Obtener el código HTML de una página web a partir del DOM

PROBLEMA

Si queremos ver el código HTML de una página web podemos hacerlo a través del navegador web, sin embargo el código mostrado es el de la página web tal cual fue escrito. Es decir:

* no se atraviesan frames que contienen otras páginas
* no se formatea el código de modo que esté bien tabulado
* no se muestran los cambios actuales en la página tras la interacción del usuario

Todo esto puede verse sin embargo con las herramientas de inspección de código que ofrecen los navegadores web o a través de extensiones de terceros para los navegadores web. Estas herramientas muestran el DOM de la página web atravesando los frames que contenga, sin embargo hay que ir haciendo clic en cada nodo para ver sus nodos hijo. Por tanto si se quiere ver el código completo de una página web cargada con todos sus frames hay que perder mucho tiempo expandiendo cada uno de los nodos.

SOLUCIÓN

Para resolver esto he creado un código javascript que puede ser ejecutado desde la consola de javascript incluida en las herramientas de inspección de código de los navegadores. La función que actúa de interfaz con el usuario es getHTMLFromDOM(depth, tabString). Esta función hace lo siguiente:

* recibe como entrada la profundidad (a la que se quiere bajar atravesando nodos en el DOM) y una cadena que será usada para tabular (por ejemplo ‘ ‘ para dos espacios).
* recorre el DOM partiendo del nodo correspondiente a la etiqueta html de la página web de partida.
* si un nodo es de tipo elemento mostrará su etiqueta con sus atributos y valores y recorrerá sus nodos hijo. Si dicho elemento es un frame o iframe en vez de recorrer sus nodos hijo recorrerá el nodo de la etiqueta html de la página a la que apunta.
* si un nodo no es de tipo elemento simplemente se mostrará su valor y se recorrerán sus nodos hijo.
* el resultado de todo esto es mostrado en una textarea en una nueva página popup que se abre, de modo que puede ser copiado a un archivo de texto (debes permitir que se abra la pestaña/ventana emergente para poder verlo).

Ejemplo de uso:

Abrir una página web en algún navegador (por ejemplo, abrir facebook con Firefox). Ir a la consola de javascript de las herramientas para desarrolladores del navegador web, pegar todo el código javascript del archivo [getHtmlFromDomJs](https://github.com/Aliuken/getHtmlFromDomJs/blob/master/getHtmlFromDomJs.js).js en la consola javascript y presionar la tecla Intro.

Se abrirá en una nueva pestaña emergente. La nueva pestaña contendrá el código fuente de la página web cargada atravesando sus nodos hasta la profundidad indicada (50 por defecto), entrando en sus frames y tabulándolo de la forma indicada (con dos espacios por defecto).