Educopia Institute, News on the Margins Data Model (https://educopia.org/research/margins)

Collection Description			
Title	On the Margins Newspapers and Periodicals Registry		
Compiler	Katherine Skinner, Educopia Institute; Emily Gore, DPLA; Franky Abbott, DPLA		
Summary	The purpose of this data collection is to create an open registry of African American newspapers and periodicals (print, microfische, digitized, and born digital) to make more visible the corpus of stories told by African American communities in their own words		
Extent	2 tables		
Format	xlsx		
Parts			
Name	**Purpose**	**Extent**	
codebook	documenting the dataset	1 sheet	
African_American_Newspapers	registry documenting titles, locations, formats, publishers, collectors, and condition of African American newspapers/periodicals	1 sheet	
Field Name	**Variable**	**Values or Explanation**	**Field**
collecting_institution	institution name	name	a
title	publication title	name	b
additional_title	Has the publication used additional title(s)?	yes, no, I don't know	c
additional_title_list	Please specify any known additional titles	name (if multiple, separate by commas)	d
tag_line	Has the publication used a tagline?	yes, no, I don't know	e
tag_line_list	text description	tag line (if multiple, separate by commas)	f
LCCN	identifier	year-serial# or unknown	g
ISSN	identifier	acronym ISSN followed by two groups of four digits separated by a hyphen or unknown	h
city_of_publication	city name	name (if multiple, separate by commas; if unknown, specify unknown)	i
state_of_publication	state or territory name	abbreviation (if multiple, separate by commas; if unknown, specify "unknown")	j
frequency_of_publication	how often published (if variable, select the most representative descriptor)	daily, weekly, biweekly, monthly, quarterly, annually, other (please specify)	k
first_publisher	name of first known publisher	org name or first last	l
last_publisher	name of last known publisher	org name or first last	m
format_print	does the collecting institution have print copies?	yes, no, unknown	n
print_range_description	what date ranges does the collecting institution have in print?	open text (separate date ranges using commas, E.g., from 1900-01-04 to 1930-05-30, from 1932-02-16 to 1932-02-18, from 1944-06-23 to 1944-12-31)	o
format_microfilm	does the collecting institution have microfilm copies?	yes, no, unknown	p

microfilm_range_description	what date ranges does the collecting institution have in microfilm?	open text (separate date ranges using commas, E.g., from 1900-01-04 to 1930-05-30, from 1932-02-16 to 1932-02-18, from 1944-06-23 to 1944-12-31)	q
format_digitized	does the collecting institution have digitized copies?	yes, no, unknown	r
digitized_range_description	what date ranges does the collecting institution have in digital form?	open text (separate date ranges using commas, E.g., from 1900-01-04 to 1930-05-30, from 1932-02-16 to 1932-02-18, from 1944-06-23 to 1944-12-31)	s
digitized_single_PDF	is the digitized issue a single pdf?	binary	t
digitized_individual_pages	are individual pages saved separately as pdfs?	binary	u
digitized_individual_articlesare individual articles saved separately as pdfs?		binary	v
digitized_page_TIFF	is each page saved as a single TIFF?	binary	w
digitized_article_TIFF	are individual articles saved separatly as TIFFs?	binary	x
digitized_page_JPEG	is each page saved as a single JPEG?	binary	y
digitized_article_JPEG	are individual articles saved separately as JPEGs	binary	z
digitized_page_TIFF	is each page saved as a single JP2000?	binary	aa
digitized_article_JP2000	are individual articles saved separately as JP2000s	binary	ab
digitized_content_browsableis the digitized content browsable by date?		binary	ac
digitized_text_searchable	is the digitized content searchable by keyword?	binary	ad
digitized_fulltext_searchable	is the digitized content full-text searchable?	binary	ae
digitized_base	what format was the content digitized from?	print, microfiche, unknown	af
digitized_available	is the digitized content freely available?	binary	ag
digitized_paywall	is the digitized content available behind a paywall?	binary	ah
OCR_type	what type of OCR/indexing was used?	none, raw OCR, corrected OCR, double key indexing, triple key indexing, other (specify)	ai
metadata_schema	what descriptive metadata schema was used?	DC, METS, ALTO, MODS, METS-ALTO, Other (specify)	aj
format_born-digital	does the collecting institution collect/maintain born-digital copies of this title?	yes, no, unknown	ak
format_born-digital_describe	describe born-digital collection process and formats collected	open text, describe 1) how it is acquired, 2) how it is processed, 3) what formats are included, and 4) if you maintain the content in renderable form.	al
url(s)	home page of the collection (if digital)	URL(s) separated by commas or "none"; if multiple, please separate by comma.	am
focus_geographical	description of any geographical focus	city, state, region, national, international, none, other (please specify)	an

focus_topical	description of any topical focuses	keywords separated by commas	ao
copyright_status	status	protected by copyright, public domain, variable, unknown	ap
copyright_status_description	who is/are the copyright owner(s)	open text, if unknown write "unknown"	aq
condition_print	condition of print content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care); Good (e.g., robust and stable, usable for any purpose); N/A	ar
condition_print	condition of print content	open text	as
condition_microfiche	condition of microfiche content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care); Good (e.g., robust and stable, usable for any purpose); N/A	at
condition_microfiche	condition of microfiche content	open text	au
condition_digitized	condition of digitized content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care); Good (e.g., robust and stable, usable for any purpose); N/A	av
condition_digitized	condition of digitized content	open text	aw
condition_born-digital	condition of born-digital content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care); Good (e.g., robust and stable, usable for any purpose); N/A	ax
condition_born-digital	condition of born-digital content	open text	ay
record_source	person doing data entry	first last	az
record_source_institution	institutional affiliation of person doing data entry	org name	ba
record_source_city	institutional city	city name	bb
record_source_state	institutional state or territory	abbreviation, see https://en.wikipedia.org/wiki/List_of_U.S._state_abbreviations	bc
record_source_email	contact information	email address	bd
OCLC_control_number	identifier	eight-digit or nine-digit number	be

Collection Description			
Title	On the Margins Newspapers and Periodicals Registry		
Compiler	Katherine Skinner, Educopia Institute; Emily Gore, DPLA; Franky Abbott, DPLA		
Summary	The purpose of this data collection is to create an open registry of LGBT newspapers and periodicals (print, microfische, digitized, and born digital) to make more visible the corpus of stories told by LGBT communities in their own words		
Extent	2 tables		
Format	xlsx		
Parts			
Name	**Purpose**	**Extent**	
codebook	documenting the dataset	1 sheet	
LGBT_Newspapers	registry documenting titles, locations, formats, publishers, collectors, and condition of LGBT newspapers/periodicals	1 sheet	
Field Name	**Variable**	**Values or Explanation**	**Field**
collecting_institution	institution name	name	a
title	publication title	name	b
additional_title	Has the publication used additional title(s)?	yes, no, I don't know	c
additional_title_list	Please specify any known additional titles	name (if multiple, separate by	d
tag_line	Has the publication used a tagline?	yes, no, I don't know	e
tag_line_list	text description	tag line (if multiple, separate by	f
LCCN	identifier	year-serial# or unknown	g
ISSN	identifier	acronym ISSN followed by two groups of four digits separated by a hyphen or unknown	h
city_of_publication	city name	name (if multiple, separate by commas; if unknown, specify	i
state_of_publication	state or territory name	abbreviation (if multiple, separate by commas; if	j
frequency_of_publication	how often published (if variable, select the most representative descriptor)	daily, weekly, biweekly, monthly, quarterly, annually,	k
first_publisher	name of first known publisher	org name or first last	l
last_publisher	name of last known publisher	org name or first last	m
format_print	does the collecting institution have print copies?	yes, no, unknown	n
print_range_description	what date ranges does the collecting institution have in print?	open text (separate date ranges using commas, E.g., from 1900-01-04 to 1930-05-30, from 1932-02-16 to 1932-02-18, from 1944-	o
format_microfilm	does the collecting institution have microfilm copies?	yes, no, unknown	p
microfilm_range_description	what date ranges does the collecting institution have in microfilm?	open text (separate date ranges using commas, E.g., from 1900-01-04 to 1930-05-30, from 1932-02-16 to 1932-02-18, from 1944-	q
format_digitized	does the collecting institution have digitized copies?	yes, no, unknown	r

digitized_range_description	what date ranges does the collecting institution have in digital form?	open text (separate date ranges using commas, E.g., from 1900-01-04 to 1930-05-30, from 1932-02-16 to 1932-02-18, from 1944-	s
digitized_single_PDF	is the digitized issue a single pdf?	binary	t
digitized_individual_pages	are individual pages saved separately as pdfs?	binary	u
digitized_individual_articles are individual articles saved separately as pdfs?		binary	v
digitized_page_TIFF	is each page saved as a single TIFF?	binary	w
digitized_article_TIFF	are individual articles saved separatly as TIFFs?	binary	x
digitized_page_JPEG	is each page saved as a single JPEG?	binary	y
digitized_article_JPEG	are individual articles saved separately as JPEGs	binary	z
digitized_page_TIFF	is each page saved as a single JP2000?	binary	aa
digitized_article_JP2000	are individual articles saved separately as JP2000s	binary	ab
digitized_content_browsabl eis the digitized content browsable by date?		binary	ac
digitized_text_searchable	is the digitized content searchable by keyword?	binary	ad
digitized_fulltext_searchabl e	is the digitized content full-text searchable?	binary	ae
digitized_base	what format was the content digitized from?	print, microfiche, unknown	af
digitized_available	is the digitized content freely available?	binary	ag
digitized_paywall	is the digitized content available behind a paywall?	binary	ah
OCR_type	what type of OCR/indexing was used?	none, raw OCR, corrected OCR, double key indexing, triple key	ai
metadata_schema	what descriptive metadata schema was used?	DC, METS, ALTO, MODS, METS-ALTO, Other (specify)	aj
format_born-digital	does the collecting institution collect/maintain born-digital copies of this title?	yes, no, unknown	ak
format_born-digital_describe	describe born-digital collection process and formats collected	open text, describe 1) how it is acquired, 2) how it is processed, 3) what formats are included, and 4) if you maintain the	al
url(s)	home page of the collection (if digital)	URL(s) separated by commas or "none"; if multiple, please	am
focus_geographical	description of any geographical focus	city, state, region, national, international, none, other	an
focus_topical	description of any topical focuses	keywords separated by commas	ao
copyright_status	status	protected by copyright, public domain, variable, unknown	ap
copyright_status_descriptio n	who is/are the copyright owner(s)	open text, if unknown write "unknown"	aq
condition_print	condition of print content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care);	ar
condition_print	condition of print content	open text	as

condition_microfiche	condition of microfiche content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care);	at
condition_microfiche	condition of microfiche content	open text	au
condition_digitized	condition of digitized content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care);	av
condition_digitized	condition of digitized content	open text	aw
condition_born-digital	condition of born-digital content	Unusable (e.g., fragile, damaged, unrenderable formats); Poor (e.g., deteriorating, low quality dpi); Fair (e.g., minor damage but stable enough to use with care);	ax
condition_born-digital	condition of born-digital content	open text	ay
record_source	person doing data entry	first last	az
record_source_institution	institutional affiliation of person doing data entry	org name	ba
record_source_city	institutional city	city name	bb
record_source_state	institutional state or territory	abbreviation, see https://en.wikipedia.org/wiki/List_of_U.S._state_abbreviations	bc
record_source_email	contact information	email address	bd
OCLC_control_number	identifier	eight-digit or nine-digit number	be