



# BICAM Data Codebook

## Citation

Use of BICAM data for academic purposes (presentations, working and conference papers, theses and dissertations, journal articles, software packages) requires that you cite the following paper:

## Committee Prints

Committee Prints are internal research documents created by committees during their deliberations and represent one of the primary ways professional staff contribute to the policymaking process. BICAM indexes 4,487 committee prints from the 67th to the 118th Congress. Committee print data is linked to other key components and related sub-components via `print_id`. Sub-components include associated bills, associated committees, and texts. Committee prints are less common than reports but provide an equally rich source of text for relevant legislation. For example, three prints were published about the Build Back Better Act (`bill_id` = `hr5376-117`) with respective `print_ids` of `hpvt46233-117`, `hpvt46234-117`, and `hpvt46235-117`. Each was printed by the House Rules Committee (`committee_code` = `hsru00`) and describes proposed amendments and changes to the bill for the committee's internal reference.

Each CSV file in this component is described below, one per page.



# BICAM Data Codebook

## Committee Prints

Row Count: 4,277

### Variables

Column Name	Description	Data Type
print_id	Primary key, unique identifier capturing an individual Committee Print across all Committee Print files.	CHAR
print_jacketnumber	Number on published Committee Print document. This is the middle portion of print_id	CHAR
congress	The term of U.S. Congress for a record	NUMERIC
session	Session of a given Congress term for a record	NUMERIC
chamber	Chamber of Congress for a record. Most often “house” or “senate”, but may also be “joint” depending on component type.	CHAR
title	Text title associated with record	CHAR
pages	Integer value of how many pages consist of text from a given record	NUMERIC
print_number	Internal ordinal value of a Committee Print	CHAR
citation	Internal citation for reference for a record	CHAR
updated_at	Date associated with last update of document (from Congress, not to be confused with BICAM version dates)	DATE



# BICAM Data Codebook

## Committee Prints-Associated Bills

Row Count: 689

### Variables

Column Name	Description	Data Type
print_id	Primary key, unique identifier capturing an individual Committee Print across all Committee Print files.	CHAR
bill_id	Primary key, unique identifier capturing an individual Bill across all bill files.	CHAR



# BICAM Data Codebook

## Committee Prints-Committees

Row Count: 1,608

### Variables

Column Name	Description	Data Type
print_id	Primary key, unique identifier capturing an individual Committee Print across all Committee Print files.	CHAR
committee_code	Primary key, unique identifier capturing an individual Committee across all Committee files. This comes from internal THOMAS ID identifiers	CHAR



# BICAM Data Codebook

## Committee Prints-Metadata

Row Count: 4,193

### Variables

Column Name	Description	Data Type
print_id	Primary key, unique identifier capturing an individual Committee Print across all Committee Print files.	CHAR
govinfo_package_id		CHAR
su_doc_class_number		CHAR
migrated_doc_id		CHAR
govinfo_collection_code		CHAR



# BICAM Data Codebook

## Committee Prints-Metadata Granules

Row Count: 4,265

### Variables

Column Name	Description	Data Type
print_id	Primary key, unique identifier capturing an individual Committee Print across all Committee Print files.	CHAR
govinfo_package_id		CHAR
govinfo_granule_id		CHAR



# BICAM Data Codebook

## Committee Prints-Texts

Row Count: 1,778

### Variables

Column Name	Description	Data Type
print_id	Primary key, unique identifier capturing an individual Committee Print across all Committee Print files.	CHAR
raw_text		CHAR
formatted_text		CHAR
pdf		CHAR
html		CHAR
xml	XML formatted document text, imported directly from Congress	CHAR
png		CHAR