
Office of the Clerk
User Guide and Data Dictionary
House Member Data in XML

Prepared by: Legislative Computer Systems

Office of the Clerk, U.S. House of Representatives

February, 2016

Revision History

| Revision | Date | Description |
|----------|----------------|---|
| 1.0 | April, 2015 | Version 1.0 |
| 1.1 | February, 2016 | Version 1.1, added a note about the code for the Select Investigative Panel of the Committee on Energy and Commerce |

Table of Contents

| | |
|---|---|
| 1. Introduction | 2 |
| 1.1. Scope of the Data | 2 |
| 2. Data Set | 2 |
| 3. XML Descriptions | 2 |
| 3.1. Elements and Attributes | 3 |
| 3.2. Unique IDs for Committees and Subcommittees..... | 6 |

1. Introduction

The Office of the Clerk makes available membership lists for the U.S. House of Representatives. These lists are available in PDF format, and starting with the 114th Congress, the data is available in XML format. The document and data are available at <http://clerk.house.gov>.

The purpose of this document is to describe the XML data now available.

1.1. Scope of the Data

The XML file contains current members of the House only. It does not contain committee and subcommittee membership and committee rankings. It contains very little historic information. *See XML descriptions for more information.*

2. Data Set

The House Member data set is compiled by the Office of the House Clerk. In general, there are no restrictions on re-use of the information in the Member data set because U.S. Government works are not subject to copyright protections and are in the public domain. The Office of the House Clerk cannot vouch for the authenticity of data that is not under its control nor do we endorse third party applications. We do not evaluate how our original content is displayed on other sites. Consumers should form their own conclusions as to whether the downloaded data can be relied upon within an application or mash-up.

3. XML Descriptions

The following conventions are used in this document:

- XML element names are denoted with angled brackets and in *courier*. For example, `<title>` is an XML element.
- XML attribute names are denoted with a “@” prefix and in *courier*. For example, `@href` is an XML attribute.

The XML file of members is well-formed. A simple schema has been created to help illustrate the parent, child and sibling relationships in the data. It can be found on <http://xml.house.gov>.

3.1. Elements and Attributes

| Element | Description |
|-----------------|---|
| <MemberData> | Root element |
| @publish-date | Attribute on the root element indicating the publication date. The date will change when the data has been updated. When the file is updated throughout the Congress, a new publish-date will be issued. It will correspond with the date on the printed publication. |
| <title-info> | Elements used for publishing the printed cover page. |
| <congress-num> | The Congress in numeric format (the two-year time period). A Congress begins at noon, January 3 of each odd-numbered year following a general election, unless by law a different day is designated. A Congress lasts for two years, with each year normally constituting a separate session. |
| <congress-text> | The Congress in sentence format (the two-year time period) |
| <session> | The session in numeric format (the one-year period) |
| <majority> | The party that is in the majority. Values are "R" for Republican and "D" for Democrat. |
| <minority> | The party that is in the minority. Values are "R" for Republican and "D" for Democrat. |
| <clerk> | The full name of the Clerk of the House |
| <weburl> | The web site address for the Clerk of the House |
| <members> | Parent container |
| <member> | Parent container for an individual member |
| <statedistrict> | Contains the identifier for the Congressional district using the two (2) character postal code plus the two (2) digit district number. In states with only one district, the two digit district code is "00" indicating an at-large seat. |
| <member-info> | Parent container |
| <namelist> | Name of member in the following format: Last name followed by a comma followed by first name. |
| <bioguideID> | Unique identification for each member of Congress. This is the BioGuide ID from the Biographical Directory of Congress: http://bioguide.congress.gov . It is alphanumeric and begins with the first letter of the Member's last name, followed by six (6) numeric digits. Once an individual is assigned a BioGuide ID, the ID stays with the member if the member resigns and returns to Congress in a non-consecutive Congress, changes his/her name, or is elected to the Senate or the Office of Vice-President |
| <lastname> | Surname or family name. Will not be empty. |
| <firstname> | First name or initial(s) of first name. Will not be empty. |
| <middlename> | Middle name or middle initial. May be maiden name or another additional surname. May be empty. |
| <suffix> | Generational family suffix. Possible values: "Jr.", "Sr." "II" "III" et.al. May be empty. Note: Suffixes (and courtesies) that indicate that the individual holds a position, educational degree, accreditation or the like are not used. |
| <courtesy> | A form of address to show respect and it is placed immediately before a person's name. Possible values: "Mr.", "Mrs.", "Ms.", "Miss" |

| | |
|--------------------------------|--|
| <official-name> | The Member's full name. It should match the Member's frank signature. It will contain a first name and a last name. It may contain a middle name and/or suffix. It does not include the words, "The Honorable" or a courtesy. |
| <formal-name> | <p>A Member's name as it read aloud on the House floor. This variation of a Member's name appears in the <i>Congressional Record</i>, on bills for sponsors and cosponsors, and other legislative documents. If the Member's name is the same or similar to another Member, the state digraph will be included and if needed, the first name of the Member.</p> <p>Samples: Mrs. Miller of Michigan, Mr. George Miller of California, Mr. Young.</p> <p>A Member's formal name may change from Congress to Congress. The publication "Official List of Duplicate And Similar Names of Members" indicates who is on the duplicate/similar list for a particular Congress. Once a Member has the state digraph for a particular Congress it stays for the entire Congress even if the other duplicate or similarly-named Member leaves office.</p> |
| <prior-congress> | Indicates the last Congress (in numeric format) that the Member served prior to the current Congress. A value of zero (0) indicates that the Member has not served in a prior Congress. |
| <elected-date> | The date the Member was elected to office for the current Congress. If there was a run-off election, the date of the run-off election is used. Example of the format: November 6, 2014 |
| <sworn-date> | The date the Member took the oath of office (sworn-in) for the current Congress. Example of the format: November 6, 2013 |
| @date | The date in the following format: YYYYMMDD |
| <party> | The political party of the Member. Possible values are "R", "D", and "I". |
| <caucus> | The party the Member chooses to affiliate with if the Member is an independent or a member of a third party. Possible values are "R" and "D". |
| <state> | Parent container for the state of the Member. |
| @postal-code | The two (2) character postal code for the state on the <state> element. |
| <state-fullname> | The full name of the state of the Member. |
| <district> | <p>The congressional district number with the ordinal indicator. Examples: 12th, 1st, 2nd</p> <p>For at-large districts, the value "At Large" is used. The terms "Delegate" and "Resident Commissioner" are used for the territories and the District of Columbia.</p> |
| <town> | Hometown as declared by the Member |
| <office-building> | The elements used for the Member's Washington DC office location. |
| <office-room> | |
| <office-zip> | |

| | |
|--------------------------------------|---|
| <office-zip-suffix> | <p><office-building> contains the building abbreviations.</p> <p>RHOB = Rayburn House Office Building, LHOB = Longworth House Office Building, CHOB = Cannon House Office Building, FHOB = Ford House Office Building OFOB = O’Neill Federal Office Building</p> <p><office-room> contains the room number</p> <p><office-zip> contains the zip code. For all Members, it should be 20515. <office-zip-suffix> contains the +4 for the zip code.</p> |
| <phone> | <p>The Member’s phone number in Washington DC.</p> <p>Format: (202) 225-XXXX</p> |
| <footnote-ref> | Used for footnotes. |
| <footnote> | |
| <predecessor-info> | <p>Parent container for information on the predecessor of the current seat. Applies to the current Congress only.</p> |
| @cause | <p>The reason why the predecessor vacated the seat. Possible values: “R” for resignation, “D” for death, “E” for expulsion.</p> |
| <pred-lastname> | Last name of the Member who preceded the current Member. |
| <pred-firstname> | First name of the Member who preceded the current Member. |
| <pred-party> | Political party of the Member who preceded the current Member. |
| <pred-vacate-date> | The date when the predecessor vacated the office. |
| <committee-assignments> | <p>Parent container for individual member committee assignments.</p> <p>Note: At the start of a Congress, there may be little or no committee information in the file until the committees hold their organizational meetings.</p> |
| <committee> | Element indicating current member’s committee assignment and member’s rank in committee |
| <subcommittee> | Element indicating member’s subcommittee assignment and member’s rank in subcommittee |
| @leadership | Title of position, if the Member has a committee or subcommittee leadership position. |
| @comcode | Committee and subcommittee codes. See <i>section about unique committee and subcommittee codes.</i> |
| @subcomcode | |
| <committees> | Parent container |
| <committee> | Element containing information about the committee |
| @type | The type of committee. Values include “standing” for standard committees, “select” for select committees, “joint” for joint committees. |
| @com-header-text | Text used in the printed publication. The value could change. The most common value is “The chairman and ranking minority member are ex officio members of all subcommittees.” |
| @com-room | The room number for the committee |
| @subcom-room | The room number for the subcommittee. |
| @com-zip | The zip code for the committee. |
| @subcom-zip | The zip code for the subcommittee. |
| @com-zip-suffix | The +4 for the zip code for committee. |
| @subcom-zip-suffix | The +4 for the zip code for subcommittee. |

| | |
|---|--|
| @com-building-code @subcom-building-code | The building code RHOB = Rayburn House Office Building, LHOB = Longworth House Office Building, CHOB = Cannon House Office Building, FHOB = Ford House Office Building OFOB = O'Neill Federal Office Building |
| @com-phone @subcom-phone | The phone number for the committee The phone number for the subcommittee |
| <committee-fullname> <subcommittee-fullname> | The full name of the committee The full name of the subcommittee |
| <ratio> | Parent container for the ratio of majority and minority members of the committee or subcommittee |
| <majority> | Child element of <ratio>. Indicates the number of members the majority party could appoint to serve on the committee or subcommittee. |
| <minority> | Child element of <ratio>. Indicates the number of members the minority party could appoint to serve on the committee or subcommittee. |

3.2. Unique IDs for Committees and Subcommittees

The unique id for the committee is the two letter committee code plus two zeros. This unique id has been used internally by the House for years. For the standing¹ committees listed below, the committee code is unique across Congresses, and it does not change if there are slight changes in the name from Congress to Congress. Select committees are generally temporary and are not included in this list. Select committees do receive codes and will be found in the appropriate XML files. Codes for select committees may not be unique across Congresses.

The id for a subcommittee is unique for the current congress. It is the two letter committee code plus two digits. It is not unique across Congresses as subcommittees can be different across Congresses.

¹ Select committees are established by the House and/or the Senate usually for limited time periods and for strictly limited purposes. Standing committees are permanent panels with jurisdiction over measures and laws in certain areas of public policy. (*Our American Government*, House Document 108-94, Pg.33-36).

| Committee Name | Code |
|--|------|
| Committee on Agriculture | AG00 |
| Committee on Appropriations | AP00 |
| Committee on Armed Services | AS00 |
| Committee on Financial Services | BA00 |
| Committee on the Budget | BU00 |
| Committee on Education and the Workforce | ED00 |
| Committee on Foreign Affairs | FA00 |
| Committee on Oversight and Government Reform | GO00 |
| Committee on House Administration | HA00 |
| Committee on Homeland Security | HM00 |
| Committee on Energy and Commerce | IF00 |
| Permanent Select Committee on Intelligence | IG00 |
| Committee on the Judiciary | JU00 |
| Committee on Transportation and Infrastructure | PW00 |
| Committee on Rules | RU00 |
| Committee on Small Business | SM00 |
| Committee on Ethics | SO00 |
| Committee on Science, Space, and Technology | SY00 |
| Committee on Veterans' Affairs | VR00 |
| Committee on Ways and Means | WM00 |
| Joint Economic Committee | EC00 |
| Joint Committee on Taxation | IT00 |
| Joint Committee on the Library | JL00 |
| Joint Committee on Printing | JP00 |
| Joint Select Committee on Deficit Reduction | JR00 |

Note: The Select Investigative Panel of the Committee on Energy and Commerce (established by H. Res. 461 of the 114th Congress) has two codes: IF04 that is in use on docs.house.gov and IE00 which is included in the Member and Committee data XML file.