

IHT Electronic library

- ▶ In 1992, we stopped clipping the paper manually and began our online archive. However, this collection only includes articles. Pages and photos are not available. All articles from 1992 are available to the public through the electronic libraries such as Nexis, Proquest, Europresse, etc.



But the offline collection.....

Conservation and storage were becoming more pressing problems.

Although housed in a secure location in our basement, we always ran the risk of fire or water damage.

- ▶ The unrestored volumes continued to deteriorate and the collection remained **dormant and inaccessible** to researchers, their monetary potential in an Internet world, unexploited.
- ▶ Since the advent of the Internet, we received an increasing number of requests for access to our collection, people believe it is available online.....



Search for a partner

- ▶ Contacts with many companies in Europe, Asia and the U.S., failed to provide a global and financially acceptable solution. It became clear we would have to approach the project at a gradual pace, step by step.



PLANMAN
TECHNOLOGIES



olive
OLIVE SOFTWARE



ProQuest



∞ doc



X-CAGO
MEDIA
INTELLIGENCE
SOLUTIONS



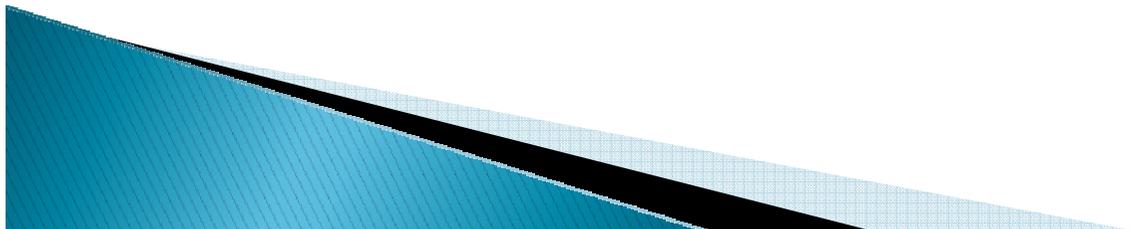
Spigraph



Irish Newspaper
Archives

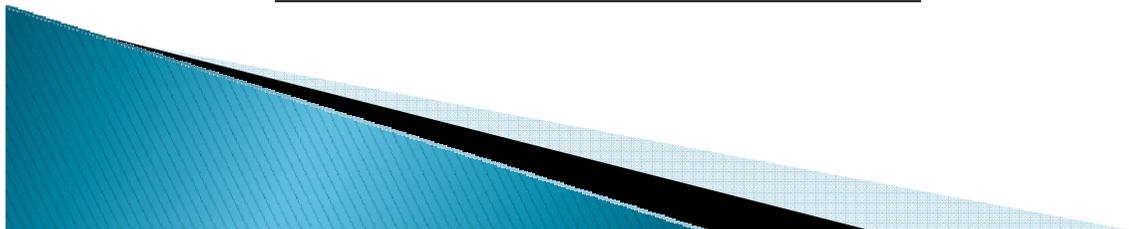
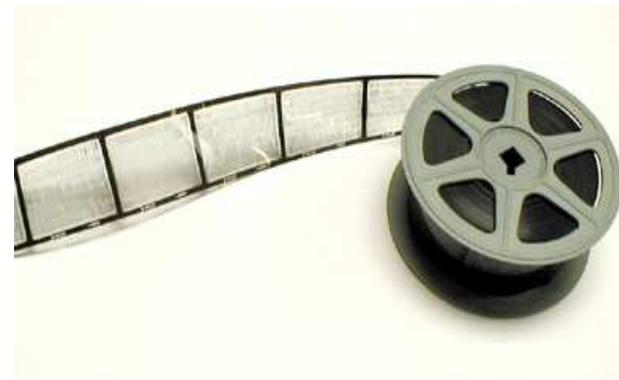
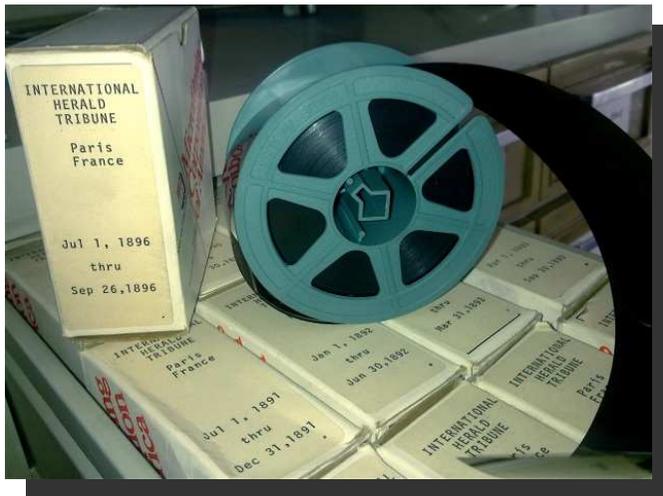
The lay of the land

- ▶ Most companies proposed scanning the microfilm with the second option of scanning the paper archive, to compensate for poor quality microfilm
- ▶ We doubted our microfilm would meet the full criteria for success
- ▶ We did not want the bound volumes to leave Neuilly to be scanned off site
- ▶ Cost of scanning in house exceeded budgets
- ▶ Further on, the IHT didn't have facilities to commercialise the scanned content



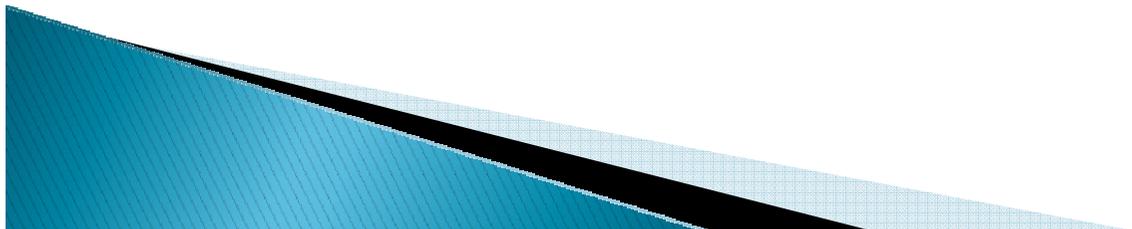
The microfilm route

- ▶ But, we have the full collection of microfilm
- ▶ It was financially feasible and a serious beginning
- ▶ We found a partner experienced in the field, willing to guide us through the initial process, without OCR or clipping as part of the project



ADVANTAGES

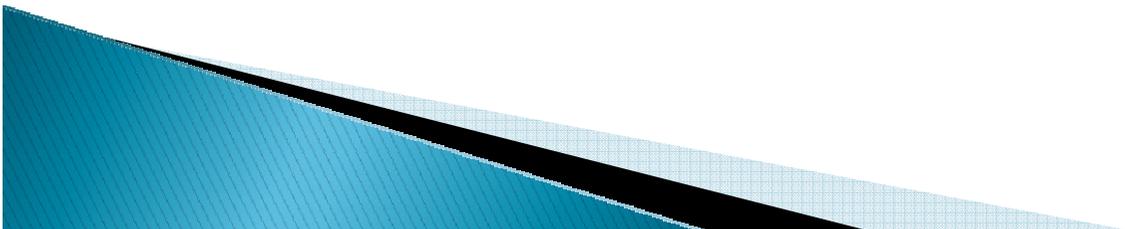
- ▶ Microfilm can be transported and scanned off site
- ▶ Conservation of the pages was ensured
- ▶ Possibility to OCR and clip the content at a later date was assured
- ▶ The financial investment was defensible
- ▶ Immediate research possibilities available
- ▶ We can move the volumes to a secure, adapted location for future conservation
- ▶ The Pandora's box was opened.....



DISADVANTAGES

We did not have the master reels and some of the film had suffered damage through use over time

- ▶ A survey of all the reels revealed the pre-1933 content to be so uneven in quality that we refrained from scanning them.
- ▶ Cost of cleaning the pages exceeded the budget
- ▶ Some editions were missing from the reels
- ▶ We have an incomplete collection online, with some of the most beautiful pages still inaccessible.



How was it done?

- ▶ We were charged by page so we had to estimate the number of pages to be scanned
- ▶ On 400 reels, we estimated we had 1000 pages per reel, giving us 400,000 pages to scan
- ▶ Some reels had less than 1000 pages, others had more, editions varied throughout the century from 4 to 24 pages a day
- ▶ We sampled a trimester from every decade and drew up a table, giving averages per month, per year and per decade.
- ▶ Total estimate: from 1887 to 2003: 411,000 pages

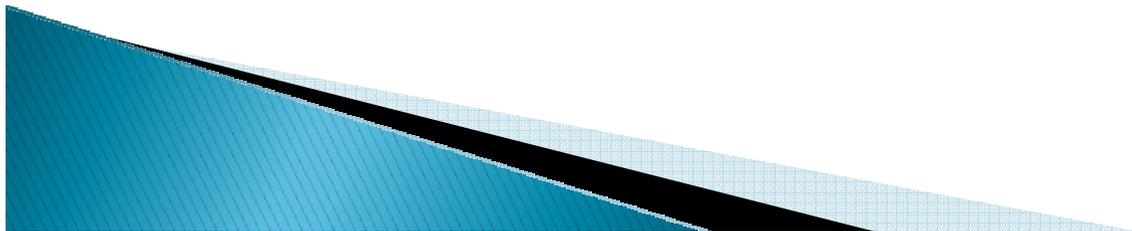


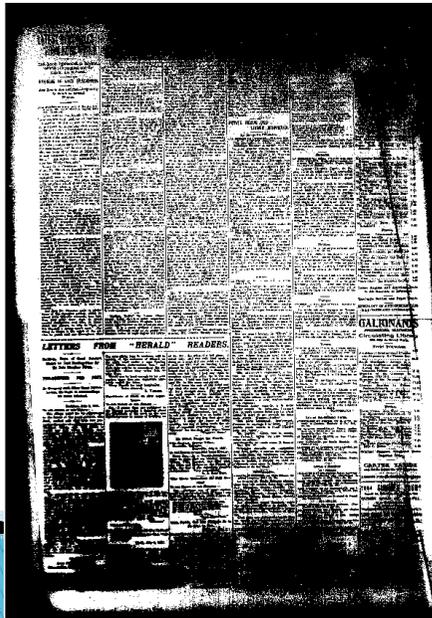
Table estimating pages published

	A	B	C	D	E	F	G	H	I	J	
1	Estimate of the number of pages published										
2		Sample	No. of pages	Monthly average	Annual estimation	10 years' estimation					
3		Oct/1887	124								
4		Nov/1887	124								
5	1880s*	Dec/1887	124	124	1488	3348	*2 years + 3 months				
6		Oct/1896	194								
7		Nov/1896	170								
8	1890s	Dec/1896	149	171	2052	20520					
9	1900s	Jul/1907	360	147	1764	17640					
10		Oct/1914	126								
11	1910s	Nov/1914	120	123	1476	14760					
12		Apr/1925	246								
13	1920s	Jun/1925	230	238	2856	28560	81480 Total 1887-1929 included)				
14		Jan/1937	268								
15		Feb/1937	210								
16	1930s	Mar/1937	238	239	2864	28640					
17		Jul/1946	123								
18		Nov/1946	139								
19	1940s	Dec/1946	165	142	1708	17080					
20		Jan/1955	258								
21		Feb/1955	207	233							
22		Jan/1959	247								
23	1950s	Mar/1959	245	246	2871	28710					
24		Jul/1963	402								
25		Jul/1965	284								
26	1960s	Sep/1965	280	227	2728	27280					
27		Jan/1977	341								
28	1970s	Feb/1977	338	340	4074	40740					
29		Apr/1984	475								
30	1980s	May/1984	432	454	5442	54420					
31		Jul/1993	487								

Quality control of microfilm

- ▶ We surveyed every reel on the microfilm reader and produced a table appraising the quality.
- ▶ We raised the bar as high as possible and gave draconian appraisals, from poor, to average to good.
- ▶ The end result was we agreed with the scanning company to only scan reels with a “good” rating which meant from 1933 to present

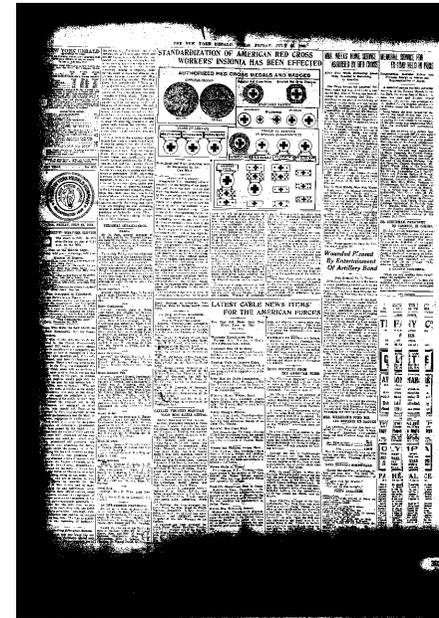
1907



1910



1918



Quality control table of microfilm, example 1887 to 1895

Microsoft Excel - IHT Microfilm Vs Archives.xls

File Edit View Insert Format Tools Data Window Help

Type a question for help

Arial 10 B I U

K124

	A	B	C	D	E	G	I	J	K	L	M	N	O	P	Q
1															
2	MICROFILM REEL												CORRESPONDING VOLUMES IN ARCHIVES		
3															
4	Reel Date						Condition		Remarks				No. of volumes		Date Range
5															
6	1st drawer: Oct 4, 1887 – Dec 31, 1896 (19 boxes)														
7															
8	Oct 4, 1887 – Mar 31, 1888						Very Poor		Issues and Pages Lacking; Speckled				2		October-December
9															January-March
10	Apr 1, 1888 – Sep 30, 1888						Average		Dark				2		April-June
11															July-September
12	[Oct 1, 1888 – Mar 31, 1889]						Reel missing - volume in archives						2		October-December
13															January-March
14	Apr 1, 1889 – Sep 30, 1889						Very Poor		Very Dark				2		April-June
15															July-September
16	Oct 1, 1889 – Mar 30, 1890						Good		Light				2		October-December
17															January-March
18	Apr 1, 1890 – Jun 30, 1890						Very Poor		Very Dark				1		April-June
20	Jul 1, 1890 – Dec 31, 1890						Very Poor		Uneven; Some Very Good				2		July-September
21															October-December
22	Jan 1, 1891 – Jun 30, 1891						Average		Speckled				2		January-March
23															April-June
24	Jul 1, 1891 – Dec 31, 1891						Very Poor		Very Dark				2		July-September
25															October-December
26	Jan 1, 1892 – Jun 30, 1892						Average		Some Dark				2		January-March
27															April-June
28	July 1, 1892 – Mar 31, 1893						Average		Dark				3		July-September
29															October-December
30															January-March
31	Apr 1, 1893 – Sep 30, 1893						Good						2		April-June
32															July-September
33	Oct 1, 1893 – Dec 31, 1893						Good						1		oct-dec
34	Jan 1, 1894 – Jun 30, 1894						Good						2		January-March
35															April-June
36	Jul 1, 1894 – Dec 31, 1894						Good						2		July-September
37															October-December
38	Jan 1, 1895 – Jun 29, 1895						Good		Light				2		January-March
39															April-June
40	Jul 1, 1895 – Dec 31, 1895						Average		Speckled				2		July-September
41															October-December

Sheet1 / Sheet2 / Sheet3

Ready NUM

Page 1

Quality control table of microfilm, 1933 “Good”

Microsoft Excel - IHT Microfilm Inven sorted supplier.xls [Read-Only]

Type a question for help

Arial 10 B I U

L161

173	(16 boxes)													
174	Jan 1, 1933 – Mar 31,	Europe	Very Good	X										
175	Apr 1, 1933 – Jun 30,	Europe	Good	X										
176	Jul 1, 1933 – Sep 30,	Europe	Good	X										
177	Oct 1, 1933 – Dec 31,	Europe	Good	X										
178	Jan 1, 1934 – Mar 31,	Europe	Average	Dark										
179	Apr 1, 1934 – Jun 30,	Europe	Poor	Dark										
180	Jul 1, 1934 – Sep 30,	Europe	Good	X										
181	Oct 1, 1934 – Dec 31,	Europe	Good	Speckled										
182	Jan 1, 1935 – Jul 30,	Europe	Good	Speckled										
183	Jul 1, 1935 – Dec 31,	Europe	Average	Dark										
184														
185	Jan 1, 1936 – Jul 30,	Europe	Good	Speckled										
186	Jul 1, 1936 – Dec 31,	Europe	Good	Light; Speckled										
187	Reel Date	Edition Number	Condition	Remarks										Manufacturer
188	Jan 1, 1937 – Mar 31,	Europe	Good	X										IM PRESS
189	Apr 1, 1937 – Jun 30,	Europe	Good	X										IM PRESS
190	Jul 1, 1937 – Sep 30,	Europe	Good	X										IM PRESS
191	Oct 1, 1937 – Dec 31,	Europe	Good	X										IM PRESS
192														
193	9th drawer: Jan 1, 1938 – Jun 3, 1940													
194	(7 boxes)													
195	Jan 1, 1938 – Mar 31,	Europe	Good	X										IM PRESS
196	Apr 1, 1938 – Jun 30,	Europe	Good	X										IM PRESS
197	Jul 1, 1938 – Sep 30,	Europe	Good	Speckled										IM PRESS
198	Oct 1, 1938 – Dec 31,	Europe	Good	Speckled										IM PRESS
199	Jan 1, 1939 – Mar 31,	1939												IM PRESS
200	Apr 1, 1939 – Jun 30,	Europe	Good	Speckled										IM PRESS
201	Jul 1, 1939 – Sep 30,	Europe	Good	X										IM PRESS
202	Oct 1, 1939 – Dec 31,	Europe	Good	X										IM PRESS

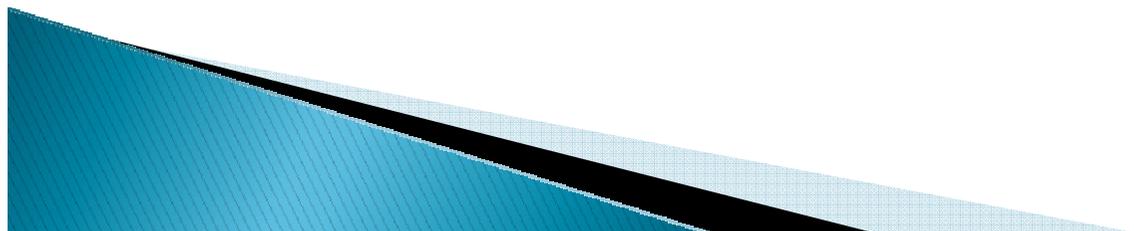
Page 7

Page 8

Ready NUM

Specifications

- ▶ We went with the industry standard to permit future OCR and clipping at a later date
- ▶ 300 DPI greyscale TIFF images and PDF of each page
- ▶ Rotated, cropped, despeckled and deskewed
- ▶ Each issue was delivered in a single folder with metadata including : Name, year, month, day, page number
- ▶ We made server space available: requirement 5.232 terabytes, to accommodate 70 years of TIFFs and PDFs based on the size of each object
- ▶ We received delivery of TIFFs and PDFs of each page accompanied by their vignette on disks and transferred them to our server



IHT quality control of pages

Microsoft Excel - MICROFILM SCAN REPORT FINAL.xls [Read-Only]

File Edit View Insert Format Tools Data Window Help

Type a question for help

Arial 8 B I U

A1 S

	A	B	C	D	E	F	G	H	I	J
	S	YEAR	EDITIONS	MISSING ISSUES	PDF	MISSING/CORRUPTED/EMPTY PDF	NON-PRINTING DAYS	FIRST EDITION	LAST EDITION	NOTES
1		1933	365	0	3104	2	0	Sun, January 1	Sun, December 31	Empty Pages 6 and 7 from Edition Dated Mon, October 2.
2		1934	364	0	3004	0	1	Mon, January 1	Mon, December 31	Mon, February 12, 1935 was a non-printing day
3		1935	365	0	2932	1	0	Tue, January 1	Tue, December 31	Empty Page 6 from Edition Dated Mon, October 21
4		1936	363	2	2892	0	1	Wed, January 1	Thu, December 31	Fri, June 5, 1936 was a non-printing day
5										Missing Edition Dated Thu, August 6, 1936 - No.17,840 Missing Edition Dated Sun, August 9, 1936 - No.17,843
6		1937	364	0	3206	3	1	Fri, January 1	Fri, December 31	Non-printing day: Sat, May 1 Empty Pages 8 from Edition Dated Fri, April 2, 1937; 10 from Edition Dated Tue, May 4, 1937; 11 from Edition Dated Sun, May 9, 1937.
7		1938	364	0	2998	0	1	Sat, January 1	Sat, December 31	Non-printing day: Sun, May 1
8		1939	364	1	2438	2	0	Sun, January 1	Sun, December 31	Missing Edition Dated Sun, August 20, 1939 - No.18,947 Empty Pages 2 and 3 from Edition Dated Mon, September 11, 1939
9		1940	152	3	596	0	0	Mon, January 1	Mon, June 3	Production ceased on Tue, 4 June Missing Edition Dated Mon, April 1, 1940 - No.19,172 Missing Edition Dated Thu, May 30, 1940 - No.19,231 Missing Edition Dated Fri, May 31, 1940 - No.19,232 Last Edition dated June 3rd is No.19,235
		1944	9	0	36	0	1	Fri, December 22	Sun, December 31	First Edition dated December 22nd is No.19,245 (missing 10 Edition No.?)

GENERAL REPORT BY YEAR MISSING EDITIONS MISSING PAGES

Ready NUM

Scanned pages uploaded to our database

Thumbnail	Date	Page	Edition(s)
	Sat, March 30, 1935	01	
	Sat, March 30, 1935	02	
	Sat, March 30, 1935	03	
	Sat, March 30, 1935	04	
	Sat, March 30, 1935	05	

For Help, press F1

Item 80 of 3704 (1:0)

Scanned pages of "gravure" edition

The screenshot shows a software window titled "The Shell - [Archived pages search results]". The interface includes a menu bar (File, Edit, View, Search, Basket, Pages, Pictures, Archive, Window, Help) and a toolbar with various icons. Below the toolbar is a table with the following columns: Thumbnail, Date, Page, Edition(s), and No. The table contains five rows of search results. The first four rows are for the date "Sun, January 15, 1933" and page numbers 09, 10, 11, and 12. The fifth row is for "Sat, January 14, 1933" and page 01. The row for page 12 is highlighted in a darker green color. A tooltip is visible over the row for page 10, displaying the following information: Date: Sun, January 15, 1933; Page: 10; Edition(s); Notes: The status bar at the bottom right of the window shows "Item 15830 of 1587".

Thumbnail	Date	Page	Edition(s)	No.
	Sun, January 15, 1933	09		
	Sun, January 15, 1933	10		
	Sun, January 15, 1933	11		
	Sun, January 15, 1933	12		
	Sat, January 14, 1933	01		

Foreign Exchange Rates
London to New York 100/15
New York to London 100/15

Old Year No. 16433
European Edition

NEW YORK
Herald Tribune

PARIS, SUNDAY, MARCH 10, 1938

Hitler at Linz Proclaims Austria Part of Germany
As Reich Army Sweeps Through Country to Brenner

Blum's Plan For National Government Abandoned
Premier-Dignitate Opps Consultations With Left For Formation Of Popular Front Group

Nicholas Bukharin Hits Back At Prosecutor Vishinsky with Slashing Protest Over Charges
Nazis Turn Down Paris, London Note

Protest Against Invasion Held 'Inadmissible' as German-Occupied Control Key Points in Austria
By Ralph Barson
From the Soviet Press Bureau

German Soldiers Cross Frontier Before Dawn, Reach Brenner At 1 p.m., Fraternize With Italian Border Guards, Vienna Is Entered by Evening 15,000, Planes, Tanks in Country; Paris and London Protest and Map Action for Future

CHANCELIER ADOLF HITLER yesterday proclaimed Austria to be part of Greater Germany. Austria is to be given a status that will make it a kind of associate of the Reich.

Creeping the border north of Munich in the middle of the afternoon, the Fuhrer arrived in the evening at Linz on the Danube. He was met by a large number of Austrian officials and was escorted to a hotel in Linz. He then proceeded to Vienna, where he was met by a large number of Austrian officials and was escorted to a hotel in Vienna.

CHANCELLOR ADOLF HITLER yesterday proclaimed Austria to be part of Greater Germany. Austria is to be given a status that will make it a kind of associate of the Reich.

Creeping the border north of Munich in the middle of the afternoon, the Fuhrer arrived in the evening at Linz on the Danube. He was met by a large number of Austrian officials and was escorted to a hotel in Linz. He then proceeded to Vienna, where he was met by a large number of Austrian officials and was escorted to a hotel in Vienna.

CHANCELLOR ADOLF HITLER yesterday proclaimed Austria to be part of Greater Germany. Austria is to be given a status that will make it a kind of associate of the Reich.

Creeping the border north of Munich in the middle of the afternoon, the Fuhrer arrived in the evening at Linz on the Danube. He was met by a large number of Austrian officials and was escorted to a hotel in Linz. He then proceeded to Vienna, where he was met by a large number of Austrian officials and was escorted to a hotel in Vienna.

Blum's Plan For National Government Abandoned
Premier-Dignitate Opps Consultations With Left For Formation Of Popular Front Group

Nicholas Bukharin Hits Back At Prosecutor Vishinsky with Slashing Protest Over Charges
Nazis Turn Down Paris, London Note

Protest Against Invasion Held 'Inadmissible' as German-Occupied Control Key Points in Austria
By Ralph Barson
From the Soviet Press Bureau

German Soldiers Cross Frontier Before Dawn, Reach Brenner At 1 p.m., Fraternize With Italian Border Guards, Vienna Is Entered by Evening 15,000, Planes, Tanks in Country; Paris and London Protest and Map Action for Future

CHANCELLOR ADOLF HITLER yesterday proclaimed Austria to be part of Greater Germany. Austria is to be given a status that will make it a kind of associate of the Reich.

Creeping the border north of Munich in the middle of the afternoon, the Fuhrer arrived in the evening at Linz on the Danube. He was met by a large number of Austrian officials and was escorted to a hotel in Linz. He then proceeded to Vienna, where he was met by a large number of Austrian officials and was escorted to a hotel in Vienna.

CHANCELLOR ADOLF HITLER yesterday proclaimed Austria to be part of Greater Germany. Austria is to be given a status that will make it a kind of associate of the Reich.

Creeping the border north of Munich in the middle of the afternoon, the Fuhrer arrived in the evening at Linz on the Danube. He was met by a large number of Austrian officials and was escorted to a hotel in Linz. He then proceeded to Vienna, where he was met by a large number of Austrian officials and was escorted to a hotel in Vienna.

CHANCELLOR ADOLF HITLER yesterday proclaimed Austria to be part of Greater Germany. Austria is to be given a status that will make it a kind of associate of the Reich.

Creeping the border north of Munich in the middle of the afternoon, the Fuhrer arrived in the evening at Linz on the Danube. He was met by a large number of Austrian officials and was escorted to a hotel in Linz. He then proceeded to Vienna, where he was met by a large number of Austrian officials and was escorted to a hotel in Vienna.

SAILINGS & NEWS
S.S. ...

No. 26,509

INTERNATIONAL
Herald Tribune

PARIS, SATURDAY-SUNDAY, APRIL 6, 1968

Martin Luther King Slaying Brings Rising Tide of Violence; Johnson Cancels His Talks in Hawaii

Civil Rights Leader Shot By a Sniper

Troops Called as Riots Hit 41 Cities To Seek New Rights Action

U.S. General Diplomatic Moves For Talks With Hanoi Envoys

Wilson Shuffles Cabinet; 16 Changes Are Announced

Raf Jet Buzzes Parliament, Flies Under Tower Bridge

Russia Offers New Pact to U.S.

Fuehrer Proclaims 'New Phiblicis'

Ann Cooper Hewitt Gets Divorce From Ronald Gay

Wall Street

The Weather

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

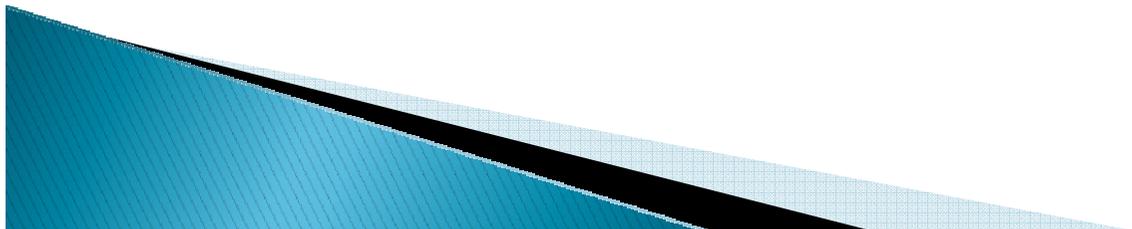
WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

WASHINGTON, April 5 (AP)—The Federal Government yesterday announced that it will send a large number of troops to Hawaii to help maintain order in the islands.

RESULTS

- ▶ We came in under budget and have access to pages from 1933 to present online for research
- ▶ And the Pandora's box:
We have partnered with a French company in the Paris region to scan the volumes, completing the missing pages from 1887 to 1933, to the same specifications as the microfilm, but in colour
- ▶ We plan to OCR all the pages very soon
- ▶ We are seeking a partner to convert the pages applying METS/ALTO xml throughout.
- ▶ We will commercialise the whole collection making it accessible to the public, creating a new revenue stream





Thank you
Merci

▶ bmurphy@iht.com