site:repository.tamu.edu "agricultural experiment station“
- Yahoo
  - Advanced Search – limit to only search in this domain/site

Google Search Results Example



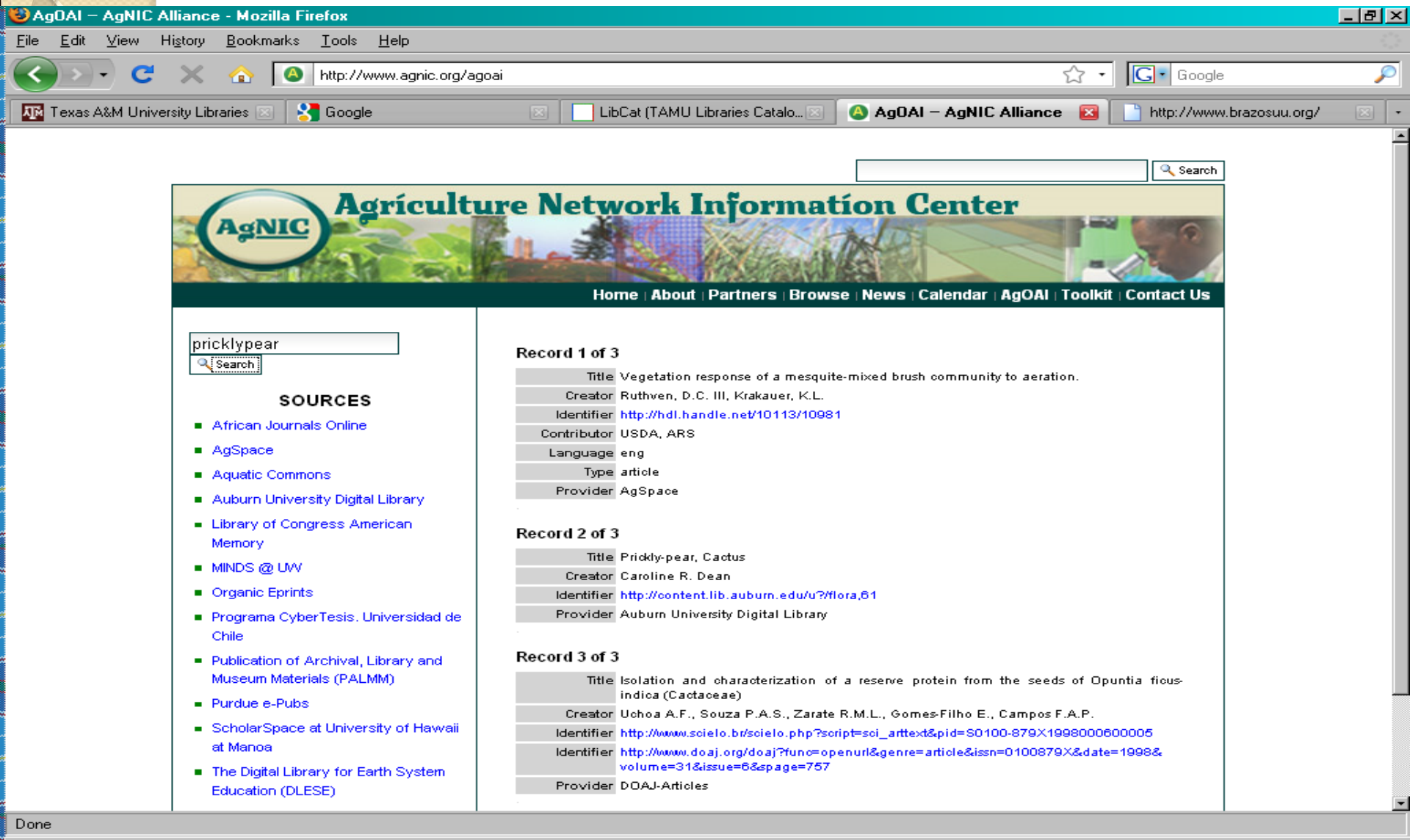
Yahoo Search Results Example



Agriculture OAI Searches

- AgOAI
  - http://www.agnic.org/agoai
- National Digital Library for Agriculture
  - http://www.nal.usda.gov/ndla/

AgOAI Example

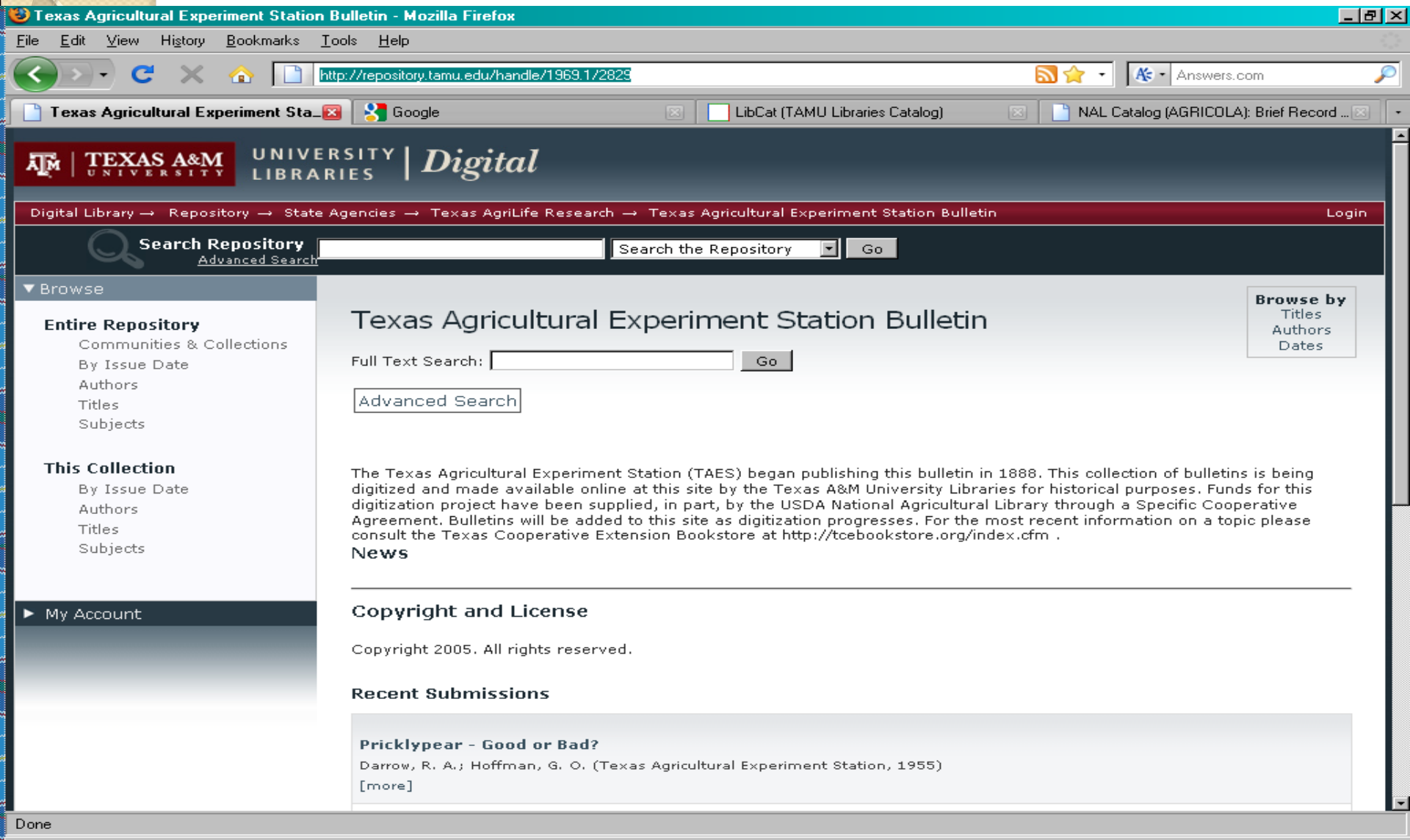


National Digital Library for Agriculture Example

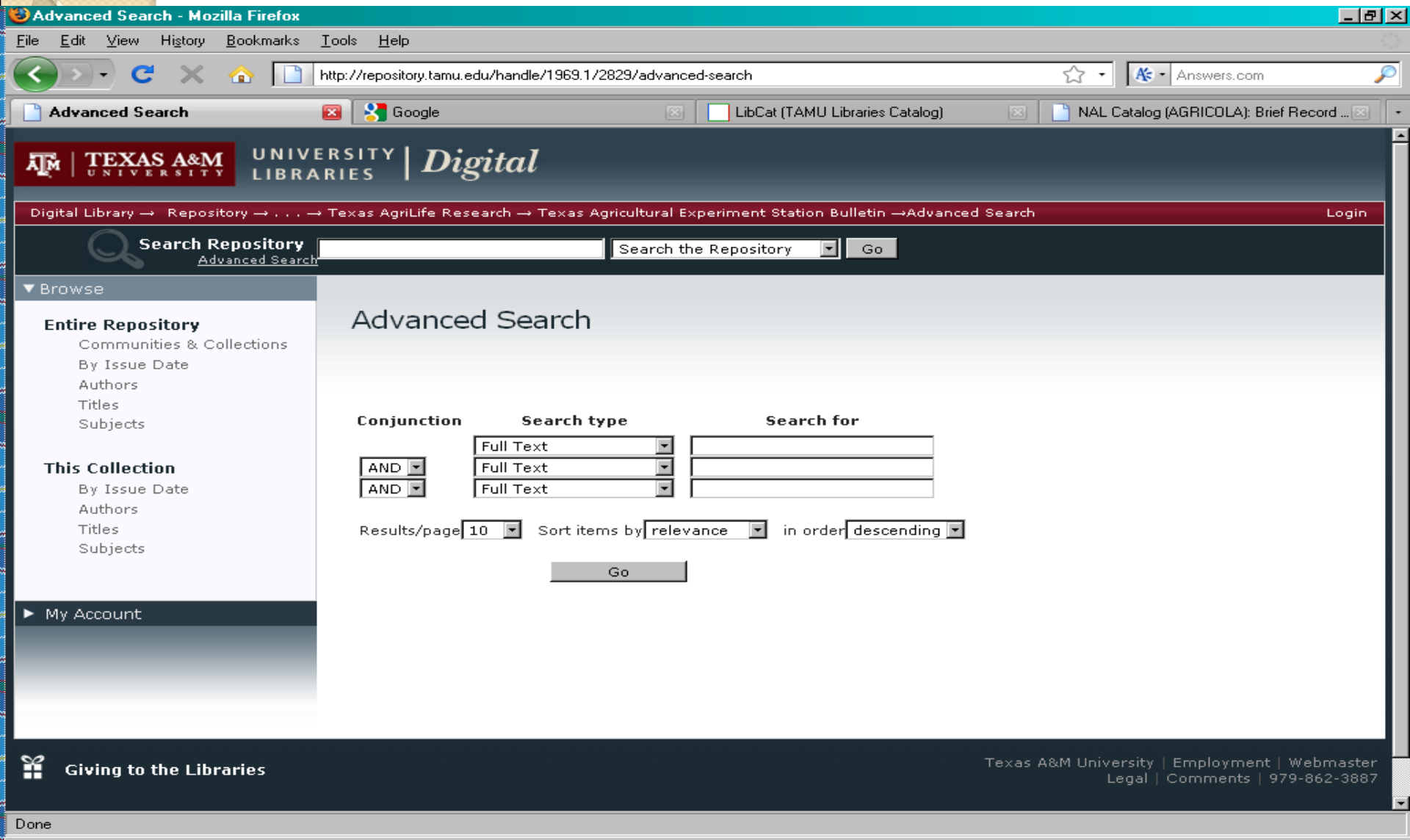
Access to Items in Repository

- Repository search and browse
  - Title
  - Collection
  - Date
  - Author
  - Subjects
- Repository assigns a unique URI for perpetual access
  - http://repository.tamu.edu/handle/1969.1/2829

Full Text Search in a Repository



Advanced Search of Repository



Summary Points

- Many libraries are becoming "publishers" of their public domain and archival materials in digital repositories.
- Knowledge of techniques to scan, digitize, preserve, and provide digital access to our materials is becoming a part of "information literacy" for librarians.

Questions about Repositories or Digitization?