

DPLA plus

Justin Clark, Dustin Hysinger, James Staub, Becky Yoose, and April Mitchell

<https://github.com/jdcc/dpla-dedupe>

Why?

Discovery

Customers Who Bought This Item Also Bought

Page 1 of 3



The Complete Adventures of Huckleberry Finn and Tom Sawyer
Mark Twain
★★★★★ (12)
Kindle Edition
\$0.99



The Adventures of Tom Sawyer (Cambridge World of Literature)
Mark Twain
★★★★★ (5)
Kindle Edition
\$0.99



Oliver Twist
Charles Dickens
★★★★★ (240)
Kindle Edition
\$0.00



Unbroken: A World War II Story of Survival, ...
Laura Hillenbrand
★★★★★ (3,049)
Kindle Edition
\$12.99



A Tale of Two Cities
Charles Dickens
★★★★★ (691)
Kindle Edition
\$0.00



Alice's Adventures in Wonderland
Lewis Carroll
★★★★★ (370)
Kindle Edition
\$0.00



Crime and Punishment
Fyodor Dostoyevsky
★★★★★ (504)
Kindle Edition
\$0.00

Are any of these items inappropriate for this page? [Let us know](#)



Watch TV shows & movies anytime, anywhere. Only \$7.99 a month.

[Member Sign In](#)

[Start Your Free Month](#)

More Like Captain America: *The First Avenger*



How is this different from just regular search?

Why we can't do "Customers who bought this also bought"

Customers Who Bought This Item Also Bought

Page 1 of 3

The Complete Adventures of Huckleberry Finn and Tom ...
Mark Twain
★★★★★ (12)
Kindle Edition
\$0.99

The Adventures of Tom Sawyer (Cambridge World ...
Mark Twain
★★★★★ (5)
Kindle Edition
\$0.99

Oliver Twist
Charles Dickens
★★★★★ (240)
Kindle Edition
\$0.00

Unbroken: A World War II Story of Survival, ...
Laura Hillenbrand
★★★★★ (3,049)
Kindle Edition
\$12.99

A Tale of Two Cities
Charles Dickens
★★★★★ (691)
Kindle Edition
\$0.00

Alice's Adventures in Wonderland
Lewis Carroll
★★★★★ (370)
Kindle Edition
\$0.00

Crime and Punishment
Fyodor Dostoyevsky
★★★★★ (504)
Kindle Edition
\$0.00

Are any of these items inappropriate for this page? [Let us know](#)



De-Duplication

What?

Possible Duplicates:

Record A

Title: [Photograph of Panola Shoals covered bridge, DeKalb County, Georgia, ca. 1940]

Subject: Architecture--Georgia--DeKalb County

Desc: "DeKalb County, ca. 1940. Panola Shoals covered bridge (second built at that location), located on Snapfinger Road (Highway 155). Robert and Wynette Park of Lithonia."--from field notes

Record B

Title: [Photograph of Panola Shoals covered bridge, DeKalb County, Georgia, ca. 1940]

Subject: Architecture--Georgia--DeKalb County

Desc: "DeKalb County, ca. 1940. Panola Shoals Covered Bridge. Second covered bridge built there. Located on Snapfinger Road (Highway 155)."--from field notes

Possible Recommendations

Title : WALB newsfilm clip of police chief Laurie Pritchett speaking about the arrest of kneel-in participants earlier in the day in Albany, Georgia, 1962 July 27

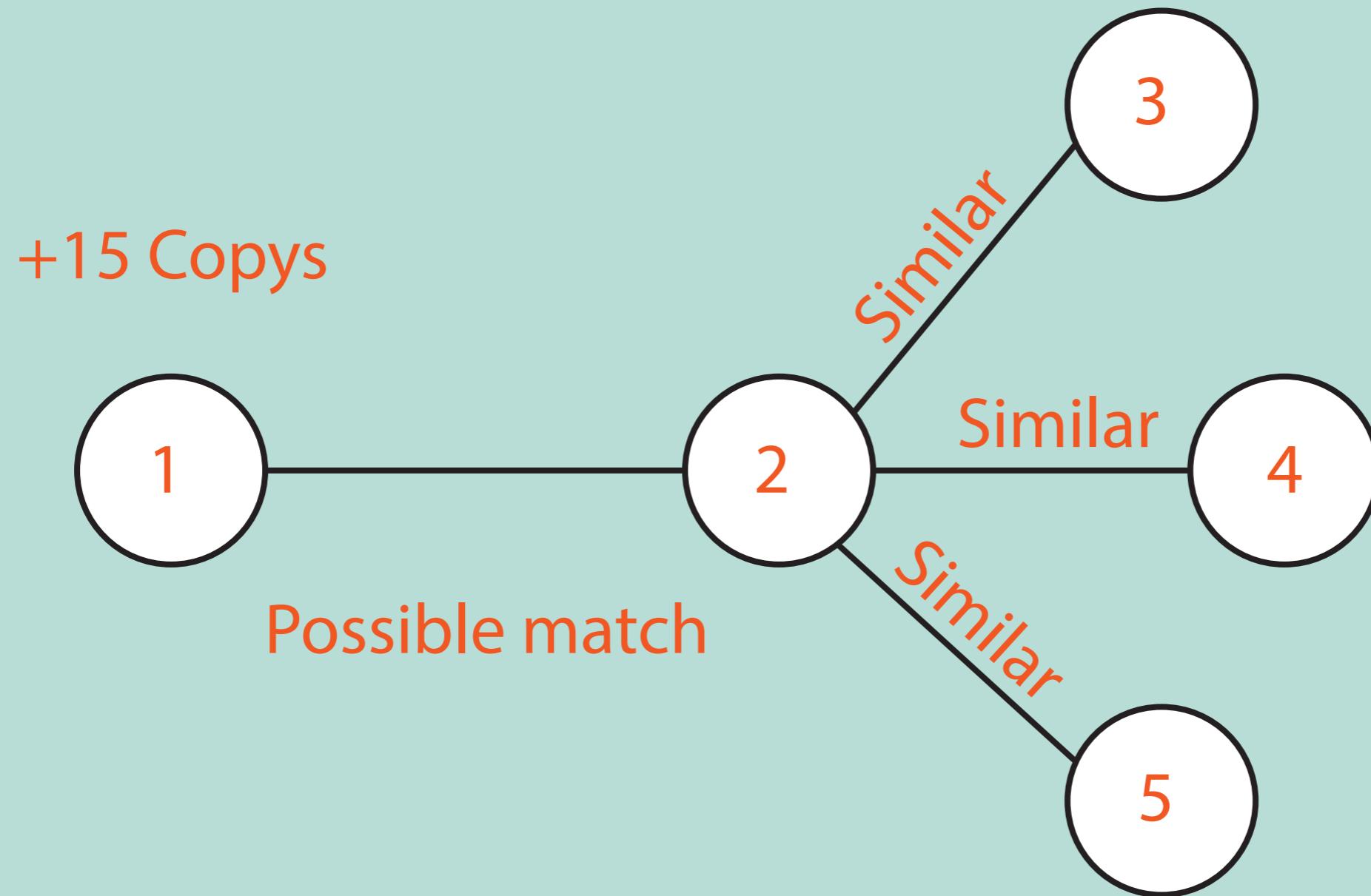
Subject : African American civil rights workers--Georgia--Albany

Creator : WALB (Television station : Albany, Ga.)

Title : WSB-TV newsfilm clip of Archie "Bullfrog Millhouse" Campbell waiting for a bus ride as part of the White Citizens' Council "Freedom Ride North" to Detroit from Macon, Georgia, 1962 May 16

Subject : African American men--Georgia--Macon

Creator : WSB-TV (Television station : Atlanta, Ga.)



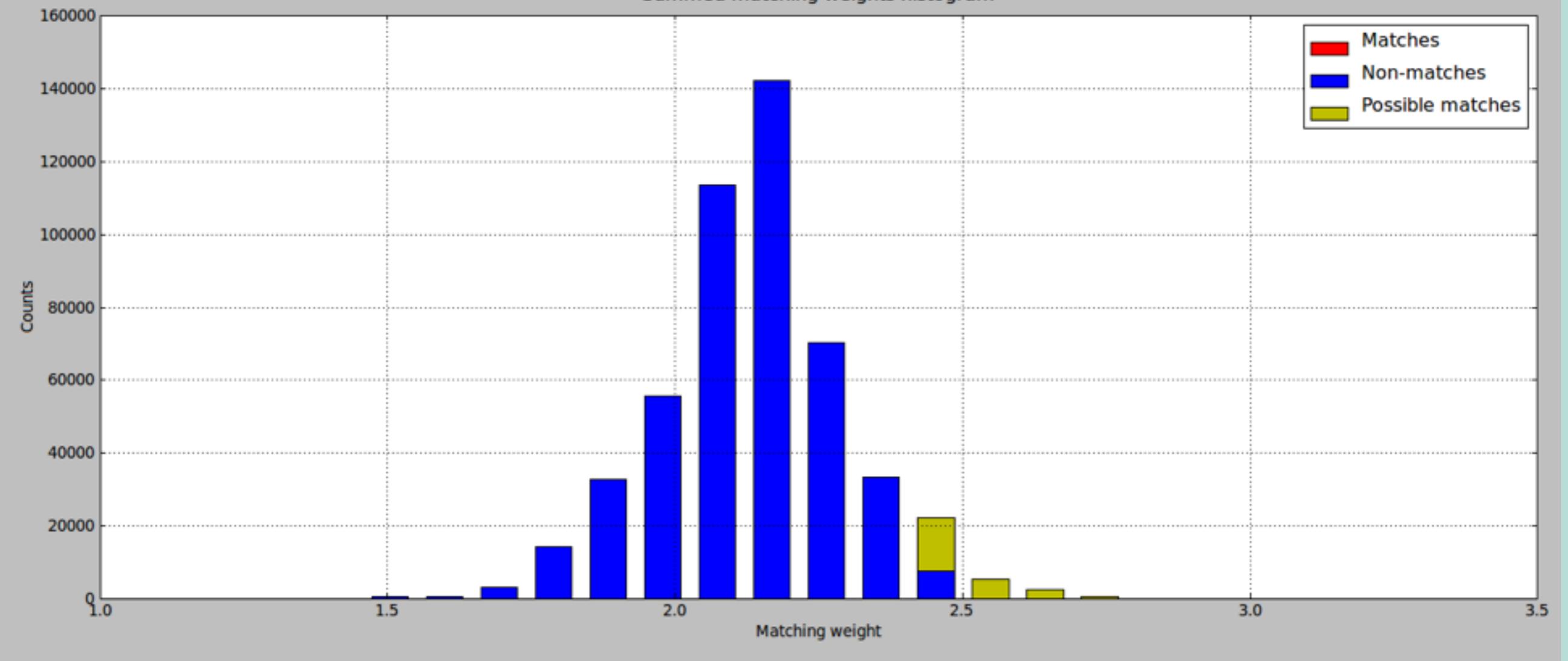
How?

Couch to CSV Converter

FEBRL - Open source python tool with
GUI for record normalization and
de-duplication

Weights CSV

Summed matching weights histogram



How Could this Work?





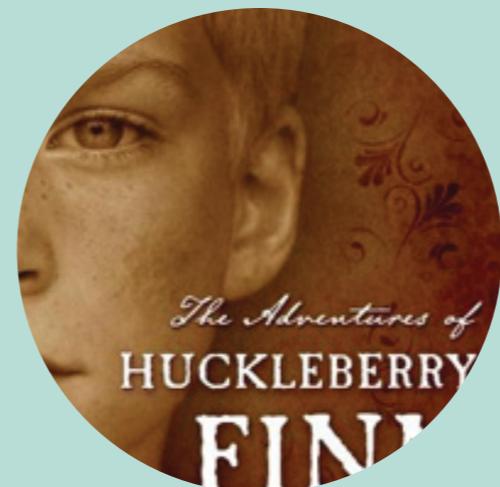
Basic Search

Title

Subject

Keyword

[Search](#)



[Advanced Search ▲](#)



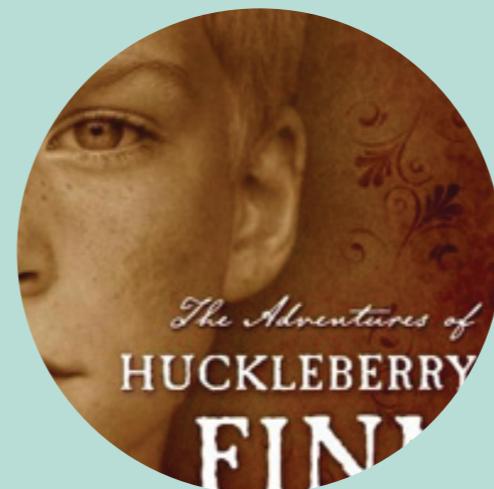
Basic Search

Title

Subject

Keyword

[Search](#)



Advanced Search ▾

Title

DPLAContributor

Subject

Creator

Keyword

Type

Publisher

Spatial

Format

Contributor

Rights

IsPartOf

Description

[Search](#)



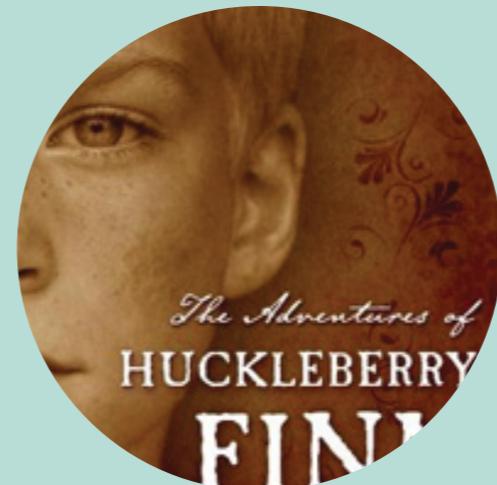
Basic Search

Title

Subject

Keyword

[Search](#)



[Advanced Search ▲](#)

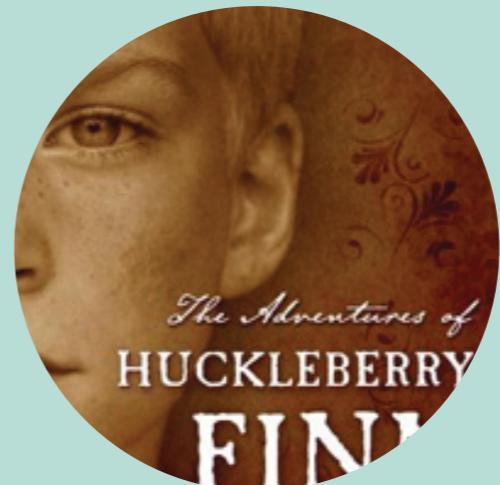


You Searched for:

Title The Adventures of Huckleberry Finn

Subject N/A

Keyword N/A



Top Results

Copys	Title	Subject	Description
55	Adventures of Huckleberry...	Fiction	Adventures of Huck ...
35	Adventures of Huckleberry...	Fiction	Adventures of Huck ...
22	Adventures of Huckleberry...	Fiction	Adventures of Huck ...
02	Adventures of Huckleberry...	Fiction	Adventures of Huck ...

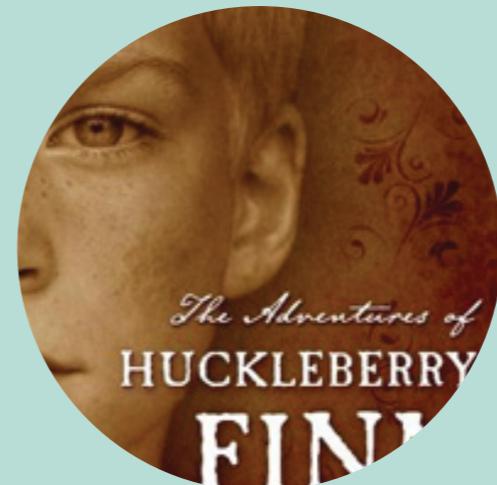


You Searched for:

Title The Adventures of Huckleberry Finn

Subject N/A

Keyword N/A



Top Results

Copys	Title	Subject	Description
55	Adventures of Huckleberry...	Fiction	Adventures of Huck ...
35	Adventures of Huckleberry...	Fiction	Adventures of Huck ...
22	Adventures of Huckleberry...	Fiction	Adventures of Huck ...
02	Adventures of Huckleberry...	Fiction	Adventures of Huck ...

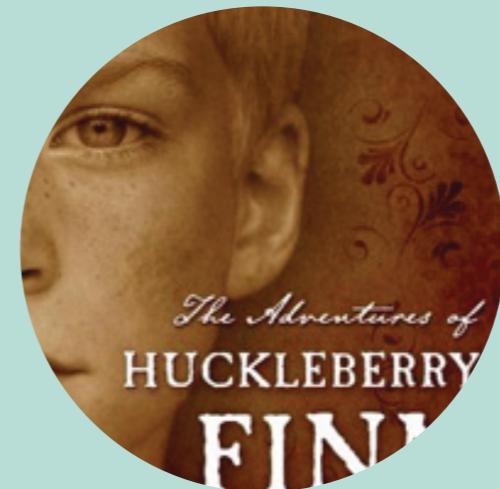


Title The Adventures of Huckleberry Finn

Author Twain, Mark

Publisher American Publishing Company

Description Adventures of Huckleberry Finn is a novel by Mark Twain, first published in England in December 1884 and in the United States in February 1885. Commonly named among the Great American Novels, the work is among the first in major American literature to be written in the vernacular, characterized by local color regionalism. It is told in the first person by Huckleberry "Huck" Finn, a friend of Tom Sawyer and narrator of two other Twain novels (Tom Sawyer Abroad and Tom Sawyer, Detective). It is a sequel to The Adventures of Tom Sawyer.



[Find Nearest Copy](#)

You Might Also Be Interested In:

Sort By: Relevance

Copys	Title	Subject	Description
45	Adventures of Tom Sawyer ...	Fiction	Adventures of Tom ...
25	A Connecticut Yankee in Ki ...	Fiction	A Connecticut Yank ...
12	The Innocents Abroad	Fiction	The Innocents Abro ...
06	Roughing It	Fiction	Roughing It was ori ...

Who?



Users:
Library Patron
Catalogers/
Metadata Librarians

Questions?