

From Chronicling America to Newspaper Navigator:
*Improving Access to Historic Newspaper Photos at
the Library of Congress through Machine Learning*

Benjamin Charles Germain Lee &
Nathan Yarasavage

LIBRARY
OF CONGRESS
LABS

chroniclingamerica.loc.gov
news-navigator.labs.loc.gov

PAUL G.
ALLEN
SCHOOL
W

National Digital Newspaper Program (NDNP)

<https://www.loc.gov/ndnp>

- United States newspaper digitization pre-1963
- Collaboration since 2005
 - 50 state partners
 - Library of Congress (LC)
 - National Endowment for the Humanities (NEH)
- Shared technical specification
- Open access



Chronicling America

<https://chroniclingamerica.loc.gov/>

- 17+ million pages
- 3,000+ newspaper titles
- Content coverage:
 - 48 states, 2 territories, and D.C.
 - 23 languages
 - 1777-1963

The screenshot displays the Chronicling America website interface. At the top, there is a navigation bar with the Library of Congress logo and search options. Below this, the site's title "CHRONICLING AMERICA" is prominently displayed, along with a brief description of the project. A search bar is present, and below it, there are filters for "Search Pages", "Advanced Search", and "All Digitized Newspapers 1777-1963". A sidebar on the left contains a list of links for "About Chronicling America", "About the Site and API", "Recommended Topics", "Chronicling America Data Visualizations", and "Help". The main content area features a "100 Years Ago Today" section with three newspaper thumbnails: "The Cordova Daily Times", "Tloveritar", and "Bonners Ferry Herald". Each thumbnail includes a caption with the newspaper's name, volume information, and location.

LIBRARY OF CONGRESS ASK A LIBRARIAN DIGITAL COLLECTIONS LIBRARY CATALOGS Search Search Loc.gov GO

The Library of Congress > Chronicling America

NATIONAL ENDOWMENT FOR THE Humanities CHRONICLING AMERICA Historic American Newspapers Search America's historic newspaper pages from 1777-1963 or use the U.S. Newspaper Directory to find information about American newspapers published between 1690-present. Chronicling America is sponsored jointly by the National Endowment for the Humanities and the Library of Congress. [Learn more](#)

Search Pages Advanced Search All Digitized Newspapers 1777-1963 US Newspaper Directory, 1690-Present

All states from 1777 to 1963 enter one or more search words GO

Pages Available: 17,759,709 Print Subscribe Share/Save Give Feedback

100 Years Ago Today: 3/8/1921 (93 issues)

The Cordova Daily Times
AMERICANS LEAD REVOLT AGAINST SOVIET RUSSIA
Three Bands of Men Controlled by Different Apartments
[volume] (8pp.)
Cordova, Alaska

Tloveritar (12pp.)
Astoria, Or.

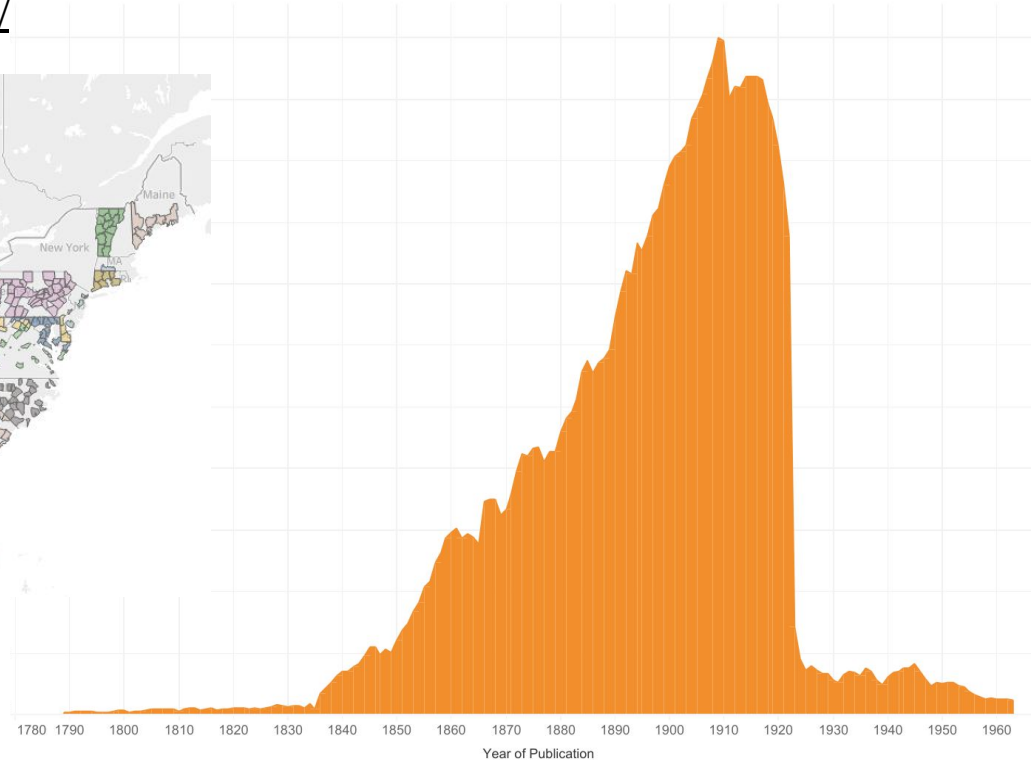
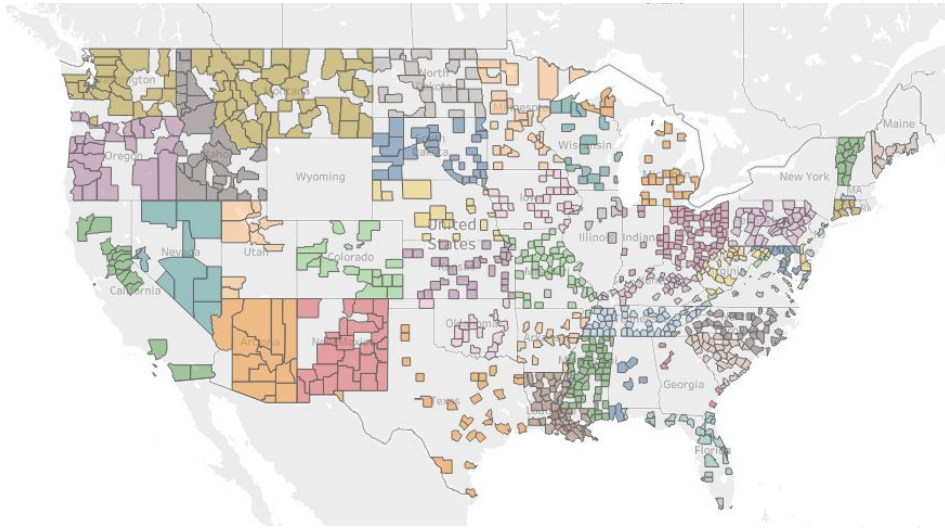
Bonners Ferry Herald
[volume] (10pp.)
Bonners Ferry, Idaho

More Resources

- > [National Digital Newspaper Program](#)
- > [NDNP Award Recipients](#)
- > [Newspaper and Current Periodicals Reading Room](#)
- > [Ask LC Newspaper & Current Periodicals Librarian](#)
- > [Historic Newspapers on Flickr](#) (part of the LC Flickr Commons photostream)

Chronicling America Data Visualizations

<https://www.loc.gov/ndnp/data-visualizations/>



Chronicling America Data Access

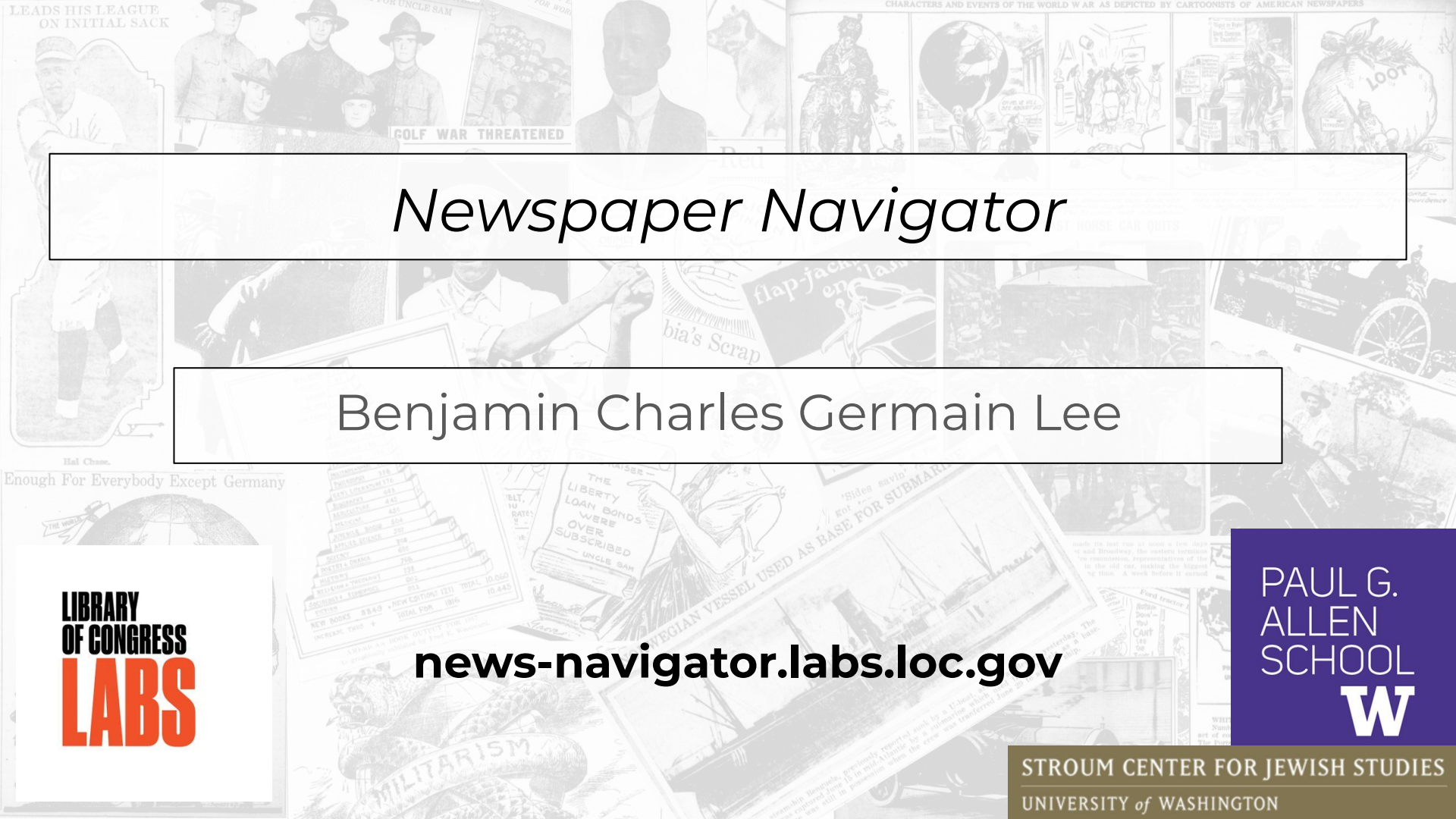
<https://chroniclingamerica.loc.gov/>

- Open access / public domain
- API & bulk data access
- Metadata & search results from site can also be returned in JSON format

Advancing Research and Scholarship

- Projects using Chronicling America data:
<https://www.loc.gov/ndnp/extras/>

The screenshot displays the Chronicling America website interface. At the top, there is a navigation bar with the Library of Congress logo and links for 'ASK A LIBRARIAN', 'DIGITAL COLLECTIONS', and 'LIBRARY CATALOGS'. A search bar is located on the right. Below this, the site title 'The Library of Congress > Chronicling America' is visible. The main content area features the 'National Endowment for the Humanities' logo and the 'CHRONICLING AMERICA' title. A search bar is present with filters for 'All states', 'from 1777', and 'to 1963'. A 'GO' button is on the right. Below the search bar, it indicates 'Pages Available: 17,759,709'. A '100 Years Ago Today: 3/8/1921 (93 issues)' section is highlighted, showing three newspaper thumbnails: 'The Cordova Daily Times', 'Tovoritir', and 'Bonners Ferry Herald'. Each thumbnail has a caption below it. On the left side, there is a sidebar with a list of links: 'About Chronicling America', 'About the Site and API', 'Recommended Topics', 'Chronicling America Data Visualizations', and 'Help'. Below this is a 'More Resources' section with links to 'National Digital Newspaper Program', 'NDNP Award Recipients', 'Newspaper and Current Periodicals Reading Room', 'Ask LC Newspaper & Current Periodicals Librarian', and 'Historic Newspapers on Flickr'.



Newspaper Navigator

Benjamin Charles Germain Lee

**LIBRARY
OF CONGRESS
LABS**

news-navigator.labs.loc.gov

**PAUL G.
ALLEN
SCHOOL**
W

STROM CENTER FOR JEWISH STUDIES
UNIVERSITY of WASHINGTON

The Washington Times

FINAL EDITION

REBELS, IN BATTLE, THREATEN TO BESIEGE MEXICO CITY

JOHNSON FIRES FIRST MD. GUN

SENATOR KILIAN JOHNSON, photographed today at Rockville as he began his speech in a crowd of voters at the reception who had assembled from all parts of Maryland...



Senator Kilian Johnson, photographed today at Rockville as he began his speech in a crowd of voters at the reception who had assembled from all parts of Maryland...

Enthusiastic Audiences Greeted Senator Today at Rockville. JOHNSON 'PUSHTFOOTER' California Plans to Make American Just America Again.

MORE BRITISH LAND IN ERIN

Britain Grabs Lion's Share of Big Booty at San Remo

WAT TRENCHARD. The victors at the table of San Remo. Control of the Caucasus. Control of the Dardanelles, the gateway to Asia...

BRITAIN MOVES TO WILSON CONSIDERS ARMENIA MANDATE

WILSON CONSIDERS ARMENIA MANDATE. President Wilson's attitude toward the Armenians...

CARRANZAMEN DESERT TO FOE

U. S. WARESHIP WATCHFUL. Carranzamen desert to foe. Carranzamen desert to foe. Carranzamen desert to foe.

WILSON CONSIDERS ARMENIA MANDATE

WILSON CONSIDERS ARMENIA MANDATE. President Wilson's attitude toward the Armenians...

The Evening Herald

Fourth Year—No. 3905. KLAMATH FALLS, OREGON, MONDAY, APRIL 26, 1920. Price, Five Cents

MODOC LUMBER CO. BUYS 40,000 ACRES OF PINE

MODOC LUMBER CO. BUYS 40,000 ACRES OF PINE. The Modoc Lumber Co. has purchased 40,000 acres of pine land in the Modoc National Forest...

SACRAMENTO MEN BUY MAIN STREET BUSINESS SITE

SACRAMENTO MEN BUY MAIN STREET BUSINESS SITE. A group of Sacramento men have purchased a business site on Main Street...

MOTORISTS KICK AT CONDITION OF CROSSINGS

MOTORISTS KICK AT CONDITION OF CROSSINGS. Motorists are expressing their dissatisfaction with the conditions of the crossings...

CROOK COUNTY IN BAD STRAIT. PLEA TO VOTERS

CROOK COUNTY IN BAD STRAIT. PLEA TO VOTERS. Crook County is in a bad strait and pleads to voters for assistance...

DISASTER FACES STATE COLLEGE

DISASTER FACES STATE COLLEGE. A disaster is facing the State College due to financial difficulties...

MAN Y CONTENDS OHIO ELECTION

MAN Y CONTENDS OHIO ELECTION. A man is contending for the Ohio election...

TWO FINED ON LIQUOR CHARGES

TWO FINED ON LIQUOR CHARGES. Two individuals have been fined on liquor charges...

FAIR SITE NOT ON MAY BALLOT

FAIR SITE NOT ON MAY BALLOT. A fair site is not on the May ballot...

PREPARING FOR PRIMARY MAY 21

PREPARING FOR PRIMARY MAY 21. Preparations are being made for the primary election on May 21...

KAMATHI RECORD TO MOVE DOWN TOWN

KAMATHI RECORD TO MOVE DOWN TOWN. The Kamathi record is to be moved down town...

SHAPPI CARD FOR BOXING EXHIBITION

SHAPPI CARD FOR BOXING EXHIBITION. A Shappi card is for the boxing exhibition...

MICKIE SAYS

MICKIE SAYS. Mickie says... (Cartoon text)

CORNER MAN MORGAN

CORNER MAN MORGAN. Corner man Morgan... (Cartoon text)

TOUR CROPS ESTIMATE... 6,073,000 000 BUSHELS... SAVE... 1000... 1000...

'6 Tubs at Home,' Says 'Gault, Denying Wife's Charge: 'Never Bathes'

'6 Tubs at Home,' Says 'Gault, Denying Wife's Charge: 'Never Bathes'. A man is accused of having six tubs at home...

WOMAN FACES JURY D.C. SUFFRAGE FIGHT ON SLAYING CHARGE. TO BE NATION-WIDE

WOMAN FACES JURY D.C. SUFFRAGE FIGHT ON SLAYING CHARGE. TO BE NATION-WIDE. A woman faces jury duty in D.C. over a suffrage fight...

GOING TO NEW YORK?

GOING TO NEW YORK? Are you going to New York? (Advertisement)

FACTORY GIVES RECESS FOR SMOKERS AND CHATS

FACTORY GIVES RECESS FOR SMOKERS AND CHATS. A factory gives a recess for smokers and chats...

RACE OPEN TO ALL—WILSON

RACE OPEN TO ALL—WILSON. President Wilson's policy on race...

RAIL BOARD PLANS ADJUSTMENT SIONS

RAIL BOARD PLANS ADJUSTMENT SIONS. The rail board plans adjustments...

WOMAN WEIGHING 190 KIDNAPED IN AUTO

WOMAN WEIGHING 190 KIDNAPED IN AUTO. A woman weighing 190 lbs. was kidnaped in an auto...

WOMAN WEIGHING 190 KIDNAPED IN AUTO

WOMAN WEIGHING 190 KIDNAPED IN AUTO. A woman weighing 190 lbs. was kidnaped in an auto...

Presidents and Pies

Presidents and Pies. Presidents and pies... (Cartoon text)

WOMAN WEIGHING 190 KIDNAPED IN AUTO

WOMAN WEIGHING 190 KIDNAPED IN AUTO. A woman weighing 190 lbs. was kidnaped in an auto...

DRY. SLA'S—Wet people wanted by Congress for production in the inland empire to a referendum. Including the dry territory in wet states more than 80 per cent of the population are under prohibition. Eight states are in submission campaigns. Two of them—New York and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, too, will vote on the question next November unless in November, 1921. Utah is a wet-and-dry state. The initiative.

MODOC LUMBER CO. BUYS 40,000 ACRES OF PINE

RECORD, April 29.—The Modoc Lumber Co. has purchased 40,000 acres of pine land in the Modoc National Forest, near the town of Madras, in the Willamette valley of Oregon.

MANY CONTENT OVER ELECTION

RECORD, April 29.—The election of the Modoc Lumber Co. directors has caused much content among the stockholders.

SACRAMENTO MEN MOTORISTS KICK BUY MAIN STREET BUSINESS SITE

RECORD, April 29.—The Sacramento Motorists' Club has kicked the purchase of a business site on Main street.

CRUOK COUNTY IN BAD STRAIT, PLEA TO VOTERS

RECORD, April 29.—The Cruok County Board of Commissioners has issued a plea to voters regarding the county's financial situation.

DISASTER FACES STATE COLLEGES

RECORD, April 29.—The state colleges are facing a disaster due to the loss of funds.

PREPARING FOR PRIMARY MAY 21

RECORD, April 29.—The primary election for May 21 is being prepared for.

TWO FINED ON LIQUOR CHARGES

RECORD, April 29.—Two individuals were fined for liquor charges.

FAIR SITE NOT ON MAY BALLOT

RECORD, April 29.—The fair site is not on the May ballot.

MOUSE BATS

RECORD, April 29.—A mouse was seen in a bat's nest.

KAMATH RECORD TO MOVE DOWN TOWN

RECORD, April 29.—The Kamath record is to be moved down town.

SHAWNY CAND FOR BOXING EXHIBITION

RECORD, April 29.—Shawny is a candidate for a boxing exhibition.

MARION MARRIES

RECORD, April 29.—Marion has just married.

SAVY BUSINESS KIDNAP

RECORD, April 29.—A savy business kidnap was reported.

Going to New York

RECORD, April 29.—A group is going to New York.

WOMAN FACES JURY

RECORD, April 29.—A woman is facing a jury.

FACTORY GIVES RECESS FOR SMOKERS AND CHATS

RECORD, April 29.—A factory gives a recess for smokers and chats.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 29.—A photography exhibition is planned.

REBELS, IN BATTLE, THREATEN TO BESIEGE MEXICO

JOHNSON FIRES FIRST MD. GUN

RECORD, April 26.—Senator Johnson has fired the first Maryland gun.

CREATOR HILAN JOHNSON

RECORD, April 26.—Hilan Johnson is the creator of the gun.

WOMAN FACES JURY

RECORD, April 26.—A woman is facing a jury.

FACTORY GIVES RECESS

RECORD, April 26.—A factory gives a recess.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.



RACE OPEN TO ALL—WILSON

RECORD, April 26.—President Wilson has declared the race open to all.

'6 Tubs at Home,' Says Gould, Denying Wife's Charge: 'Never Bathes'

RECORD, April 26.—Gould denies his wife's charge that he has six tubs at home.

WOMAN FACES JURY

RECORD, April 26.—A woman is facing a jury.

FACTORY GIVES RECESS

RECORD, April 26.—A factory gives a recess.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

WOMAN FACES JURY

RECORD, April 26.—A woman is facing a jury.

FACTORY GIVES RECESS

RECORD, April 26.—A factory gives a recess.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.

PHOTOGRAPHY

RECORD, April 26.—A photography exhibition is planned.



Wednesday, July 4, 1917 --THE TACOMA TIMES-- Page Seven

Doings of the Duffs

A Flivver Will Pass Anything on the Road.



By Allman

U. S. TROOPS GIVEN
OVATION IN PARIS

(United Press Special Wire)
PARIS, July 4.—Paris went wild with enthusiastic demonstrations to greet the arrival of the American Expeditionary Force. The millions cheered them as they entered the city. The American commander, General Pershing, was the center of attention. He was surrounded by troops, massed four deep. The grand marshal of France was assembled in the center, and they, like the populace, were hurrying in an automobile throng.

First on the program of the day was the presentation of the American Expeditionary Force. The presentation of the American Expeditionary Force was the presentation of the American Expeditionary Force. The presentation of the American Expeditionary Force was the presentation of the American Expeditionary Force.

Prince Insects
Captured Shell

SEATTLE ROUTE
Steamers Tacoma and Indianapolis for Seattle
Leave Tacoma 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.
Leave Seattle 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.
Return to Tacoma 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.

CENTRAL BUS STATION
Leaving Times in Connection
Bellingham 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.
Carnation 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.
Natchez 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.

DEPOW-TACOMA
Leave Depow 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.
Leave Tacoma 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.

GENERAL-TACOMA
Leave General 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.
Leave Tacoma 11:00 a. m. 1:00 p. m. 3:00 p. m. 5:00 p. m.

Choose the best transcription or enter your own!

Original prompt: Enter picture's caption, illustrator or photographer and category.

Caption Doings of the Duffs—A Flivver Will Pass Anything on the Road.

Artist W. R. Allman

Category Comics/Caroon

Caption Doings of the Duffs—A Flivver Will Pass Anything on the Road.

Artist F. W. Allman

Category Comics/Caroon

ED
OF
OUSE

What
July
man
the
crowd
at
gates.
Eight
bitter
other
hand
by
the
pocket.

L
IS
LESS

What
—Wash
or
trally
to
the
most
more
than
five
of
were
de
vessels
made
were

ICES

A wicker basket, especially prepared for carrying large shells safely, is the way the Germans are transferring their shells to the navy. Prince Arthur of Connaught, governor-general of Canada and now at the French front, is now shown taking off the basket to inspect the shells.

View A Tutorial

View the original page



Wednesday, July 4, 1917 --THE TACOMA TIMES-- Page Seven

Doings of the Duffs

A Flivver Will Pass Anything on the Road.



By Allman

U. S. TROOPS GIVEN
OVATION IN PARIS

(United Press Special Wire) PARIS, July 4.—Paris went wild with enthusiastic demonstrations today in celebrating the grand triumph of the American Expeditionary Force in France. The millions cheered them as they entered the city, and the American commander, General Pershing, was hailed as a hero. The parade, which was the grandest ever seen in Paris, lasted for four days. The grand marshal of France was assembled in the center, and troops themselves, cheered as they marched, were hurrying in an automobile thru

the streets. First on the program of the day was the presentation of the American Expeditionary Force to the American commander, General Pershing, at the Invalides. These troops, which were surrounded by troops, marched for three days. The grand marshal of France was assembled in the center, and troops themselves, cheered as they marched, were hurrying in an automobile thru

Prince H's sects
Captured She

Editorial Cartoon
Comics/Cartoon
Illustration
Photograph
Map

Choose the best transcription or enter your own!

Original prompt: *Enter picture's caption, illustrator or photographer and category.*

Caption **Doings of the Duffs—A Flivver Will Pass Anything on the Road.**

Artist **W. R. Allman**

Category **Comics/Cartoon**

Caption **Doings of the Duffs—A Flivver Will Pass Anything on the Road.**

Artist **F. W. Allman**

Category **Comics/Cartoon**

VED
OF
OUSE

Wires
July 4
man were
the White
crowd of
and when
the dead
ations at
gates.
might two
litter was
other han-
sides by
sum the
jacket.

IS
LESS

Wires
Warily
or miser-
trally to
the past
more than
five or six
were de-
voted to
the
matters were

ICES

A wicker basket, especially prepared for carrying large shells safely, is the way the Germans are transferring action airplanes to the navy. Prince Arthur of Connaught, governor general of Canada and now at the French front, is now shown taking off the basket to inspect the shell



LE. Tacoma 4:00, 10:30 a. m.; 5:00, 11:10 p. m. 5:00 a. m., 1:00, 4:00, 8:10 p. m. 1:00, 4:00, 8:10 p. m. and 11:00 p. m.

ROBERTS-TACOMA
Nelson Truck and Stage Bus
Leave Tacoma 7:00, 9:30 a. m.
1:00, 4:00, 8:10 p. m. 1:00, 4:00, 8:10 p. m. and 11:00 p. m.

ORING-TACOMA
Leave Tacoma 7:00, 9:30, 11:00 a. m.; 1:00, 4:00, 8:10 p. m. 1:00, 4:00, 8:10 p. m. and 11:00 p. m. Leave Oring 4:45.

GENERAL-TACOMA
Elio, Alvo, La Grange, M. M. 1:00, 4:00, 8:10 p. m. and 11:00 p. m.

ASHFORD-TACOMA
1:00, 4:00, 8:10 p. m. and 11:00 p. m. 1:00, 4:00, 8:10 p. m. and 11:00 p. m.

BUCKE AUTO STAGE
Hood-Coff Hotel, 10:30 Pacific av. and Central Bay Station. Tacoma and Port George. Billings, Gresham, Jay, McKenna, N. J. Hill.

The Prohibition amendment bill, passed by the Sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 90 per cent of the population are under prohibition.

Eight states are in submission campaigns. Two of these—New York and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November, where in November, 1918. Ohio is in a very early fight through the initiative.



Wednesday, July 4, 1917 --THE TACOMA TIMES-- Page Seven

Doings of the Duffs

A Flivver Will Pass Anything on the Road.

By Allman

U. S. TROOPS GIVEN OVATION IN PARIS



(United Press Special Wire) PARIS, July 4.—Paris went wild with enthusiasm today in celebrating the grand triumph of the American Expeditionary Force. The millions cheered their heroes as a battalion of the American commander, Hal Pershing, reviewed the veterans every step. They cheered to get at the top appearing in every group. They cheered again, so thunderously that the whole city seemed to rock as they cheered the marching heroes, and they, like the populace of the front, were hurrying in an automobile thru the streets.

Prince H's sects Captured She

View A Tutorial

View the original page

Editorial Cartoon
Comics/Cartoon
Illustration
Photograph
Map

Choose the best transcription or enter your own!

Original prompt: *Enter picture's caption, illustrator or photographer and category.*

Caption **Doings of the Duffs—A Flivver Will Pass Anything on the Road.**

Artist **W. R. Allman**

Category **Comics/Cartoon**

Caption **Doings of the Duffs—A Flivver Will Pass Anything on the Road.**

Artist **F. W. Allman**

Category **Comics/Cartoon**

VED OF ROUSE

Wired July 4.—men were the White crowd of and when the scaled rations at gates. might two litter was other hands by am the jacket.

IS LESS

Wired July 4.—or miss, rally to the most sore than five of were de- g vessels were were

ICES

A wicker basket, suspended from the ceiling, is the way the Germans are inflicting their victims in the navy. Prince Arthur of Connaught, governor general of Canada and now at the French front, is now shown taking off the hat to inspect the shell



6. Fremont—4:45, 10:30 a. m.; 1:30, 7:15 p. m.	9. Duane—4:30 a. m., 1:30, 7:15 p. m.
7. Franklin—4:45, 10:30 a. m.; 1:30, 7:15 p. m.	10. Duane—4:45, 10:30 a. m.; 1:30, 7:15 p. m.
8. Franklin—4:45, 10:30 a. m.; 1:30, 7:15 p. m.	11. Duane—4:45, 10:30 a. m.; 1:30, 7:15 p. m.
12. Duane—4:45, 10:30 a. m.; 1:30, 7:15 p. m.	13. Duane—4:45, 10:30 a. m.; 1:30, 7:15 p. m.
14. Duane—4:45, 10:30 a. m.; 1:30, 7:15 p. m.	15. Duane—4:45, 10:30 a. m.; 1:30, 7:15 p. m.

Thousands of Annotations!

Enough



The Prohibition amendment bill, passed by the Sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 90 per cent of the population are under prohibition.

LEADS HIS LEAGUE ON INITIAL SACK



Hal Chase.



GOLF WAR THREATENED



REV. E. W. HARTHORNE, D. D. Nashville, Tenn., and ardent Congress worker.

QUIPET



Red



bias Scrap

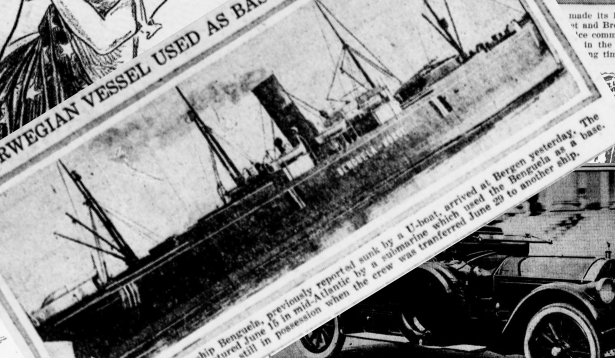


NEW YORK'S LAST HORSE CAR QUITS



made its last run at about a few days ago on Broadway, the eastern terminus of the commission, representatives of the in the old car, making the biggest push of the week before it earned

NORWEGIAN VESSEL USED AS BASE FOR SUBMARINE



steamship Benguela, previously captured sunk by a U.S. boat, arrived at Bergen yesterday. The captured vessel is in possession when the crew was transferred

AMERICAN BOOK OUTLET FOR 1918 to graphically exhibited by Fred E. Woodman.

BOOKS	377	
NOVELS	87	
COMICS	117	
THE PRESS	143	
NON-LEADS	171	
SINGLE 1918		
LAW	216	
BUSINESS	199	
TELEGRAPH	296	
PHILOLOGY	336	
MILITARY	338	
STYL LITERATURE	226	
AGRICULTURE	426	
MEDICINE	452	
MUSIC	504	
MATHS	504	
ATLANTIC	504	
SCIENCE	504	
ITALY'S DIARIES	798	
MILITARY	128	
MILITARY	632	
MILITARY	632	
NEW BOOKS 2849		NEW EDITIONS 1211
TOTAL 10,660		TOTAL 10,660
NEW BOOKS 2849		TOTAL FOR 1918 10,660
INCREASE 1708 +		

Enough For Everybody Except Germany



THE NATION'S GOING DRY.

WHITE=Dry. BLACK=Wet.
Number of prohibition states, 20. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people ratified by congress. The Prohibition citizenship bill, passed by the Sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 60 per cent of the population are under prohibition.
Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico also will vote on the question next November; others in November, 1918. Ohio is in a wet-and-dry fight through the initiative.



IN TOWN FOR THE BRICE P. CASHES;—STERLING HIS STUFF—BROUGHT—THE GOOD NEWS ABOUT VICTORY, FROM THE SAN FRANCISCO CHRONICLE;—THE BURNING CLUB—PIZZERELLO; IN THE ST. LOUIS POST-DEMOCRAT;—ROBINSON BAR SHELLS FOR HIS COMPANY—WILSON;—THE DAYTON (OHIO) DAILY NEWS—DOUTRIE—(2)

LEADS HIS LEAGUE ON INITIAL SACK

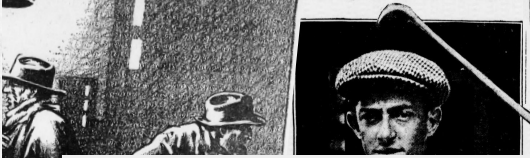


Hal Chase.

Enough For Everybody Except



FOR UNCLE SAM



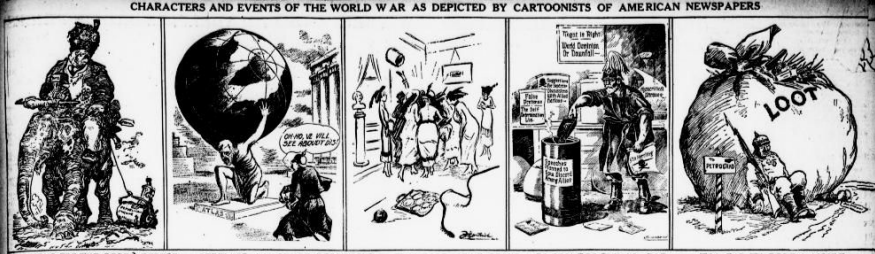
GOLF WAR THREATENED



REV. E. W. HARTHORNE, D. D. Nashville, Tenn., and ardent Congress worker.



-Red



IN TOWN FOR THE RICE... STERLING HAS STUFF-BROUGHT... THE GOOD NEWS ABOUT... THE BROWNING BUN... THE BROWNING CLUB... FOR UNCLE SAM SHELLS FOR... TOO BIG TO CARRY HOME... THE PRESIDENT... (The Daily News...)

NEW YORK'S LAST HORSE CAR QUITS

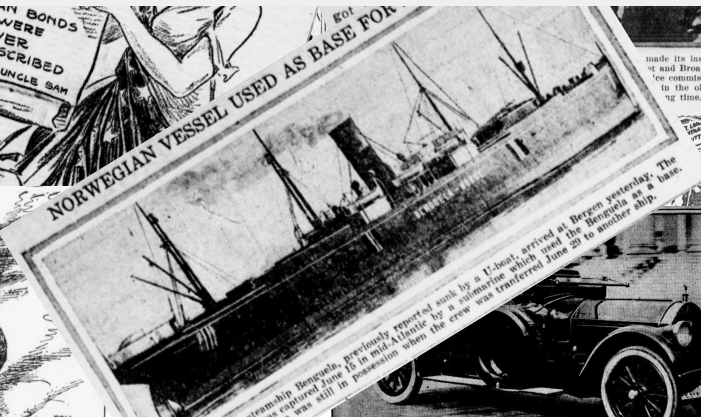


Can we "throw open the treasure chest" by training a machine learning algorithm to process all of Chronicling America?



AGRICULTURE	456
MEDICINE	554
AMERICAN BOOK	554
ATLANTA	554
SCIENCE	554
INDUSTRIAL DRAMA	758
RELIGION	605
MYSTERY	758
RELIGION & THEOLOGY	605
SCIENCE & ECONOMY	605
MYSTERY	758
NEW BOOKS	2849
NEW EDITIONS	1211
TOTAL FOR 1917	10,650
TOTAL FOR 1916	10,445
INCREASE	THOS +

AMERICAN BOOK OUTPUT FOR 1917 to graphically exhibited by Fred E. Woodman.



NORWEGIAN VESSEL USED AS BASE FOR

stream-ship Benguela, previously reported sunk by a U-boat, arrived at Bergen yesterday. The vessel was captured June 15 in mid-Atlantic by a submarine which was towed to another ship. The crew was transferred.

made its last run at about a few days... and Broadway, the eastern terminal... commission, representatives of the... in the old car, making the biggest... time. A week before it earned



THE NATION'S GOING DRY.

WHITE—Dry. BLACK—Wet. Number of prohibition states, 20. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people ratified by congress. The Porto Rican citizenship bill, passed by the Sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 90 per cent of the population are under prohibition. Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November; others in November, 1918. Ohio is in a wet-and-dry fight through the initiative.

LEADS HIS TEAM ON INITIAL SACK



Hal Chase.

Enough For Everybody Except



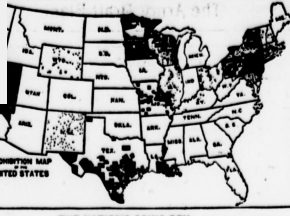
-Red



CHARACTERS AND EVENTS OF THE WORLD WAR AS DEPICTED BY CARTOONS OF AMERICAN NEWSPAPERS

3 SHELLS FOR... TOO BIG TO CARRY HOME...

- facilitate research in computer science, library science, and the digital humanities

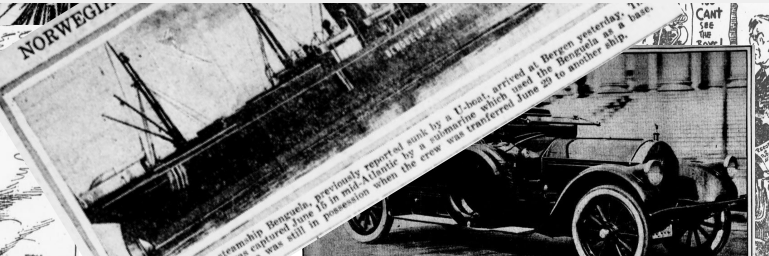


THE NATION'S GOING DRY.

WHITE—Dry. **BLACK**—Wet.

Number of prohibition states, 20. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people ratified by congress. The Prohibition citizenship bill, passed by the Sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 60 per cent of the population are under prohibition.

Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico also, will vote on the question next November; others in November, 1918. Ohio is in a wet-and-dry fight through the initiative.



stream-ship Benguela, previously reported sunk by a U-boat, arrived at Bergen yesterday. The ship captured June 15 in mid-Atlantic by a submarine which on June 20 to another ship. The crew was transferred.



LEADS HIS LEAGUE ON INITIAL SACK



Hal Chase.

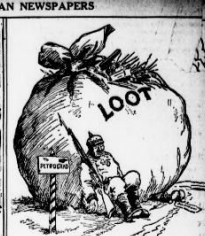
Enough For Everybody Except



GOLF WAR THREATENED



-Red



5 SHELLS FOR... TOO BIG TO CARRY HOME...

- facilitate research in computer science, library science, and the digital humanities
- Engage the public and enable a wide range of creative computing and public humanities projects



THE NATION'S GOING DRY.

WHITE—Dry. BLACK—Wet. Number of prohibition states, 20. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people ratified by congress. The Porto Rican citizenship bill, passed by the Sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 60 per cent of the population are under prohibition. Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico also, will vote on the question next November; others in November, 1918. Ohio is in a wet-and-dry fight through the initiative.



Weather Forecast
Partly cloudy with showers
Temperature 70-80
Wind light S.W.

The Evening Star

Guide for Readers
Subscription rates
Single copies 5c

NAZI COUNTERATTACKS REPULSED NEAR CAEN

Allied Infantry Pushes 10 Miles North of Rome



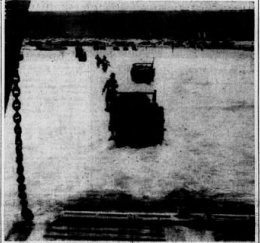
ANKA OVERLOOKS ON THE WEST... (Caption describing the photograph)

Allies Clear Enemy From Landing Areas; Battle Raging Inland

8th Army Fights Break Action; 5th Sees Little of Foe

General Sent Home, Demoted For D-Day Tip

Lehman Tells People Of Conquered Nations Supplies Are Near



One of the open lines of a leading craft American tank group on the beach at the Normandy invasion...

450 Keep Wright Plan Idle Despite Appeal To Help Invasion

U.S. Compromise Seen On Ball-Bearing Exports

Patotic Appeal Ends Stake at Lockport

260 Billion Debt Limit Bill Sent to White House

Help Conserve Copies of The Star

Cleveland Opa Hearings Call 3 More Food Dealers

ODI Hits Civilian Travel; Wants Rails for War Use

U.S. Compromise Seen On Ball-Bearing Exports

Patotic Appeal Ends Stake at Lockport

260 Billion Debt Limit Bill Sent to White House

Help Conserve Copies of The Star

Cleveland Opa Hearings Call 3 More Food Dealers

ODI Hits Civilian Travel; Wants Rails for War Use

49 Billions for Army Voted by House For Next Fiscal Year

New Food Shortages Tantalize War Changers

260 Billion Debt Limit Bill Sent to White House

Bauch Ending Job Of Postwar Advisor Will Continue as Special Consultant to Barnes

260 Billion Debt Limit Bill Sent to White House

U.S. Compromise Seen On Ball-Bearing Exports

Patotic Appeal Ends Stake at Lockport

260 Billion Debt Limit Bill Sent to White House

Help Conserve Copies of The Star

Cleveland Opa Hearings Call 3 More Food Dealers

ODI Hits Civilian Travel; Wants Rails for War Use

Supreme Headquarters Allied Expeditionary Force...

London (AP)—The British broadcast today that an entire Allied air-borne division has been captured on the western coast of the Netherlands.

New York (AP)—Broadcast by the anti-Nazi Atlantic underground station in Germany heard by ABC and today had been heard by the British, says a radio broadcast on the Chongqing Postoffice.

Supreme Headquarters Allied Expeditionary Force...

Headquarters and staff were captured by the Germans...

The German radio report of the capture...

London (AP)—The British broadcast today that an entire Allied air-borne division has been captured on the western coast of the Netherlands.

London (AP)—The British broadcast today that an entire Allied air-borne division has been captured on the western coast of the Netherlands.

Ship Built for Chinese To Be Launched Today

Six Airmen Feared Dead, One Safe in Plane Crash

Red Public Opinion... (Advertisement for Red Public Opinion)

Swedish Vessel Used... (Advertisement for a Swedish vessel)

Comic strip panels with dialogue bubbles.

Take My Word for It... (Text block for a comic strip)

Pvt. Bregar Abroad... (Text block for a comic strip)

Comic strip panel titled 'Nature's Children'.

The World of My West... (Text block for a comic strip)

Comic strip panel titled 'Private Lives'.

IF YOU KNOW EM WELL ENOUGH... (Text block for a comic strip)

Comic strip panel titled 'Modern Maidens'.

Comic strip panel titled 'Ship Built for Chinese'.

Comic strip panel titled 'Six Airmen Feared Dead'.

RADIO PROGRAM table listing various radio stations and their schedules.

Medicine Cabinet advertisement with an image of a cabinet.

Perma-Stone advertisement.

Quaker City advertisement.

Wonder Paint advertisement.

Great Moments advertisement.

Templing, lanky, tasty... advertisement.

Great War News advertisement.

Radio Service advertisement.

Radio Service advertisement.

Radio Service advertisement.

Weather Forecast
Partly cloudy with showers
Temperature 60-70
Wind light S.W. breeze

The Evening Star

Guide for Readers
Published daily except on Sundays and public holidays
Subscription rates: \$1.00 per month, \$10.00 per year

NAZI COUNTERATTACKS REPULSED NEAR CAEN

Allied Infantry Pushes 10 Miles North of Rome

8th Army Fights Break Action; 5th Sees Little of Fighting

ROME (AP)—The 8th Army's offensive against the German line north of Rome has pushed the front 10 miles in places, according to a report from the American command.



Allies Clear Enemy From Landing Areas, Battle Raging Inland

Beachheads Extend Over 50 Miles; Both Sides Scold in Air Warfare

BRITAIN (AP)—The Allies' beachheads in Normandy have been extended to a total of over 50 miles, according to a report from the British command.

General Sent Home, Demote For D-Day Tip

8th Army Fights Break Action; 5th Sees Little of Fighting

ROME (AP)—General Alexander Patch, commander of the 8th Army, has been demoted for a tip-off about the D-Day invasion.

Man reits People if Conquered Nation Applies Are Near

UNION REITS PEOPLE if Conquered Nation Applies Are Near

WASHINGTON (AP)—The State Department is considering a plan to re-educate the people of a conquered nation.

Supreme Headquarters Allied Expeditionary Force

Supreme Headquarters Allied Expeditionary Force

BRITAIN (AP)—The Supreme Headquarters Allied Expeditionary Force has been established in Britain.

50 Keep Wright Plan Despite Appeal to Help Invasion

50 Keep Wright Plan Despite Appeal to Help Invasion

WASHINGTON (AP)—Fifty members of the House of Representatives have voted to keep the Wright plan for the invasion of Europe.



London 27-28 and 29 in Battle

London 27-28 and 29 in Battle

LONDON (AP)—London was in a state of battle for three consecutive days, with heavy fighting and air raids.

9 Billions for Army to Next Fiscal Year

9 Billions for Army to Next Fiscal Year

WASHINGTON (AP)—The House of Representatives has approved a bill for \$9 billion for the Army for the next fiscal year.

3,000 Sorries Flow in Invasion Area

3,000 Sorries Flow in Invasion Area

WASHINGTON (AP)—Over 3,000 letters of sympathy have been received for the invasion area.

Attack Ending Job Postwar Advisor Will Continue as Special Consultant to Bernes

Attack Ending Job Postwar Advisor Will Continue as Special Consultant to Bernes

WASHINGTON (AP)—A postwar advisor will continue as a special consultant to the State Department.

S. Compromise Seen in Bull-Beating Fight

S. Compromise Seen in Bull-Beating Fight

WASHINGTON (AP)—A compromise is seen in the fight over the bull-beating bill.

60 Billion Debt Limit Will Seal in White House

60 Billion Debt Limit Will Seal in White House

WASHINGTON (AP)—The 60 billion debt limit bill is expected to be signed in the White House.

Ship Built for China to Be Launched Today

Ship Built for China to Be Launched Today

WASHINGTON (AP)—A ship built for China is to be launched today.

Help Center Comes of The Story

Help Center Comes of The Story

WASHINGTON (AP)—A help center has been established for the story.

Levelled OPA Hearings all 3 More Food Dealers

Levelled OPA Hearings all 3 More Food Dealers

WASHINGTON (AP)—Three more food dealers have been levelled by the OPA.

New National Cemetery in D.C. or 50,000 Is Sought by Army

New National Cemetery in D.C. or 50,000 Is Sought by Army

WASHINGTON (AP)—The Army is seeking a new national cemetery in D.C. for 50,000 soldiers.

DT His Civilian Travel Vans Rails for War Use

DT His Civilian Travel Vans Rails for War Use

WASHINGTON (AP)—The DT is requesting that civilian travel vans be used for war purposes.

Ship Built for China to Be Launched Today

Ship Built for China to Be Launched Today

WASHINGTON (AP)—A ship built for China is to be launched today.

Ship Built for China to Be Launched Today

Ship Built for China to Be Launched Today

WASHINGTON (AP)—A ship built for China is to be launched today.

Ship Built for China to Be Launched Today

Ship Built for China to Be Launched Today

WASHINGTON (AP)—A ship built for China is to be launched today.

Ship Built for China to Be Launched Today

Ship Built for China to Be Launched Today

WASHINGTON (AP)—A ship built for China is to be launched today.

Ship Built for China to Be Launched Today

Ship Built for China to Be Launched Today

WASHINGTON (AP)—A ship built for China is to be launched today.

Red Public Opinion

Red Public Opinion

Advertisement for Red Public Opinion, featuring a woman's face and the text 'Red Public Opinion'.

Radio Program

Radio Program

Advertisement for a radio program, featuring a cartoon illustration of a man and a woman.

Special This Week

Special This Week

Advertisement for a special offer, featuring a medicine cabinet and other products.

Take My Word for It

Take My Word for It

Advertisement for a product, featuring a cartoon illustration of a man and a woman.

My Uncle Asks

My Uncle Asks

Advertisement for a product, featuring a cartoon illustration of a man and a woman.

Perma-Stone

Perma-Stone

Advertisement for Perma-Stone, featuring a photograph of a stone surface.

World of Its West

World of Its West

Advertisement for a product, featuring a cartoon illustration of a man and a woman.

Great Moments

Great Moments

Advertisement for Great Moments, featuring a photograph of a man and a woman.

Modern Maidens

Modern Maidens

Advertisement for Modern Maidens, featuring a cartoon illustration of a woman.

Radio Service

Radio Service

Advertisement for Radio Service, featuring a photograph of a radio set.

Advertisement for a product, featuring a photograph of a woman's face and the text 'Advertisement for a product'.

The Evening Star

WASHINGTON, D. C., WEDNESDAY, JUNE 7, 1944, FORTY-FIVE PAGES

NAZI COUNTERATTACKS PEDIJ CED NEAR CAEN

Allied Infantry Pushes 10 Miles North of Rome

General Sent Home, Demot For D-Day Tri

General Montgomery's decision to delay the invasion of France until after the capture of Rome was criticized by some military leaders. However, Montgomery's strategy was to use the capture of Rome as a diversion, drawing German forces away from the main invasion site in Normandy.

150 Keep Wright Plie Despite Appeal to Help Invasion

Air Force Officer's Plan To 'Kick The Market Business' Ignored

A plan by an Air Force officer to use a market business as a cover for military operations was ignored by the command. The officer's plan involved using a legitimate business to transport supplies and personnel across enemy lines.

S. Compromise Seen in Bull-Beating Exports

Exports of bull-beating equipment from the South are seen as a compromise between military needs and civilian interests. The equipment is used for both military operations and civilian agricultural activities.

Help Center Comes of The Star

A new help center has been established to assist readers with their inquiries and provide information on various topics. The center is staffed by experienced personnel and is open to the public.

DT His Civilian Travel Vans Ralls for War Use

DT's civilian travel vans are being converted for use in the war effort. The vans will be used to transport supplies and personnel across the front lines.

Atlantic Appeal Ends in Lockport

The Atlantic Appeal for war materials has ended in Lockport. The appeal was a successful campaign to raise funds and collect supplies for the war effort.

Ship Built for Chinese To Be Launched Today

A ship built for the Chinese is scheduled to be launched today. The ship is being built in the United States and will be used to transport supplies and personnel to the front lines.

14th AVN Rearranged To Be Safe in Plane Crash

The 14th AVN is being rearranged to ensure safety in the event of a plane crash. The new arrangement will allow for a faster and more efficient evacuation of the crew.

Help Center Comes of The Star

Help center established to assist readers with their inquiries. The center provides information on various topics, including military operations and civilian life.

Atlantic Appeal Ends in Lockport

Atlantic Appeal for war materials ends in Lockport. The appeal was a successful campaign to raise funds and collect supplies for the war effort.

Ship Built for Chinese To Be Launched Today

Ship built for the Chinese is scheduled to be launched today. The ship is being built in the United States and will be used to transport supplies and personnel to the front lines.

14th AVN Rearranged To Be Safe in Plane Crash

The 14th AVN is being rearranged to ensure safety in the event of a plane crash. The new arrangement will allow for a faster and more efficient evacuation of the crew.

7 Different categories:

- Photographs
- Illustrations
- Comics
- Editorial Cartoons
- Maps
- Headlines
- Advertisements



RADIO PROGRAM	
10:00-10:30	Blue Network
10:30-11:00	Blue Network
11:00-11:30	Blue Network
11:30-12:00	Blue Network
12:00-12:30	Blue Network
12:30-1:00	Blue Network
1:00-1:30	Blue Network
1:30-2:00	Blue Network
2:00-2:30	Blue Network
2:30-3:00	Blue Network
3:00-3:30	Blue Network
3:30-4:00	Blue Network
4:00-4:30	Blue Network
4:30-5:00	Blue Network
5:00-5:30	Blue Network
5:30-6:00	Blue Network
6:00-6:30	Blue Network
6:30-7:00	Blue Network
7:00-7:30	Blue Network
7:30-8:00	Blue Network
8:00-8:30	Blue Network
8:30-9:00	Blue Network
9:00-9:30	Blue Network
9:30-10:00	Blue Network

SPECIAL THIS WEEK

Bechler Co.

Medicine Cabinet

Bechler's four doors are made of the finest materials and are available in a wide variety of colors and finishes.

Bechler's Home Radio

WENX-1240

PERMA-STONE

These are all made with the finest materials and are available in a wide variety of colors and finishes.

Taylor 6420

Tonight

SYMPHONY

SYMPHONY

SYMPHONY

SYMPHONY

SYMPHONY

SYMPHONY

SYMPHONY

Ruth Crane

WYMA

WYMA

WYMA

WYMA

WYMA

WYMA

RADIO TROUBLE!

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

LUNCH

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

MID CITY RADIO SHOP

GREAT WAR NEWS IS BREAKING!

ELIOT

CALMER

Tomorrow and every Thursday

WTOP 6:15 P.M.

RADIO SERVICE

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

WMAJ

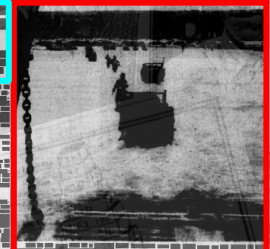
WMAJ



BLAUG-Wat of Columbia is city by the mouth of congress. fourth negative, growth of the area of the relation are under gov of them--long and tember will vote to on the question. ready light. Students

LEADS HIS LEAGUE
ON I

The Evening Star

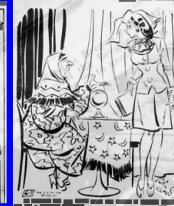
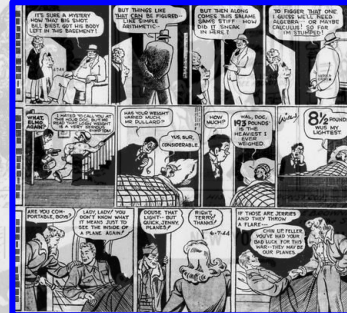


Grid of text blocks, mostly illegible due to low resolution and small font size. Some blocks are highlighted with cyan borders.

Red
PUBLIC OPINION

bia's Scrap

WEGIAN VESSEL USED



THE WONDER PAINT
Quarrel City
SPREAD
A GOLDEN product

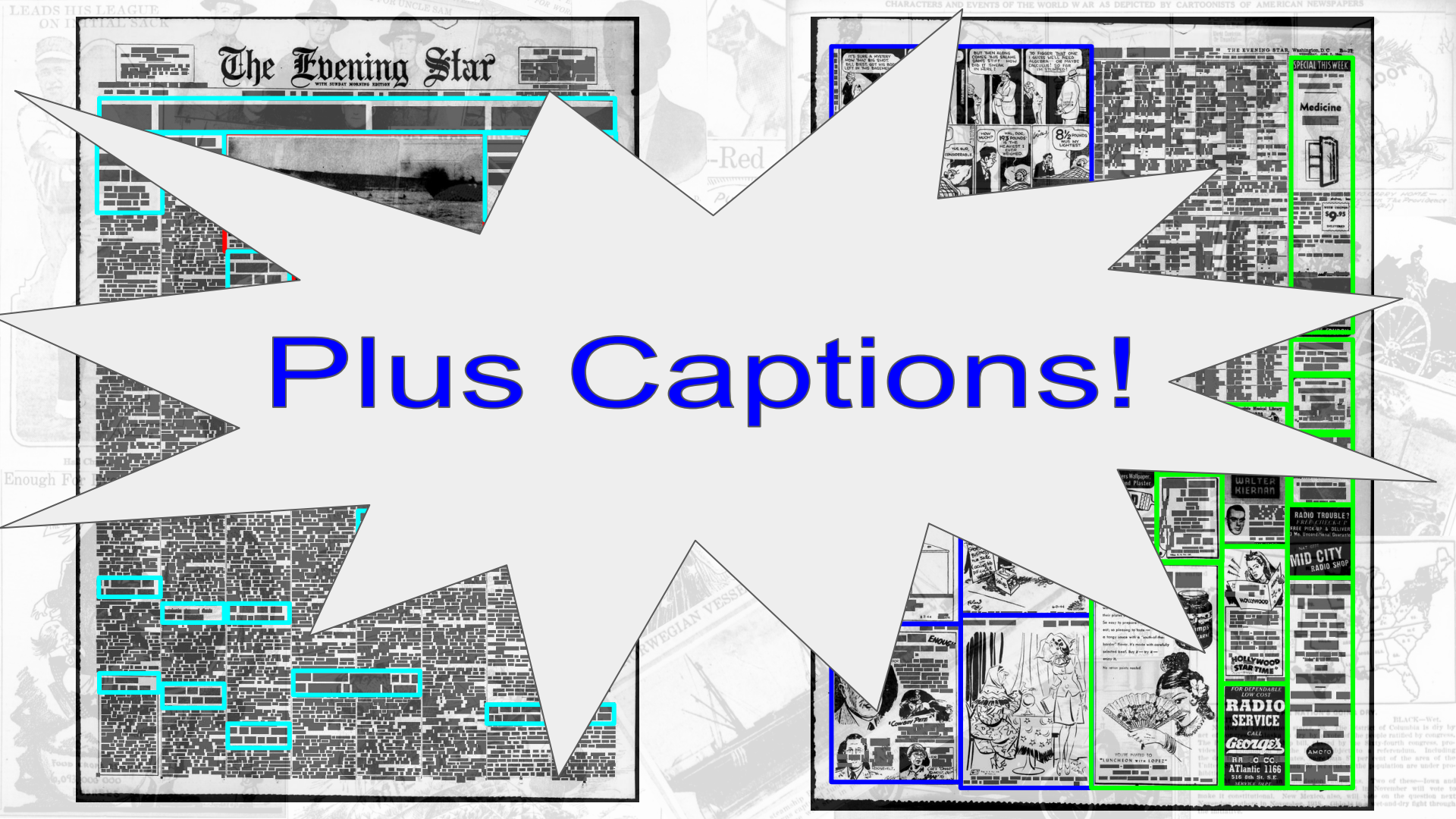
FOR DEPENDABLE SERVICE
CALL
Georges
RADIO SERVICE
11th & CC
Atlantic 1166
516 8th St. S.E.
WASH. D.C.

SPECIAL THIS WEEK
Medicine
\$9.95

FOR DEPENDABLE SERVICE
CALL
Georges
RADIO SERVICE
11th & CC
Atlantic 1166
516 8th St. S.E.
WASH. D.C.

BLA'N-Was
of Columbia is dry by
is needed by congress
through negative pro
information. Including
that of the area of the
relation are under pro

of three-John and
November will vote to
on the question pres
country light through



The Evening Star

Red

SPECIAL THIS WEEK

Medicine

Plus Captions!

WALTER MIERNER

MID CITY RADIO SHOP

HOLLYWOOD STARTIME

FOR DEPENDABLE SERVICE CALL **Geigey's** In. 4-66 Atlantic 1166 516 8th St. S.E. NEWBY 2177

AMCO

NORWEGIAN VESSEL

ENOUGH

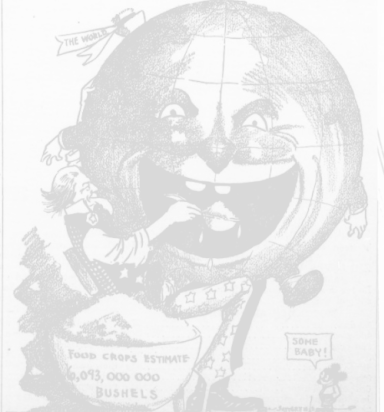
BLACK-White of Columbia is dry by the... of the area of the... of these... and... will vote to... on the... and... light... through...

LEADS HIS LEAGUE ON INITIAL SACK



Hal Chase

Enough For Everybody Except Germany



GOLF WAR THREATENED



RRV. E. W. HAWTHORNE, D. D. Nashville, Tenn., and ardent Congress worker.

Red



bia's Scrap



BELT TO RATES

MASSACHUSETTS	137
NEW YORK	127
ILLINOIS	117
PENNSYLVANIA	107
CALIFORNIA	97
TEXAS	87
INDIANA	77
MICHIGAN	67
OHIO	57
PAID UP	47
MISSOURI	37
WISCONSIN	27
CONNECTICUT	17
NEW JERSEY	7
NEW MEXICO	2
ARIZONA	1
ALABAMA	1
MISSISSIPPI	1
LOUISIANA	1
FLORIDA	1
GEORGIA	1
SOUTH CAROLINA	1
NORTH CAROLINA	1
VIRGINIA	1
MARYLAND	1
DELAWARE	1
MONTANA	1
WYOMING	1
COLORADO	1
NEBRASKA	1
KANSAS	1
MINNESOTA	1
IOWA	1
MORELLO	1
NEW HAMPSHIRE	1
VERMONT	1
NEW ENGLAND	1
NORTHEAST	1
SOUTH	1
WEST	1
CENTRAL	1
TOTAL	10,000
REPRESENTED	853
UNREPRESENTED	10,845

A GRAND ANNUAL OUTLET FOR THE year is annually initiated by J. E. Watson



THE HONORABLE THE HONORABLE THE HONORABLE THE HONORABLE THE HONORABLE



NEW YORK'S LAST HORSE CAR QUILTS



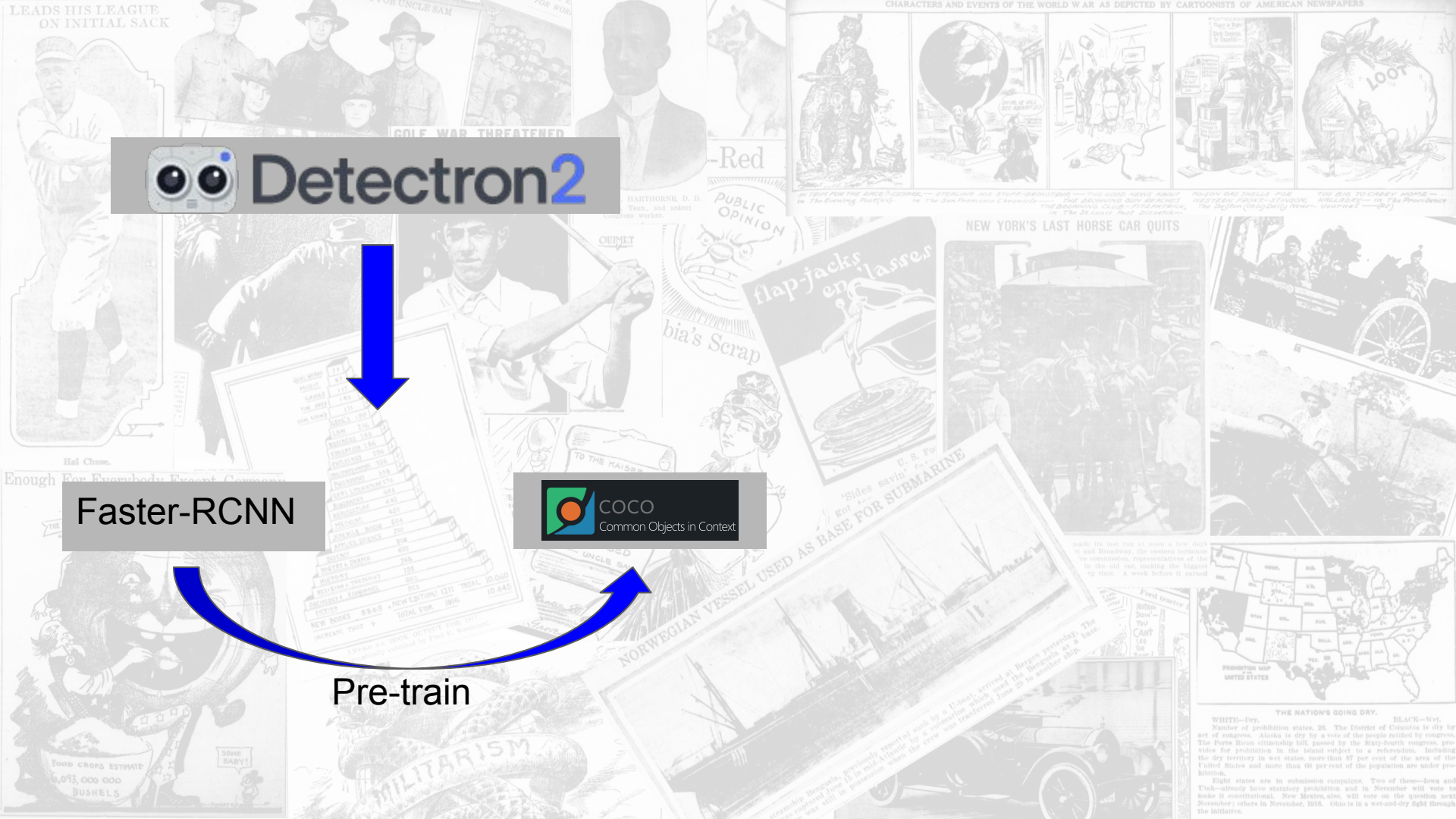
NORWEGIAN VESSEL USED AS BASE FOR SUBMARINE

Norwegian vessel used by a team, arrived at Bergen yesterday. The vessel was used as a base for submarine operations...



THE NATION'S GOING DRY.

WHITE-Ty. BLACK-Wet. Number of prohibition states, 23. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people...



Detectron2



COCO
Common Objects in Context

Faster-RCNN



Pre-train



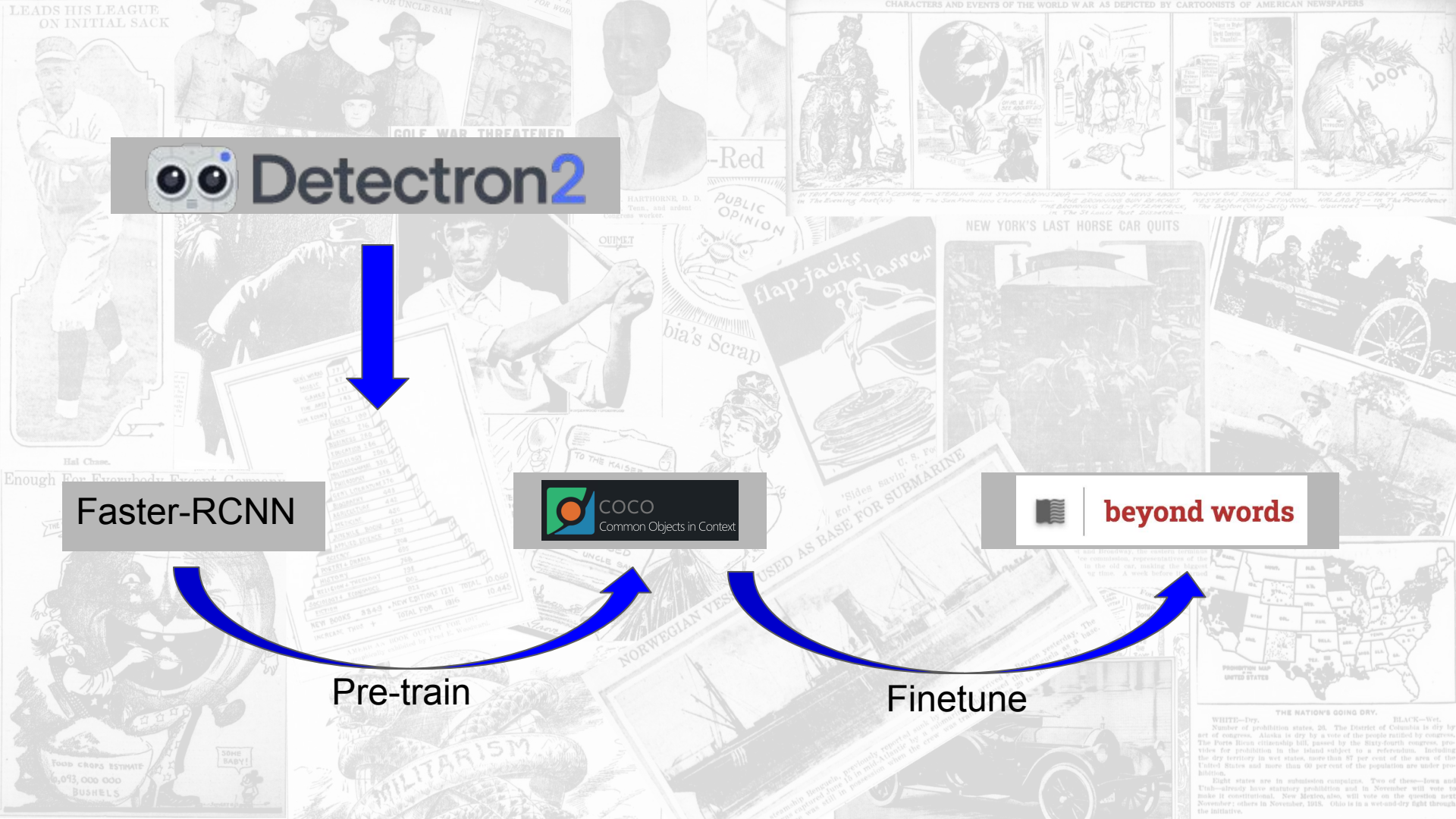
Enough For Everybody Except Company

WHEAT	177
CORN	111
BARLEY	10
RYE	10
CLAY	10
COAL	10
IRON	10
STEEL	10
BRASS	10
COPPER	10
ZINC	10
LEAD	10
SILVER	10
GOLD	10
DIAMONDS	10
JEWELRY	10
TEXTILES	10
FOODS	10
DRUGS	10
TOBACCO	10
ALCOHOL	10
RELIABLE PROPERTY	10
UNRELIABLE PROPERTY	10
RENTALS	10
RENTALS - FOREIGN	10
RENTALS - DOMESTIC	10
RENTALS - TOTAL	10
NEW BOOKS	10
NEW EDITIONS	10
TOTAL	10,645

DEAD AS DOGS OUTLET FOR CITY
Specially compiled by Fred L. Wadsworth

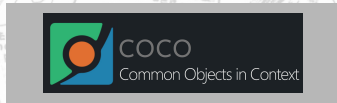


THE NATION'S GOING DRY.
WHITE—Dry. BLACK—Wet.
Number of prohibition states, 21. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people passed by congress. The Porto Rico citizenship bill, passed by the sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 97 per cent of the area of the United States and more than 90 per cent of the population are under prohibition.
Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico will vote on the question next November where in November, 1915. This is a wet and dry fight through the initiative.



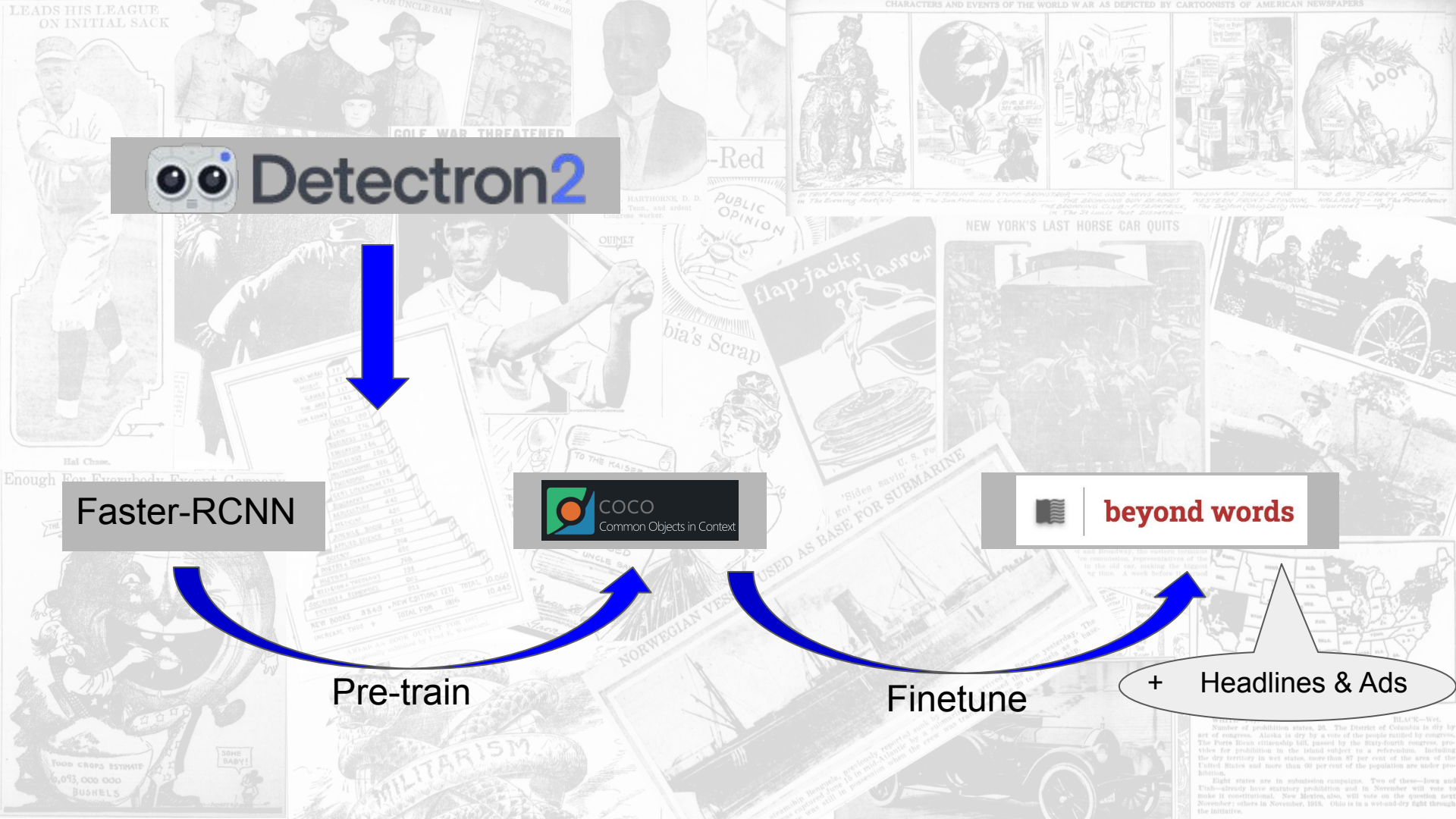
Detectron2

Faster-RCNN

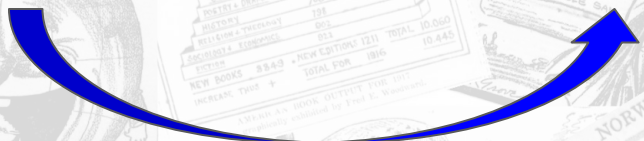
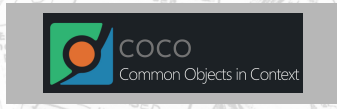


Pre-train

Finetune



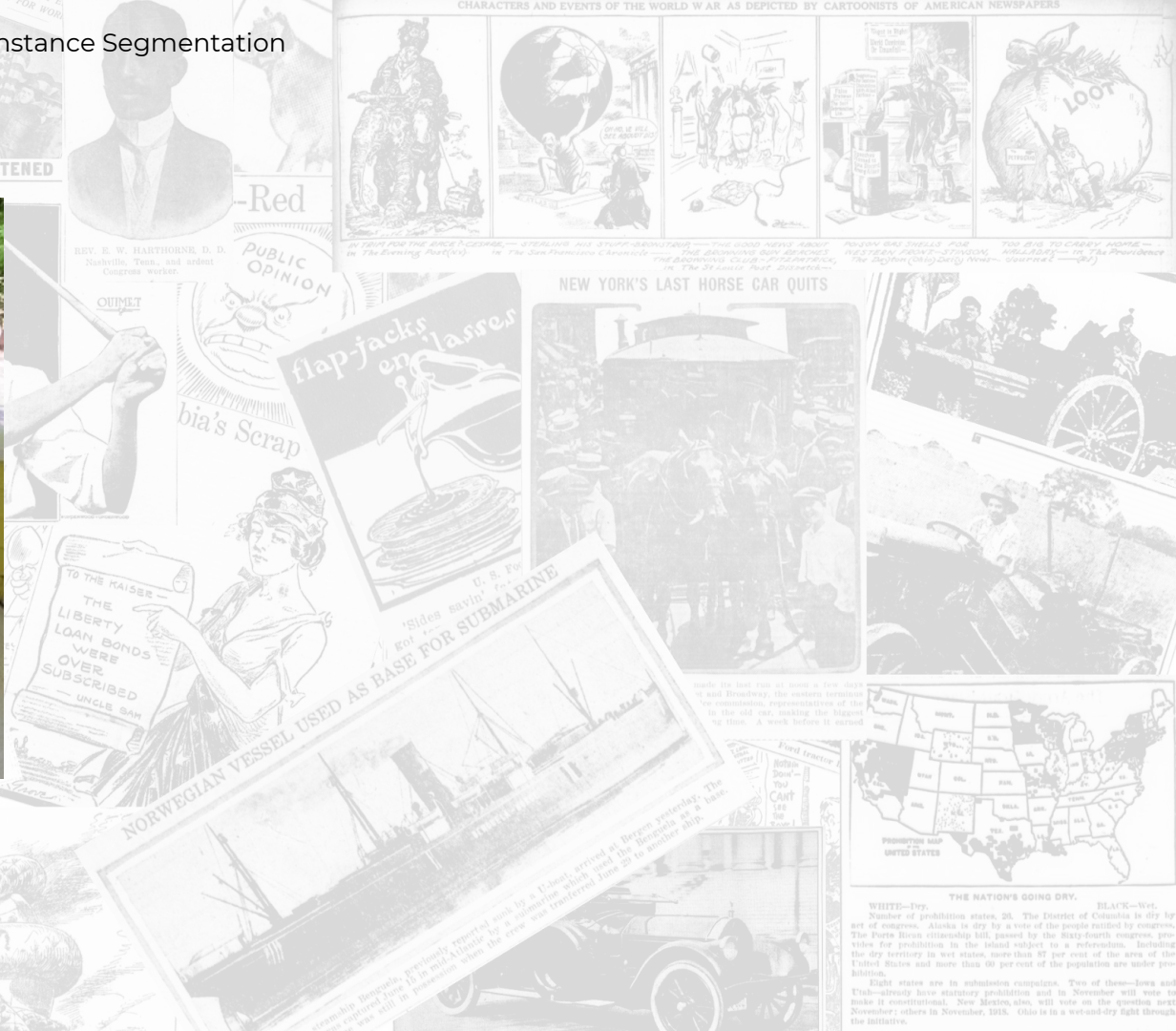
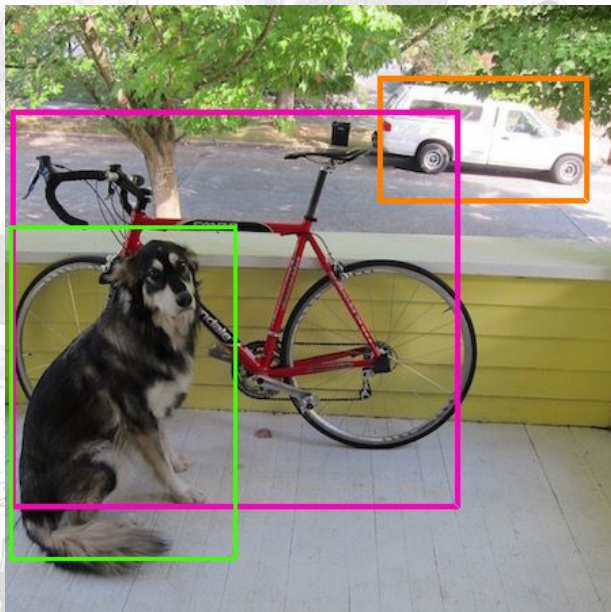
Faster-RCNN



+ Headlines & Ads

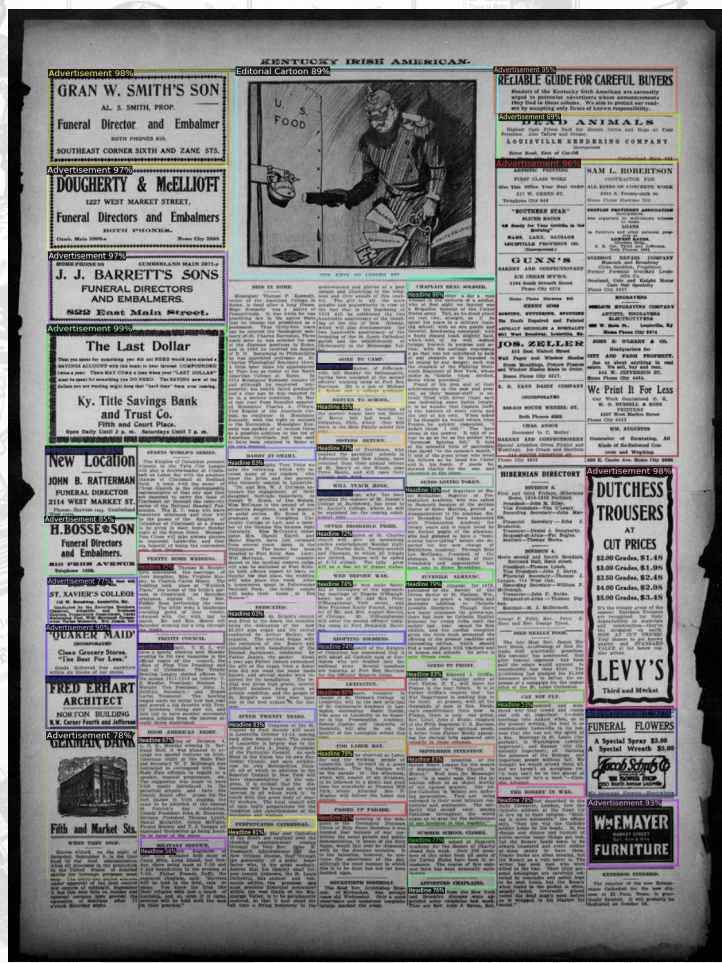
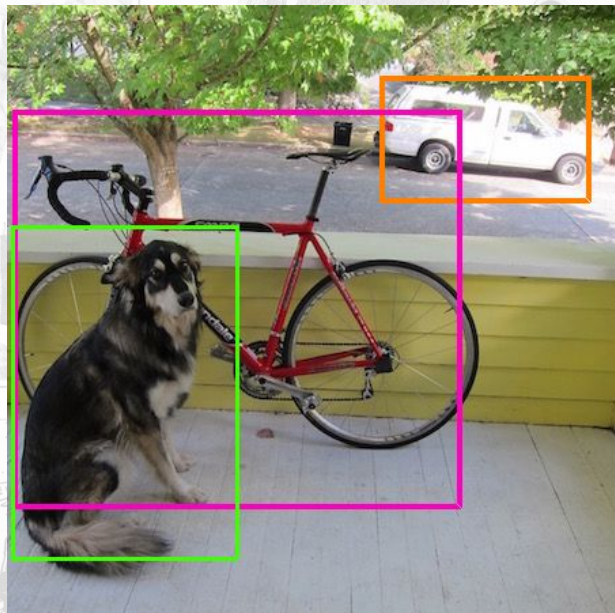
Source: Joe Redmon, Lecture 18: Detection and Instance Segmentation

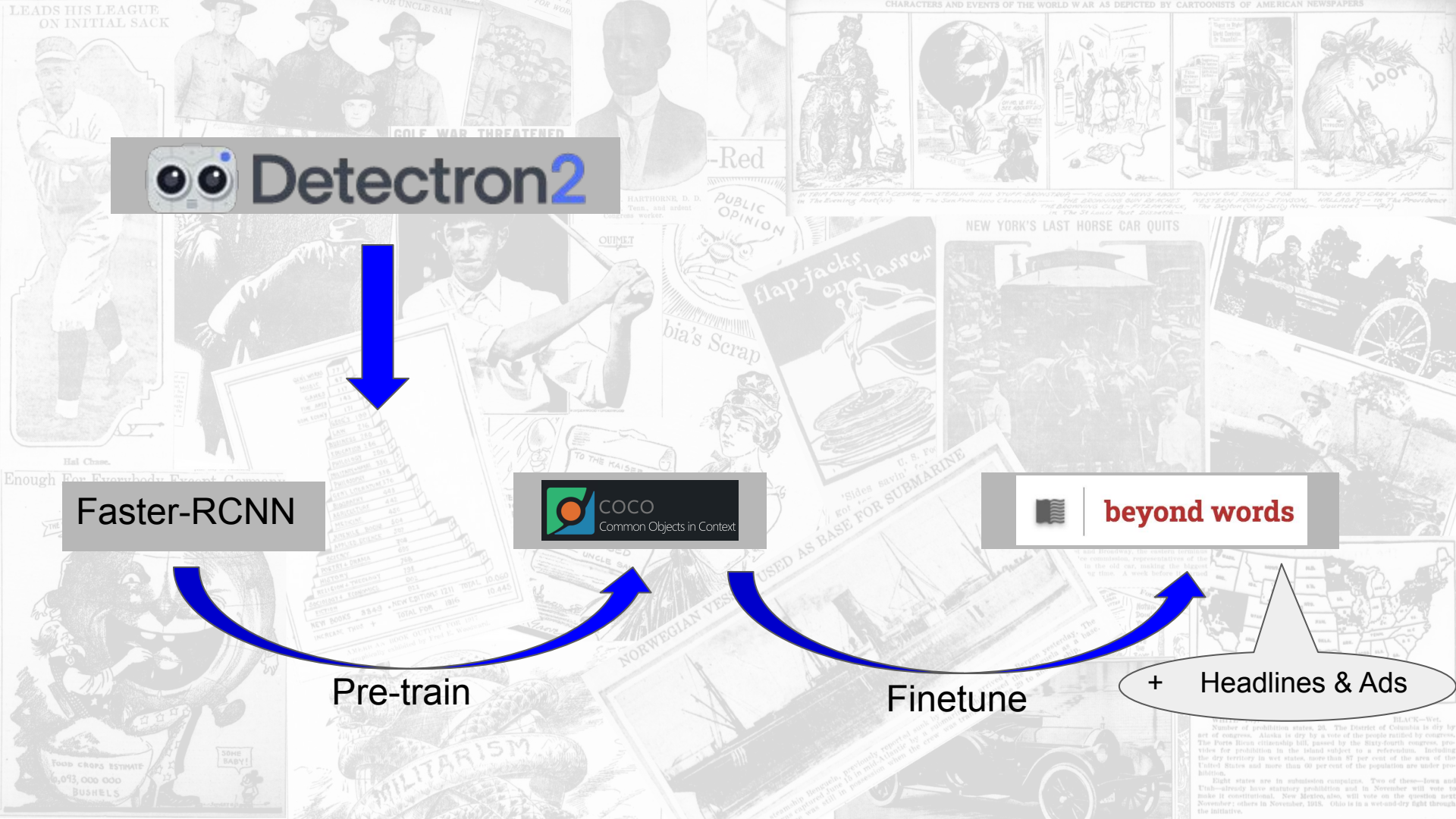
<https://pjreddie.com/courses/computer-vision/>



Source: Joe Redmon, Lecture 18: Detection and Instance Segmentation

<https://pjreddie.com/courses/computer-vision/>






Detectron2



Faster-RCNN

 **COCO**
Common Objects in Context

 **beyond words**

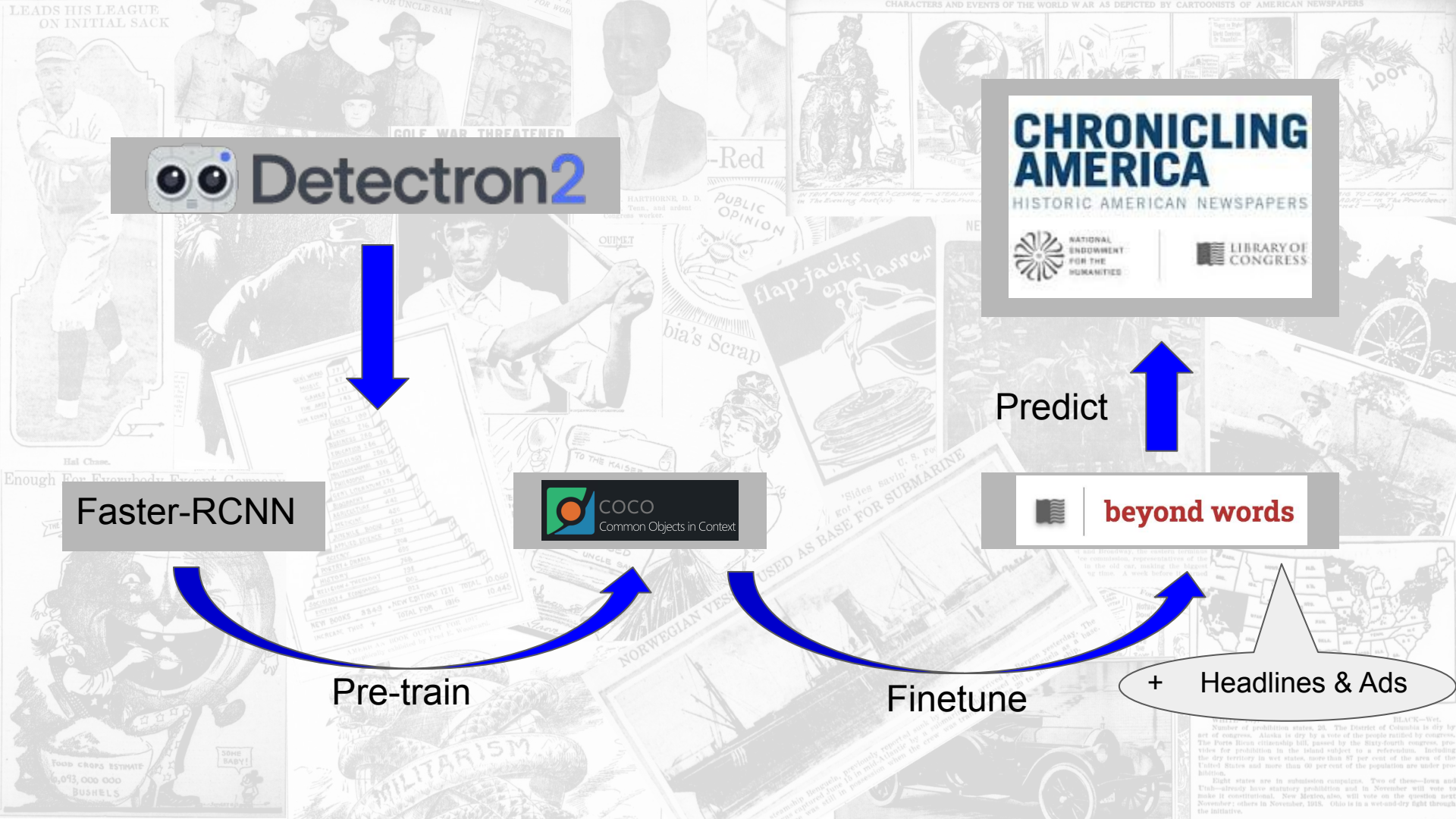


Pre-train



Finetune

+ **Headlines & Ads**



 **Detectron2**


CHRONICLING AMERICA
HISTORIC AMERICAN NEWSPAPERS

 NATIONAL ENDOWMENT FOR THE HUMANITIES

 LIBRARY OF CONGRESS



Faster-RCNN

 **COCO**
Common Objects in Context

 **beyond words**

Predict



Pre-train



Finetune

+ Headlines & Ads

Download Image, METS/ALTO OCR

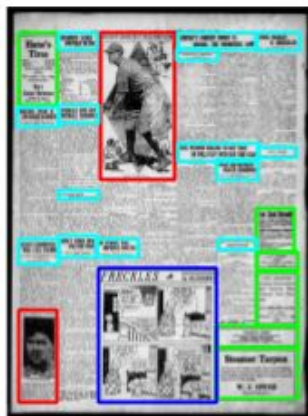


```
<alto xmlns="http://www.loc.gov/standards/alto/ns-v2#"  
  <description>  
    <MeasurementUnit>inch1200//MeasurementUnit>  
    <sourceImageInformation>.../sourceImageInformation>  
  </OCRProcessing ID="OCR_0">  
    <ocrProcessingStep>  
      <processingStepSettings>.../processingStepSettings>  
      <processingSoftware>.../processingSoftware>  
    </ocrProcessingStep>  
  </OCRProcessing>  
</Description>  
<Styles>.../Styles>  
<Layout>.../Layout>  
</alto>
```

OCR

Download Image,
METS/ALTO OCR

Run Visual Content
Recognition



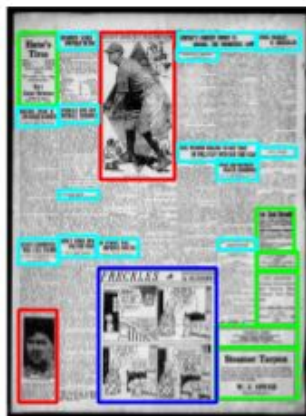
```
<?xml version="1.0" encoding="UTF-8" ?>
<alio xmlns="http://www.loc.gov/standards/alto/ns-v2#"
  <description>
    <MeasurementUnit>inch1200//MeasurementUnit>
    <sourceImageInformation>...</sourceImageInformation>
  </OCRProcessing ID="OCR-0" />
  <?xmlProcessingStep>
    <processingStepSettings>...</processingStepSettings>
    <processingSoftware>...</processingSoftware>
  </OCRProcessingStep>
  </Description>
  <?xml>...</?xml>
  <?xml>...</?xml>
  </alio>
```



Download Image, METS/ALTO OCR



Run Visual Content Recognition



Crop and Save Visual Content



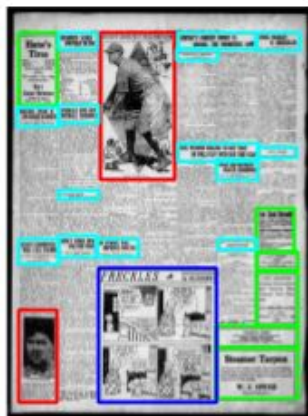
```
<?xml version="1.0" encoding="UTF-8" ?>
<ALTO xmlns="http://www.loc.gov/standards/alto/ns-v2#">
  <Description>
    <MeasurementUnit>inch1200</MeasurementUnit>
    <SourceImageInformation>
      <SourceImageID>OCR_01</SourceImageID>
    </SourceImageInformation>
  </Description>
  <ALTOProcessing>
    <ProcessingStep>
      <ProcessingStepSettings>
        <ProcessingSoftware>ALTO</ProcessingSoftware>
      </ProcessingStepSettings>
    </ProcessingStep>
  </ALTOProcessing>
  <Description>
    <Style>...</Style>
    <Layout>...</Layout>
  </Description>
</ALTO>
```



Download Image, METS/ALTO OCR



Run Visual Content Recognition



Crop and Save Visual Content



Extract OCR within Predicted Bounding Boxes

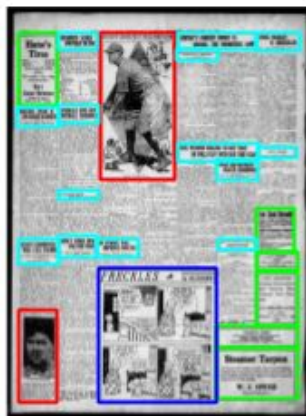
```
<?xml version="1.0" encoding="UTF-8" ?>
<ALTO xmlns="http://www.loc.gov/standards/alto/ns-v2#">
  <Description>
    <MeasurementUnit>Inch1200</MeasurementUnit>
    <SourceImageInformation>...</SourceImageInformation>
  </Description>
  <ALTOProcessingID>OCR-0</ALTOProcessingID>
  <ALTOProcessingStep>
    <ProcessingStepSettings>...</ProcessingStepSettings>
    <ProcessingSoftware>...</ProcessingSoftware>
  </ALTOProcessingStep>
  </ALTOProcessing>
  </Description>
  <Styles>...</Styles>
  <Layout>...</Layout>
</ALTO>
```



Download Image,
METS/ALTO OCR



Run Visual Content
Recognition



Crop and Save
Visual Content



Generate
Image
Embeddings

```
<?xml version="1.0" encoding="UTF-8" ?>
<alio xmlns="http://www.loc.gov/standards/alto/ns-v2#"
  >
  <description>
    <MeasurementUnit>Inch1200</MeasurementUnit>
    <sourceImageInformation>...</sourceImageInformation>
  </alioProcessing ID="OCR-0" ?>
  <alioProcessingStep>
    <processingStepSettings>...</processingStepSettings>
    <processingSoftware>...</processingSoftware>
  </alioProcessingStep>
  </alioProcessing>
  </description>
  <style>...</style>
  <layout>...</layout>
</alio>
```

OCR

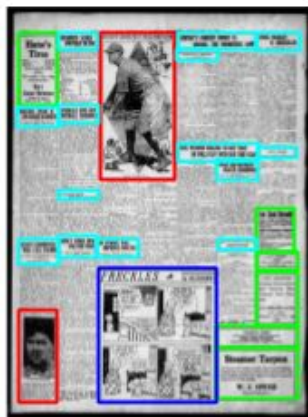
Extract OCR within
Predicted Bounding
Boxes



Download Image,
METS/ALTO OCR



Run Visual Content
Recognition



Crop and Save
Visual Content



Generate
Image
Embeddings

```
<?xml version="1.0" encoding="UTF-8" ?>
<alio xmlns="http://www.loc.gov/standards/alto/ns-v2#"
  <description>
    <MeasurementUnit>Inch1200</MeasurementUnit>
    <sourceImageInformation>...</sourceImageInformation>
  </alio>
```

OCR

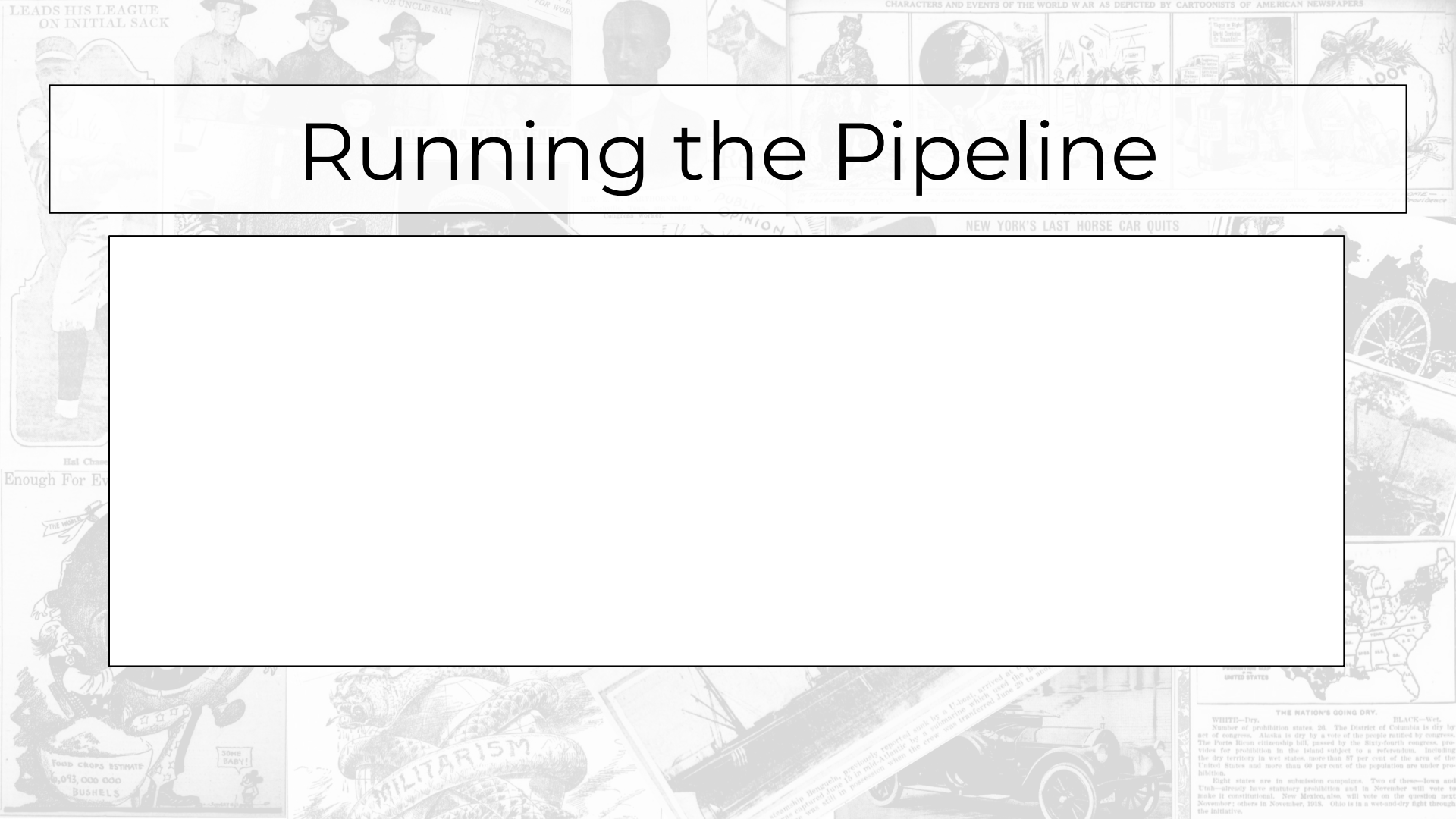
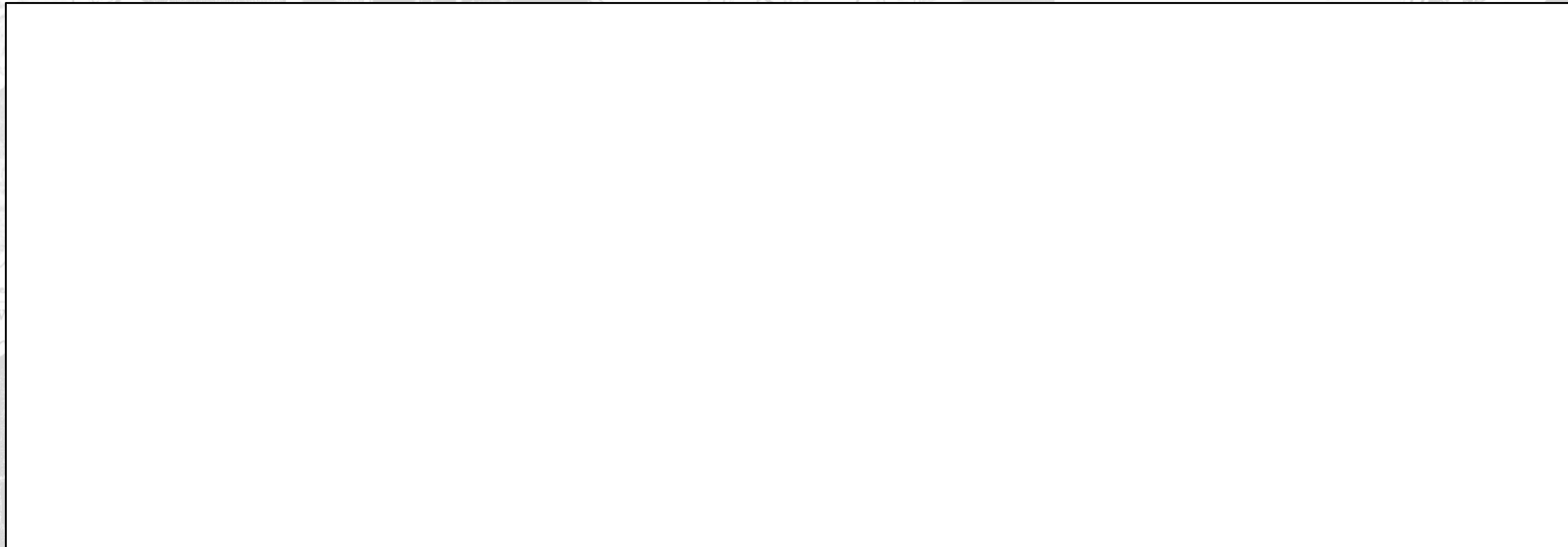


Extract OCR within
Predicted Bounding
Boxes

Save Extracted
Metadata in
JSON Format



Running the Pipeline



Running the Pipeline

- 16,368,041 *Chronicling America* pages processed in total

Running the Pipeline

- 16,368,041 *Chronicling America* pages processed in total
 - ~100 TB of image & XML data

Running the Pipeline

- 16,368,041 *Chronicling America* pages processed in total
 - ~100 TB of image & XML data
- ~19 days on 2 g4dn.12xlarge Amazon AWS EC2 instances

Running the Pipeline

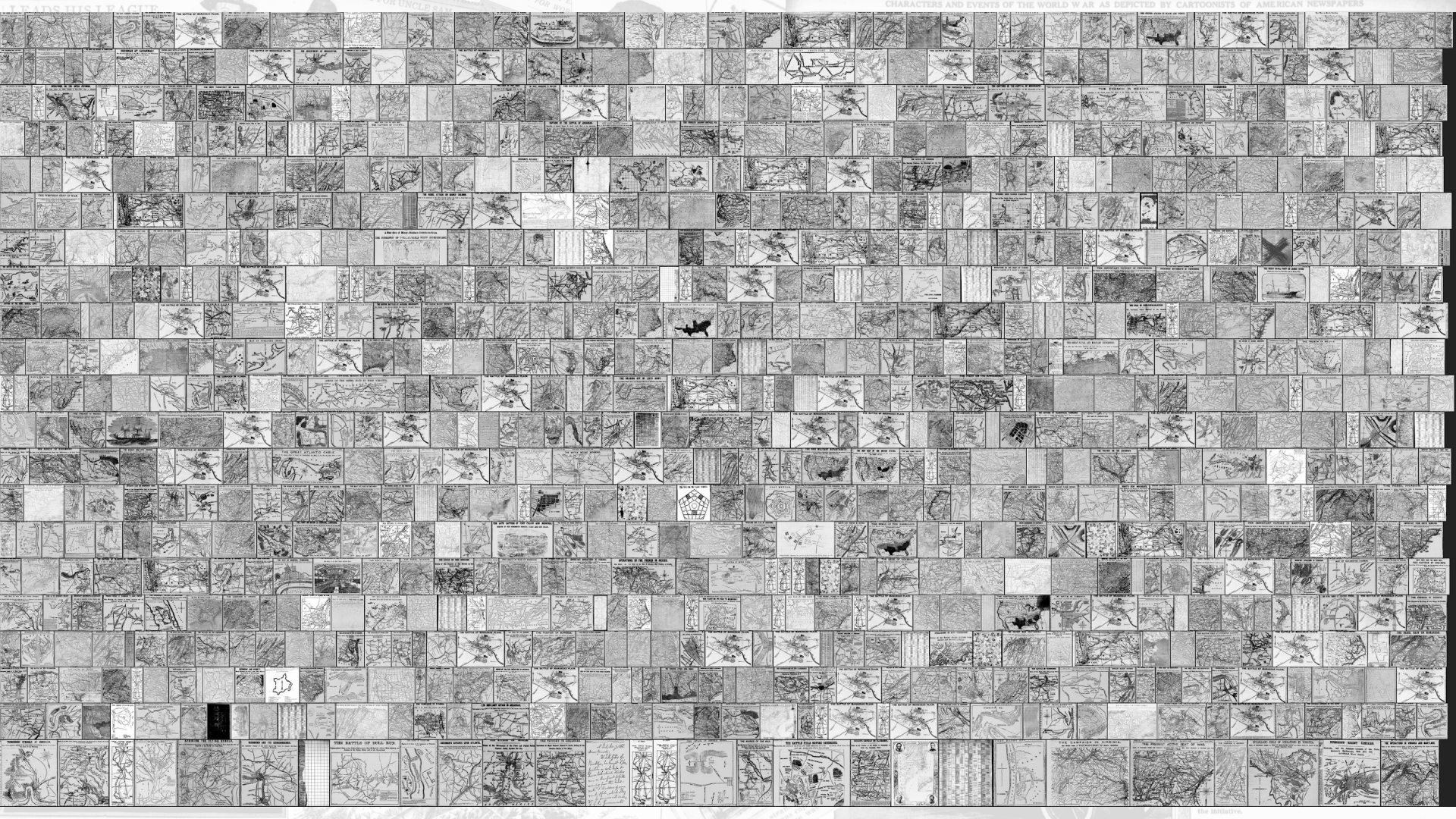
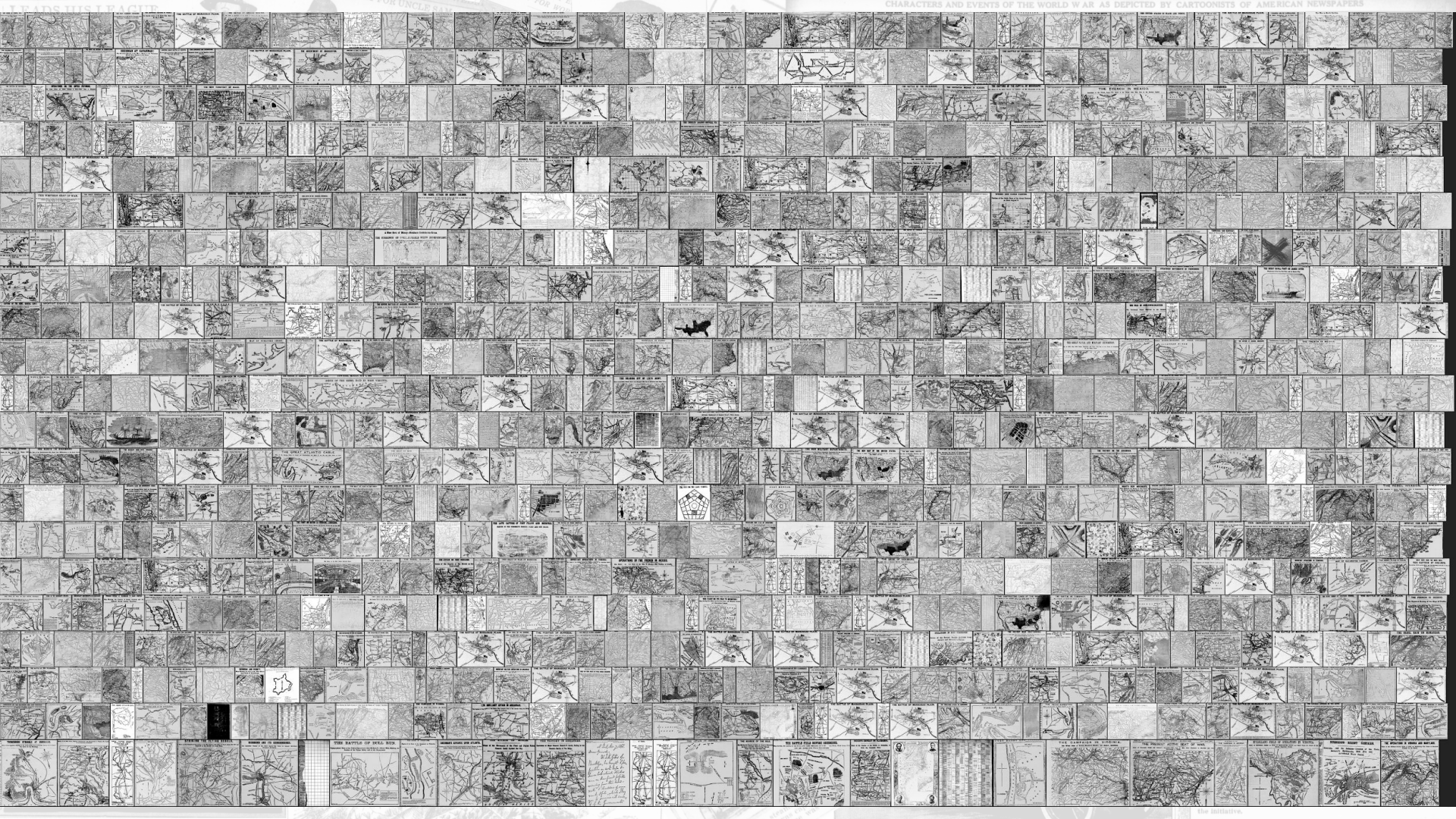
- 16,368,041 *Chronicling America* pages processed in total
 - ~100 TB of image & XML data
- ~19 days on 2 g4dn.12xlarge Amazon AWS EC2 instances
 - 96 vCPUs & 8 NVIDIA T4 GPUs

Running the Pipeline

- 16,368,041 *Chronicling America* pages processed in total
 - ~100 TB of image & XML data
- ~19 days on 2 g4dn.12xlarge Amazon AWS EC2 instances
 - 96 vCPUs & 8 NVIDIA T4 GPUs
- All pipeline code in Python, PyTorch

Running the Pipeline

- 16,368,041 *Chronicling America* pages processed in total
 - ~100 TB of image & XML data
- ~19 days on 2 g4dn.12xlarge Amazon AWS EC2 instances
 - 96 vCPUs & 8 NVIDIA T4 GPUs
- All pipeline code in Python, PyTorch
- notebooks/process_chronam_pages.ipynb in Newspaper Navigator Repo



LIBRARY OF CONGRESS LABS

Experiments

Newspaper Navigator

Explore the visual and textual content within the *Chronicling America* digitized newspaper collection in new ways using machine learning!

Overview

Welcome to the Newspaper Navigator dataset! This dataset consists of extracted visual content for 16.3 million historic newspaper pages in *Chronicling America*. The visual content was identified using an object detection model trained on annotations of diverse text in *Chronicling America* pages. Volunteering annotators made by volunteers as part of the Beyond Words crowdsourcing project. The resulting visual content recognition model detects the following types of visual content:

1. Photograph
2. Illustration
3. Map
4. Cartoon/Caricature
5. Editorial Cartoon
6. Illustration
7. Advertisement

The dataset also includes text corresponding to the visual content, identified by extracting the Optical Character Recognition, or OCR, within each annotated bounding box. For example, if the visual content recognition model produced a bounding box around a headline for corresponding textual content, provides a machine-readable version of the headline. Likewise, for a photograph, illustration, or map, the textual representation often contains descriptive captions.

QR code to Github repo:



beyond words

COCO
Common Objects in Context

arXiv:2005.01583v1 [cs.LG] 4 May 2020

The Newspaper Navigator Dataset: Extracting And Analyzing Visual Content from 16 Million Historic Newspaper Pages in *Chronicling America*

Benjamin Lee University of Washington Library of Congress bgl@cs.washington.edu	Julian Mears LC Labs Library of Congress	Ellen Jakeway LC Labs Library of Congress
Meghan Ferriter LC Labs Library of Congress	Chris Adams IT Design & Development Library of Congress	Nathan Yarassava National Digital Newspaper Program Library of Congress
Deborah Thomas National Digital Newspaper Program Library of Congress	Kate Zward Digital Strategy & LC Labs Library of Congress	Daniel Weld University of Washington

ABSTRACT

Chronicling America, a product of the National Digital Newspaper Program, is a partnership of the Library of Congress and the National Endowment for the Humanities to digitize historic newspapers. Over 16 million pages of historic American newspapers have been digitized for *Chronicling America* to date, complete with high-resolution images and machine-readable METS/SALTO OCR. Of considerable interest in *Chronicling America* users is a semi-annotated corpus, complete with extracted visual content and headlines. To accomplish this, we introduce a visual content recognition model trained on bounding box annotations of photographs, illustrations, maps, comics, and editorial cartoons collected as part of the Library of Congress's Beyond Words crowdsourcing initiative and augmented with additional annotations including those of headline advertisements. We describe our pipeline that utilizes this deep learning model to extract 7 classes of visual content: headlines, photographs, illustrations, maps, comics, editorial cartoons, and advertisements, complete with textual content such as captions derived from the METS/SALTO OCR, as well as image boudings for fast work conducted with an InnoScale ResNet at the Library of Congress and P.D. studies in the Paul H. Hays School for Computer Science and Engineering at the University of Washington.

INTRODUCTION

Chronicling America, an initiative of the National Digital Newspaper Program, is a partnership of the Library of Congress and the National Endowment for the Humanities - an invaluable resource for academic, local, and public historians, educators and students, genealogists, journalists, and members of the public to explore American history through uniquely rich content preserved within historic local newspapers. Over 16 million pages of newspapers published between 1789 to 1963 are publicly available online through a search portal as well as via a public API. Among the page-level data are 400 EPW images, as well as METS/SALTO OCR, a standard maintained by the Library of Congress that includes text localization [2].

The 16.3 million *Chronicling America* pages included in the Newspaper Navigator cover 174 years of American history, inclusive of 47 states, Washington, D.C., and Puerto Rico. In Figure 1, we show overlaid maps displaying the geographic coverage of the 16.3 million *Chronicling America* newspaper pages included in the Newspaper Navigator dataset. In Figure 2, we show the temporal coverage of these pages. The coverage reflects the selection process for determining which newspapers to include in *Chronicling America*, for as in-depth

17 of 18

Download History

This notebook trains a visual content recognition model on *Beyond Words* bounding box annotations (plus additional headline + advertisement annotations).

designer: chris-adams (see COCO metadata at <https://cocodataset.org>)

This notebook features a pre-trained object detection model (Faster-RCNN R50FPN7) to predict bounding boxes around photographs, illustrations, maps, comics, editorial cartoons, headlines, and annotations in historical newspaper images. The notebook is described in more detail in the newspaper-navigator repo.

This notebook makes the following assumptions:

- The Newspaper Navigator repo has been cloned, so the local file structure is intact.
- Dockerhub has been installed. For Amazon EC2 instances with a Deep Learning Ubuntu AMI, the bash script for installing Dockerhub (install-dockerhub.sh) can be run to bypass the installation (see the repo for more info).

This notebook is based on the following notebook:

<https://github.com/alexanderjones/notebook/blob/master/15%20a%20COCO%20metadata%20COCO.ipynb>

First, we handle imports and data formatting.

This cell imports libraries and constructs a COCO instance using the training and validation JSON files; essentially, this enables the model to handle the data loading using the COCO standard:

```

(1) from object_detection.config import CfgProto
import sys
import os
import cv2
import random
import glob

```

news-navigator.labs.loc.gov

#NewspaperNavigator

LIBRARY OF CONGRESS LABS

Experiments

Newspaper Navigator

Explore the visual and textual content within the *Chronicling America* digitized newspaper collection in new ways using machine learning!

Overview

Relative to the *Newspaper Navigator* dataset: This dataset consists of extracted visual content for 16,526,161 historic newspaper pages in *Chronicling America*. The visual content was identified using an image detection model based on annotations of diverse text in *Chronicling America* pages, including annotations made by volunteers as part of the *Beyond Words* crowdsourcing project. The resulting visual content recognition model detects the following types of visual content:

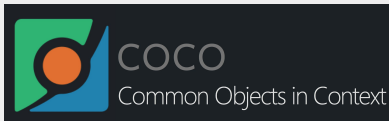
1. Photograph
2. Illustration
3. Map
4. Cartoon/Caricature
5. Editorial Cartoon
6. Illustration
7. Advertisement

The dataset also includes text corresponding to the visual content, identified by extracting the Optical Character Recognition, or OCR, within each annotated bounding box. For example, if the visual content recognition model produced a bounding box around a headline for corresponding textual content, provides a machine-readable version of the headline. Likewise, for a photograph, illustration, or map, the textual representation often contains descriptive captions.

- HTTPS, S3 requests
- Pre-packaged datasets



 **beyond words**



The Newspaper Navigator Dataset: Extracting And Analyzing Visual Content from 16 Million Historic Newspaper Pages in *Chronicling America*

Benjamin Lee University of Washington Library of Congress bgl@cs.washington.edu	Julian Mears LC Labs Library of Congress	Eileen Jakeway LC Labs Library of Congress
Meghan Ferriter LC Labs Library of Congress	Chris Adams IT Design & Development Library of Congress	Nathan Yarassavage National Digital Newspaper Program Library of Congress
Deborah Thomas National Digital Newspaper Program Library of Congress	Kate Zward Digital Strategy & LC Labs Library of Congress	Daniel Weld University of Washington

arXiv:2005.01583v1 [cs.LG] 4 May 2020

ABSTRACT

Chronicling America is a product of the National Digital Newspaper Program, a partnership between the Library of Congress and the National Endowment for the Humanities to digitize historic newspapers. Over 16 million pages of historic American newspapers have been digitized for *Chronicling America* to date, complete with high-resolution images and machine-readable METS/ALTO OCR. Of considerable interest in *Chronicling America* users is a semi-annotated corpus complete with extracted visual content and headlines. To accomplish this, we introduce a visual content recognition model trained on bounding box annotations of photographs, illustrations, maps, comics, and editorial cartoons collected as part of the Library of Congress's *Beyond Words* crowdsourcing initiative and augmented with additional annotations including those of headline advertisements. We describe our pipeline that utilizes this deep learning model to extract 7 classes of visual content: headlines, photographs, illustrations, maps, comics, editorial cartoons, and advertisements, complete with textual content such as captions derived from the METS/ALTO OCR, as well as image embeddings for fast work conducted within an InnoScale instance at the Library of Congress and D.D. studies in the Paul H. Hays School for Computer Science and Engineering at the University of Washington.

INTRODUCTION

Chronicling America, an initiative of the National Digital Newspaper Program, is a partnership of the Library of Congress and the National Endowment for the Humanities - an invaluable resource for academic, local, and public historians, educators and students, genealogists, journalists, and members of the public to explore American history through uniquely rich content preserved within historic local newspapers. Over 16 million pages of newspapers published between 1789 to 1963 are publicly available online through a search portal as well as via a public API. Among the page-level data are 400 EPW images, as well as METS/ALTO OCR, a standard maintained by the Library of Congress that includes text localization [2].

The 16.3 million *Chronicling America* pages included in the *Newspaper Navigator* cover 174 years of American history, inclusive of 47 states, Washington, D.C., and Puerto Rico. In Figure 1, we show choropleth maps displaying the geographic coverage of the 16.3 million *Chronicling America* newspaper pages included in the *Newspaper Navigator* dataset. In Figure 2, we show the temporal coverage of these pages. The coverage reflects the selection process for determining which newspapers to include in *Chronicling America*, for as-in-depth

17.4 MB Download History

This notebook trains a visual content recognition model on *Beyond Words* bounding box annotations (plus additional headline + advertisement annotations).

designer: chris-adams (see COCO metadata at cocodataset.org)

This notebook features a pre-trained object detection model (Faster-RCNN R50FPN) to predict bounding boxes around photographs, illustrations, maps, comics, editorial cartoons, headlines, and annotations in historical newspaper images. The notebook is described in more detail in the *newspaper-navigator* repo.

This notebook makes the following assumptions:

- The *Newspaper Navigator* repo has been cloned, so the local file structure is intact.
- Detectron2 has been installed. For Amazon EC2 instances with a Deep Learning Ubuntu AMI, the bash script for installing Detectron2 (install-detectron2.sh) can be run to leverage the installation (see the repo for more info).

This notebook is based on the following notebook:

- <https://github.com/robertknight/visual-content-recognition>

First, we handle imports and data formatting.

This cell imports libraries and constructs a COCO instance using the training and validation JSON files; essentially, this enables the model to handle the data loading using the COCO standard.

```
In [1]: # To do: don't change the line
import os
if os.getenv('PWD') is None:
    os.chdir('..')
import sys
import cv2
import random
import glob
```

news-navigator.labs.loc.gov

#NewspaperNavigator

Newspaper Navigator

Explore the visual and textual content within the *Chronicling America* digitized newspaper collection in new ways using machine learning!

Overview

Belonging to the Newspaper Navigator dataset? This dataset consists of extracted visual content for 16.3 million historic newspaper pages in *Chronicling America*. The visual content was identified using an object detection model trained on annotations of diverse text in *Chronicling America* pages, including annotations made by volunteers as part of the *Beyond Words* crowdsourcing project. The resulting visual content recognition model detects the following types of visual content:

1. Photograph
2. Illustration
3. Map
4. Cartographic
5. Editorial Cartoon
6. Headline
7. Advertisement

The dataset also includes text corresponding to the visual content, identified by extracting the Optical Character Recognition, or OCR, within each annotated bounding box. For example, if the visual content recognition model produced a bounding box around a headline for corresponding textual content, provides a machine-readable version of the headline. Likewise, for a photograph, illustration, or map, the textual representation often contains a title and caption.

QR code to Github repo:



beyond words

COCO
Common Objects in Context

The Newspaper Navigator Dataset: Extracting And Analyzing Visual Content from 16 Million Historic Newspaper Pages in *Chronicling America*

Benjamin Lee University of Washington Library of Congress bgl@cs.washington.edu	Julian Mears LC Labs Library of Congress	Ellen Jakeway LC Labs Library of Congress
Meghan Ferriter LC Labs Library of Congress	Chris Adams IT Design & Development Library of Congress	Nathan Yarassava National Digital Newspaper Program Library of Congress
Deborah Thomas National Digital Newspaper Program Library of Congress	Kate Zward Digital Strategy & LC Labs Library of Congress	Daniel Weld University of Washington

arXiv:2005.01583v1 [cs.LG] 4 May 2020

ABSTRACT

Chronicling America is a product of the National Digital Newspaper Program, a partnership between the Library of Congress and the National Endowment for the Humanities to digitize historic newspapers. Over 16 million pages of historic American newspapers have been digitized for *Chronicling America* to date, complete with high-resolution images and machine-readable METS/SALTO OCR. Of considerable interest to *Chronicling America* users is a semi-annotated corpus, complete with extracted visual content and headlines. To accomplish this, we introduce a visual content recognition model trained on bounding box annotations of photographs, illustrations, maps, comics, and editorial cartoons collected as part of the Library of Congress's *Beyond Words* crowdsourcing initiative and augmented with additional annotations including those of other user advertisements. We describe our pipeline that utilizes this deep learning model to extract 7 classes of visual content: headlines, photographs, illustrations, maps, comics, editorial cartoons, and advertisements, complete with textual content such as captions derived from the METS/SALTO OCR, as well as image embeddings for fast work conducted with an InnoScale instance at the Library of Congress and D.D. studies in the Paul H. Hays School for Computer Science and Engineering at the University of Washington.

INTRODUCTION

Chronicling America, an initiative of the National Digital Newspaper Program, is a partnership of the Library of Congress and the National Endowment for the Humanities - an invaluable resource for academic, local, and public historians, educators and students, genealogists, journalists, and members of the public. Over 16 million pages of newspapers published between 1789 to 1963 are publicly available online through a search portal as well as via a public API. Among the page-level data are 400 EP3 images, as well as METS/SALTO OCR, a standard maintained by the Library of Congress that includes text localization [2].

The 16.3 million *Chronicling America* pages included in the Newspaper Navigator cover 174 years of American history, inclusive of 47 states, Washington, D.C., and Puerto Rico. In Figure 1, we show choropleth maps displaying the geographic coverage of the 16.3 million *Chronicling America* newspaper pages included in the Newspaper Navigator dataset. In Figure 2, we show the temporal coverage of these pages. The coverage reflects the selection process for determining which newspapers to include in *Chronicling America*, for as-in-depth

This notebook trains a visual content recognition model on *Beyond Words* bounding box annotations (plus additional headline + advertisement annotations).

This notebook features a pre-trained object detection model (Faster-RCNN R50FPN) to predict bounding boxes around photographs, illustrations, maps, comics, editorial cartoons, headlines, and annotations in historical newspaper images. The notebook is described in more detail in the newspaper-navigator repo.

This notebook makes the following assumptions:

- The Newspaper Navigator repo has been cloned, so the local file structure is intact.
- Dockerhub has been installed. For Amazon EC2 instances with a Deep Learning Ubuntu AMI, the bash script for installing Dockerhub (install-dockerhub.sh) can be run to bypass the installation (see the repo for more info).

This notebook is based on the following notebook:

<https://colab.research.google.com/github/USICA/colabCOCO/blob/master/COCO.ipynb>

First, we handle imports and data formatting.

This cell imports libraries and constructs a COCO instance using the training and validation JSON files; essentially, this enables the model to handle the data loading using the COCO library:

```

In [1]: # To do this on Colab, set the
import os
if os.getenv('IS_COLAB') == 'true':
    import os
    import sys
    import sys
    import sys

```

Experiments

Newspaper Navigator



Explore the visual and textual content within the *Chronicling America* digitized newspaper collection in new ways using machine learning!

- Overview**
- Welcome to the Newspaper Navigator dataset! This dataset consists of extracted visual content for 16.3 million historic newspaper pages in *Chronicling America*. The visual content was identified using an object detection model trained on annotations of diverse text in *Chronicling America* pages. Volunteering annotators made by volunteers as part of the *Beyond Words* crowdsourcing project. The resulting visual content recognition model detects the following types of visual content:
1. Photograph
 2. Illustration
 3. Map
 4. Cartographic
 5. Editorial Cartoon
 6. Headline
 7. Advertisement

The dataset also includes text corresponding to the visual content, identified by extracting the Optical Character Recognition, or OCR, within each annotated bounding box. For example, if the visual content recognition model produced a bounding box around a headline, the corresponding textual content provides a machine-readable version of the headline. Likewise, for a photograph, illustration, or map, the textual representation often contains a title and caption.

QR code to Github repo:



The Newspaper Navigator Dataset: Extracting And Analyzing Visual Content from 16 Million Historic Newspaper Pages in *Chronicling America*

- | | | |
|---|--|--|
| Benjamin Lee
University of Washington
Library of Congress
bgl@cs.washington.edu | Julian Mears
LC Labs
Library of Congress | Elihen Jakeway
LC Labs
Library of Congress |
| Meghan Ferriter
LC Labs
Library of Congress | Chris Adams
IT Design & Development
Library of Congress | Nathan Yarassavage
National Digital Newspaper Program
Library of Congress |
| Deborah Thomas
National Digital Newspaper Program
Library of Congress | Kate Zward
Digital Strategy & LC Labs
Library of Congress | Daniel Weld
University of Washington |

arXiv:2005.01583v1 [cs.LG] 4 May 2020

ABSTRACT

Chronicling America is a product of the National Digital Newspaper Program, a partnership between the Library of Congress and the National Endowment for the Humanities to digitize historic newspapers. Over 16 million pages of historic American newspapers have been digitized for *Chronicling America* to date, complete with high-resolution images and machine-readable METS/SALTO OCR. Of considerable interest in *Chronicling America* users is a semi-annotated corpus, complete with extracted visual content and headlines. To accomplish this, we introduce a visual content recognition model trained on bounding box annotations of photographs, illustrations, maps, comics, and editorial cartoons collected as part of the Library of Congress's *Beyond Words* crowdsourcing initiative and augmented with additional annotations including those of other advertisements. We describe our pipeline that utilizes this deep learning model to extract 7 classes of visual content: headlines, photographs, illustrations, maps, comics, editorial cartoons, and advertisements, complete with textual content such as captions derived from the METS/SALTO OCR, as well as image embeddings for fast work conducted with an Inception v3 ResNet at the Library of Congress and D.P. Anderson in the Paul H. Hays School for Computer Science and Engineering at the University of Washington.

INTRODUCTION

Chronicling America, an initiative of the National Digital Newspaper Program, recently a partnership of the Library of Congress and the National Endowment for the Humanities - is an invaluable resource for academic, local, and public historians, educators and students, genealogists, journalists, and members of the public. It offers a unique and historically rich content preserved within historic local newspapers. Over 16 million pages of newspapers published between 1789 to 1963 are publicly available online through a search portal as well as via a public API. Among the page-level data are 400 DPI images, as well as METS/SALTO OCR, a standard maintained by the Library of Congress that includes text localization [2].

The 16.3 million *Chronicling America* pages included in the Newspaper Navigator cover 174 years of American history, inclusive of 47 states, Washington, D.C., and Puerto Rico. In Figure 1, we show eleven different maps displaying the geographic coverage of the 16.3 million *Chronicling America* newspaper pages included in the Newspaper Navigator dataset. In Figure 2, we show the temporal coverage of these pages. The coverage reflects the selection process for determining which newspapers to include in *Chronicling America*; for an in-depth

beyond words

COCO
Common Objects in Context

17.4 MB Download History

This notebook trains a visual content recognition model on *Beyond Words* bounding box annotations (plus additional headline + advertisement annotations).

This notebook features a pre-trained object detection model (Faster-RCNN R50FPN) to predict bounding boxes around photographs, illustrations, maps, comics, editorial cartoons, headlines, and annotations in historical newspaper images. The notebook is described in more detail in the newspaper-navigator repo.

This notebook makes the following assumptions:

- The newspaper-navigator repo has been cloned, so the local file structure is intact.
- Docker has been installed. For Amazon EC2 instances with a Deep Learning AMI, the bash script for installing Docker (install-docker.sh) can be run to bypass the installation (see the repo for more info).

This notebook is based on the following notebook:

https://github.com/alexmjordan/coco-research/blob/master/15_coco_research.ipynb

First, we handle imports and data formatting.

This cell imports libraries and constructs a COCO instance using the training and validation JSON files; essentially, this enables the model to handle the data loading using the COCO libraries:

```
In [1]: # To do: don't change! Set the
import sys
import os
if os.getenv('HOME') is None:
    os.environ['HOME'] = os.path.expanduser('~')
import cv2
import random
```


Newspaper Navigator

Explore the visual and textual content within the *Chronicling America* digitized newspaper collection in new ways using machine learning!



Overview

Welcome to the Newspaper Navigator dataset! This dataset consists of extracted visual content for 1,325,611 historic newspaper pages in *Chronicling America*. The visual content was identified using an object detection model trained on annotations of diverse types of *Chronicling America* pages, including advertisements made by volunteers as part of the *Beyond Words* crowdsourcing project. The resulting visual content recognition model detects the following types of visual content:

1. Photograph
2. Illustration
3. Map
4. Cartoon/Caricature
5. Editorial Cartoon
6. Headline
7. Advertisement

The dataset also includes text corresponding to the visual content, identified by extracting the Optical Character Recognition, or OCR, within each annotated bounding box. For example, if the visual content recognition model predicted a bounding box around a headline, the corresponding textual content provides a machine-readable version of the headline. Likewise, for a photograph, illustration, or map, the textual representation often contains a title and caption.

QR code to Github repo:



The Newspaper Navigator Dataset: Extracting And Analyzing Visual Content from 16 Million Historic Newspaper Pages in *Chronicling America*

Benjamin Lee University of Washington Library of Congress bgl@cs.washington.edu	Julian Mears LC Labs Library of Congress	Elihen Jakeway LC Labs Library of Congress
Meghan Ferriter LC Labs Library of Congress	Chris Adams IT Design & Development Library of Congress	Nathan Yarassavage National Digital Newspaper Program Library of Congress
Deborah Thomas National Digital Newspaper Program Library of Congress	Kate Zward Digital Strategy & LC Labs Library of Congress	Daniel Weld University of Washington

arXiv:2005.01853v1 [cs.LG] 4 May 2020

ABSTRACT
Chronicling America is a product of the National Digital Newspaper Program, a partnership of the Library of Congress and the National Endowment for the Humanities to digitize historic newspapers. Over 16 million pages of historic American newspapers have been digitized for *Chronicling America* to date, complete with high-resolution images and machine-readable METS/SALTO OCR. Of considerable interest in *Chronicling America* users is a semi-annotated corpus, complete with extracted visual content and headlines. To accomplish this, we introduce a visual content recognition model trained on bounding box annotations of photographs, illustrations, maps, comics, and editorial cartoons collected as part of the Library of Congress's *Beyond Words* crowdsourcing initiative and augmented with additional annotations including those of headlines and advertisements. We describe our pipeline that utilizes this deep learning model to extract 7 classes of visual content: headlines, photographs, illustrations, maps, comics, editorial cartoons, and advertisements, complete with textual content such as captions derived from the METS/SALTO OCR, as well as image descriptions for fast

INTRODUCTION
Chronicling America, an initiative of the National Digital Newspaper Program, is a partnership of the Library of Congress and the National Endowment for the Humanities - an invaluable resource for academic, local, and public historians, educators and students, genealogists, journalists, and members of the public to explore American history through uniquely rich content preserved within historic local newspapers. Over 16 million pages of newspapers published between 1789 to 1963 are publicly available online through a search portal as well as a public API. Among the page-level data are 400 EPW images, as well as METS/SALTO OCR, a standard maintained by the Library of Congress that includes text localization [2].

The 16.3 million *Chronicling America* pages included in the Newspaper Navigator cover 174 years of American history, inclusive of 47 states, Washington, D.C., and Puerto Rico. In Figure 1, we show eleven-thick maps displaying the geographic coverage of the 16.3 million *Chronicling America* newspaper pages included in the Newspaper Navigator dataset. In Figure 2, we show the temporal coverage of these pages. The coverage reflects the selection process for determining which newspapers to include in *Chronicling America*; for an in-depth

*Work conducted while an InnoCentive or Residence at the Library of Congress and D.D. Anderson in the Paul G. Allen School for Computer Science and Engineering at the University of Washington.

beyond words

COCO
Common Objects in Context

This notebook trains a visual content recognition model on *Beyond Words* bounding box annotations (plus additional headline + advertisement annotations).

This notebook features a pre-trained object detection model (Faster-RCNN R50FPN) to predict bounding boxes around photographs, illustrations, maps, comics, editorial cartoons, headlines, and annotations in historical newspaper images. The notebook is described in more detail in the newspaper-navigator repo.

This notebook makes the following assumptions:

- The Newspaper Navigator repo has been cloned, so the local file structure is intact.
- Detectors have been installed. For Amazon EC2 instances with a Deep Learning AMI, the bash script for installing detectors (install-detectors.sh) can be run to handle the installation (see the repo for more info).

This notebook is based on the following notebook:

<https://coco.common-objects.org/notebooks/visual-content-recognition.html>

First, we handle imports and data formatting.

This cell imports libraries and constructs a COCO instance using the training and validation JSON files; essentially, this enables the model to handle the data loading using the COCO library:

```

In [1]: # To allow images to be
import cv2
import numpy as np
import os
import sys
import glob

```

LEADS HIS LEAGUE
ON INITIAL SACK

LIBRARY OF CONGRESS L.A.B.S.

Experiments

Newspaper Navigator

Explore the visual and textual content within the *Chronicling America* newspaper collection in new ways using machine learning!

Overview

Relative to the *Newspaper Navigator* dataset? This dataset consists of extracted visual content from 16,326,241 historic newspaper pages in *Chronicling America*. The visual content was identified using a computer vision model based on annotations of bounding boxes in *Chronicling America* pages. Annotations made by volunteers as part of the *Beyond Words* crowdsourcing project. The visual content extraction algorithm must detect the following types of visual content:

1. Photograph
2. Illustration
3. Map
4. Cartoon/Caricature
5. Editorial Cartoon
6. Reading
7. Advertisement

The dataset also includes text corresponding to the visual content, identified by extracting the Optical Character Recognition, or OCR, within each annotated bounding box. For example, if the visual content recognition model produces a bounding box around a headline, the corresponding OCR text is also included.

**The Newspaper Navigator Dataset:
Extracting And Analyzing Visual Content from 16 Million
Historic Newspaper Pages in Chronicling America**

Benjamin Lee University of Washington Library of Congress bclg@cs.washington.edu	Julie Mears LC Labs Library of Congress	Eileen Jakeway LC Labs Library of Congress
Megan Ferris LC Labs Library of Congress	Chris Adams Design & Development Program Library of Congress	Nathan Yarassavage National Digital Newspaper Program Library of Congress
	Kate Zward Digital Strategy & LC Labs Library of Congress	Daniel Weld University of Washington

All in the public domain!



17.4 MB

Download History

This notebook trains a visual content recognition model on *Beyond Words* bounding box annotations (plus additional headline + advertisement annotations).

benjamin-lee@cs.washington.edu

This notebook features a pre-trained object detection model (Faster-RCNN R50FPN) to predict bounding boxes around photographs, illustrations, maps, comics, editorial cartoons, headlines, and annotations in historical newspaper images. The notebook is described in more detail in the *newspaper-navigator* repo.

This notebook makes the following assumptions:

- The *Newspaper Navigator* repo has been cloned, so the local file structure is intact.
- Dockerized has been installed. For Amazon EC2 instances with a Deep Learning Ubuntu AMI, the bash script for installing dependencies (install-dependencies.sh) can be run to handle the installation (see the repo for more info).

This notebook is based on the following notebook:

- https://colab.research.google.com/github/USICA/loc02CFA26810c12e5d07286f1_01

First, we handle imports and data formatting.

This cell requires datasets and constructs a COCO dataset using the training and validation .json files; essentially, this enables the model to handle the data loading using the COCO libraries:

```

In [1]: # To do: load images, set line
import os
import cv2
import random
import glob
  
```

news-navigator.labs.loc.gov

#NewspaperNavigator

Data Archaeology

Iowa state bystander.
(Des Moines, Iowa)
14 Oct. 1910



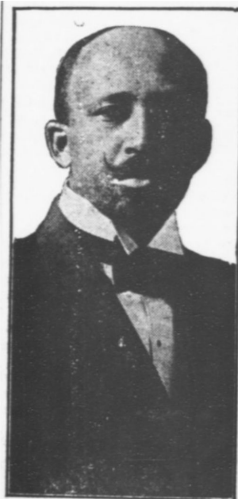
W. E. B. DU BOIS, PH. D.

Franklin's paper the statesman.
(Denver, Colorado)
15 Oct. 1910



W. E. B. DU BOIS, PH. D.

The broad ax.
(Salt Lake City, Utah)
15 Oct. 1910



W. E. B. DU BOIS, PH. D.

The broad ax.
(Salt Lake City, Utah)
26 Nov. 1910



W. E. B. DU BOIS, PH. D.



THE NATION'S GOING DRY.

WHITE—Dry. **BLACK**—Wet.
Number of prohibition states, 23. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people passed by congress. The Porto Rican citizenship bill, passed by the sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 97 per cent of the area of the United States and more than 90 per cent of the population are under prohibition.
Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November; where in November, 1915. This is a wet-and-dry fight through the initiative.

Data Archaeology

Iowa state bystander.
(Des Moines, Iowa)
14 Oct. 1910



Franklin's paper the statesman.
(Denver, Colorado)
15 Oct. 1910



The broad ax.
(Salt Lake City, Utah)
15 Oct. 1910

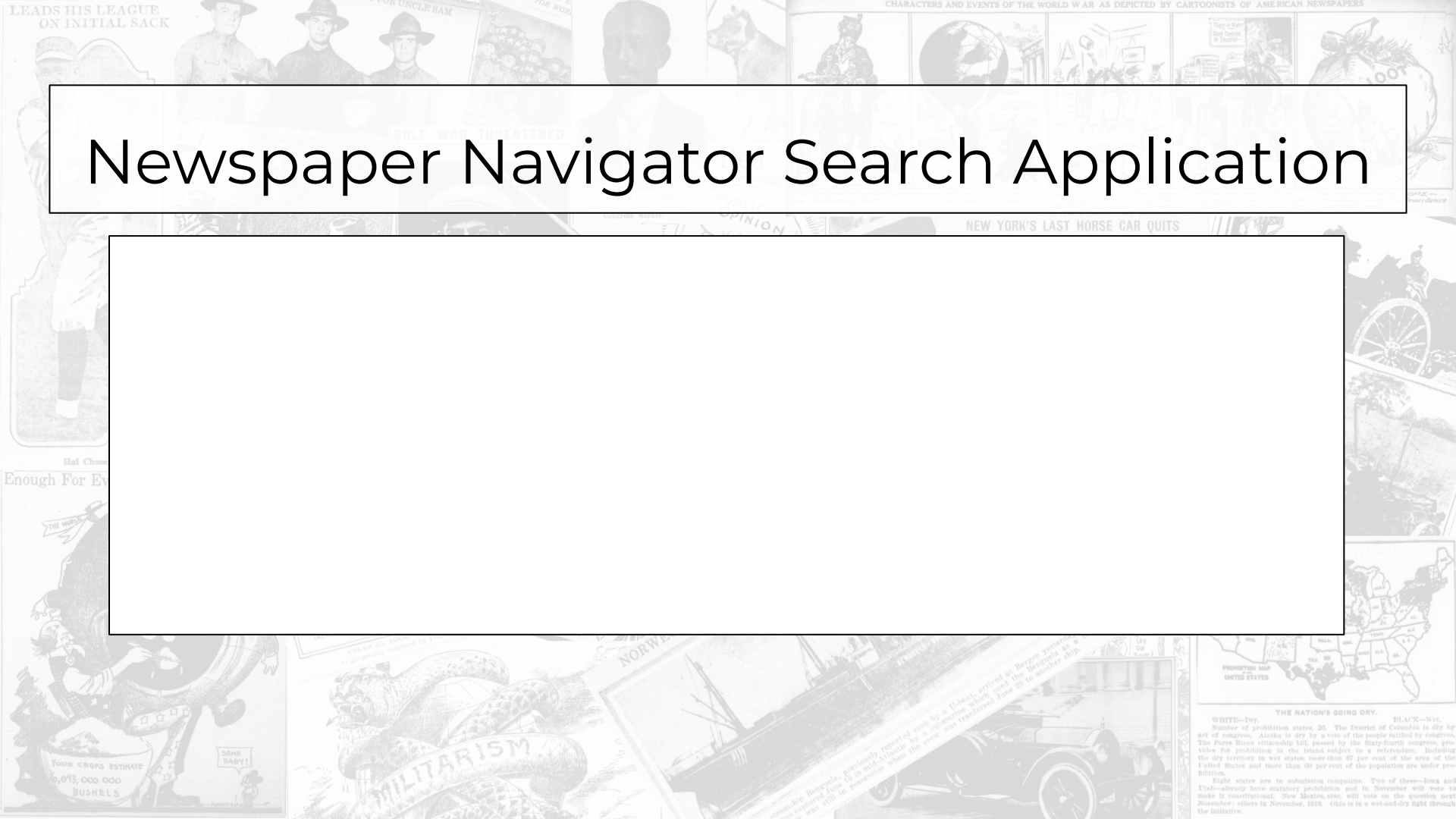
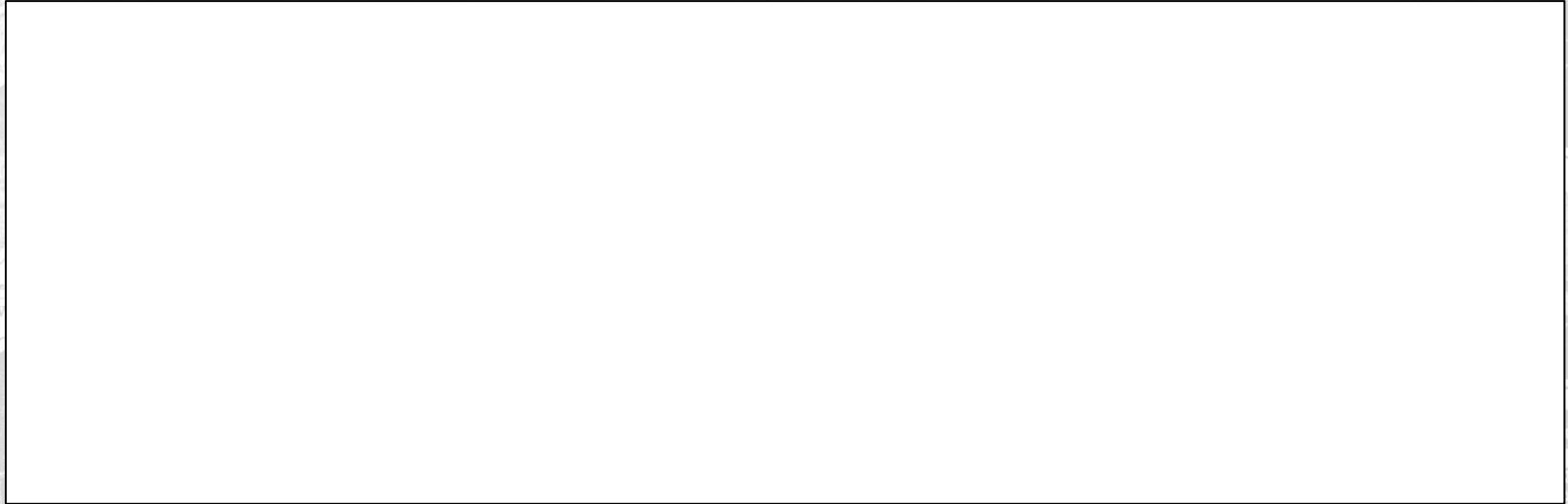


The broad ax.
(Salt Lake City, Utah)
26 Nov. 1910



<http://dx.doi.org/10.17613/k9gt-6685>

Newspaper Navigator Search Application



Newspaper Navigator Search Application

- A search user interface for 1.5 million photos from the Newspaper Navigator dataset

Newspaper Navigator Search Application

- A search user interface for 1.5 million photos from the Newspaper Navigator dataset
- Empowers users to train AI navigators to retrieve visually similar content

Newspaper Navigator Search Application

- A search user interface for 1.5 million photos from the Newspaper Navigator dataset
- Empowers users to train AI navigators to retrieve visually similar content
- Exposes the machine learning training to the user

Newspaper Navigator Search Application

- A search user interface for 1.5 million photos from the Newspaper Navigator dataset
- Empowers users to train AI navigators to retrieve visually similar content
- Exposes the machine learning training to the user
- Training and predicting on all 1.5 million photos takes just a couple of seconds

Search

My Collection

Train My AI Navigators

Chronicling America

LC Labs

Data Archaeology

About

Save

Clear & Restart

Gallery

List

Filters

Location: Start Year: End Year:

Your AI Navigators

None defined yet

Search with Newspaper Navigator!



displaying 1 - 100 records in total 5199

Update

Reset Filters

Sorting by:

BIG DOG TO DEATH.



ALL PAGES FOR THE NEWSPAPER TO SEE.

"MAIR," THE NITED PREPPY WAR DOG,
NOW IN SCOTLAND, BORN FOR ENGLAND.

Prof. Albert Marshall.

Who will exhibit his act of trained dogs at Barnum, Iowa, July 6th.

Trick Dogs That Hail From
Anaconda.

Barnum, Iowa, and in various other cities.



The "trick dog of the moment" was the subject of a newspaper clipping from the Anaconda, Montana, which had the dog performing a trick on a stage. The dog was a small, dark-colored dog, and the man was a man in a suit and tie.



TAKED BY GREAT DOBERMAN.



The dog was a Great Dane, and the man was a man in a suit and tie.



A dog on a newspaper clipping.



A MOST REMARKABLE DOG.



A dog on a newspaper clipping.



A woman on a newspaper clipping.



INHUMAN DOG CATCHER



Search My Collection

Train My AI Navigators

Chronicling America

IC Labs

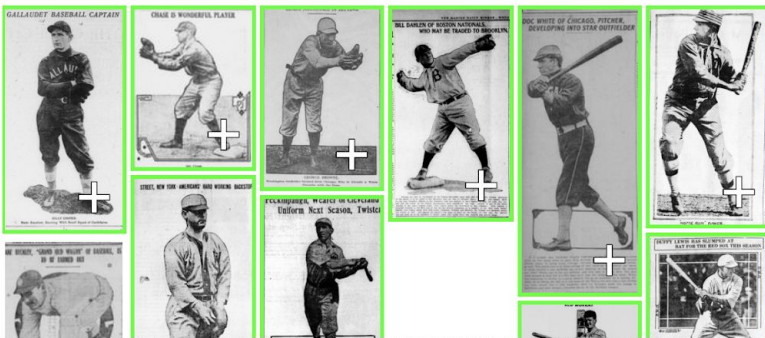
Data Archaeology

About

Save

Select an AI
navigator:

(+) The AI navigator will try to find photos
like the ones selected below!



Train Your AI Navigator!

(-) The AI navigator will try to avoid photos
like the ones selected below!



baseball players

Below, your AI navigator has found visually similar photos based on the selections on the left. You can re-train your AI navigator by adjusting your selections on the left and making new selections below.



Search My Collection

Train My AI Navigators

Chronicling America

LC Labs

Data Archaeology

About

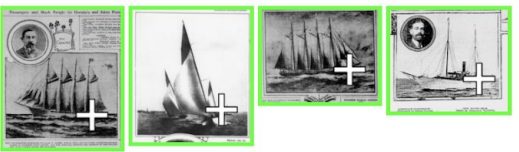
Select an AI navigator:

sailboats

+ New AI Navigator

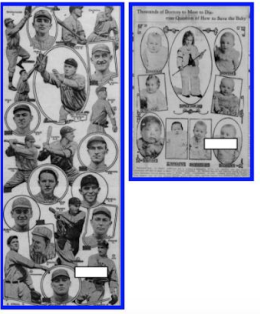
Save

(+) The AI navigator will try to find photos like the ones selected below!



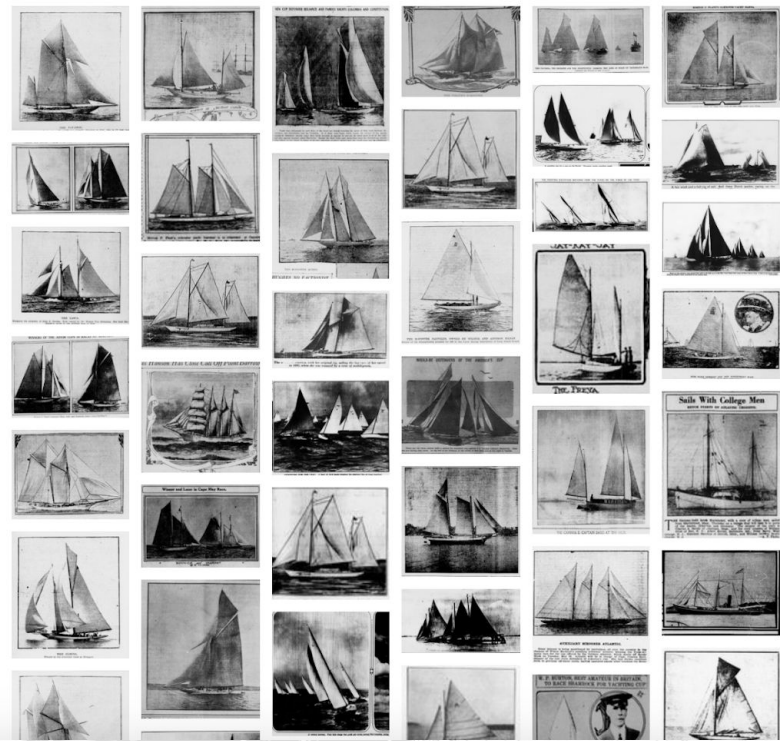
Train Your AI Navigator!

(-) The AI navigator will try to avoid photos like the ones selected below!

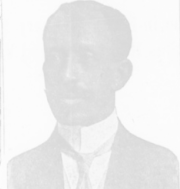


sailboats

Below, your AI navigator has found visually similar photos based on the selections on the left. You can re-train your AI navigator by adjusting your selections on the left and making new selections below.



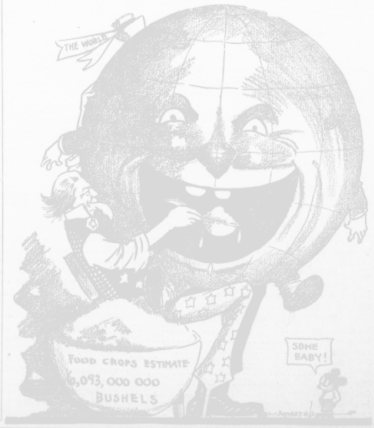
LEADS HIS LEAGUE ON INITIAL SACK



<https://news-navigator.labs.loc.gov/search>



Enough For Everybody Except Germany



WORLD	177
INDIA	117
EUROPE	143
AMERICA	105
ASIA	105
AFRICA	105
AUSTRALIA	105
PACIFIC	105
ANTARCTICA	105
ARABIA	105
AFRICA	105
ASIA	105
AMERICA	105
EUROPE	105
INDIA	105
WORLD	105
NEW BOOKS	3849
NEW EDITIONS	1211
TOTAL	5060
NEW BOOKS	3849
NEW EDITIONS	1211
TOTAL	5060



THE NATION'S GOING DRY. **BLACK**—Wet. **WHITE**—Dry.

Number of prohibition states, 23. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people ratified by congress. The Porto Rican citizenship bill, passed by the sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 97 per cent of the area of the United States and more than 90 per cent of the population are under prohibition.

Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November; where in November, 1915. This is in a wet and dry fight through the initiative.



Newspaper Navigator + the Digital Humanities





Q1: How can Newspaper Navigator help us “see” editors?

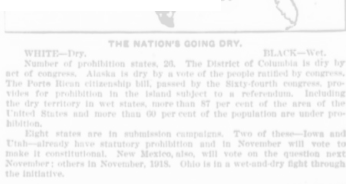
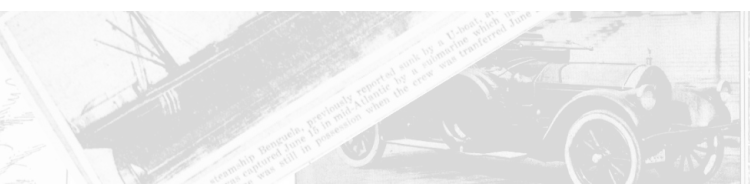
Enough For Everybody Except Germany

WHEAT	115
BARLEY	115
RYE	115
OATS	115
CORN	115
SOYBEANS	115
WHEAT	115
BARLEY	115
RYE	115
OATS	115
CORN	115
SOYBEANS	115

THE KAISER - THE LIBERTY LOAN BONDS WERE

FOR SUBMARIN

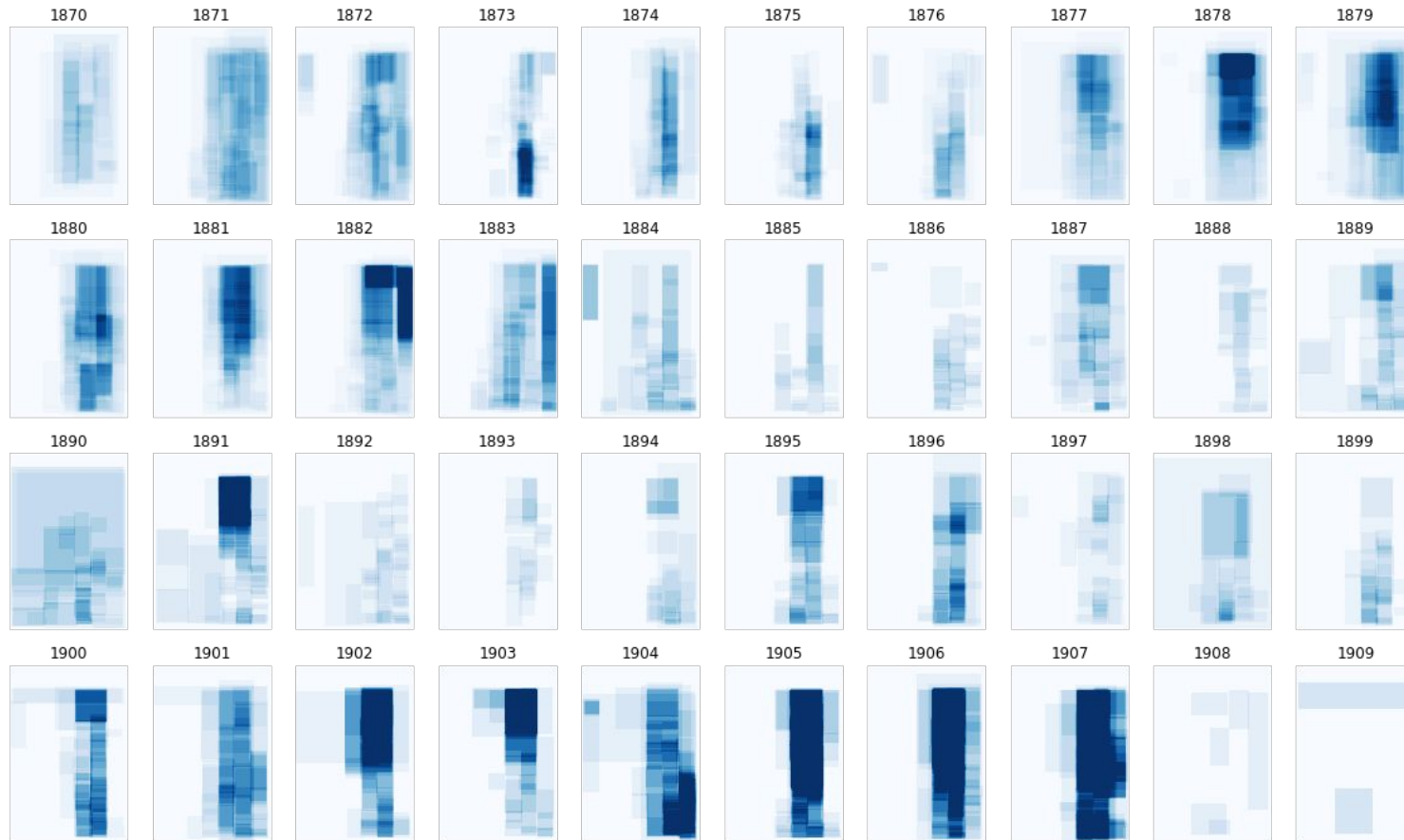
In collaboration with: Jim Casey, Molly Hardy, Joshua Ortiz Baco, Sarah Salter



1907 Front Pages



A Heat Map of Advertisements on the Front Pages of *The Opelousas Courier*, 1870-1909

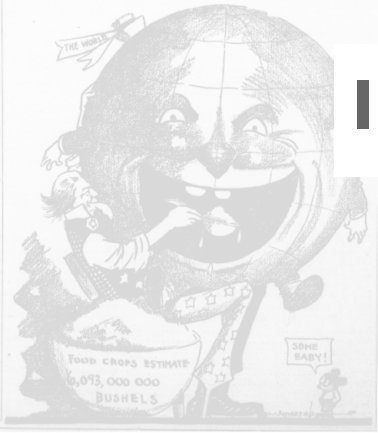


LEADS HIS LEAGUE ON INITIAL SACK



Hal Chase.

Enough For Everybody Except Germany



Q2: How can Newspaper Navigator help us study Ladino newspapers?

In collaboration with: Devin Naar

PROVISIONS	100
STRENGTH	110
ALLENANCE	120
ARMY	130
NAVY	140
INDUSTRY	150
AGRICULTURE	160



THE NATION'S GOING DRY.

WHITE—Ivy. Number of prohibition states, 23. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people passed by congress. The Porto Rico citizenship bill, passed by the sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 90 per cent of the population are under prohibition.

BLACK—Wet. Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November; where in November, 1918. This is in a wet and dry fight through the initiative.

Lo'ara, 1930

MEYER LONDON
494-68 Grand Street
New York

אחדר לה פ"אמיאה
אויס יו אויבנה

ARISTOCRATIC

אויטומאטיק
ליקאס און דריבאס

LEKAS & DRIVAS
100 BROADWAY STREET
NEW YORK, N.Y.

ה'תש"א
ה'תש"ב
ה'תש"ג
ה'תש"ד
ה'תש"ה
ה'תש"ו
ה'תש"ז
ה'תש"ח
ה'תש"ט
ה'תש"ע

ה'תש"א
ה'תש"ב
ה'תש"ג
ה'תש"ד
ה'תש"ה
ה'תש"ו
ה'תש"ז
ה'תש"ח
ה'תש"ט
ה'תש"ע

ה'תש"א
ה'תש"ב
ה'תש"ג
ה'תש"ד
ה'תש"ה
ה'תש"ו
ה'תש"ז
ה'תש"ח
ה'תש"ט
ה'תש"ע

ה'תש"א
ה'תש"ב
ה'תש"ג
ה'תש"ד
ה'תש"ה
ה'תש"ו
ה'תש"ז
ה'תש"ח
ה'תש"ט
ה'תש"ע

**קומינטוס סוב"רי אלגונאם
אלקוניאם איבריאם**

**אין לום אולטימוס טוימפוס, נוס אה
פירסיב"ימוס קי קאדה פירסונה
נון סולאמינטי טייני און נומברי, מה
טאמביין אי און נומברי די פ"אמיאה
(אלקונייה, טומאדו דיל אראבו דיל
ביירב"ו "קיניאה") מה איסטי אוזו
די פאסאר איל נומברי יו פ"אמיאה
די פאדרי אה אויז", פ"ין און טיימפו
אינדיטרימינאדו, נון איגזיסטיאה, אין**

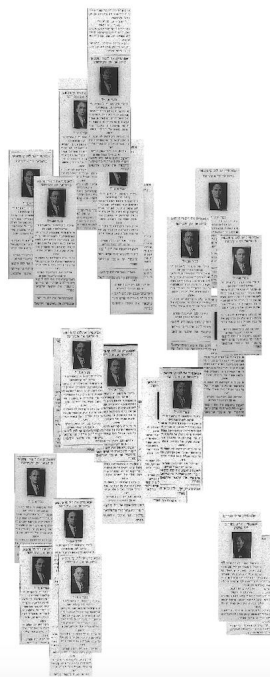
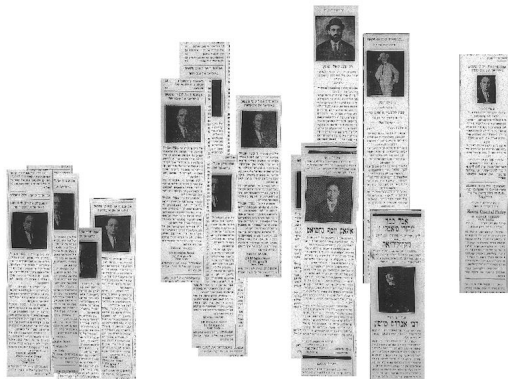
DR. BLAUG—Wet-ports of Columbia is dry by...
The city...
The United States are more than 90 per cent of the population are under prohibition.
Eight states are in submission...
Two of three—New and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, will vote on the question next December unless in November, 1933. Ohio is a necessary right State, the initiative.

Newspaper Title	Number of Images	Number of Pages
<i>El Instruktor revista siyentifika i literaria</i>	331	331
<i>El jugueton, Jurnal umoristiko</i>	4	4
<i>El Kirbatch Americano</i> (1915 - 1917)	206	206
<i>El Luzero Sefaradi</i> (October, 1926)	28	28
<i>El Luzero Sefaradi</i> (May, 1927)	28	28
<i>El Luzero Sefaradi</i> (Total)	56	56
<i>El Progreso / Yosef Daat</i>	332	332
<i>El Tiempo</i> (1890-1891)	926	926
<i>El Tiempo</i> (1896-1897)	1,138	1,138
<i>El Tiempo</i> (1900)	584	584
<i>El Tiempo</i> (1900-1901)	981	981
<i>El Tiempo</i> (Total)	3,629	3,629
<i>La boz de Oriente</i> (April, 1931 - April, 1932)	860	860
<i>La Vara</i> (1922 - 1941)	5,071	10,142
Total	10,489	15,560


Visual Content Type	# > 90%	# > 70%	# > 50%
Photographs	348	770	1060
Illustrations	52	378	960
Maps	27	182	300
Comics	10	39	184
Editorial Cartoons	8	31	111
Advertisements	18,381	31,523	42,505

T-SNE embedding visualization of visual content with confidence scores > 90% (2,518 images in total)





Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.



SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

אגודת ישראל
תחנת הארבעים שנה
הנהלת האגודה ודרכי העבודה



S.S. A.
S.S. A.
S.S. A.

S.S. EUROPA
S.S. EUROPA
S.S. EUROPA

Weiser's Exchange
ALBANY, N.Y. NEW YORK
SOLE IMPORTERS OF THE GREAT EASTERN LINE
NEW YORK
MADE IN U.S.A.

אגודת ישראל
תחנת הארבעים שנה
הנהלת האגודה ודרכי העבודה

אגודת ישראל
תחנת הארבעים שנה
הנהלת האגודה ודרכי העבודה

Acknowledgments

BCGL would like to thank:

- Jaime Mears, Eileen Jakeway, Meghan Ferriter, Laurie Allen, Abbey Potter, Leah Weinryb Grohsgal, Kate Zwaard, Deb Thomas, Nathan Yarasavage, Robin Butterhof, Chris Adams, Tong Wang, John Foley, Megan Halsband, Brian Foo, Mark Sweeney, the entire NDNP staff at the Library of Congress, and everyone who has made *Chronicling America + Beyond Words* possible
- Dan Weld at the University of Washington
- Gabe Pizzorno (Harvard), Michael Haley Goldman (USHMM), Robert Ehrenreich (USHMM), Jim Casey (Penn State), Molly Hardy (NEH), Joshua Ortiz Baco (UT Austin), Sarah Salter (TAMUCC), Devin Naar (University of Washington)

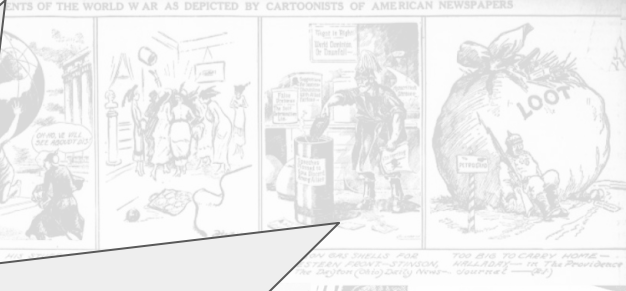
Acknowledgments

- This research is based upon work supported by the **National Science Foundation Graduate Research Fellowship under Grant DGE-1762114**, the **Library of Congress Innovator in Residence position**, and the **Richard Willner Memorial Fellowship** as part of the University of Washington's Jewish Studies Graduate Research Fellowship program.



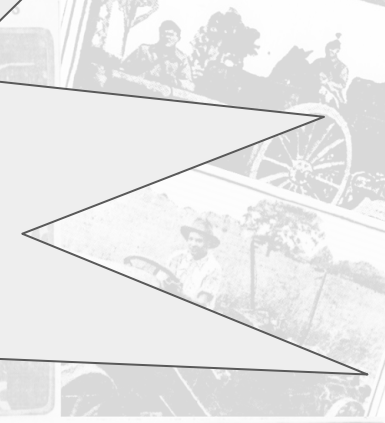
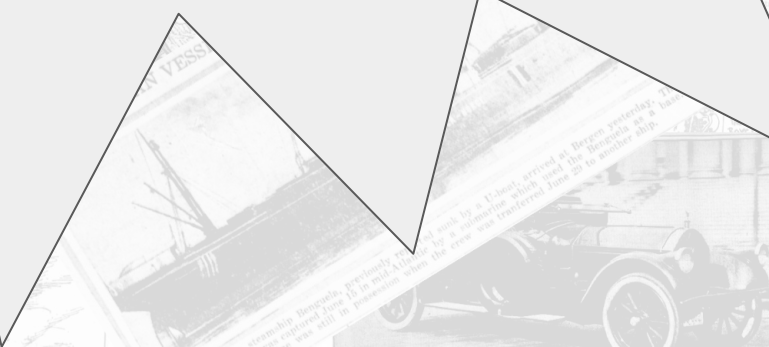
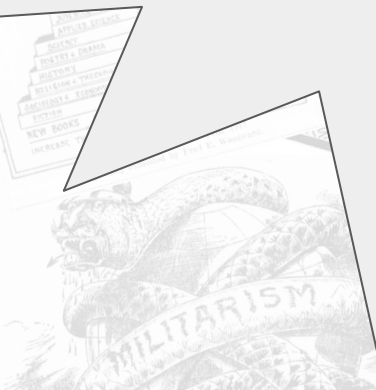
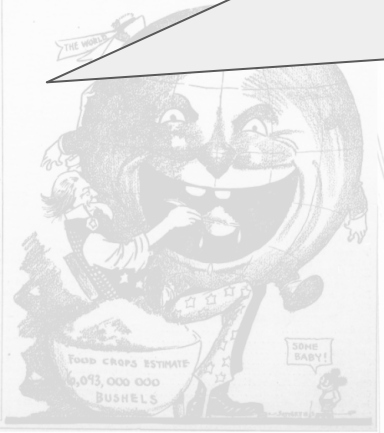


LEADS HIS LEAGUE ON INITIAL SACK



Feel free to reach out!
bcgl@cs.washington.edu

Enough For Everybody Else

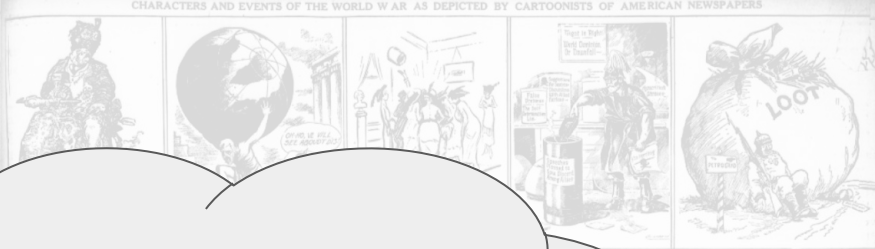
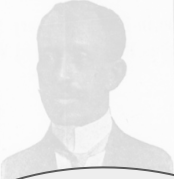
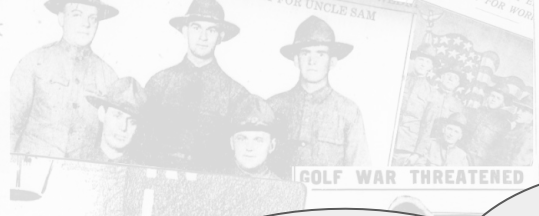


THE NATION'S GOING DRY. WHITE-Ivy. Number of prohibition states, 21. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people passed by congress. The Porto Rico citizenship bill, passed by the sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 60 per cent of the population are under prohibition. Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November; where in November, 1918. This is in a wet and dry fight through the initiative.

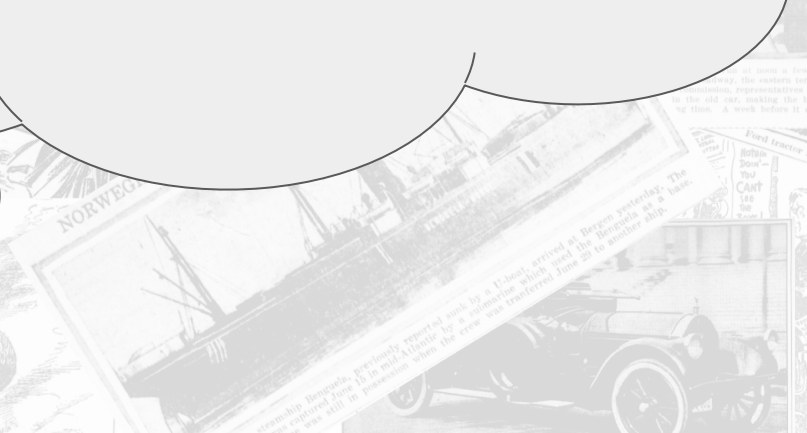
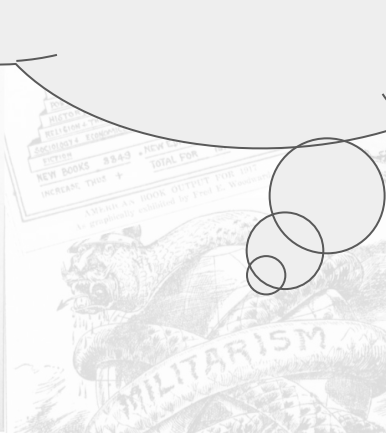
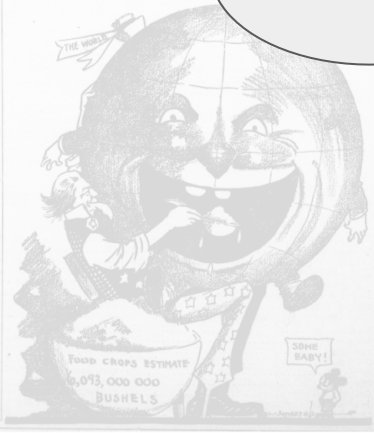
LEADS HIS LEAGUE ON INITIAL SACK



Hal Chase.
Enough For Everybody.



Questions?



THE NATION'S GOING DRY. **BLACK**—Wet. **WHITE**—Dry. Number of prohibition states, 23. The District of Columbia is dry by act of congress. Alaska is dry by a vote of the people passed by congress. The Porto Rican citizenship bill, passed by the sixty-fourth congress, provides for prohibition in the island subject to a referendum. Including the dry territory in wet states, more than 87 per cent of the area of the United States and more than 90 per cent of the population are under prohibition. Eight states are in submission campaigns. Two of these—Iowa and Utah—already have statutory prohibition and in November will vote to make it constitutional. New Mexico, also, will vote on the question next November; others in November, 1918. This is in a wet and dry fight through the initiative.