

Est-ce que je me suis trompé avec mon fichier UTF-8 ?

Afin de pouvoir utiliser Voozadoo avec n'importe quel jeu de caractères, Voozadoo est basé sur la norme internationale UTF-8. Il permet ainsi de créer des systèmes d'information qui affichent fidèlement les textes en français, en chinois, en hindi, en hébreu et en arabe et qui acceptent également des entrées dans ces autres jeux de caractères. Parfois, vous avez besoin de téléverser des fichiers sur Voozadoo. Ce guide a pour but de vous aider à vous assurer que les fichiers que vous téléversez sont encodés en UTF-8.

téléverser des varsets

téléverser des dictionnaires

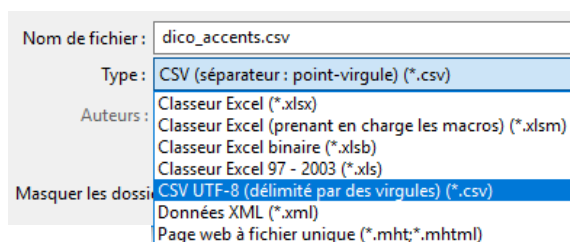
téléverser des utilisateurs

Comment puis-je m'assurer que mon fichier est en UTF-8 ?

Tout d'abord, comprenez que vous ne pouvez pas savoir si votre fichier est UTF-8 en le regardant. Il peut sembler correct jusqu'au moment où vous l'importez dans votre projet Voozadoo, puis, une fois importé, vous remarquerez des caractères étranges à la place de vos caractères accentués (comme le titre en plaisanterie de ce document).

Enregistrer sous

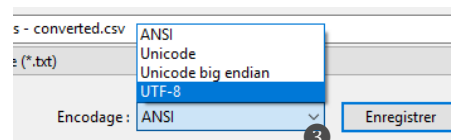
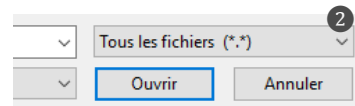
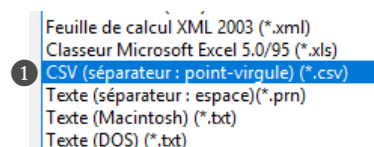
Heureusement, Microsoft Excel (le tableur que beaucoup d'entre nous utilisent pour créer les varsets, dictionnaires et listes d'utilisateurs) a finalement inclus l'option d'exportation UTF-8 en 2017. Cela signifie donc que si vos textes ont l'air bien dans votre tableur, ils seront exportés correctement en UTF-8 lorsque vous utiliserez l'option **CSV UTF-8 (délimité par des virgules)(*.csv)**.



Pas mon vieux Excel !

Tout le monde n'a pas la dernière version de MS Excel. Il se peut que vous utilisiez encore un ancien Excel qui n'a pas l'option d'exportation UTF-8. Ne désespérez pas. Au lieu d'essayer d'afficher la méthode pour chaque version pour obtenir un fichier UTF-8 correct, faites ce qui suit.

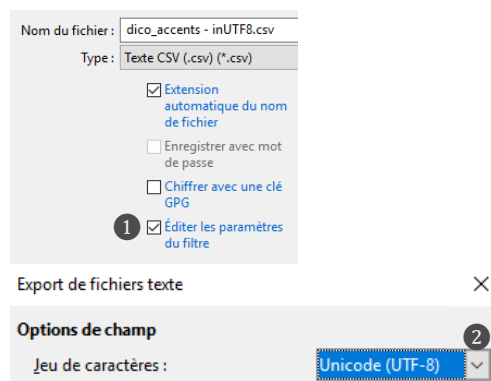
1. Sauvegardez votre feuille Excel au format .csv avec le type **CSV (séparateur : point-virgule) (*.csv)** comme indiqué à droite.
2. Ouvrir Bloc-notes. Ensuite, avec le bloc-note vide, sélectionnez **Ouvrir** et assurez-vous de sélectionner **Tous les fichiers (*.*)** pour pouvoir voir tous les types, pas seulement le type .txt et ouvrir le fichier .csv que vous venez de sauvegarder.
3. Vous verrez le texte de la feuille de calcul. Il y aura beaucoup de points-virgules (ou virgules) séparant les valeurs. Vérifiez que le texte est correct. Sélectionnez maintenant **Enregistrer sous ...** et choisissez l'encodage UTF-8 comme illustré à droite. Voilà, votre fichier .csv est maintenant encodé en UTF-8 et prêt à être téléversé sur Voozadoo.



Mais j'utilise Libre/Open Office

Félicitations. Il fait tout ce que fait MS Excel mais avec des mises à jour gratuites. Un très bon choix !

1. Une fois votre feuille de calcul est prête à être sauvegardée, sélectionnez **Enregistrer sous ...** puis dans la boîte de dialogue, cochez **Editer les paramètres du filtre**.
2. Sauvegarder le fichier. Lorsqu'il est demandé, sélectionnez **Unicode (UTF-8)** dans le menu. Voilà, votre fichier .csv est maintenant encodé en UTF-8 et prêt à être téléversé sur Voozadoo.



Culture générale : En 1963, lorsque la première version du code de caractères ASCII est apparue, il avait les lettres majuscules et minuscules de l'alphabet américain. Il n'y avait pas de caractères accentués. Lorsque les ordinateurs personnels ont commencé à se répandre parmi les chercheurs dans les années 70, l'incompatibilité des caractères accentués a posé un réel problème. Les experts ont donc créé un codage de caractères standardisé pour chacun des principaux types d'alphabet. La norme européenne s'appelait ISO 8859-1, ce qui rendait l'Europe et les Amériques heureuses, et permettait d'échanger des documents avec des caractères accentués. Mais alors l'asiatique, le cyrillique, et une foule d'autres jeux de caractères ont leur propre encodage et n'ont jamais pris la peine d'inclure des lettres latines (enfin certaines versions l'ont fait). Enfin, certains malins ont pensé que nous avions besoin d'un encodage de caractères pour chaque jeu de caractères possible dans le monde. Il s'avère que tout le monde a aimé l'UTF-8 (une variante de l'Unicode) et 93% des sites web l'utilisent maintenant (selon Wikipedia). [fondu sur des utilisateurs d'ordinateur au coucher du soleil]

Is my file correctly encoded in UTF-8 ?

In order to be able to use Voozanoo with any language's character set, Voozanoo is based on the UTF-8 international standard. This way, it can be used to create information systems that faithfully display texts in French, Chinese, Hindi, Hebrew and Arabic as well as accept inputs in these other character sets as well. Sometimes you need to upload files to Voozanoo. This guide is to help you make sure your uploaded files are UTF-8 encoded.

Upload varsets

Upload dictionaries

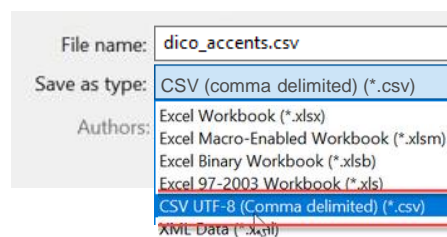
Upload Users

How do I make sure my file is in UTF-8?

First of all, understand that you can't tell if your file is UTF-8 just by looking at it. It can look correct right up to the moment you get it imported to your Voozanoo project and then, once imported, you'll notice some strange characters in place of your accented characters (like the joke-title above).

Save as ...

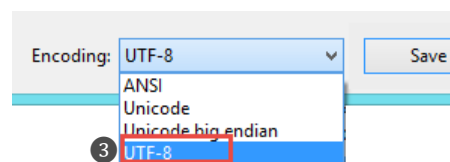
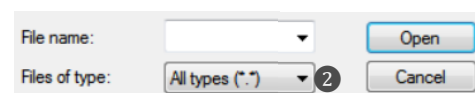
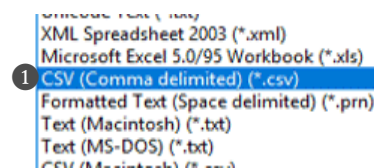
Thankfully, Microsoft Excel (the spreadsheet that many of us use for creating the varsets, dictionaries and user lists) has of 2017 finally included the UTF-8 export option. So that means, if your texts look good in your spreadsheet, then they will export correctly to UTF-8 when you use the option type **CSV UTF-8 (Comma delimited)(*.csv)**.



Not my old clunky Excel!

Not everyone has the latest version of MS Excel. You may still be using an older Excel that doesn't have the UTF-8 export option. Don't despair. Instead of trying to show each version's method of getting a nice UTF-8 file, do the following.

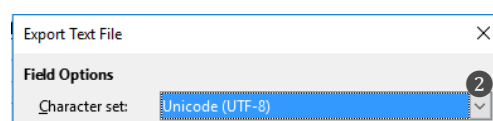
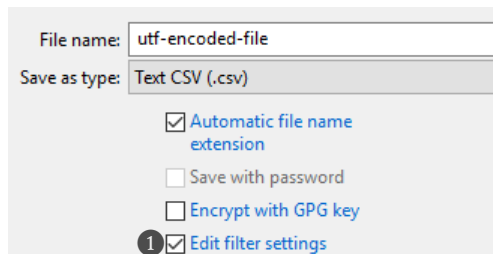
1. Save your finished Excel sheet as a .csv with the type **CSV (comma delimited) (*.csv)** as shown on the right.
2. Create a new Bloc-note file. Then with an empty bloc-note, select **Open** and make sure you select **(all files)** so you can see all types, not just .txt type and open the .csv file that you just saved.
3. You will see the text of the spreadsheet file. It will have lots of commas (or semi-colons) separating the values. Verify the text is OK. Now select **Save as ...** and choose the encoding **UTF-8** as shown to the right. That's it, your .csv file is now encoded in UTF-8 and ready to be uploaded to Voozanoo.



But I use Libre/Open Office

Congratulations. It does everything MS Excel does but with free updates. A very good choice!

1. Once you have your spreadsheet ready to save, select **Save as ...** and then in the dialog box, check **Edit filter settings**.
2. Save your file. When you are asked, select **Unicode (UTF-8)** from the menu. That's it, your .csv file is now encoded in UTF-8 and ready to be uploaded to Voozanoo.



History: Back in 1963 when the first version of the ASCII character code appeared, it had the upper and lowercase letters of the American alphabet. It did not have accented characters. When personal computers started to spread among researchers in the 70s, the incompatibility of accented characters posed a real problem. So experts made a standardised character coding for each of the major alphabet types. The European one was called ISO 8859-1 which made Europe and the Americas happy, and able to exchange documents with accented characters. But then the Asian, Cyrillic, and a host of other character sets has their own encoding and never bothered to include latin letters (well some versions did). Finally, some smart guys figured that we need one character encoding for every possible character set in the world. Turns out that everyone liked UTF-8 (a variation of Unicode) and 93% of websites now use it (so says Wikipedia). [fade to happy computer users in the sunset]