

Greg Wiedeman

University Archivist
University at Albany, SUNY
@gregwiedeman

Born-Digital Photography at UAlbany

- Campus Photographer in Digital Media Department
 - 134 events in 2014, around 3-300 images per event
- Camera raw files (.NEF, .CR2)
- JPG derivatives
- Images go back to 1999

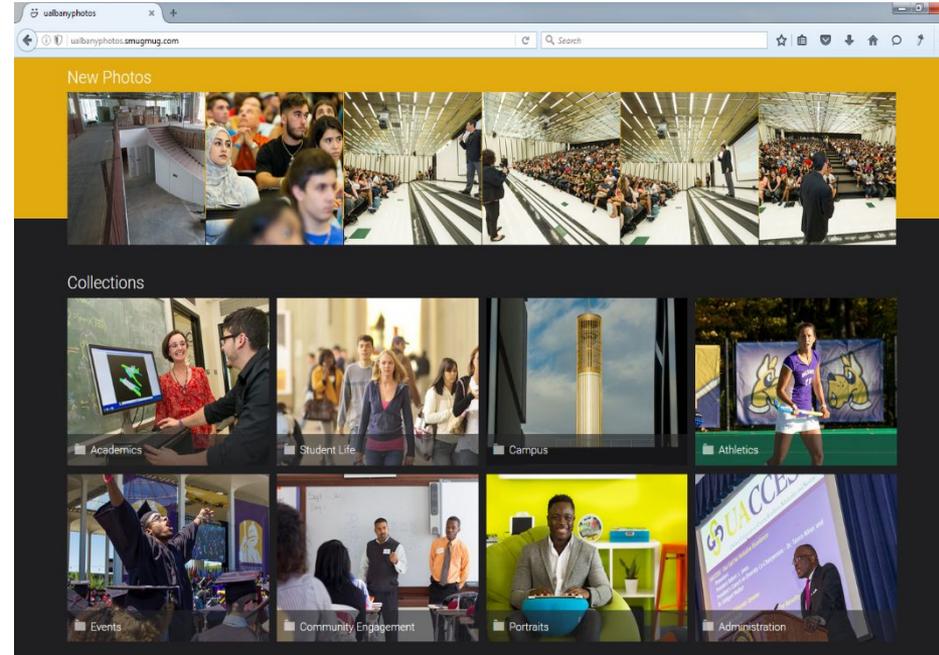
Disks in Boxes

- 4 boxes, 598 DVDs and CD-Rs
- 1.8 TB
- In folders by Job Number
- Subfolders have minimal description
- 1999-2008 Access Database
 - Has descriptions
- 2008-2012 REST DB
 - Dates, no descriptions



Born-Digital Photography at UAlbany

- Implemented SmugMug service in 2012
 - Online public photo database
 - Over 19,000 images
- Uploads and enters metadata in SmugMug



Immigration Symposium at Rockefeller Institute of Government

U.S. Immigration Reform and State-level Immigration Policymaking symposium at the Rockefeller Institute of Government.

This symposium considers opportunities and challenges presented by immigration and national immigration policy reforms with a focus on states and localities, particularly NYS and New York's cities. Featured guests include Michele Waslin, Pew Center; Neil Ruiz, Brookings Institution; and Jorge Montalvo, NYS Office of New Americans.



Principles

- Automation
 - Need to scale
 - No metadata creation, must describe themselves
- Standardization
 - Format-independent tools and utilities for born-digital records
- Transparency
 - Researchers need context
- Access
 - No restrictions, immediate public access

SmugMug API

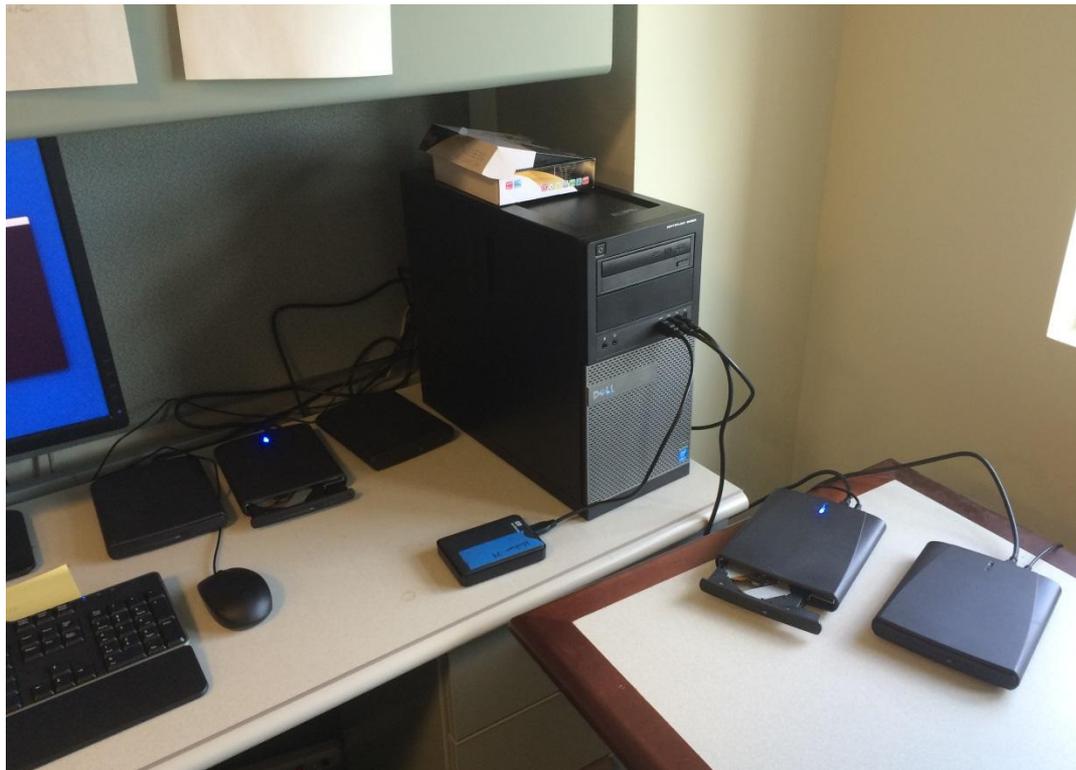
```
<album name="Collins Circle Flag Half Mast" uri="/api/v2/album/tXw8pw">
  <image>
    <uri>/api/v2/album/tXw8pw/image/GbPX8qd-0</uri>
    <filename>Collins Circle_Flag Half Mast_0025.JPG</filename>
    <date>2015-09-11T18:16:41+00:00</date>
    <weburi>http://www.ualbanyphotos.com/Campus/Aerials/Collins-Circle-Flag-Half-Mast
    <lastupdated>2015-09-11T18:17:41+00:00</lastupdated>
    <archivedmd5>194a181a01b7d9ff388ad2b4c7a09970</archivedmd5>
    <thumbnailurl>https://photos.smugmug.com/Campus/Aerials/Collins-Circle-Flag-Half-Mast
    <caption>DCIM\100MEDIA\DJI_0025.JPG Photo by Brian Busher</caption>
    <keywords>aerial; collins circle; 911; flag; ualbany; 2015</keywords>
    <archiveduri>https://photos.smugmug.com/Campus/Aerials/Collins-Circle-Flag-Half-Mast
    <downloadTime type="posix">1467146745.27</downloadTime>
    <hash type="md5">success</hash>
  </image>
  <image>
    <uri>/api/v2/album/tXw8pw/image/CJTPg5H-0</uri>
```

Crawling SmugMug

- Develop crawler for SmugMug
 - Download all images
 - Periodically crawl for updates
 - Hash index to see if already downloaded
 - Package into standard SIPs with metadata
 - After approval, automatically incorporate into EAD files and make publically available

Mass Image DVDs

```
22 def driveListen(drivePath, disk):
23
24     while disk == False:
25
26         try:
27             #check if disk has size|
28             get_file_size(drivePath)
29             disk = True
30         except:
31             disk = False
32
33     #do stuff with disk
34     print "Disk found, ripping disk"
35
36     #image disk
37     start = time.time()
38     subprocess.call(["dd", "if=/dev/" + driveName,
39                     end = time.time()
40                     elapsed = end - start
41                     print str(elapsed) + " seconds"
42                     print str(elapsed / 60) + " minutes"
43
44     print "eject disk"
45     subprocess.call(["eject", driveName], shell=False)
46
47     print "listen again"
48     disk = False
49     driveListen(drivePath, disk)
50
51 driveListen(drivePath, disk)
52
```



Issues with Disk Imaging at Scale

- Carve files with fiwalk and icat (TSK)
- Audit against fiwalk output
 - Batch 1: 49646 of 50212 – 98.87%
 - Batch 2: 47574 of 48030 – 99.05%
 - Batch 3: 22436 of 24530 – 91.46%
 - Batch 4: 49646 of 50212 – 98.87%
 - Total: 169302 of 172984 – 97.87%
- Convert with ImageMagik

```
bcadmin@S0582: ~/Documents/diskRip
8817280+0 records out
4514447360 bytes (4.5 GB) copied, 1963.44 s, 2.3 MB/s
1965.16476512 seconds
32.7527460853 minutes
eject disk
listen again
Disk found, ripping disk
8719424+0 records in
8719424+0 records out
4464345088 bytes (4.5 GB) copied, 1892.56 s, 2.4 MB/s
1894.15984018 seconds
31.569182363 minutes
eject disk
listen again
Disk found, ripping disk
8574880+0 records in
8574880+0 records out
4389928960 bytes (4.4 GB) copied, 1962.42 s, 2.2 MB/s
1963.51571107 seconds
32.7252618512 minutes
eject disk
listen again
Disk found, ripping disk
9072256+0 records in
9072256+0 records out
4644995072 bytes (4.6 GB) copied, 1855.07 s, 2.5 MB/s
1856.08229804 seconds
30.9347049673 minutes
eject disk
listen again
Disk found, ripping disk
8756480+0 records in
8756480+0 records out
448317760 bytes (4.5 GB) copied, 1872.61 s, 2.4 MB/s
1872.814013 seconds
31.213668834 minutes
eject disk
listen again
Disk found, ripping disk
8447232+0 records in
8447232+0 records out
4324982784 bytes (4.3 GB) copied, 1526.87 s, 2.8 MB/s
1538.78855999 seconds
25.5118093332 minutes
eject disk
listen again
Disk found, ripping disk
8152064+0 records in
8152064+0 records out
4173856768 bytes (4.2 GB) copied, 1833.42 s, 2.3 MB/s
1836.8689518 seconds
30.61448253 minutes
eject disk
listen again
Disk found, ripping disk
8309056+0 records in
8309056+0 records out
4254236672 bytes (4.3 GB) copied, 1872.51 s, 2.3 MB/s
1878.17832303 seconds
31.3029720505 minutes
eject disk
listen again
Disk found, ripping disk
8993216+0 records in
8993216+0 records out
4604526592 bytes (4.6 GB) copied, 2094.6 s, 2.2 MB/s
2095.3096509 seconds
34.921827515 minutes
eject disk
listen again
Disk found, ripping disk
```

```
bcadmin@S0582: ~/Documents/diskRip
9127536+0 records out
4670226432 bytes (4.7 GB) copied, 1834.07 s, 2.5 MB/s
1837.40557808 seconds
30.624468013 minutes
eject disk
listen again
Disk found, ripping disk
8976000+0 records in
8976000+0 records out
4307488000 bytes (4.3 GB) copied, 1674.08 s, 2.3 MB/s
1675.67888392 seconds
27.917946732 minutes
eject disk
listen again
Disk found, ripping disk
8375872+0 records in
8375872+0 records out
4288046464 bytes (4.3 GB) copied, 1783.39 s, 2.4 MB/s
1783.88978189 seconds
29.7313463648 minutes
eject disk
listen again
Disk found, ripping disk
9111744+0 records in
9111744+0 records out
4655212928 bytes (4.7 GB) copied, 2006.77 s, 2.3 MB/s
2008.15444303 seconds
33.4692407171 minutes
eject disk
listen again
Disk found, ripping disk
8732544+0 records in
8732544+0 records out
4471862528 bytes (4.5 GB) copied, 1958.87 s, 2.3 MB/s
1959.16627407 seconds
32.6527712345 minutes
eject disk
listen again
Disk found, ripping disk
```

```
bcadmin@S0582: ~/Documents/diskRip
9027840+0 records out
4622254080 bytes (4.6 GB) copied, 2027.47 s, 2.3 MB/s
2031.4105539 seconds
33.850497551 minutes
eject disk
listen again
Disk found, ripping disk
8439808+0 records in
8439808+0 records out
4273078072 bytes (4.3 GB) copied, 1792.27 s, 2.4 MB/s
1793.93482302 seconds
29.8985770504 minutes
eject disk
listen again
Disk found, ripping disk
7955136+0 records in
7955136+0 records out
428029632 bytes (4.1 GB) copied, 1917.11 s, 2.1 MB/s
1918.11013103 seconds
31.9685021838 minutes
eject disk
listen again
Disk found, ripping disk
8864896+0 records in
8864896+0 records out
4538826752 bytes (4.5 GB) copied, 1771.31 s, 2.6 MB/s
1771.8801599 seconds
29.5313352664 minutes
eject disk
listen again
Disk found, ripping disk
9041088+0 records in
9041088+0 records out
4629037056 bytes (4.6 GB) copied, 2058.43 s, 2.2 MB/s
2058.64531898 seconds
34.3107553164 minutes
eject disk
listen again
Disk found, ripping disk
```

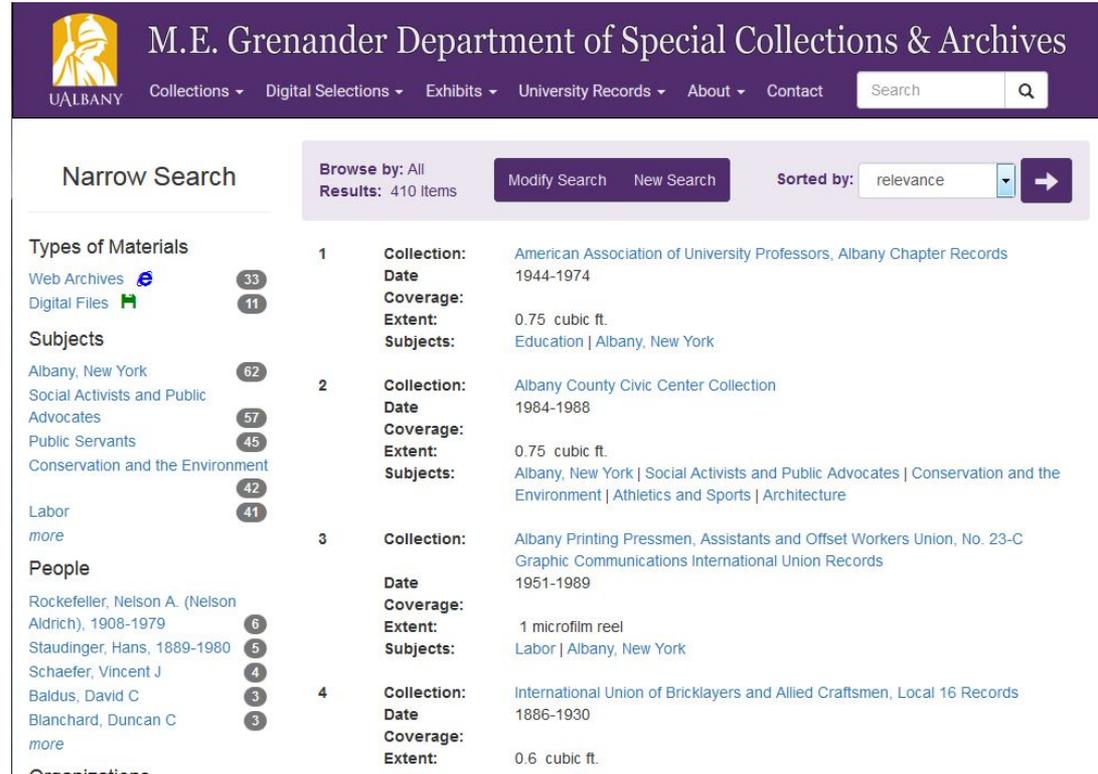
```
bcadmin@S0582: ~/Documents/diskRip
8689984+0 records out
4449271808 bytes (4.4 GB) copied, 1854.81 s, 2.4 MB/s
1856.14916801 seconds
30.9339194669 minutes
eject disk
listen again
Disk found, ripping disk
8222848+0 records in
8222848+0 records out
4210994176 bytes (4.2 GB) copied, 1729.68 s, 2.4 MB/s
1732.33043388 seconds
29.8754923979 minutes
eject disk
listen again
Disk found, ripping disk
9163504+0 records in
9163504+0 records out
4691755008 bytes (4.7 GB) copied, 2096.82 s, 2.2 MB/s
2097.48840785 seconds
34.9581401308 minutes
eject disk
listen again
Disk found, ripping disk
8968384+0 records in
8968384+0 records out
4862336 bytes (4.6 GB) copied, 1768.49 s, 2.6 MB/s
1769.07452893 seconds
29.5012421489 minutes
eject disk
listen again
Disk found, ripping disk
8968896+0 records in
8968896+0 records out
4587978752 bytes (4.6 GB) copied, 2032.73 s, 2.3 MB/s
2033.53113103 seconds
33.8921855172 minutes
eject disk
listen again
Disk found, ripping disk
```

Appraisal Decisions

- Not accept camera raw
 - Large, hard to make available
 - Proprietary
- Convert all files to JPG prior to accessioning
 - .CR2 Canon raw lossless or lossy JPG compression
 - .NEF Nikon proprietary lossless or lossy
 - 1.8 TB to 274 GB
 - Not using compression is not a preservation strategy
- Not spend time recovering files

Access

- New public access system
- Drupal, XTF, and static pages
- Bootstrap 3
- Schema.org
- Public domain
- Over 180,000 images



The screenshot displays the website interface for the M.E. Grenander Department of Special Collections & Archives. The header includes the department name and navigation links for Collections, Digital Selections, Exhibits, University Records, About, and Contact. A search bar is located in the top right corner.

The main content area features a 'Narrow Search' section on the left, which lists various categories with their respective item counts:

- Types of Materials
 - Web Archives (33)
 - Digital Files (11)
- Subjects
 - Albany, New York (62)
 - Social Activists and Public Advocates (57)
 - Public Servants (45)
 - Conservation and the Environment (42)
 - Labor (41)
- People
 - Rockefeller, Nelson A. (Nelson Aldrich), 1908-1979 (6)
 - Staudinger, Hans, 1889-1980 (5)
 - Schaefer, Vincent J (4)
 - Baldus, David C (3)
 - Blanchard, Duncan C (3)

On the right side, there are search controls including 'Browse by: All Results: 410 Items', 'Modify Search', 'New Search', and 'Sorted by: relevance'. Below these controls, four search results are displayed, each with a numbered list item, collection name, date, coverage, extent, and subjects.

Item #	Collection	Date	Coverage	Extent	Subjects
1	American Association of University Professors, Albany Chapter Records	1944-1974		0.75 cubic ft.	Education Albany, New York
2	Albany County Civic Center Collection	1984-1988		0.75 cubic ft.	Albany, New York Social Activists and Public Advocates Conservation and the Environment Athletics and Sports Architecture
3	Albany Printing Pressmen, Assistants and Offset Workers Union, No. 23-C Graphic Communications International Union Records	1951-1989		1 microfilm reel	Labor Albany, New York
4	International Union of Bricklayers and Allied Craftsmen, Local 16 Records	1886-1930		0.6 cubic ft.	

<http://meg.library.albany.edu:8080/archive/view?docId=ua395.xml>



Narrow Search

Search: vonnegut in keywords [\[Clear Search\]](#)

Results: 9 Items

[Modify Search](#)

[New Search](#)

Sorted by:



Types of Materials

Digital Files **1**

Subjects

Albany, New York **2**

Conservation and the Environment **2**

Schenectady, New York **2**

Athletics and Sports **1**

Literature **1**

[more](#)

People

Blanchard, Duncan C **3**

Schaefer, Vincent J **3**

Vonnegut, Bernard **3**

Langmuir, Irving, 1881-1957 **2**

Moore, Charles B., 1920-2010 **2**

[more](#)

Organizations

State University of New York. Atmospheric Sciences Research Center **4**

General Electric Company **3**

State University of New York at Albany **3**

Arthur D. Little, Inc **2**

- | | |
|----------|--|
| 1 | Collection: Bernard Vonnegut Papers
Date Coverage: 1828-1997
Extent: 39.04 cubic ft.
Subjects: Conservation and the Environment
Matches: ...Vonnegut, Bernard - Tornado Articles...
...Bernard Vonnegut Papers 1828-1997...
...Tribute to Bernard Vonnegut... |
| 2 | Collection: Vincent J. Schaefer Papers
Date Coverage: 1891-1993
Extent: 135 cubic ft.
Subjects: Schenectady, New York Conservation and the Environment Athletics and Sports
Matches: ...Vonnegut/Florida...
...Magdalena Vonnegut/Moore McLean...
...Magdalena Vonnegut/Moore McLean... |
| 3 | Collection: Duncan Blanchard Papers
Date Coverage: 1789-2003
Extent: 35.8 cubic ft.
Matches: ...2010 Schaefer, Vincent J. Vonnegut, Bernard Woodcock, Alfred...
...Schaefer, and Bernard Vonnegut. While Blanchard enjoyed...
...Schaefer and Bernard Vonnegut as a senior research associate... |
| 4 | Collection: Eugene McLaren Papers |

<http://meg.library.albany.edu:8080/archive/view?docId=ua395.xml>



Narrow Search

Search: vonnegut in **keywords** [\[Clear Search\]](#)
Digital Files in **genreform** [\[Clear Search\]](#)
Results: 1 Item

[Modify Search](#) [New Search](#)

Sorted by:

Types of Materials

[Digital Files](#)

Organizations

[State University of New York at Albany. Office of Communication and Marketing](#)

Repositories

[University at Albany M. E. Grenander Department of Special Collections and Archives](#)

1 **Collection:** [Digital Media Department Collection](#)
Date Coverage: 2002-2016
Extent: 219.75 GB 164854 Digital Files
Types: [Digital Files](#)
Matches: ...Bldg/ Red Room / Kurt [Vonnegut](#) 1/22/00 @ 5 PM reception...
1 hit

Open to the public Monday-Friday, 9am-5pm
Closed for state holidays and winter intersession
Located on the top floor of the Science Library on the Uptown Campus
[Directions](#)

M. E. Grenander Department of Special Collections & Archives
Science Library 350
1400 Washington Avenue
Albany, NY 12222

<http://meg.library.albany.edu:8080/archive/view?docId=ua395.xml>



1 occurrence of "vonnegut" [Clear Hits]

- Digital Media Department
- About the Collection ▾
- Collection History ▾
- Access and Use ▾

- Contents
- Academics ▾
- Administration ▾
- Athletics ▾
- Campus ▾
- CommunityEngagement ▾
- Events ▾ 1
- Events 1
- 2015 Welcome Week Events

<input type="checkbox"/>	<input type="checkbox"/>	NYS Writer's Institute: 1/22/00 @ 3:30 PM State Author/Poet Award Ceremony Capitol Bldg/ Red Room / Kurt Vonnegut 1/22/00 @ 5 PM reception Eida's on Lark St. Photo session 1/22/00 @ 7 PM Public Reading/Reception Page Hall color print ▾	2001 January 5
<input type="checkbox"/>	<input type="checkbox"/>	University Relations: 1/31/01 @ 11 AM Lecture Center Concourse Community Service Fair digital ▾	2001 February 9
<input type="checkbox"/>	<input type="checkbox"/>	University Relations: 3/15/01 @ 4:15 PM LC-23 Carl H. Rosner at the podium digital ▾	2001 March 29
<input type="checkbox"/>	<input type="checkbox"/>	University Relations: 3/19/01 @ 3 PM MSB-200 Roslyn Jefferson digital ▾	2001 April 3
<input type="checkbox"/>	<input type="checkbox"/>	University Relations: 9/12/01 @ Collins Circle/CC Fountain Unity Walk ▾	2001 September 26
<input type="checkbox"/>	<input type="checkbox"/>	University Relations: 9/21/01 @ 12:30 RACC Memorial Service digital and color print ▾	2001 October 5
<input type="checkbox"/>	<input type="checkbox"/>	Advancement: 10/30/01 @ 8 AM Standish Room Business Matter debut taping w/ Dick Highfield digital? ▾	2001 November 5
<input type="checkbox"/>	<input type="checkbox"/>	Advancement: 1/9/02 @ 7:15 AM Crowne Plaza Hotel Albany Boosting Business in NY State w/ Elliot Spitzer digital ▾	2002 January 1
<input type="checkbox"/>	<input type="checkbox"/>	Health Management: 1/15/02 @ 8 AM NYS Nurses Assoc. Latham digital imagery ▾	2002 January 22
<input type="checkbox"/>	<input type="checkbox"/>	Media and Marketing: 3/6/02 @ 2 - 3 PM CC Ballroom Public Interest Career Fair digital ▾	2002 March 18
<input type="checkbox"/>	<input type="checkbox"/>	Annual Fund: 3/12/02 @ 6:30 PM Annual Fund Telemarketing Room web shots digital ▾	2002 March 25
<input type="checkbox"/>	<input type="checkbox"/>	Media and Marketing: 3/14/02 @ 2 PM - 3/15/02 @ 12 Noon CC Ballroom "Leave no child behind" Conference Henry Louis Taylor / Terrence Thornberry digital ▾	2002 March 29
<input type="checkbox"/>	<input type="checkbox"/>	Stewardship: 4/11/02 @ 1 - 2 PM BA 231 Towers Perrin Gift announcement digital ▾	2002 April 4
<input type="checkbox"/>	<input type="checkbox"/>	Media and Marketing: 4/3/02 @ 10:30 AM State Capital Bldg. Horizon Campus Press Conference digital ▾	2002 April 17

<http://meg.library.albany.edu:8080/archive/view?docId=ua395.xml>



- Digital Media Department
- About the Collection ▾
- Collection History ▾
- Access and Use ▾

Request

- Contents
- Academics ▾
- Administration ▾
- Athletics ▾
- Campus ▾
- CommunityEngagement ▾
- Events ▾**
- Portraits ▾
- StudentLife ▾

← **NYS Writer's Institute: 1/22/00 @ 3:30 PM State Author/Poet Award Ceremony Capitol Bldg/ Red Room / Kurt Vonnegut 1/22/00 @ 5 PM reception Elda's on Lark St. Photo session 1/22/00 @ 7 PM Public Reading/Reception Page Hall color print** →

Database Entry

2001 January 5

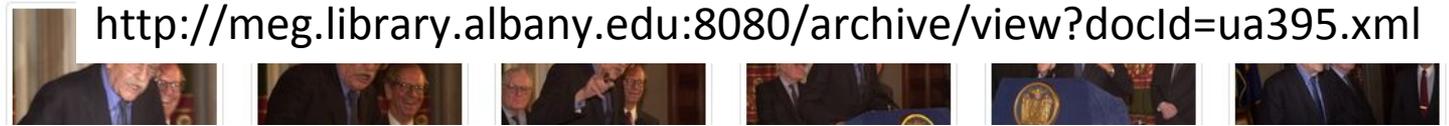
Processing Details

2016 July 21 - Imaged from optical media with dd, carved from image with fiwalk and icat, converted all raw to jpg

2016 July 21 - description record extracted from photographer's Microsoft Access database by Job number



<http://meg.library.albany.edu:8080/archive/view?docId=ua395.xml>



DSC_0033.JPG, 2001 January 23 (EXIF.DateTimeOriginal)



<http://meg.library.albany.edu:8080/archive/view?docId=ua395.xml>

