GLORIA J. LETT DEPUTY CLERK

ROBERT F. REEVES
DEPUTY CLERK

Office of the Clerk H.S. House of Representatives Washington, DC 20515–6601

April 15, 2021

The Honorable Zoe Lofgren Chairperson Committee on House Administration 1309 Longworth House Office Building Washington, DC 20515

The Honorable Rodney Davis Ranking Member Committee on House Administration 1216 Longworth House Office Building Washington, DC 20515

Dear Chairperson Lofgren and Ranking Member Davis,

As required by Section 502 of House Resolution 756 from the 116th Congress, agreed to by the House on March 10, 2020, I am attaching this quarterly report about the comparative print project.

If you or others on the Committee have questions about this report, please do not hesitate to contact me.

Sincerely,

Clerk of the House

Enclosure

Comparative Print Project

As required by Section 502 of House Resolution 756 from the 116th Congress, the Clerk of the House submits this quarterly report about the comparative print project.

As described in the previous three reports, the House is building a suite of software programs that will allow House staff and others to create on-demand, point-in-time comparative prints for three distinct types of comparisons. These comparative prints illustrate changes between

- two versions of a bill, resolution, or amendment (document to document comparisons)
- current law and current law as proposed to be changed by amendments contained in a bill, resolution, or amendment to current law (codified and non-codified law)
- a bill or resolution and the bill or resolution as proposed to be modified by amendments (amendment impact).

Project Updates

Stakeholder Engagement and the Pilot Group

Since the last report, the Clerk and the Legislative Counsel (through the project team and individually) have continued to ensure that stakeholders are engaged in this project through individual conversations and status update meetings. Since the initial report, a pilot group of Committee staff has been engaged with the project team and actively using the alpha/beta system. Throughout the pilot group period, users have given feedback about the usability, desirability, learnability, and accuracy of the system. The group's responses continue to be positive and helpful.

During this reporting period, we invited additional staff to the group and held two pilot group meetings. Currently, more than 100 House Committee staff from the following Committees have access to the software: Agriculture, Appropriations, Armed Services, Foreign Affairs, House Administration, Judiciary, Natural Resources, Oversight and Reform, Rules, Science, and Transportation. We are inviting additional Committee staff, including those serving Committees that have yet to participate, to be part of the pilot group. New participant onboarding and groupwide meetings are scheduled for April and May.

Suite of Software Programs

As described in previous reports, the Comparative Print System is designed to display legislative changes in context: how a bill might change current law; how two versions of a legislative proposal are different; and how an amendment proposes to change a bill. Understanding these changes is critical to making decisions on pending legislation. Project staff have continued to improve these tools, including the new feature called the "Bill Viewer."

Comparative Print Project Office of the Clerk · April 15, 2021 · Page 1 of 14

 $[\]frac{1}{https://cha.house.gov/sites/democrats.cha.house.gov/files/documents/Clerk\%20Comparative\%20Print\%20Report.pdf;}{https://cha.house.gov/sites/democrats.cha.house.gov/files/Clerk-QR1-StandarizedFormat.pdf.}$

In our system, comparative print results are shown in two ways: an online, dynamic, interactive report and a PDF that can be printed or downloaded to share. Users access these results through a key feature, the Toolbox View. Previous reports describe the features of the Toolbox View and how it relates to the "How a Bill Changes Current Law" tool. In this report, we will address the tool "Bill Differences." Improving this tool has been a primary focus this past quarter.

Bill to Bill Differences

The "Bill to Bill Differences" tool illustrates how two versions of a legislative proposal are different. Comparing legislative text can be difficult because of changes in structure and complexity, as well as the tendency to move legislative provisions to a different place in the document as the bill moves through the legislative process.

An example of this comparison is H.R. 2546 from the 116th Congress. H.R. 2546 was introduced with the short title, "Colorado Wilderness Act of 2019." A quick search on Congress.gov returns a bill with the same short title—H.R. 6492—from the 115th Congress.

If we are doing typical legislative research, we would assume that both bills—H.R. 6492 (115th Congress) and H.R. 2546 (116th Congress)—are the same or similar proposals. The Comparative Print System will help us quickly confirm our assumptions.

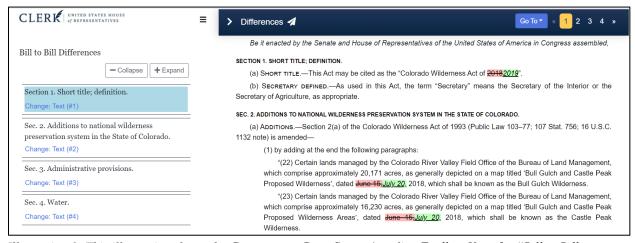


Illustration 1: This illustration shows the Comparative Print System's online Toolbox View for "Bill to Bill Differences." It is comparing the bill 115HR6492 with 116HR2546. We can see that the year in the short title was updated and descriptions of two maps were modified.

The interactive outline in Illustration 1 allows users to view an outline of the bill and navigate to each section to see changes. The differences are illustrated using colors and text styling (strikethrough and underline). Using both colors and styling aids all users, including those who may have difficulty distinguishing colors.

As we assumed, the bills are similar in nature and few revisions were made to the base text when it was reintroduced in the 116th Congress.

Commercial, off-the-shelf products, including those currently available to House staff, can accurately compare differences in federal legislative proposals that are not very complex—like the ones seen in Illustration 1. However, these products often fail in both accuracy and precision when the differences include changes in structure (e.g., a standalone bill was added to another bill) and location (e.g., a portion of the legislative text moved).

To do the same comparison with tools available to House staff, we copied the text version of both bills into separate MS Word documents. In MS Word, we used the compare function to view the differences between the two documents. Illustration 2 shows that the MS Word comparison reveals the same differences that the Comparative Print System does. Readers can view the changes by looking for the red marks on the compared document (left side) and then reading the text from the original and revised documents (right side).

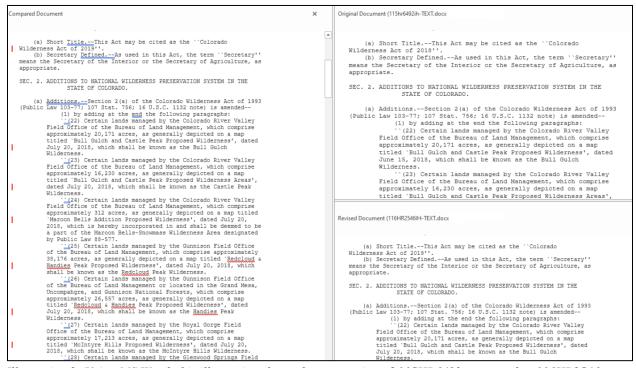


Illustration 2: Using MS Word, this illustration shows the text version of 115HR6492 compared to 116HR2546.

Viewing output with simple red marks on the left side as shown in Illustration 2 is difficult for some users. A second product created by a Washington, DC,-based company and available for use by House staff, is shown in Illustration 3. The product's downloaded PDF version is similar to what one would see in MS Word (Illustration 4).

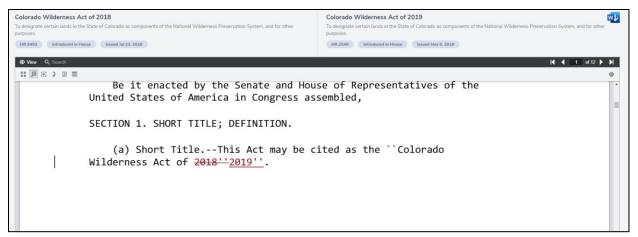


Illustration 3: Using a product currently available to House staff, this illustration shows an online comparison of 115HR6492 and 116HR2546.

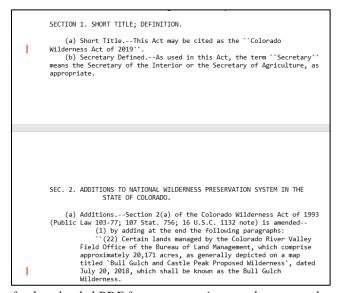


Illustration 4: This snippet of a downloaded PDF from a comparison product currently available to House staff shows a result similar to the MS Word comparison in Illustration 2.

On February 2, 2020, the Committee on Natural Resources reported H.R. 2546 to the House with amendments. Using the Toolbox View in the Comparative Print System, we can quickly see the changes between the introduced and reported versions of the bill.

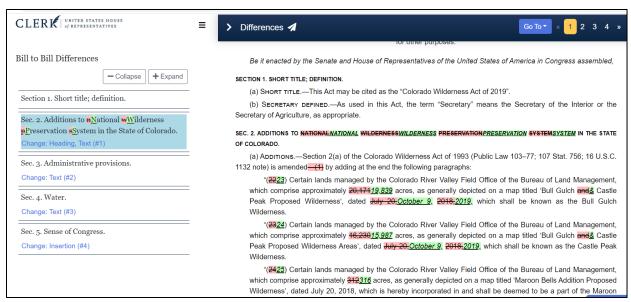


Illustration 5: This illustration shows differences between 116HR2546IH and 116HR2546RH.

After the Committee reported the bill to the House, the Rules Committee met and constructed Rules Committee Print 116-50, comprising modified base text from the reported version and the incorporation of five additional bills reported from the Committee on Natural Resources. The Comparative Print System recognizes this, and the changes are shown in Illustration 6.²

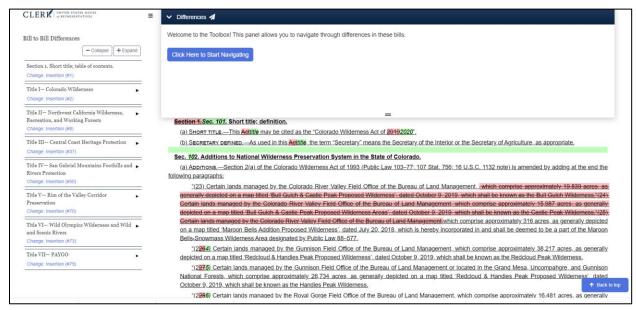


Illustration 6: This illustration shows differences between 116HR2546RH and the Rules Committee Print 116-50.

² https://rules.house.gov/bill/116/hr-2546; Rules Committee Print 116-50. Showing the text of H.R. 2546, H.R. 2250, H.R. 2199, H.R. 2215, H.R. 1708, and H.R. 2642, as ordered reported by the Committee on Natural Resources, with modifications.

As for MS Word, it "over redlines" and shows H.R. 2546 RH as being completely stricken and the material in the six Titles containing the six originally reported measures included in the Rules Committee Print as newly inserted text. The other product available to House staff does not provide the ability to compare the Rules Committee Print; however, when comparing the RH version of H.R. 2546 to the EH version, the results are the same as the MS Word comparison.

Moves and Deletions

In Illustration 6 of H.R. 2546, the Comparative Print System displays the differences in the paragraphs contained in subsection (a) of Section 201 accurately. However, it is not as precise as the project team and some of our pilot group users would like, and the project team is continuing to discuss this issue. If we were looking at the entire Section 102, the quoted matter in subsection (a) contained paragraphs 23 to 34 in the RH version, but the Committee Print deletes three paragraphs and moves and redesignates the remaining paragraphs. It is not easy to distinguish which three paragraphs have been deleted.

A more precise comparison would display the quoted paragraphs in subsection (a) as follows—

- "(25 23) Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land Management, which comprise approximately 316 acres, as generally depicted on a map titled 'Maroon Bells Addition Proposed Wilderness', dated July 20, 2018, which is hereby incorporated in and shall be deemed to be a part of the Maroon Bells-Snowmass Wilderness Area designated by Public Law 88–577.
- "(26 24) Certain lands managed by the Gunnison Field Office of the Bureau of Land Management, which comprise approximately 38,217 acres, as generally depicted on a map titled 'Redcloud & Handies Peak Proposed Wilderness', dated October 9, 2019, which shall be known as the Redcloud Peak Wilderness.
- "(27 25) Certain lands managed by the Gunnison Field Office of the Bureau of Land Management or located in the Grand Mesa, Uncompahgre, and Gunnison National Forests, which comprise approximately 26,734 acres, as generally depicted on a map titled 'Redcloud & Handies Peak Proposed Wilderness', dated October 9, 2019, which shall be known as the Handies Peak Wilderness.
- "(28 26) Certain lands managed by the Royal Gorge Field Office of the Bureau of Land Management, which comprise approximately 16,481 acres, as generally depicted on a map titled 'Table Mountain & McIntyre Hills Proposed Wilderness', dated November 7, 2019, which shall be known as the McIntyre Hills Wilderness.
- "(29 27) Certain lands managed by the Glenwood Springs Field Office of the Bureau of Land Management, which comprise approximately 10,282 acres, as generally depicted on a map titled 'Grand Hogback Proposed Wilderness', dated October 16, 2019, which shall be known as the Grand Hogback Wilderness.
- "(31 28) Certain lands managed by the Grand Junction Field Office of the Bureau of Land Management, which comprise approximately 25,624 acres, as generally depicted on a map titled 'Demaree Canyon Proposed Wilderness', dated October 9, 2019, which shall be known as the Demaree Canyon Wilderness.
- "(32 29) Certain lands managed by the Grand Junction Field Office of the Bureau of Land Management, which comprise approximately 28,279 acres, as generally depicted on a map titled 'Little Books Cliff Proposed Wilderness', dated October 9, 2019, which shall be known as the Little Bookcliffs Wilderness
- "(23 30) Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land Management, which comprise approximately 19,839 14,886

acres, as generally depicted on a map titled 'Bull Gulch & Castle Peak Proposed Wilderness', dated October 9, 2019 January 29, 2020, which shall be known as the Bull Gulch Wilderness.

"(24 31) Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land Management, which comprise approximately 15,987 12,016 acres, as generally depicted on a map titled 'Bull Gulch & Castle Peak Proposed Wilderness Areas', dated October 9, 2019 January 29,2020, which shall be known as the Castle Peak Wilderness.

[Deleted paragraphs]

"(30) Certain lands managed by the Glenwood Springs Field Office of the Bureau of Land Management or located in the White River National Forest, which comprise approximately 16,101 acres, as generally depicted on a map titled 'Flat Tops Proposed Wilderness Addition',

dated October 9, 2019, and which are hereby incorporated in and shall be deemed to be a part of the Flat Tops Wilderness designated by Public Law 94–146.

"(33) Certain lands managed by the Glenwood Springs Field Office of the Bureau of Land Management, which comprise approximately 7,376 acres, as generally depicted on a map titled 'Pisgah East & West Proposed Wilderness' and dated October 16, 2019, which shall be known as the Pisgah East Wilderness, upon being designated as wilderness as provided by section 3(h)(2) of the Colorado Wilderness Act of 2019.

"(34) Certain lands managed by the Glenwood Springs Field Office of the Bureau of Land Management, which comprise approximately 6,828 acres, as generally depicted on a map titled 'Pisgah East & West Proposed Wilderness' and dated October 16, 2019, which shall be known as the Pisgah West Wilderness, upon being designated as wilderness as provided by section 3(h)(2) of the Colorado Wilderness Act of 2019.".

Again, the project team continues to analyze whether the desired view above is achievable, particularly without degrading the results for other comparisons. The text of the quoted paragraphs in Section 201 has ~75 percent repeated language, with only a few key words changed. In the moved paragraphs, two of the key phrases (date and wilderness size) are changed along with the move. Updating the differencing algorithm to "match" these paragraphs may result in too many false positives in other bills. The point here is that accuracy can be achieved; however, the degree to which precision can be achieved is still being analyzed, designed, and implemented. Feedback from the pilot group is key in this matter.

Illustration 7 shows the improvements that the project team has made and what may be implemented in future releases of the tool.

SEC. 101. SECTION 1. SHORT TITLE; DEFINITION. (a) Short TITLE.—This Actitile may be cited as the "Colorado Wilderness Act of 2019 2020" (b) Secretary Defined.—As used in this Actitile, the term "Secretary" means the Secretary of the Interior or the Secretary of Agriculture, as appropriate. SEC. 102. ADDITIONS TO NATIONAL WILDERNESS PRESERVATION SYSTEM IN THE STATE OF COLORADO. (a) ADDITIONS.—Section 2(a) of the Colorado Wilderness Act of 1993 (Public Law 103-77; 107 Stat. 756; 16 U.S.C. 1132 note) is amended by adding at the end the following paragraphs: "(23) Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land Management, which comprise approximately 316 acres, as generally depicted on a map titled 'Maroon Bells Addition Proposed Wilderness', dated July 20, 2018, which is hereby incorporated in and shall be deemed to be a part of the Maroon Bells-Snowmass Wilderness Area designated by Public Law 88-577 Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land Management which comprise approximately 19,839 acres, as generally depicted on a map titled 'Bull Gulch & Castle Peak Proposed Wilderness', dated October 9, 2019, which shall be known as the Bull Gulch Wilderness. "(24) Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land Management, which comprise approximately 15,987 acres, as generally depicted on a map titled 'Bull Gulch & Castle Peak Proposed Wilderness Areas', dated October 9, 2019, which shall be known as the Castle Peak Wilderness "(25) Certain lands managed by the Colorado River Valley Field Office of the Bureau of Land nagement, which comprise approximately 316 acres, as generally depicted on a map titled 'Maroon Bells Addition Proposed Wilderness', dated July 20, 2018, which is hereby incorporated in and shall be deemed to be a part of the Maroon Bells-Snowmass Wilderness Area designated by Public Law 88-577. "(2624) Certain lands managed by the Gunnison Field Office of the Bureau of Land Management, which comprise approximately 38,217 acres, as generally depicted on a map titled 'Redcloud & Handies Peak Proposed Wilderness', dated October 9, 2019, which shall be known as the Redcloud Peak Wilderness "(2725) Certain lands managed by the Gunnison Field Office of the Bureau of Land Management or located in the Grand Mesa, Uncompangre, and Gunnison National Forests, which comprise approximately 26,734 acres, as generally depicted on a map titled 'Redcloud & Handies Peak Proposed Wilderness' dated October 9, 2019, which shall be known as the Handies Peak Wilderness. "(2826) Certain lands managed by the Royal Gorge Field Office of the Bureau of Land Management, which comprise approximately 16,481 acres, as generally depicted on a map titled 'Table Mountain & McIntyre Hills Proposed Wilderness', dated November 7, 2019, which shall be known as the McIntyre Hills Wildemess

Illustration 7: This illustration shows improvements that are being made to the display of the differences compared to Illustration 6. While a difference can be accurately displayed, it may not be as precise as expected.

Additionally, the Comparative Print System is handling these types of deletions, moves, and reorganization of material at the section levels and higher. Using as an example H.R. 1319, the American Rescue Plan Act of 2021, in Illustration 8, Section 2005 concerns the Bureau of Indian Education and was included in H.R. 1319, as reported from the House Committee on the Budget.

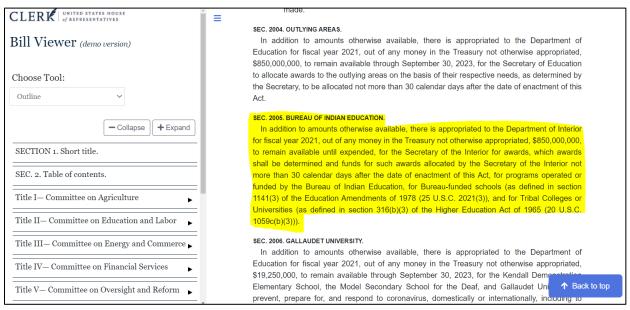


Illustration 8: The Comparative Print System's Bill Viewer displays a portion of 117HR1319RH.

In the Engrossed Amendment from the Senate, Section 2005 was modified, moved near the bottom of the bill, and redesignated as Section 11005. The Comparative Print System was able to detect this move and illustrate the modifications made to the text. This is one of the features that distinguishes this tool from MS Word and other commercially available text comparison programs. It is not that other comparison program approaches are necessarily wrong; they simply take a different approach and are not as precise as what is needed for legislative language as it is refined throughout the legislative process.

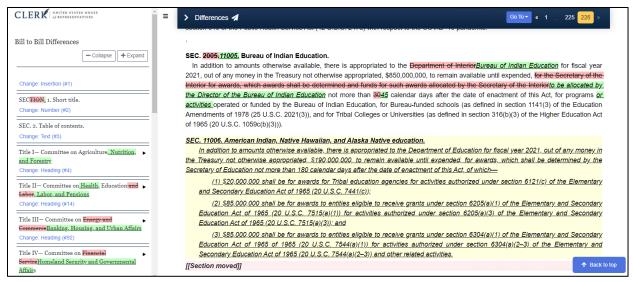


Illustration 9: The Comparative Print System's Toolbox View displays 117HR1319EH compared with 117HR1319EAS.

Another example of the accuracy and precision we are striving for can be seen when comparing different versions of H.R. 133 from the 116th Congress. H.R. 133 was introduced as the "United States–Mexico Economic Partnership Act." It passed both chambers and was enrolled as the "Consolidated Appropriations Act of 2021."

To view what happened with the original, introduced text of H.R. 133, a user would need to manually search for keywords from the text—hoping to find where those provisions ended up in the final version. With the Comparative Print System, a user can easily see that the sections were moved and renumbered and the text was modified. Illustrations 10 and 11 show that five sections were moved from the previous version and found lower in the document.

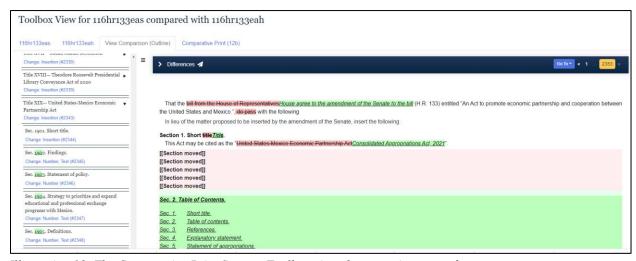


Illustration 10: The Comparative Print Systems Toolbox view shows sections moved.

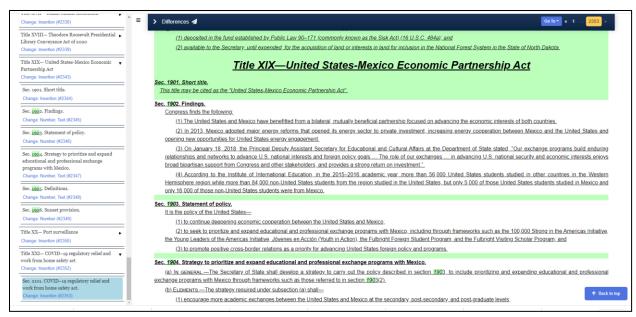


Illustration 11: The Comparative Print System's Toolbox view shows the moved sections from Illustration 10.

Again, the project team continues to improve the display of text differences, ensuring that the comparison is not only accurate but precise. For reasons discussed, we have assigned an "alpha" label to this portion of the Comparative Print System.

How a Bill Changes Current Law and the Current Law Repository

The Comparative Print System illustrates how a bill might change current law—the most complex of all the comparisons. We described it in some detail in the previous reports. Since the last report, the project team has made minor updates and enhancements to the program. We have assigned a "beta" label to this portion of the Comparative Print System.

As mentioned in previous reports, there is no central location for an up-to-date, electronic, machine-readable repository for current law and all the matters that Congress can amend. The Comparative Print System will serve as this central location and potentially provide Members and staff all current law in one centralized location. While the U.S. Code is available in its entirety online, the complete set of Statute Compilations used by House and Senate drafters is not available online. Some, but not all, Statute Compilations are available online on the Government Publishing Office's (GPO) govinfo.³ Until the legislative branch completes its current project "COMPS in USLM Project" and makes available to the public all Statute Compilations, no commercially available product could accurately and precisely offer the House a tool to see how a bill or resolution would change current law.

³ https://www.govinfo.gov/app/collection/comps.

How an Amendment Changes a Bill

Another tool in the Comparative Print System is "How an Amendment Changes the Bill," which illustrates how a Committee or floor amendment would change a bill. The project team continues to work on the initial release of this program. The current AIP (amendment impact program) has been connected to the front end of the comparative print system and was made available to the pilot group in late December 2020. The display of the amendments and the comparative print will be similar to the "How Bill Changes Law" and "Bill to Bill Differences" features described in the previous reports.

Bill Viewer

In previous reports, we described a new feature called the "Bill Viewer." This smart reading tool allows users to search or upload a single bill document and view it using the interactive outline view as shown in Illustrations 8 and 12. The Bill Viewer is a necessary feature to ensure that staff continue to review, analyze, and understand the entire bill proposal. The free-standing provisions are as important as the provisions amending current law. Simply looking at the changes proposed to current law contained in the bill or seeing differences between two versions does not give the reader a full understanding of the underlying proposal. This new feature, along with the three comparisons, will provide House Members and staff a set of powerful tools to examine bills, resolutions, and amendments.

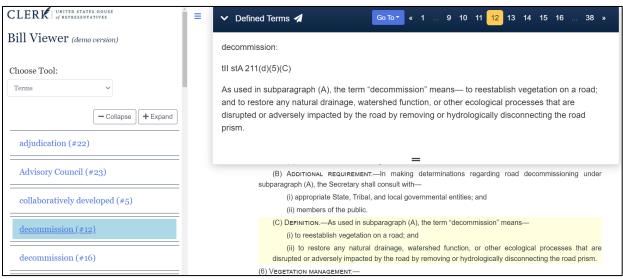


Illustration 12: Using the Bill Viewer tool, this illustration shows a list of defined terms in alphabetical order contained in 116HR2546EH.

Additionally, the project team, along with the stakeholders, continues to discuss what smart-reading capabilities can be added to the Bill Viewer that will allow users to interact with and work with matters of interest contained in the legislative proposal. For example, one of those capabilities is quickly identifying terms defined by the bill and easily navigating to those

portions of the proposal containing the defined terms (Illustration 12). The Bill Viewer has been given a "demonstration label" at this time.

Project Management

To complete this work, the project team continues to use agile methodologies, which allow for the continued refinement of the requirements and help ensure that the House obtains a product that meets its needs. Furthermore, we continue to use the principles of human-centered design to ensure that we address product development with the end user in mind. This approach assists in mitigating risk from lack of user adoption as users are involved in the design process.

Roles and Responsibilities

Defining the roles and responsibilities needed to fully support this application is key. The Legislative Computer Systems division of the Clerk's office and the Office of the Legislative Counsel are taking steps to identify staff who can serve as Tier 1 and Tier 2 help desk support.

As mentioned in the previous reports, some features of the system will be made available to users only after they receive instructor-led training. Other features, such as Bill Viewer, will be made available to all House staff without training. The project team has been working with the Chief Administrative Officer's Congressional Staff Academy to develop on-demand and instructor-led courses. We have developed draft storyboards for the on-demand course and these are under review. We anticipate that staff from House Office of Legislative Counsel (HOLC) and the Clerk's office will lead these courses during the initial rollout phases of this project. Accordingly, we have updated the following chart to reflect these on-going discussions.

Roles and Responsibilities	CLERK	HOLC	CAO
Maintain the compare.house.gov system (software, hardware, related components)	X		
Maintain the files/content of the Statute Compilations component current law dataset		X	
Maintain data exchange components	X	X	
Maintain XML schemas (maintained by the Legislative Branch XML Working Group, which is co-chaired by Clerk staff)	X	X	
Receive and route user calls to the appropriate staff	X	X	X
Answer technical helpdesk calls (e.g., access, log-in)	X		TBD
Answer content-related calls (e.g., What is this comparison showing me? What is the current law?)		TBD	TBD
Provide access to the application to users	X	TBD	TBD
Training (on-demand webinars and instructor-led courses)	X	X	X

Project Funding

Currently, the project is funded with fiscal year 2020 funds from both the Office of the Clerk and HOLC. As discussed in the initial report, fiscal year 2021 funds were requested and included in

Public Law 116-260, the Consolidated Appropriations Act, 2021. Procurement activities are underway to obligate the FY21 funds.

Related Projects

As mentioned in the initial report about adopting standardized formats for legislative documents transmitted to the Committee on May 11, 2020, this project utilizes documents formatted in XML—both USLM and the first-generation schemas.⁴ The current project transforming Statute Compilations to the USLM XML directly impacts this project. At the time of this report, legislative branch staff continue to perform work on the second phase of the COMPS in USLM project to convert legacy Statute Compilation files into Beta USLM XML and provide access to those files on GPO's govinfo system. This project is on track to be completed in 2021.

Next Steps

The project team will continue working on tasks mentioned in this report as well as those in the previous reports. We are excited about this project and the solution we are building. We see its potential impact on how staff complete their work and understand current law and legislative proposals. We will continue to update the Committee and stakeholders about the progress of this project. We are using a phased-in approach to deployment, and we plan to release a beta-version of the Comparative Print System House-wide at the end of 2021.

⁴ https://cha.house.gov/sites/democrats.cha.house.gov/files/documents/Adopting%20Standardized%20Formats%20for%20Legislative%20Documents.pdf.