

The distant future of the recent past

The outlook for newspaper collections
in the born-digital era

From the archives

JPEG
corruption,
1997-2006



From the archives

JPEG
corruption,
1997-2006



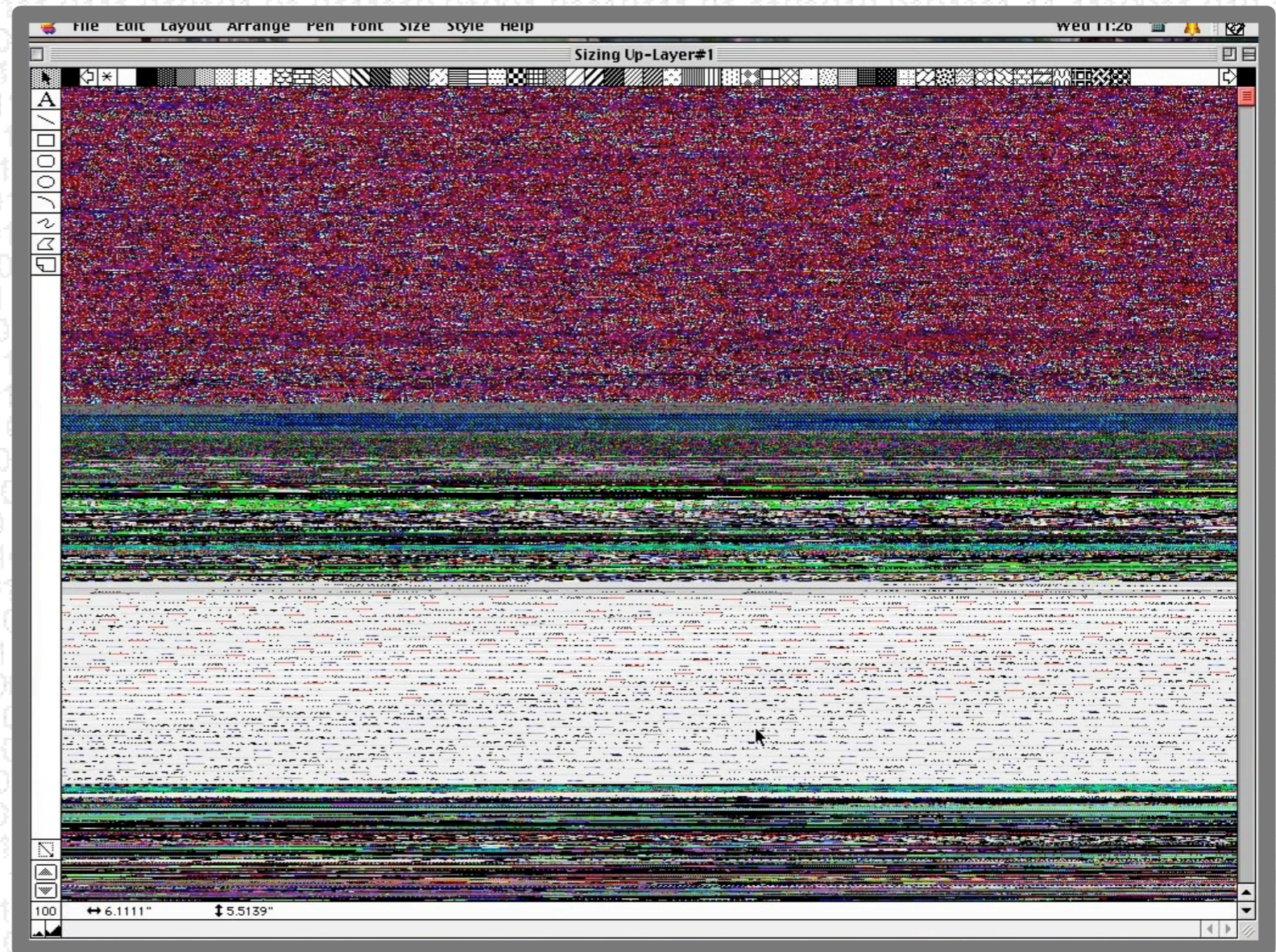
From the archives

Server
software
upgrade,
2005



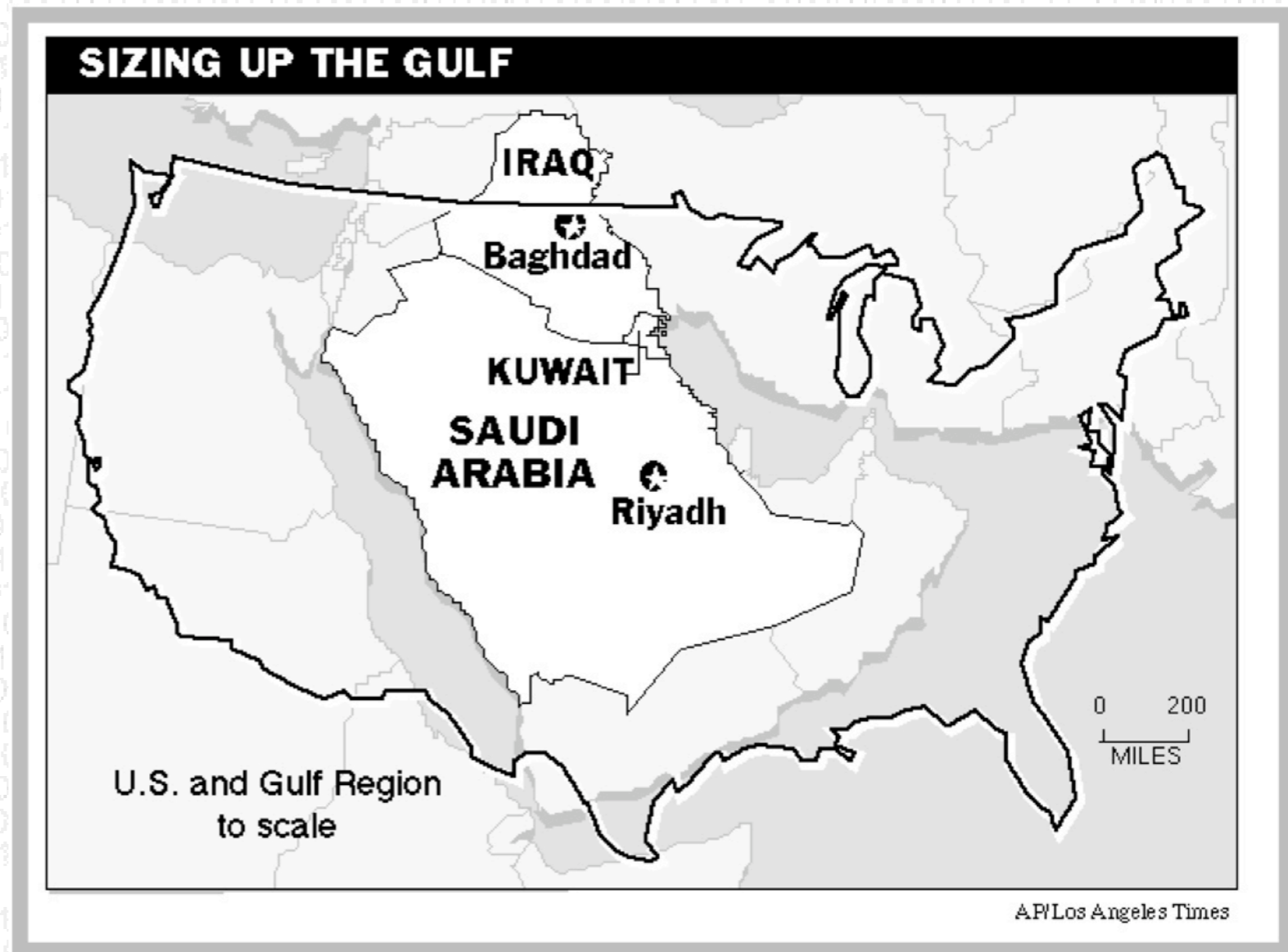
From the archives

Obsolete
software,
1991-2003



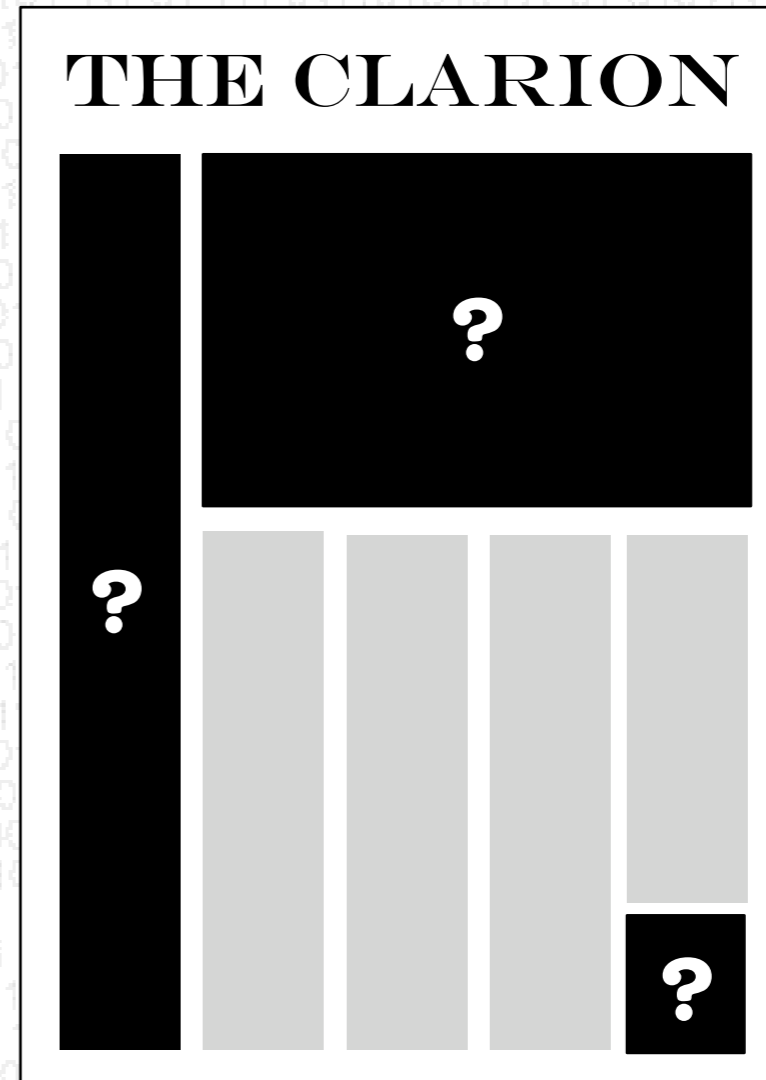
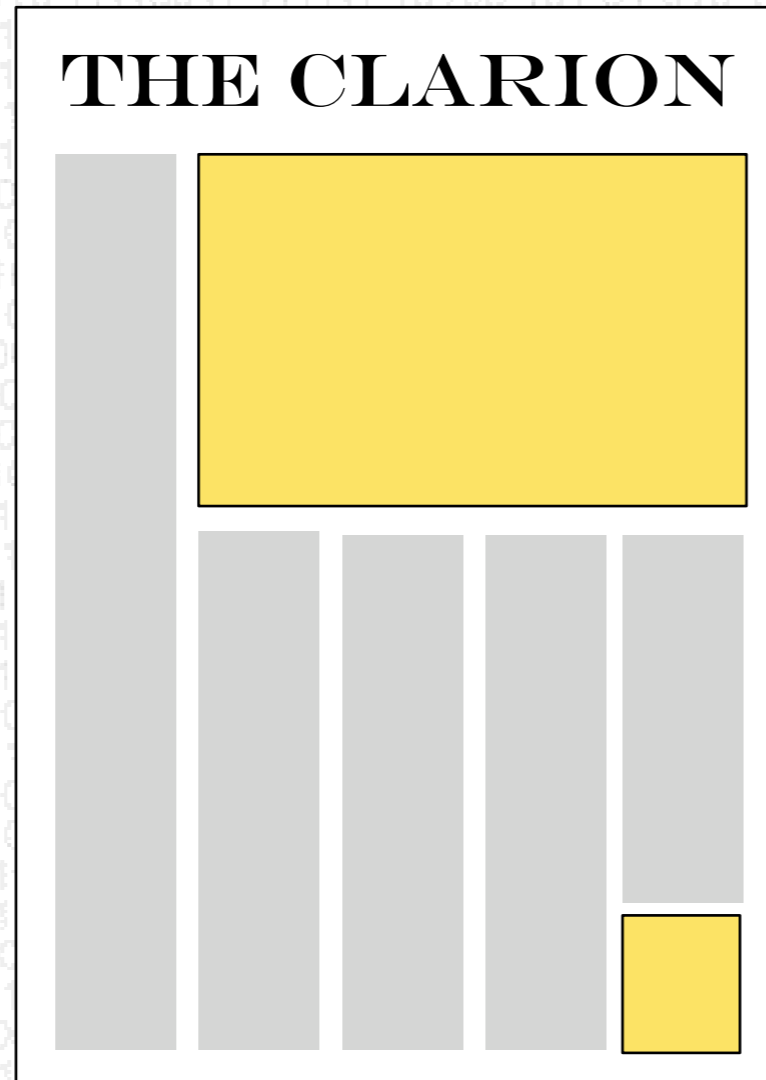
From the archives

Obsolete
software,
1991-2003



And then there's copyright

Page
facsimile
vs. access
to content



Tomorrow's newspaper tonight

- The basics (mid-1970s to present)
 - Reporters write, editors edit, artists draw, photographers shoot
 - Layout editors create pages, electronic production processes get them to press
 - Presses roll
 - Post-production processes capture text and images for storage and retrieval
 - Microfilm and text sales

Tomorrow's newspaper tonight

- But wait! (2000-present)
 - Add video, animation, web pages
 - Different, distributed staffs
 - Licensed content, not ownership
 - No more 'published version'
 - Fewer humans in the archives
 - Revenue from archives
 - Dwindling motivation to make microfilm

Tomorrow's newspaper tonight

- All adds up to:
 - Failure to capture certain types of content
 - Fewer checks and balances in data validation -- library and archives 'best practices'
 - Indexing for subject access gives way to tagging for web placement and targeted advertising content

Tomorrow's newspaper tonight

- All adds up to...
 - Proliferation of problematic formats in one-way databases
 - Loss of experienced librarians and archivists
 - Abandoning the most reliable method of secure long-term storage: Microfilm
 - A digital preservation disaster

News archives survey

- InterPARES digital preservation project
 - 2005-06 through News Division of SLA
 - 81 news libraries of all sizes and resources
 - Looked for emerging best practices in digital archives preservation
 - Growing awareness of threat to historic archives

News archives survey

- Have you lost anything during migration or upgrade? Examples:
 - 20% of stories and photos during migration
 - Entire photo archive
 - All the subject index terms for 1987
 - All the punctuation in text
 - All the photo metadata
 - Some number of JPEGs corrupted

News archives survey

- Have you lost anything during migration or upgrade? Examples:
 - We think so, but we don't know.

News archives survey

- Other takeaways
 - No standards for metadata
 - Few limits on ingesting unsustainable formats: archive it, figure it out later
 - Few on staff or in IT are knowledgeable about digital preservation issues
 - Who's responsible?
 - No vendors yet filling the void
 - Dwindling resources

It's not all bad news

- Vendor interest in commercial systems
 - Sun and Elektrofilm in digital movie preservation venture
 - Issue with definition of 'preservation'
 - Economics are still uncertain
 - Sustaining content of unknown value

It's not all bad news

- Computer output microfilm (COM)
 - Print page PDF's directly to microfilm
 - Color microfilm now a reality
 - Good for 300-plus years, properly stored
 - Continuous tone film designed for photography
 - Scan it again, scan it later
 - Film as a 'digital preservation' medium

It's not all bad news

- Academic and government models emerging for long-term preservation
 - Institutional repositories
 - ‘Dark archives’ for copyrighted material
 - TRAC (trusted repository audit and certification) was developed in part using Lexis-Nexis news databases as a model

Celebrate the challenge

Born-digital pages are more fragile than acidic newsprint, and our best chance for assuring future access is still microfilm -- and will be for the foreseeable future.

From the archives

mccargar
@mac.com

