

先進ゲノム支援講習会 事前準備資料

本資料では先進ゲノム支援講習会の事前準備として、

- 遺伝研スパコンのアカウント申請と接続
- 必要なソフトウェアのインストール方法

の手順を説明します。

スパコンアカウントの発行には通常約2週間程度かかるので、早めの準備を推奨します。

すでにスパコンアカウントをお持ちの方は新規にアカウントを申請する必要はありません。

本講習会はおもに初めてスパコンを使用する方を対象としていますので、既にスパコンを利用中の方は既存の解析環境と異なる可能性があることをご了承ください。

講習会で必要なソフトウェアは下記の3点です。

- ターミナルエミュレータ (WindowsではPowerShell、Macではターミナルappを利用)
- テキストエディタ (WindowsではTeraPad、Macではmiを利用します。すでに使い慣れたものがあればそれらを利用しても構いません)
- SFTPソフト (ファイルの転送に用いる。FileZillaを利用)

スパコンログイン後、Minicondaを使って必要なソフトウェアのインストールを行います (手順は講習会の中で説明を行います)。

1. スパコンアカウント申請および接続

遺伝研スパコンへはSSH秘密鍵・公開鍵を使った認証方法で接続を行う。秘密鍵・公開鍵はスパコンへの接続に使う端末 (Mac, Win, Linux) 上であらかじめ作成し、スパコンアカウント申請時に公開鍵の登録を行う。スパコンへの接続時には、端末上の秘密鍵とスパコン上に登録された公開鍵とを照合させることによってユーザー認証を行なっている (公開鍵認証)。

遺伝研スパコンの公式ウェブサイトの情報も参考にすること。

参考) 遺伝研スパコンの公式ウェブサイトの[利用の準備](#)および[ログイン方法](#)のページ

1.1 SSH秘密鍵・公開鍵の作成

1.1.1 ターミナルエミュレータを立ち上げる

(Mac) 標準のターミナルappを使用する。Windowsの場合にはPowerShellを用いる。以下、いずれの場合も”ターミナル”と呼ぶ。

(Windows) 最近のWindows OSであればWindows PowerShellの標準機能のみで接続が可能。少し古いOSでは接続に必要な**OpenSSHクライアント**が含まれていないことがある。その場合にはOS設定の「オプション機能の追加」でOpenSSHクライアントをインストールする必要がある。OpenSSHがインストールされているかは、`ssh` とコマンドを打って次のように使い方が表示されているかで確認できる。(参考: [OpenSSHのインストール手順](#))

```
PS C:\Users\tanizawa> ssh
usage: ssh [-46AaCfGgKkMNnqsTtVvXxYy] [-B bind_interface]
          [-b bind_address] [-c cipher_spec] [-D [bind_address:]port]
          [-E log_file] [-e escape_char] [-F configfile] [-I pkcs11]
          [-i identity_file] [-J [user@]host[:port]] [-L address]
          [-l login_name] [-m mac_spec] [-O ctl_cmd] [-o option] [-p port]
          [-Q query_option] [-R address] [-S ctl_path] [-W host:port]
          [-w local_tun[:remote_tun]] destination [command]
```

1.1.2 SSH秘密鍵・公開鍵の作成

ターミナルで下記を実行する

```
ssh-keygen
```

windowsでは `ssh-keygen.exe` と打っても良い。

以下のように画面に表示されるので指示にしたがって進める。

```
Generating public/private rsa key pair.  
Enter file in which to save the key (/Users/XXXX/.ssh/id_rsa):           (1)  
Enter passphrase (empty for no passphrase):                             (2)  
Enter same passphrase again:                                             (3)  
Your identification has been saved in /Users/XXXX/.ssh/id_rsa.  
Your public key has been saved in /Users/XXXX/.ssh/id_rsa.pub.  
The key fingerprint is:  
e5:23:f0:fc:b7:60:70:80:79:91:f2:f1:6d:a8:ae:90 xxxx@yyyyy
```

(1) では鍵の保存先を指定する。通常はホームディレクトリの下に `.ssh` という隠しディレクトリが作られ、その中に `id_rsa` という名前で秘密鍵が、`id_rsa.pub` という名前で公開鍵が作成される。保存先は通常はデフォルトのままで構わない。

注) Windowsを使っていてユーザー名にスペースや日本語の文字 (全角文字) が含まれていると鍵の場所が正しく認識されないことがある。(別の場所に鍵を作ることもできるがファイルの読み取り権限の設定が必要になることもあるので、うまくいかない場合には、英数字のみのユーザー名を作った方が無難)

(1) で指定した保存先に過去に生成された鍵が存在している場合には上書きするか確認される。その場合には、既存の鍵を使い回すことができるので上書きせずに中断する。

(2) で秘密鍵に設定するパスフレーズを指定し、(3) で確認のためパスフレーズを再入力する。パスフレーズについては秘密鍵を適切に管理されていれば (ネットワーク経由でアクセスできる場所においたりメールで送付したりしなければ)、空文字を設定することもできる。ただし、秘密鍵の盗難時の被害を軽減するためには設定しておくことが推奨される。

注) パスフレーズはスパコンアカウントのパスワードとは別のものなので注意すること。

注) スパコン公式ページでは鍵生成時に `ssh-keygen -t rsa -b 2048` というようにオプションをつけて明示的に鍵の作成方法を指定しているが、オプションなしでも構わ

ない。

1.1.3 作成した公開鍵の確認

次のコマンドで作成した公開鍵を表示させる

```
cat ~/.ssh/id_rsa.pub
```

なお、`~` はホームディレクトリを表す。先の手順でデフォルトとは別の場所に鍵を作成した場合は、上記の `~/.ssh/id_rsa.pub` の部分は差し替える必要がある。

実行結果は次のようになる。画面では改行されているように見えるが、実際には改行なしの1行になっている。

```
$ cat ~/.ssh/id_rsa.pub  
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQgQDAyZHZxx... 以下省略 ... xxxx@yyyyy
```

スパコンアカウント申請時に公開鍵を記入する欄があるので上で表示された鍵全体をコピーして貼り付ける。

1.2 スパコンアカウントの申請

利用登録申請フォームに必要事項を記入して申請を行う。記入時の諸注意はこちらの利用案内を参照すること。

アカウントを申請すると、責任者あて（申請者あてではない）に誓約書提出依頼メールが届く。誓約書の提出がないとアカウント発行はできない。

誓約書の提出後にアカウント発行がされるが、アカウント情報は郵送にて発送されるため時間がかかる。

講習会用のアカウントを作成するにあたっては特に以下の点に注意すること

- 講習会参加のみを目的としたアカウント申請は不可

講習会参加のため」という記載ではなく、「先進ゲノム支援情報解析講習会に参加後、〇〇の解析に使用するため」といった内容で記載する (150文字以上)

- 申請者の所属・住所については、研究室名まで含めた所属機関の住所を記載すること

アカウント証は記載された所属研究室宛てに郵送にて送付されるので確実に受け取れる住所を書くこと。また、住所は信頼できる機関に実際に所属することを確認するためのものであるため、自宅への郵送や講習会当日の手渡しは不可。(在宅勤務者や実家に帰省中の学生であっても自宅や帰省先の住所ではなく、所属先の住所を書くこと)

- メールアドレスについてもセキュリティ上、なるべく所属機関のアドレスでの登録をすること。
- 責任者欄には、所属研究室の責任者の情報を書くこと。ただし、教員に関しては利用者と責任者が同一人物でも可。

1.3 スパコンへの接続

ターミナルを開き

```
ssh ユーザーアカウント名@gw.ddbj.nig.ac.jp
```

と入力する。秘密鍵にパスフレーズを設定している場合には入力を促される。

初めて接続する場合には続行するか聞かれるのでその場合には **yes** と入力して先に進む。

秘密鍵をデフォルトの保存場所とは異なる場所や、別の名前で作っている場合には、下記のように **-i** オプションで秘密鍵の場所を明示的に指定する。

```
ssh -i ~/.ssh/id_rsa ytanizaw@gw.ddbj.nig.ac.jp
```

ログインに成功すると下記のように表示される。

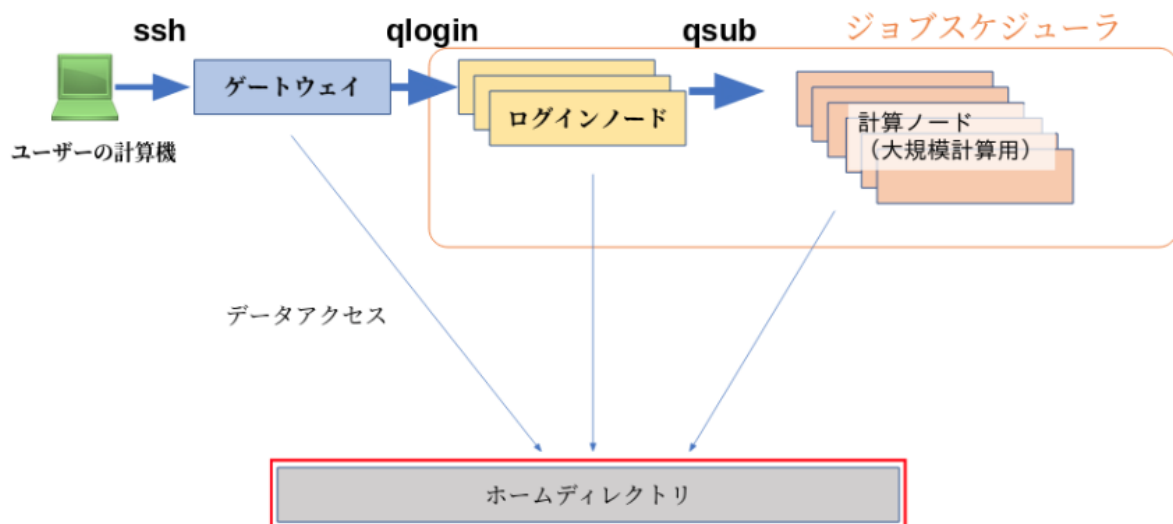
```
-----  
Thank you for using NIG supercomputer system.
```

```
This is the gateway node, do not run program here.  
Please use 'qlogin' to login to a login node.  
-----
```

メッセージに表示された通り、ログインした場所はスパコンの **gateway** と呼ばれるノード (マシン) である。これはログインやファイルの転送のために用いられるマシンであるので、負荷のかかる作業を行うことは認められていない。スパコン上で作業を行う場合には

```
qlogin
```

と打ってログインノードに移動してから操作を継続する。



ログアウトするには

```
exit
```

を実行する。(qloginした状態で **exit** するとゲートウェイノードに戻り、さらにその状態で **exit** するとスパコンからログアウトされる)

TIPS 他の端末からスパコンに接続するには

次の2つの方法がある

- 他の端末で新しく秘密鍵・公開鍵のペアを作成し、スパコンウェブサイトにご登録し公開鍵を追加登録する*1 *2。
- 作成した秘密鍵を他の端末にコピーして用いる。

*1 登録された公開鍵はスパコン上の `~/.ssh/authorized_keys` というファイルに記載される。このファイルを直接自分で編集して新たな公開鍵を追記してもよい。

*2 2022/9現在、ウェブサイトで設定した公開鍵が反映されるのに時間がかかる場合がある (参考: [申請システム注意事項](#))

2. テキストエディタのインストールと初期設定

遺伝研スパコンを初めとするLinuxシステムで動くサーバーでは文字コードUTF-8、改行コードLF形式が使用されている。

Macではデフォルトでこの形式がサポートされているので、標準のテキストエディタ (テキストエディットapp) をそのまま使用することができる。

一方、Windows標準のテキストエディタ (メモ帳) はUTF-8/LF形式の既存のファイルの閲覧や編集には対応しているが、新規でファイルを作成した場合には改行コードCRLF形式のファイルが作成されてしまう (2022/9現在では改行コードの指定は不可)。



Windowsのメモ帳で新規にファイルを作成した場合 (左) と、既存のUTF-8/LF形式のファイルを開いた場合 (右)

Mac/Windowsいずれの場合も標準のテキストエディタを使用することは可能だが、高機能なテキストエディタを使用することをお勧めする。

ここではWindowsではTeraPad、Macではmiを導入する。すでに使い慣れたテキストエディタがあればこのステップは省略して良い。

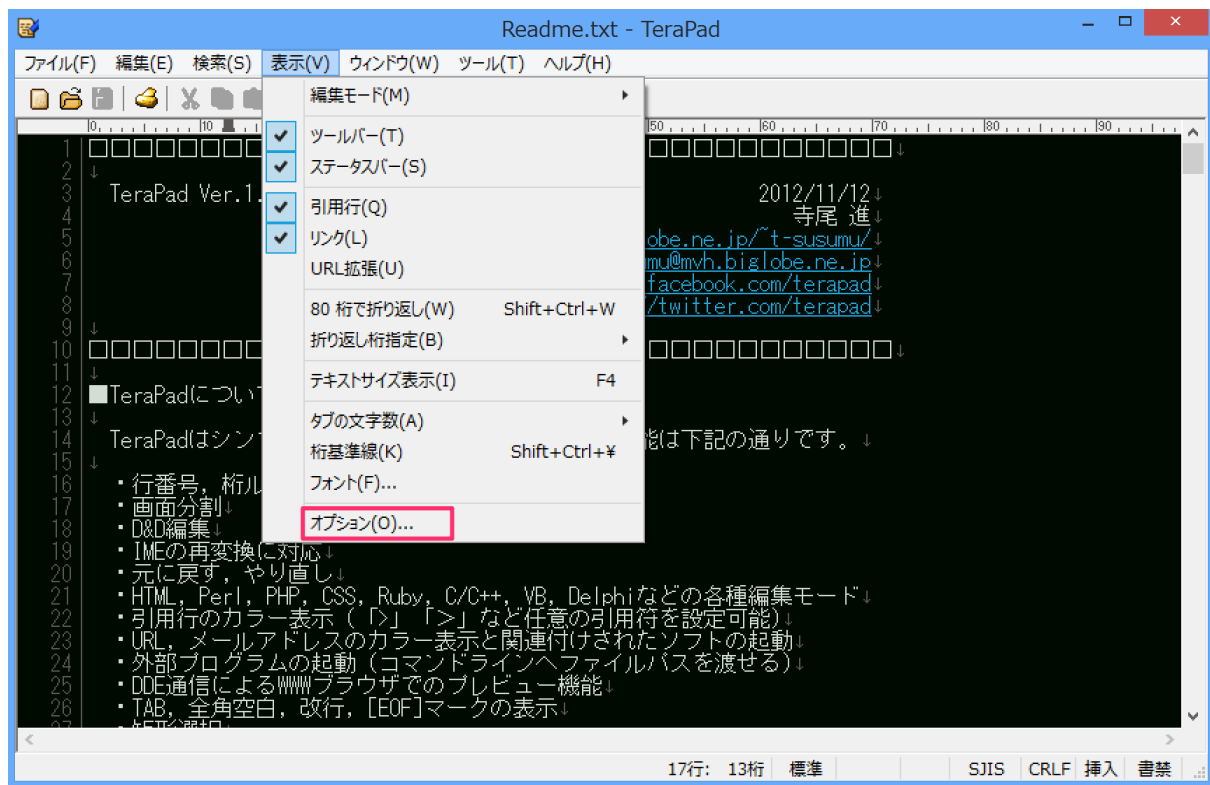
2.1 TeraPadのインストール (Windows用)

TeraPad 公式サイト (tera-net.com) からインストールを行う。インストーラーを直接ダウンロードする場合は[ここ](#)をクリック。

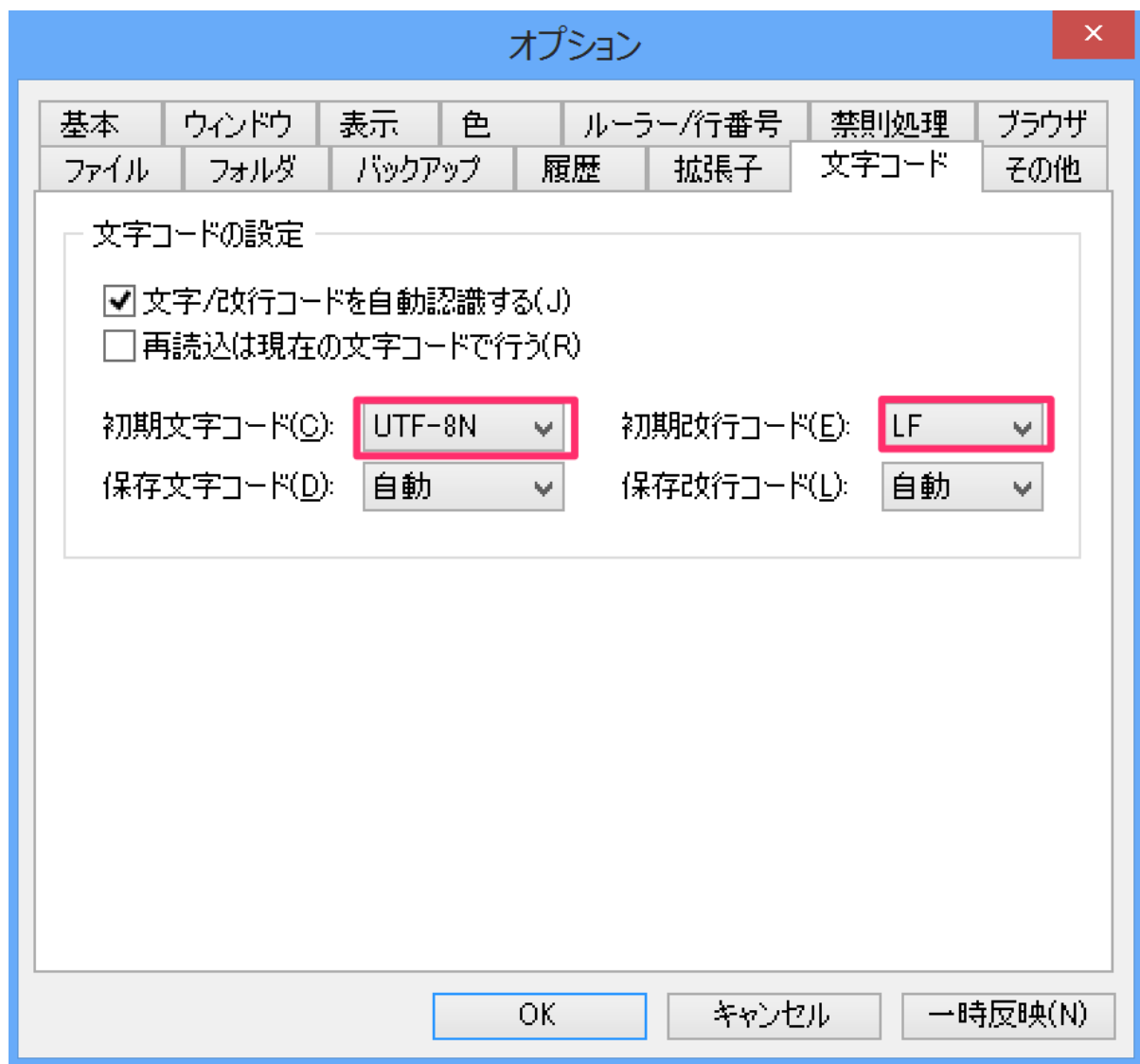
指示に従ってインストールを行う。

インストール後、文字コード・改行コードに関して初期設定を行なっておくが良い。

2.1.1 表示 → オプションを選択



2.1.2 文字コードオプションの設定



初期文字コード **UTF-8N** と初期改行コード **LF** にし、OKボタンを押す。

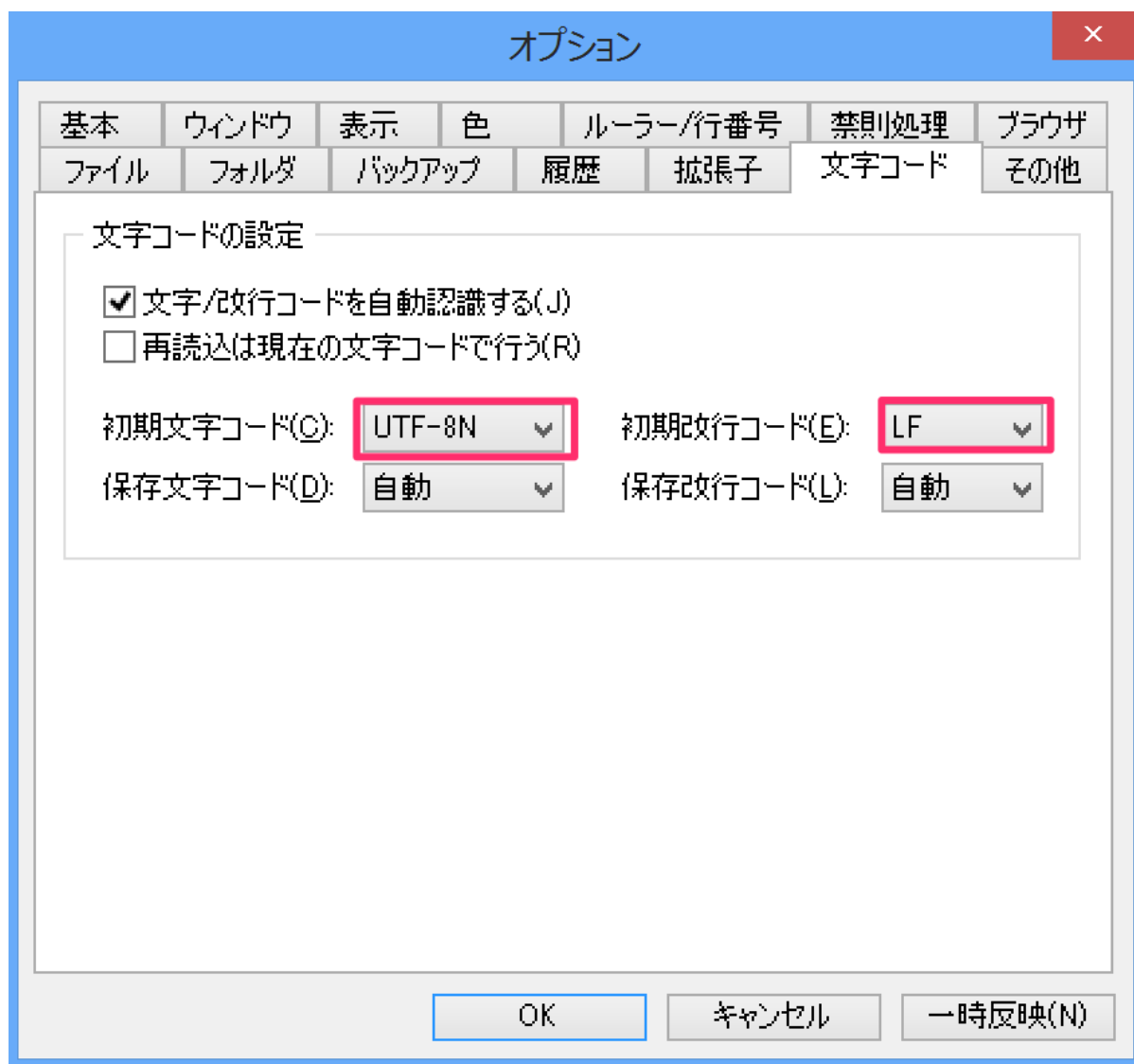
注意)

TeraPadの場合、**UTF-8** は選ばない事（ファイル先頭にBOM *1が埋め込まれてしまう）

他のテキストエディタを使用する場合にも、文字コードはBOMなしのUTF-8を指定する (BOMなしのUTF-8の方が一般的なので、単に **UTF-8** と書かれていたら通常はBOMを含まない)。

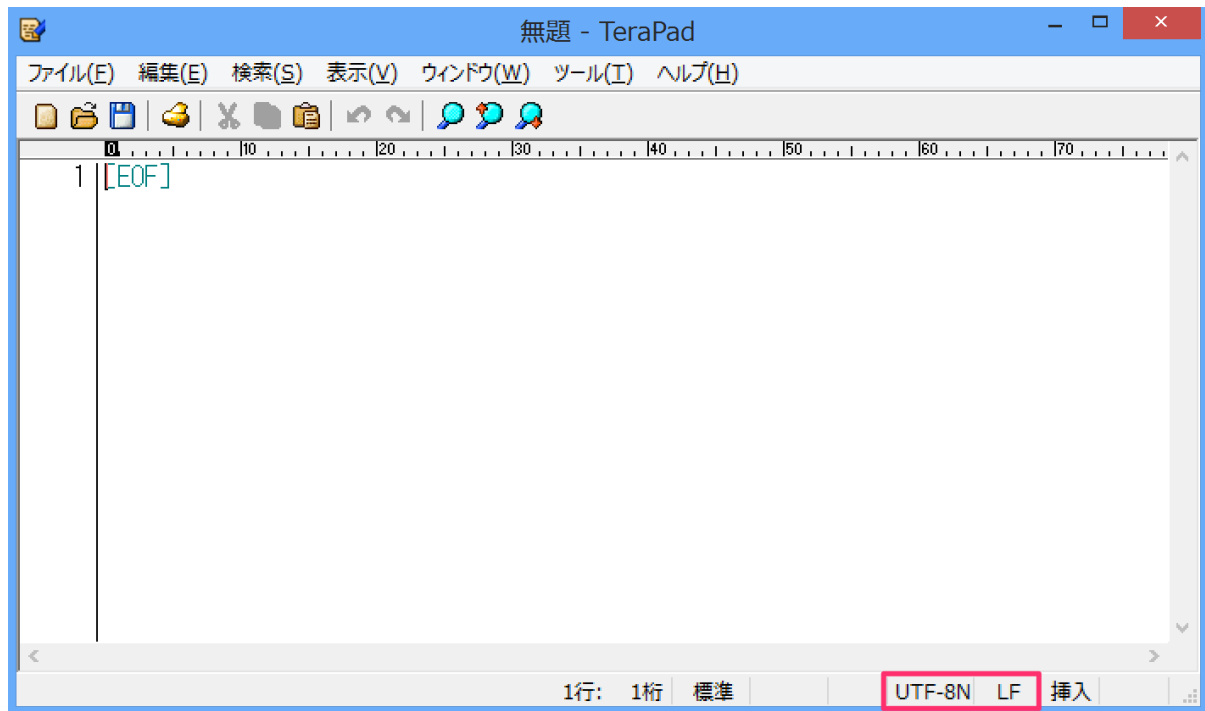
保存文字コードと保存改行コードも同じくUTF-8NとLFにしておいても構わない。
(設定した場合には別のソフトで作成したファイルも、TeraPadで開いて編集すると保存時に自動で変換される。)

*1 BOM : 文字コードを識別するための符号のようなもの (参考)。



2.1.3 新規ファイルを作成して確認

UTF-8N と LF になっていればOK。



2.2 mi のインストール (mac用)

公式ウェブサイトからダウンロードしてインストールする。特に設定は不要。

3 FileZilla のインストールと初期設定

本講習会ではSFTPソフトとして、Windows/Mac版の双方が利用できるFileZillaをインストールする。FileZilla以外のSFTPソフトウェアの例としては、

- Windows用: WinSCP
- Mac用: CyberDuck, Transmit

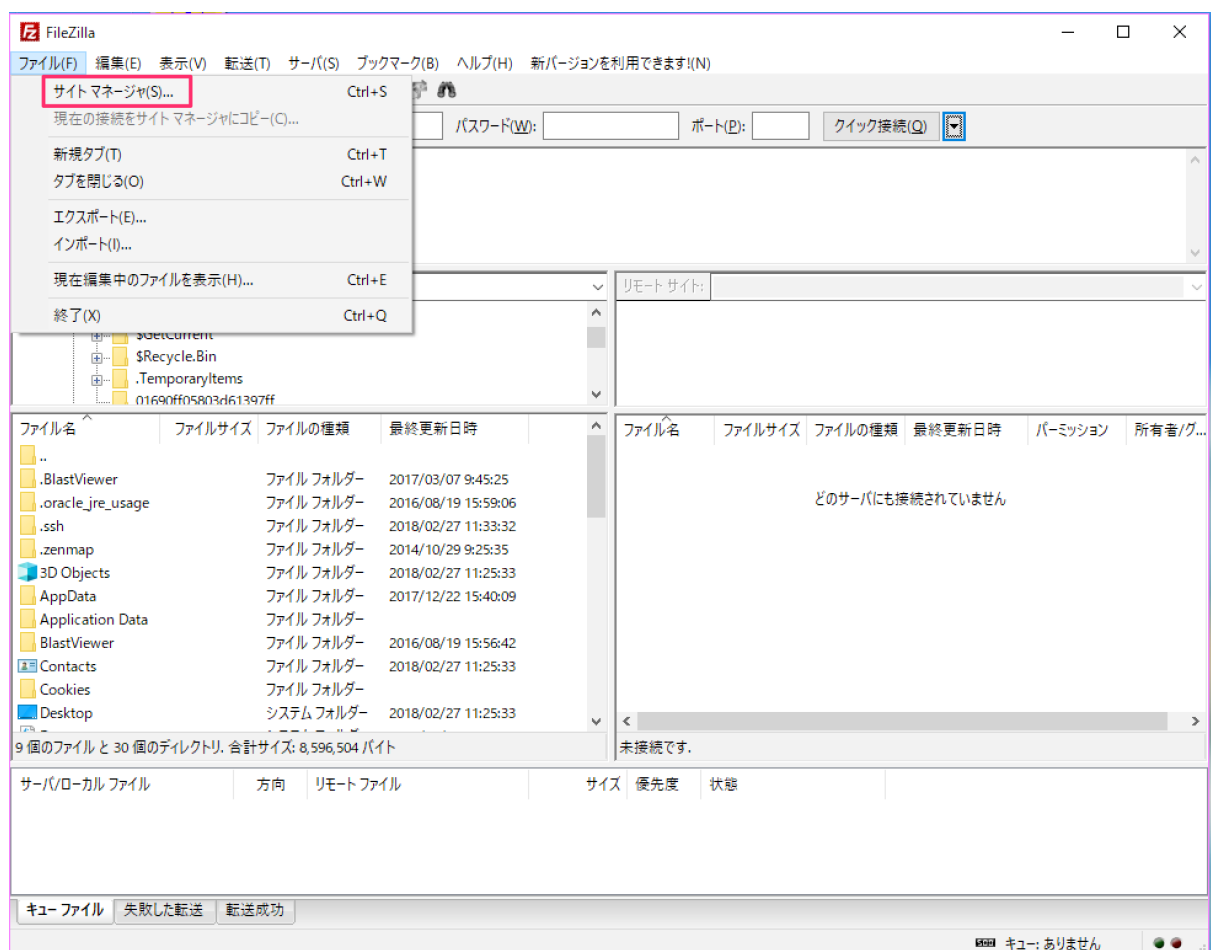
などがある。

3.1 FileZilla をインストール

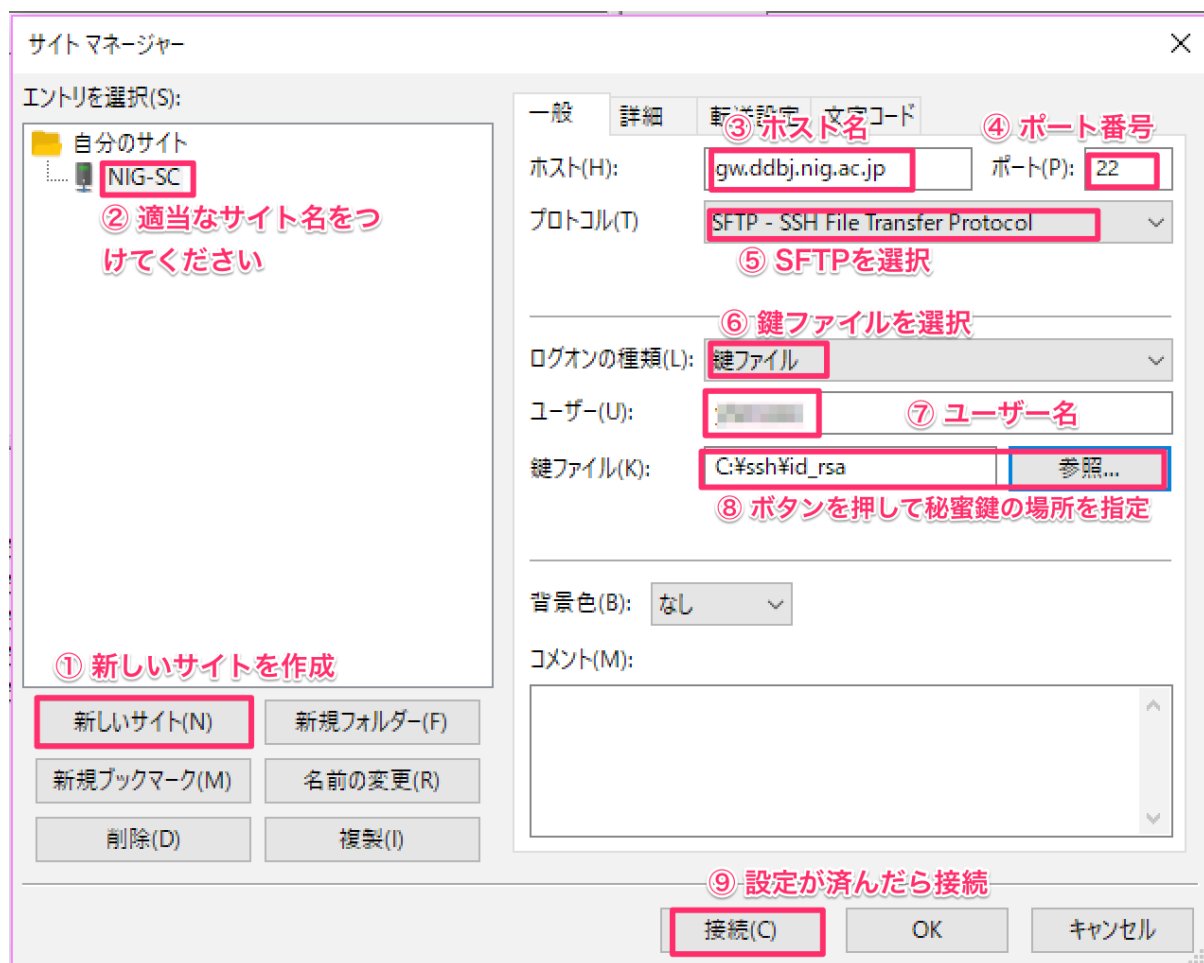
公式ページからFileZilla Clientのインストーラーをダウンロードし、指示に従ってインストールを行う。

3.2 FileZilla を使ってスパコンへ接続

FileZillaを起動し、ファイル→サイトマネージャを開く。(メニューアイコンから開いても可)



下記参考に新規の接続設定を作成



⑧の鍵ファイルの場所指定を行う際に、デフォルトの保存場所は不可視フォルダ（.ssh）になっていることに注意。また、windows版の場合、表示対象ファイルの拡張子を”すべてのファイル”にしないとファイルが表示されない(補足1参照)。

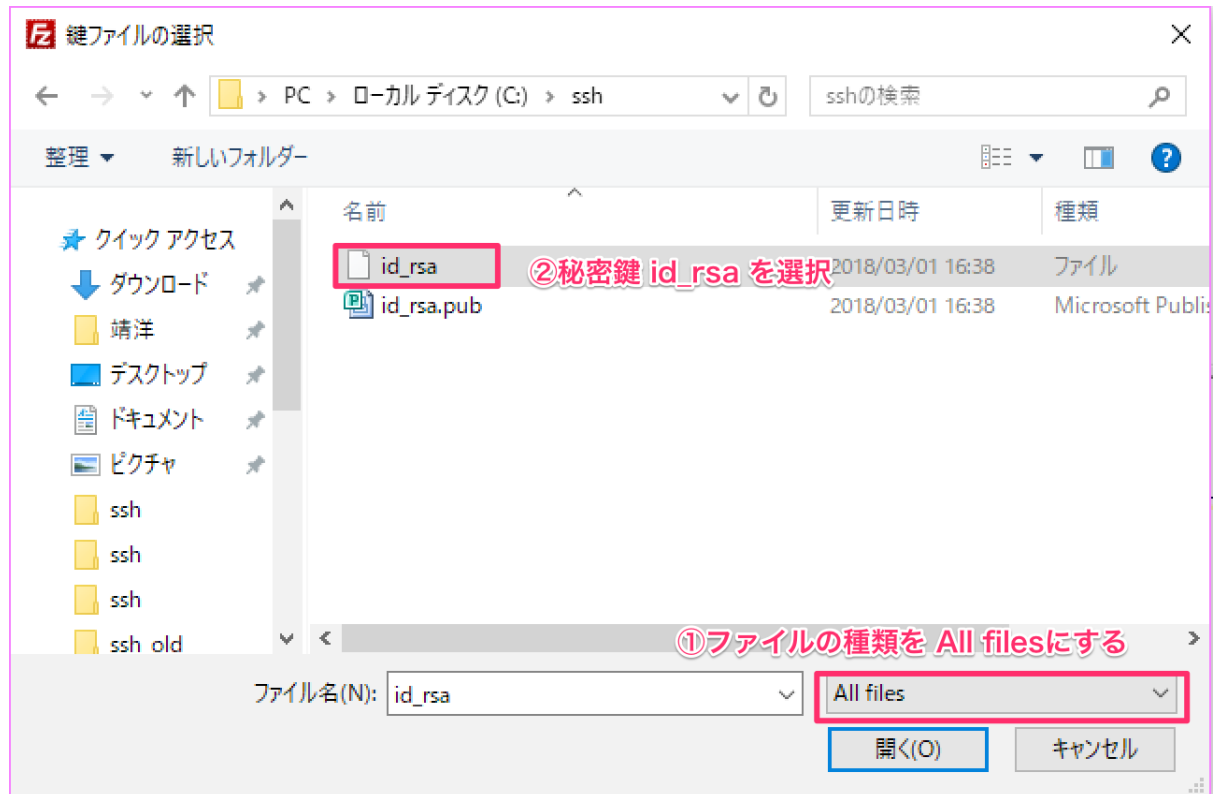
ファイルの指定がうまくいかない場合には、windows版の場合には `C:\Users\%username%\ssh\id_rsa`、Mac版の場合には `/Users/xxxxx/.ssh/id_rsa` というように直接手入力で打った方が早い。

秘密鍵にパスフレーズを設定している場合、FileZillaがサポートしている形式ではないという警告がでる場合がある(補足2参照)。その場合には、指示に従って変換した鍵を指定する。

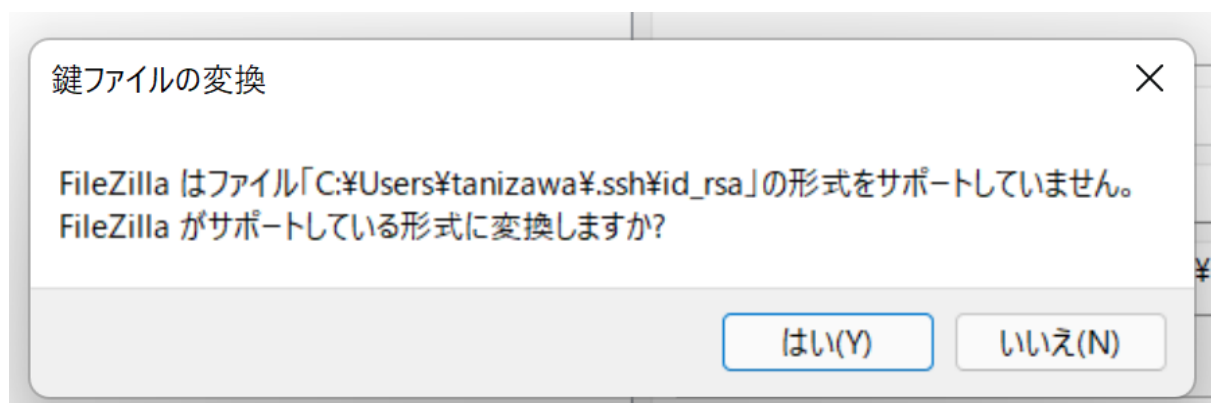
秘密鍵にパスフレーズを設定している場合、⑨を押した後にパスフレーズの入力を促される。

初めて接続する場合、”不明なホスト鍵”に関して警告が出る場合があるが、無視して構わない。

補足1)



補足2)



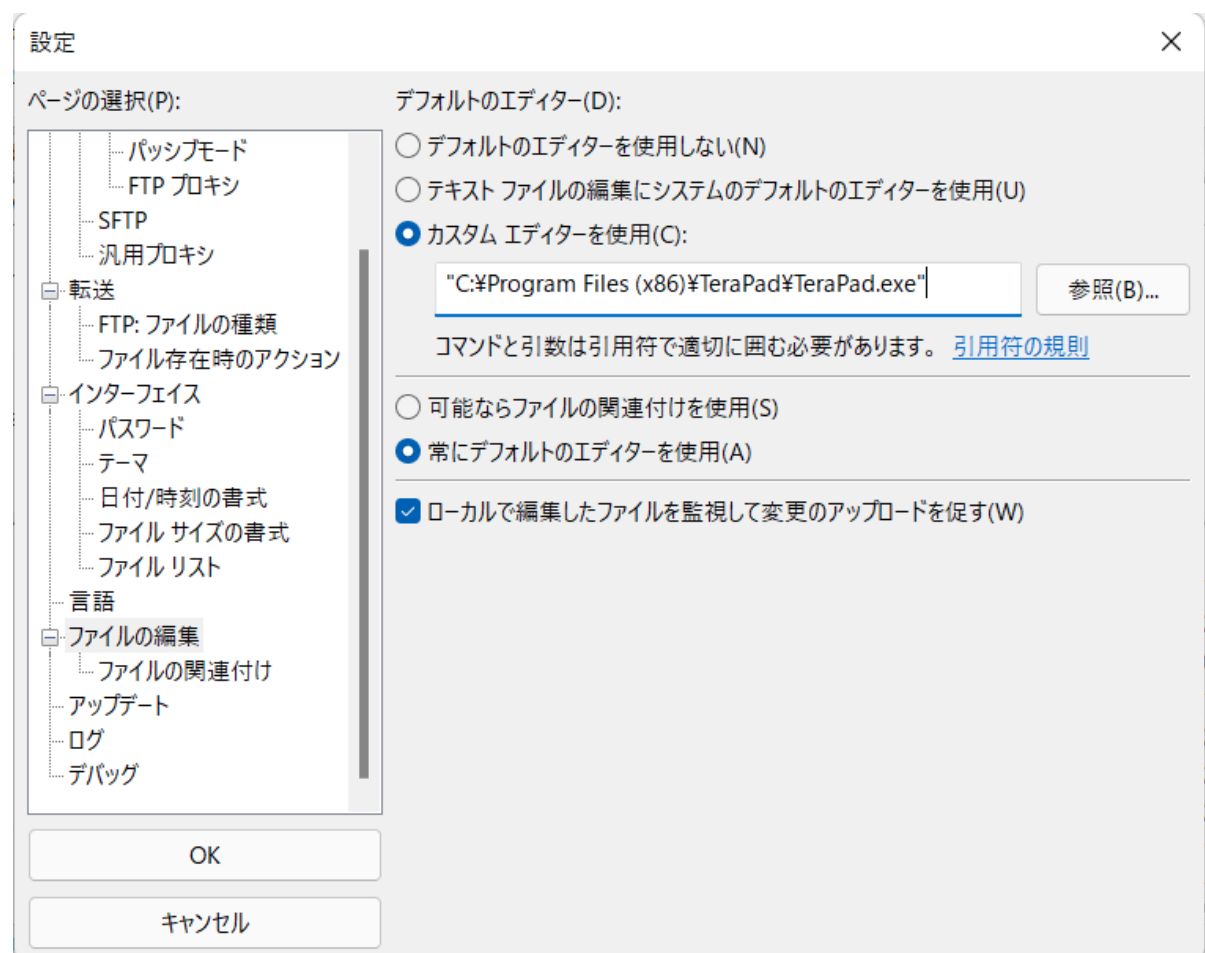
という警告が出た場合には指示にしたがって **.ppk** 形式の鍵ファイルに変換し、それを指定する。

3.3 デフォルトテキストエディタの指定

FileZillaを使うとサーバー上のファイルを直接開いて編集することができる。この時に使用されるテキストエディタを指定する。

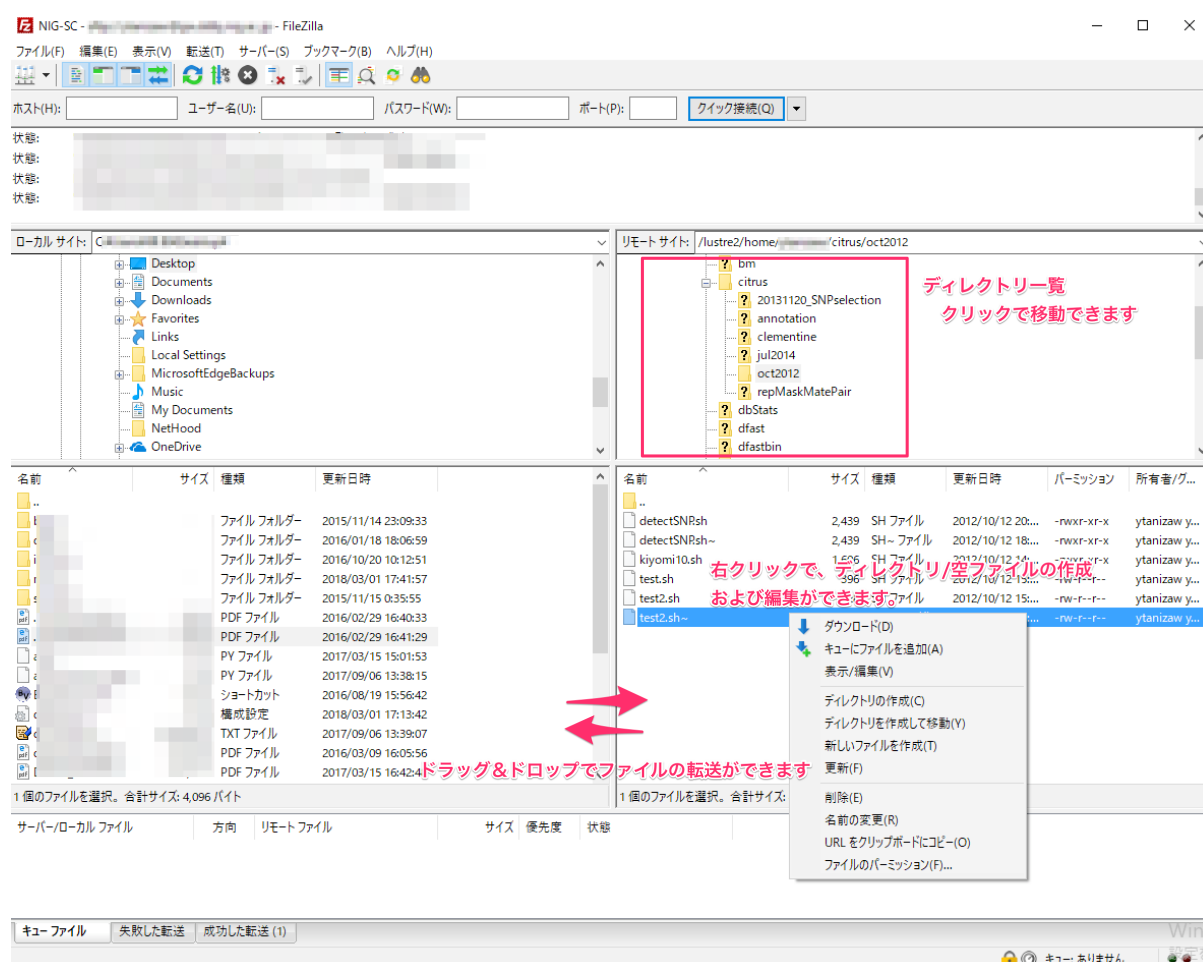
メニューの編集→設定を開いて、「ファイルの編集」のところで、

- **カスタムエディターを使用** を選び、TeraPad (winの場合)、mi (macの場合) を指定
- **常にデフォルトのエディターを使用** を選ぶ
- **ローカルで編集したファイルを監視して変更のアップロードを促す** にチェックを入れる (デフォルト状態でチェックされている)



3.4 FileZilla 基本的な使い方

左側にローカルファイル (自端末のファイル)、右側にリモートファイル (スパコンのファイル)が表示される。ドラッグアンドドロップでのファイルの転送や、右クリックメニューでディレクトリや新規ファイルの作成ができる。



適当なファイルを作成したのち、そのファイル上で右クリックメニューから”表示/編集”を選ぶと3.3で指定したテキストエディタが立ち上がることを確認する。

テキストエディタでファイルを編集して保存すると変更したファイルをサーバーにアップロードする確認のメッセージが表示される。

事前準備は以上です。

講習会では実際にスパコンにログインして操作を行います。基本的なLinuxのコマンドの使い方も説明しますが、事前に慣れておくことをおすすめしますので余力があれば下記のようなサイトで予習を行なってください。

ファイル・ディレクトリの操作コマンド (ls, pwd, cd, rm, mkdir, touch等) が使えれば今回の講習では十分です。

- Linux「基本」コマンド Quickリファレンス
- 初心者向け！基本的なLinuxコマンドと使い方！
- 【Linux入門】初心者が覚えておくべきコマンド28選+役立つコマンド5選