
Outils informatiques de gestion de projet

Archivage et partage des fichiers avec GIT

Introduction

Objectifs de ce cours

Pour un projet reposant sur un ensemble de fichiers (par exemple, du code, des images, des documents de travail etc ...), il est inévitable de se retrouver avec un certain nombres de copies (partielles ou totales, modifiées à des degrés divers) d'un même ensemble. Une des difficultés dans la gestion du projet est de manipuler ces copies avec suffisamment de rigueur pour ne pas perdre d'informations. Cela peut être réalisé manuellement, en créant des archives facilement identifiables (idéalement, en précisant au moins pour chaque version la date et l'auteur.).

La gestion manuelle d'un tel ensemble d'archive est très lourd, même pour des opérations relativement simples, comme le partage de la dernière version, l'intégration du travail des membres d'un groupe, la recherche d'un fichier d'une des versions précédentes etc..

L'objectif de ce cours est d'apprendre à utiliser un outil dédié à la gestion des fichiers d'un projet, et qui permet de simplifier toutes ces tâches.

Présentation générale de GIT

GIT est un système distribué de gestion de versions. Un système de gestion de version est un outil pour gérer les versions d'un projet, c'est à dire les archives des différents fichiers. On dit que le système est distribué (par opposition à un système centralisé) quand chaque dépôt local (c'est à dire chaque utilisateur) a l'intégralité des archives disponibles (localement, donc). Un tel système permet aussi de synchroniser les différents dépôts, c'est à dire de partager les différentes versions, et de combiner les modifications de plusieurs utilisateurs (ce qui peut être fait automatiquement, dans une certaine mesure).

La bonne utilisation de GIT vous permettra entre autre :

- D'assurer la sauvegarde des données de votre projet (sous forme d'archives, que l'on peut explorer selon de nombreux critères comme la date où est fichiers modifiés).
- D'éviter la perte d'archive : chaque utilisateur a une copie de toutes les archives, ce qui rend très improbable la perte totale des données (il faudrait que chaque membre du groupe perde ses données ...). En pratique, seules les données non partagées sont vulnérables.
- De partager les modifications apportées par chaque membre au projet. Les méthodes de synchronisation sont nombreuses (principalement réseau (serveur spécialisé, http et ssh entre autre), par clef usb ou échange de fichiers sur une même machine).
- D'intégrer, souvent de manière automatique, les différentes modifications. Autrement dit, à partir de plusieurs versions d'un même projet, il est possible de produire une version combinant les différentes modifications.
- De faciliter le déploiement du projet, en mettant facilement à jour les fichiers sur la machine de test (dans le cas d'un site web, par exemple).

- De consulter facilement les archives du projet, par exemple pour localiser un problème, ou pour trouver l'auteur d'une modification.

En revanche, l'utilisation de git ne résout pas en soi les problèmes liés à l'organisation d'un groupe autour d'un même projet ; comme tout outil de gestion de projet, il rend juste plus facile les tâches nécessaires au bon avancement du projet. Git doit donc être intégré dans vos processus de travail, il ne peut en aucun cas se substituer à une bonne organisation du projet (distribution des rôles et des tâches par exemple).

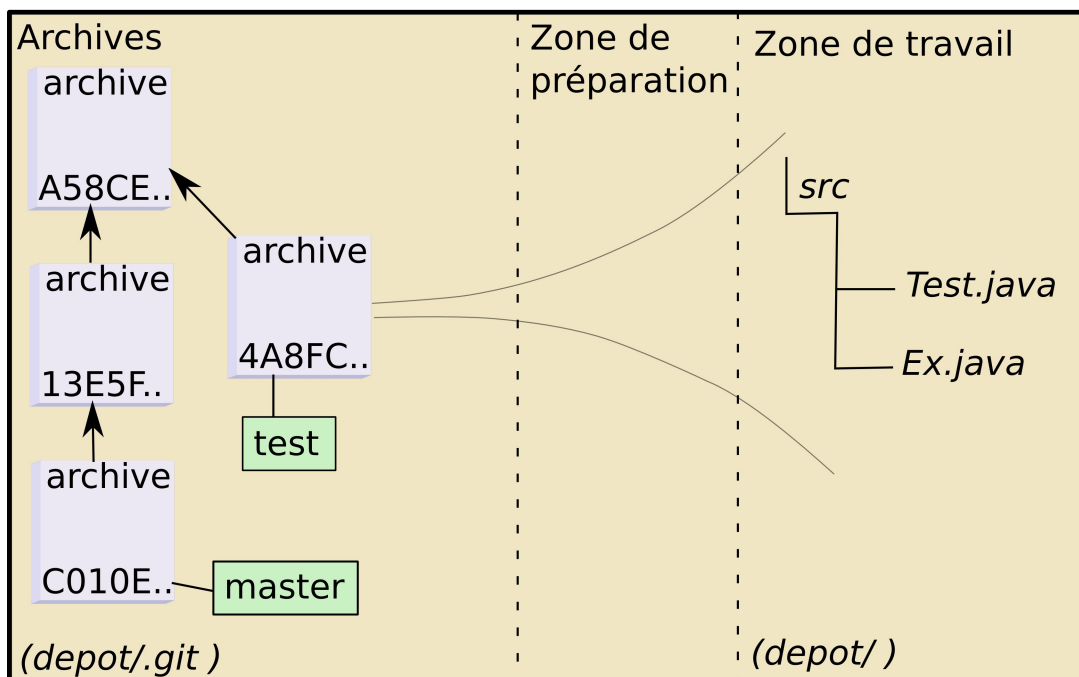
Les archives

L'élément de base géré par git est l'archive, qui représente l'ensemble des fichiers du projet à un moment donné, sur une machine donnée. On parle aussi d'instantané (au sens photographique du terme), ou en anglais de snapshot. Ce sont les archives qui sont sauvegardées, échangées, fusionnées etc.. Chaque archive se comporte comme une version compressée (même si la réalité est plus complexe – et plus efficace) du répertoire principal du projet, celui contenant tous les fichiers. Chaque archive contient les fichiers du projet, ainsi que les informations caractéristiques de la version (texte descriptif, archive(s) d'origine, date , auteur ..), ainsi qu'un identifiant unique, sous forme d'un code en hexadécimal, et calculé à partir de l'ensemble des informations de l'archive.

Ajouter la version actuelle des fichier en tant que nouvelle archive ne modifie en rien les archives existante, il n'y a donc aucun risque de perte de données en archivant (donc, faites-le souvent...).

Les entrepôts

Entrepôt (*depot*)



Ces archives sont regroupées dans des entrepôts (repository en anglais). Chaque membre du projet travaille sur (au moins) un de ces dépôts, local à sa machine. La plupart des tâches peuvent être effectuées sur ce dépôt, sans nécessiter aucun accès aux autres dépôts. Internet n'est donc pas indispensable pour travailler avec git. Même pour partager les informations, il est possible de s'en passer (clefs usb, connexion directe d'ordinateurs à ordinateur etc...)

Chaque entrepôt occupe un répertoire (par exemple depot), qui lui-même comporte plusieurs zones. Le répertoire depot/.git contient les archives, la configuration et les autres informations liées au fonctionnement de git. Le répertoire depot/ abrite la zone de travail, c'est à dire les répertoires et fichiers du projets sur lesquels le travail se porte à un moment donnés. Ces fichiers sont disponibles, utilisables et modifiables normalement et sans restriction ni outil particulier.

Les branches

Pour chaque nouvelle archive ajoutée au dépôt, git note l'archive dont elle est issue (autrement dit, la version précédente). La dernière version en date d'une archive ainsi que les versions précédentes constituent ce que l'on appelle une branche. Par défaut, la branche contenant la version actuelle d'un projet s'appelle la branche maître (master). Il est possible (et recommandé) de définir de nouvelles branches pour le travail de chacun, chaque branche correspondant à une des tâche à réaliser. Une fois le travail réalisé et le résultat vérifié, il est possible d'intégrer les améliorations d'une branche à la branche maître. Si plusieurs branches ont été développées en parallèle, il faut alors prendre en compte les modifications apportée par chacune. On appelle cette opération la fusion (merge) des branches.

Le partage d'archives entre entrepôts.

Git est conçu et optimisé pour le partage des archives entre entrepôt. De nombreuses options sont disponibles, en utilisant les protocoles réseaux standards (ssh, scp, ftp, http, mail ...), un serveur spécifique, ou simplement le partage direct des informations sur une même machine (y compris en utilisant une clef. Les opération de partage de version permettent d'envoyer ou de recevoir les archives qui n'existent que d'un des deux cotés de l'échange. Ces échange ne peuvent qu'ajouter à un entrepôt des archives qui ne sont pas déjà présentes. Il n'y a aucun risque d'altérer des archives d'une dépôt (on est donc assuré de ne pas perdre de données lors de la synchronisation des dépôt !).

il est parfaitement possible de définir un dépôt central, mais cela n'est en rien obligatoire : il est possible de travailler chacun sur son dépôt, mais cela impose de procéder à des mises à jour pour chaque dépôt. En pratique, un dépôt centralisé est recommandé. Cela peut être le dépôt d'une personne du projet jouant le rôle d'intégrateur, ou alors un dépôt extérieur facilement accessible (pour les projets open-source, il existe des services de dépôt git gratuits. Si vous disposer d'une machine accessible sur internet, elle peut assez facilement servir de dépôt, si elle dispose d'un des protocoles supportés par git).

Les commandes

Les commandes de git permettent principalement :

- d'enregistrer la zone de travail dans une archive, en passant par une étape de préparation (staging), ou au contraire d'ouvrir une archive dans la zone de travail
- de gérer les archives elles-mêmes, à l'intérieur d'un même dépôt (par exemple en les étiquetant, ou en combinant les modifications réparties dans plusieurs archives)
- de manipuler les dépôts, par exemple en les copiant, en échangeant des archives etc..

En fin de cours, vous trouverez le détails des principales commandes nécessaires à toutes ces tâches, ces commandes sont regroupées dans un aide-mémoire en fin de document. Mais l'utilisation d'un système de gestion de version ne se résume pas à une maîtrise technique de l'outil, c'est l'intégration de l'outil dans la méthode de travail du groupe qui permet de bénéficier pleinement des avantages de celui-ci. Les différentes étapes nécessaires à un usage efficace de git dans le cadre de votre projet sont détaillées dans le prochain chapitre.

Méthodes de travail

Introduction

Le développement en groupe d'un projet commun s'appuie sur une bonne organisation et de bonnes méthodes de travail. Ces éléments sont nécessaires à une bonne efficacité, c'est à dire pour obtenir les meilleurs résultats possibles pour une quantité de travail donnée. Avec une mauvaise organisation ou coordination entre groupe, le résultat produit peut ne pas être meilleur que le travail d'un seul individu, et même dans certains cas être pire (par exemple en cas de pertes de données). Même un travail individuel peut bénéficier d'une bonne méthode, en particulier en ce qui concerne la gestion des versions des différents fichiers d'un gros projet.

Le fonctionnement recommandé est le suivant :

- réunions de travail, permettant de répartir les tâches, de discuter de l'avancement des travaux, de faire les choix à sur l'orientation du projet, de distribuer les nouvelles versions.
- travail individuel, sur les aspects déterminés en réunion.

Pour les gros groupes, il peut être nécessaire de désigner un membre chargé de réaliser l'intégration des travaux individuels, afin de réduire le temps nécessaire pour les réunions de travail.

Ce fonctionnement est basé entre autre sur le fait que travailler à plusieurs sur une tâche qui pourrait être réalisée par une seule personne revient à diviser l'efficacité du groupe par le nombre de personnes mobilisée sur la tâche. Il faut donc réserver les réunions à ce qui doit être fait en groupe, comme la répartition du travail, le choix des objectifs et l'échange d'idées et de remarques. C'est l'utilisation de git dans le cadre d'une organisation de ce type qui est expliquée dans ce document.

Session de travail

Phase du travail	Etapes	Commandes git	Disponible hors ligne ?
<i>Préparation du travail</i>	Obtenir un entrepôt à jour		
	S'il n'y en a pas, le créer	create	oui
	S'il y en a un, mais pas localement	clone	non
	S'il y en a un local à mettre à jour	fetch,merge / pull	non
	(Créer une branche pour le travail à réaliser)	branch / checkout	oui
<i>Travail sur le projet (à répéter autant de fois que nécessaire)</i>	Créer et/ou éditer les fichiers		oui
	Corriger les problèmes jusqu'à ce que le projet tourne.		oui
	Sauvegarder la version qui tourne dans le dépôt	add / commit	oui
<i>Synchronisation du travail avec celui de l'équipe</i>	Mettre à jour l'entrepôt local (si en ligne)	Fetch, merge / pull	non
	Fusionner la branche créée avec la branche master	checkout / merge	oui (sans intérêt sans le fetch)
	Régler les conflits de version	add/commit	Id.
	Vérifier que le projet fusionné tourne. Sinon, retour à « travail sur le projet » pour corriger les problèmes.		Id.
	Exporter les changements.	push / bundle / format-patch	oui (sauf le push)

TP 1 Découverte de git.

L'objectif de ce TP est d'utiliser les fonctions de base de GIT pour la création et la gestion d'une page web pour le groupe. Cette page comportera un lien vers une page individuelle pour chaque étudiant.

1] Création du dépôt local.

Un dépôt a déjà été mis en place dans le répertoire `/export/src/git/groupe` (**groupe** correspond à votre groupe TP). Pour obtenir une copie de ce dépôt, créez un répertoire pour ce TP, placez-vous dedans, puis utilisez la commande

```
git clone /export/src/git/groupe
```

Cette commande crée une copie locale du dépôt. C'est sur cette copie que nous allons travailler.

- 1) Quels sont les fichiers et/ou répertoires créés par `git clone`? (`ls -al` permet de lister aussi les fichiers cachés).
- 2) Quels sont les principaux protocoles supportés par `git clone`? (le manuel est accessible par la commande `git help`)
- 3) Utilisez un navigateur ouvrir les fichiers html.

2] Configuration générale

Adaptez les commandes suivantes pour configurer votre installation de git :

```
git config --global user.name 'Votre nom'
```

```
git config --global user.mail login@ma_machine.net
```

```
git config --global core.editor nano
```

3] Travail local.

- 1) Modifiez maintenant le fichier `mon_nom.html` (sans le renommer pour l'instant) en mettant le contenu à votre nom. Pour sauvegarder les modifications dans votre entrepôt, utiliser les commandes suivante :

```
git add *.html
```

```
git commit -m 'Message expliquant ce qui a été modifié.'
```

La commande `git add` indique les fichiers à prendre compte pour la sauvegarde. La commande `git commit` réalise l'archivage de la version actuelle dans le dépôt local. L'option `-m` permet de préciser le message explicatif dans la ligne de commande, au lieu d'ouvrir un éditeur de texte.

Il est recommandé d'utiliser ces deux commandes (`git add / commit`) très souvent : seules les versions ayant fait l'objet d'un commit sont conservées.

- 2) Renommez maintenant votre fichier `mon_nom.html` en le mettant à votre nom (pensez aussi à mettre à jour le lien dans le fichier `index.html`).

Pour vérifier comment git voit vos dernières modifications (les modifications depuis le dernier commit), utilisez la commande :

```
git status
```

Pour indiquer que le nouveau fichier doit être inclus dans le commit, utilisez `git add`.
Pour faire le commit, utilisez l'option `-a` :

```
git add fichier
```

```
git commit -a
```

A quoi correspond l'option `-a` ? (pensez à utiliser le manuel)

- 3) Vérifiez maintenant que la page est toujours correcte dans le navigateur, et ajoutez quelques informations sur vous et/ou une présentation. N'oubliez pas de faire régulièrement des commits !

- 4) Que produit la commande `git log` ?

Cette commande accepte aussi l'option `-oneline`, essayez la. Il est aussi possible d'afficher ces informations (et beaucoup d'autres !) avec des outils externes, comme `gitk`.

4] Mise à jour du dépôt distant.

Votre dépôt local contient les différentes versions de votre travail. L'étape suivante est partager ces versions (commits) avec le reste du groupe.

- 1) Vérifiez avec `git status` que votre dépôt est propre, c'est à dire que le répertoire de travail contient bien la même chose que le dernier commit. **Attention**, le fichier `mon_nom.html` ne doit plus être présent dans votre répertoire (il a été renommé à la question 3.3)
- 2) La commande `git push` permet de 'pousser' vos commits vers le dépôt d'origine. Attention, avant de mettre à jour le dépôt d'origine, vérifiez que votre dépôt local est à jour :

```
git status (suivi si nécessaire de git add/commit)
```

```
git push
```

- 2) Que s'est-il passé ?
- 3) Quel est le nom de la branche dans laquelle s'effectue le travail ?

5] Mise à jour à partir du dépôt distant.

La commande pour récupérer la dernière version du dépôt d'origine est `git pull`. Cette commande réalise en fait deux actions : la récupération des commits distants (qui sont désignés par `origin/branche`), et la fusion :

```
git pull
```

ou

```
git fetch
```

```
git merge origin/master
```

Comme vous l'avez probablement constaté, il est impossible d'envoyer sur le dépôt distant un commit qui n'est pas basé sur la dernière version du dépôt. Au lieu de risquer d'égarer des données, git prévient que les données ont été modifiées depuis votre dernière mise à jour.

- 1) Réalisez un diagramme indiquant les commits impliqués dans ce problème, en indiquant par des flèches les versions précédentes.

La commande `git pull` permet de récupérer la dernière version, et de la fusionner avec votre travail.

6] Gestion des conflits

La commande `git merge` (et donc la commande `pull`) peut aboutir à un conflit, si un même fichier a été modifié de manière différente par plusieurs personnes. Vous devriez normalement vous retrouver dans cette situation (sauf si vous êtes le premier à avoir exécuté `git push`).

- 1) Comment sont indiquées les zones de conflit ? (aller voir dans le fichier incriminé) Comment voit-on sa version et la version du dépôt distant ?
- 2) Corrigez le fichier (supprimez les indicateurs de conflit, et corrigez ce qui doit l'être)
- 3) Validez le résultat par `git add/commit`

Une fois les conflits résolus et validés, vous pouvez à nouveau tenter le `git push`.

7] Corrections diverses

- 1) Une fois les différentes pages mises à jour sur le dépôt distant, récupérez la dernière version et affichez le journal des modifications :

```
git pull
```

```
git log
```

- 2) Passez en revue (par exemple avec le navigateur) les différentes pages, et corrigez les fautes d'orthographe et/ou les erreurs dans le code. Pensez à faire un `commit` et un `push` régulièrement (à chaque fichier corrigé par exemple).

8] Sauvegarde sur une clef

Il est possible de sauvegarder tout ou partie d'un dépôt dans un seul fichier compressé avec la commande `git bundle` :

```
git bundle create nom_du_fichier master
```

- 1) Quelle est la taille de ce fichier ? Comparez avec la taille des fichiers dans le répertoire de travail et le nombre d'archives (commits).

Le fichier ainsi créé peut être transmis (par exemple sur une clef, par mail, par un ftp, par un service type dropbox etc..). A partir de ce fichier, on peut cloner le dépôt d'origine :

```
git clone nom_du_fichier -b master
```

Si l'on dispose déjà d'une version du dépôt, on peut faire un `git pull` sur ce fichier :

```
git pull nom_du_fichier master
```

- 2) Faites une modification sur un de vos fichiers, puis le `commit`, puis partagez votre dépôt avec un camarade sans passer par le dépôt distant.
- 3) Mettez à jour le dépôt distant.

9] Bilan

Au vu du TP :

- 1) Quelles sont les étapes du TP qui ne nécessitent pas un accès au dépôt distant ?
- 2) Que se passe-t-il si le dépôt distant est perdu ? Et si un dépôt local est perdu ?

- 3) Est ce qu'un outil comme git permet de se dispenser de l'organisation du travail entre membres du groupe ? (justifiez)
- 4) A votre avis, pourquoi le message à chaque commit est-il obligatoire ? Même question pour l'auteur ?
- 5) Pour les projets très actifs et impliquant beaucoup de personnes, on utilise un système de patchs (fichier décrivant les modifications entre deux commits) envoyés à un membre de l'équipe appelé intégrateur. C'est l'intégrateur (et lui seul) qui ajoute les nouvelles versions au dépôt distant. A votre avis, pourquoi ce mode de fonctionnement ? (il y a plusieurs raisons valables).

TP 2 Utilisation de github

L'objectif de ce TP est de réaliser une mini revue de presse, de manière collaborative.

1] Dépôt sur github

Pour ce TP, nous allons créer un dépôt sur github, un service d'hébergement de code basé sur Git.

- 1) Créez d'un compte sur <https://github.com>. Vous choisirez un identifiant de la forme NomPrenomGroupe.
- 2) Paramétrage local de Git

Si ce n'est pas déjà fait, configurez votre nom et votre email (voir TP 1). Configurez ensuite Git avec les informations relatives à votre compte sur github :

```
git config --global github.user identifiant
```

```
git config --global github.token token
```

token se trouve dans Account Settings->Account Setting et se présente en hexadécimal.

Ces commandes sont rappelées dans l'aide de github :

<http://help.github.com/set-your-user-name-email-and-github-token/>

L'iut étant derrière un proxy, il est nécessaire de le préciser :

```
git config --global http.proxy cache-etu.univ-artois.fr:3128
```

- 3) Obtention du dépôt

Chaque étudiant va disposer de son propre dépôt github. Néanmoins, un dépôt sera considéré comme la version de référence du groupe. L'étudiant hébergeant cette version sera l'intégrateur du projet, c'est à dire celui en charge de la vérification des modifications proposées, ainsi que de la fusion. L'intégrateur est habituellement quelqu'un choisit pour sa compétence technique, son sérieux et sa fiabilité (ce n'est pas nécessairement le chef de projet, il est même préférable qu'il ne le soit pas).

L'intégrateur obtient un embranchement (fork) du dépôt indiqué par l'enseignant, les autres étudiants obtiennent ensuite un embranchement de cet embranchement. Un embranchement s'obtient en cliquant sur 'fork' à partir de la page d'un projet. Une copie du projet est alors réalisée sur votre compte github.

La copie locale (sur votre machine) est ensuite obtenue par la commande clone, en utilisant l'url donnée sur votre page du projet en cliquant sur (https) :

```
git clone url nom_local_depot
```

L'utilisation d'un embranchement est détaillé dans l'aide de github :

<http://help.github.com/fork-a-repo/>

En suivant ces instructions, ajouter un lien vers 'upstream' (en utilisant l'url http de l'intégrateur).

remarque : l'intégrateur n'a pas à configurer un 'upstream', puisqu'il est lui-même la référence du projet pour le groupe.

2] Création d'un style

L'objectif va être ici de proposer chacun une feuille de style en modifiant par exemple les couleurs ou la disposition des éléments. Pour cela, vous suivrez utiliserez une des fonctionnalités les plus puissantes de Git, qui est la possibilité de créer facilement de nouvelles branches. Une branche

désigne une version différente de la version principale (dont le nom par défaut est 'master'), qui évolue de manière indépendante. Les branches peuvent ensuite être fusionnées, combinant ainsi le travail réalisé sur chacune.

Pour créer votre feuille de style :

- 1) Créez votre branche et placez-vous dedans :

```
git checkout -b nom_de_branche
```

(En fait, il s'agit d'un raccourci pour

```
git branch nom_de_branche
```

```
git checkout nom_de_branche
```

(La commande branch crée la branche, et la commande checkout permet de placer la branche dans votre espace de travail)

pour afficher toutes les branches, et la branche dans laquelle on se trouve :

```
git branch -a
```

(La commande git status affiche aussi, entre autres, la branche dans laquelle vous êtes.)

- 2) Modifiez, testez et sauvegardez la feuille de style.
- 3) Décidez quel style est retenu pour le groupe.
- 4) Pour le style retenu, fusionnez votre branche avec la branche master :

```
git checkout master
```

```
git merge nom_de_branche
```

(merge fusionne la branche courante avec la branche donnée en paramètre)

```
git push
```

Il ne vous reste plus qu'à demander à l'intégrateur d'intégrer vos modifications

<http://help.github.com/send-pull-requests/>

Notez que github permet de commenter ces demandes avant d'accepter l'ajout.

- 5) Une fois les modifications intégrées, les autres utilisateurs doivent intégrer les modifications

```
git checkout master
```

```
git pull upstream master
```

```
git push
```

3] Création du contenu

En utilisant une méthode similaire à ce qui a été utilisé pour le style, ajoutez du contenu dans la page. Comme source d'informations, vous pourrez utiliser par exemple des sites comme

<http://www.afp.com/afpcom/fr/content/news>

<http://labs.ebuzzing.fr/top-news>

Ou tout autre site d'information, thématique ou non (ou même votre imagination).

Vous pouvez aussi choisir de changer la structure du site, par exemple en créant des pages thématiques.

Le contenu lui-même a peu d'importance pour ce TP, l'important est que chacun ait l'occasion de réaliser une contribution en passant par github (*restez tout de même dans le domaine du raisonnable, et ayez des propos corrects: il vous faudra assumer la responsabilité de ce que vous écrivez...*).