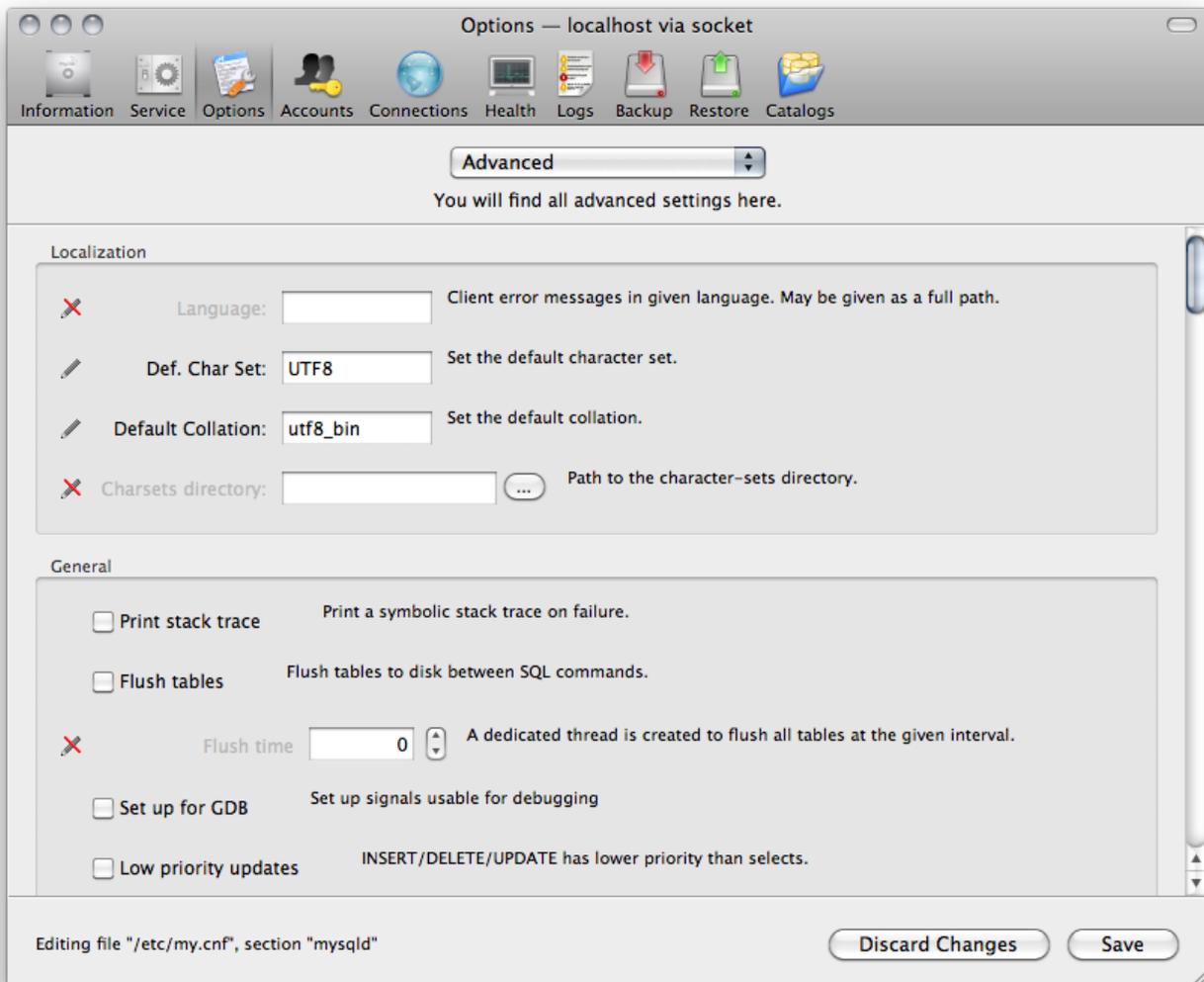# UTF-8 Encoding Tips

UTF-8 Encoding Tips

Encoding questions are asked frequently on the mailing list. This is just a collection of tips for using UTF-8. It's a checklist of sorts. Make sure you've done all the things specified here before pitching your computer into the ocean 🙂

Check your database

The database needs to be storing values in UTF-8. If it isn't, then all your effort is wasted. For example, on MySQL that means a db url like

```
jdbc:mysql://localhost/Example?
capitalizeTypenames=true&zeroDateTimeBehavior=convertToNull&useUnicode=true&characterEncoding=UTF-8
```

And setting your default charset and collation in your my.cnf file

Or you can set these things at table create as below. One can also set these on the database as a whole.

```
mysql> show create table cards \G
*************************** 1. row ***************************
       Table: cards
Create Table: CREATE TABLE `cards` (
  `pk` int NOT NULL,
  `creator_pk` int DEFAULT NULL,
  `created_date` bigint DEFAULT NULL,
  `invalid_date` bigint DEFAULT NULL,
  PRIMARY KEY (`pk`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```
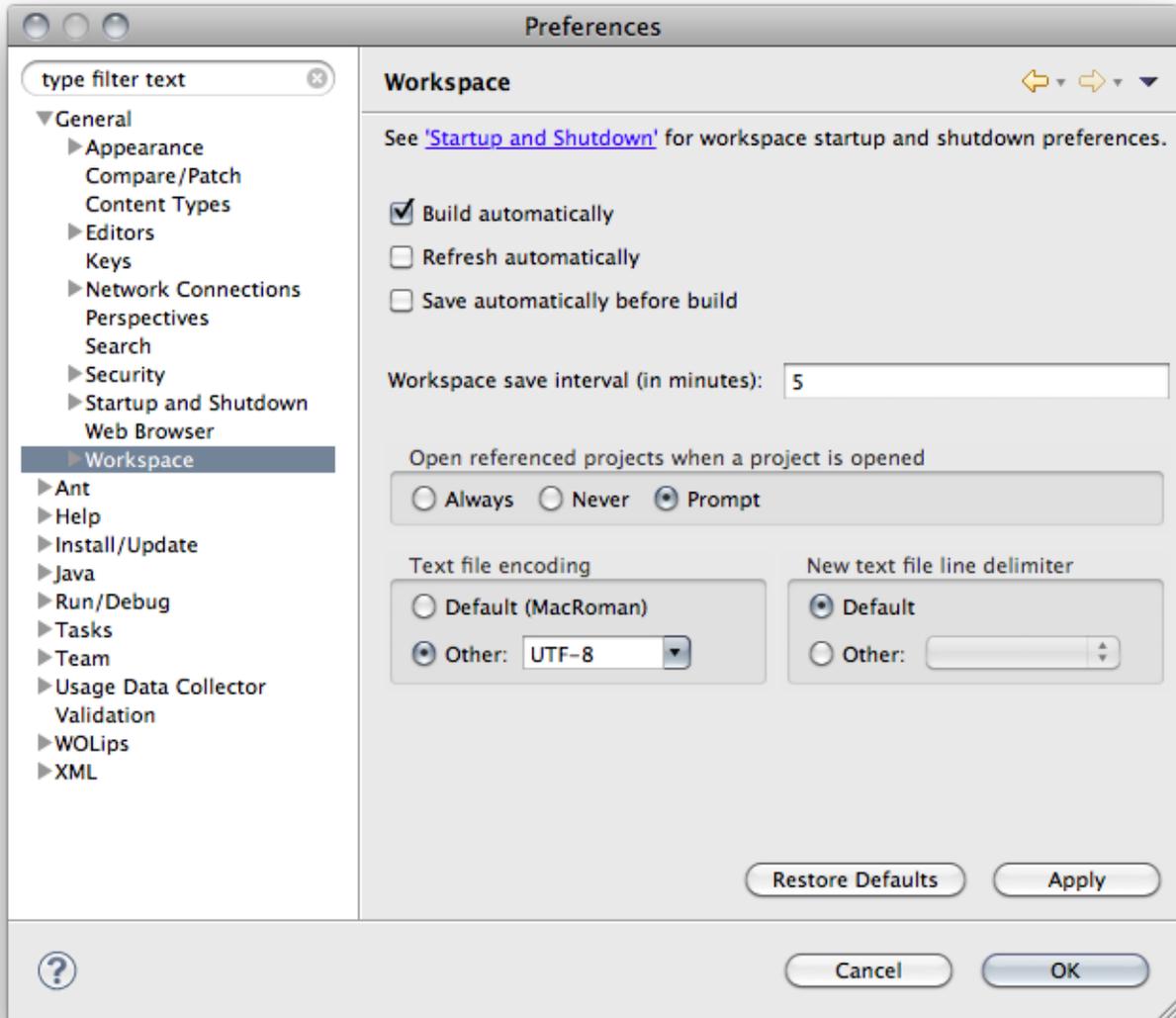
Fonts & CSS

Different fonts may not have all the characters to display the different characters. If you're using a default stylesheet, then the browsers may be displaying differently simply because of fonts. Speaking of stylesheets, you probably want to encode that in UTF-8 also. Start your stylesheet with something like

```
@charset "UTF-8";
@import url("reset.css");

/* Begin site CSS */
```

Set eclipse encoding

Use Project Wonder

I think this goes without saying but: **Use Wonder**. Set encoding in the properties file. Notice it is UTF-8 with a hyphen. It it always UTF-8 with a hyphen... well, except with the MySQL image above because they excel in doing things differently 🙂

```
# Project Encoding
er.extensions.ERXApplication.DefaultEncoding=UTF-8
```

Set encoding in your page wrapper

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE html PUBLIC
    "-//W3C//DTD XHTML 1.1 plus MathML 2.0 plus SVG 1.1//EN"
    "http://www.w3.org/2002/04/xhtml-math-svg/xhtml-math-svg-flat.dtd">
```

Localizable strings should be in UTF-16!

Localizable.strings should be encoded in UTF-16. The localizer can detect UTF-16 without error, where it can confuse UTF-8 with other encodings. Specifically, you should be using UTF-16 BE with no BOM if you are using an external text editor instead of eclipse.

Build your files in UTF-8.

If you have some special characters in your code (like '€' for exemple), then you will need to specify which encoding you want when you build your application. To do that, you have to modify your "build.xml" file by adding the property "encoding="utf-8" into your <wocompile> statement.

```
<wocompile srcdir="Sources" destdir="bin" encoding="utf-8">
```