

Projet à déposer sur **Didel** avant le 8 mai 2011, rubrique **Sciences > UFR d'Informatique > Licence**, onglet **BD6 - Base de données**. Vous amenez un exemplaire papier du rapport lors de la soutenance.

Vous ferez votre projet par **binôme**. Lors des soutenances vous passerez ensemble mais les questions et la notation seront **individualisées**. Nous vous demandons de préciser explicitement la façon dont vous vous êtes réparti le travail. Vous devez chacun pouvoir répondre aux questions sur tous les aspects du projet. En particulier l'élaboration de la base de données doit retenir toute votre attention. Organisez-vous bien. Il est indispensable de réfléchir à *tous* les aspects du projet avant de vous lancer dans la création des tables.

Il vous est demandé :

- un rapport au format **pdf** comprenant
 - une modélisation par un diagramme de la situation exposée, accompagnée éventuellement d'annotations,
 - une description des choix effectués,
 - la liste des tables utilisées,
 - sur les points techniquement significatifs, des exemples de modifications possibles des tables ;
- une interface suffisante et claire ;
- s'ils ne sont pas déjà fournis, des jeux d'exemples **significatifs**.

Surfant sur le succès des réseaux sociaux, une petite compagnie souhaite étudier la possibilité de créer un concurrent à **facebook** : vous êtes chargés de ce travail préliminaire. Vous pourrez vous inspirer de ce site ou d'autres pour enrichir le projet, mais votre modélisation devra prendre en compte au moins les éléments de notre description ci-dessous.

1 Les utilisateurs

Les utilisateurs sont tenus de s'inscrire et de donner un certain nombre d'informations les concernant. Une photo est requise, un nom, une adresse mail, la date de naissance, le sexe, le niveau d'étude, l'emploi actuel, les opinions politiques et religieuses, les sports pratiqués, les hobbies, le téléphone, l'adresse et le statut familial. Pour ce dernier, le cas échéant on demande de préciser les liens vers les personnes concernées (conjoint, enfants). Comme certains des éléments demandés relèvent assez clairement de la vie privée, il faut introduire un mécanisme de protection avec une certaine porosité. Typiquement certaines informations sont visibles par tout le monde, d'autres seulement par vos amis, ou bien par vos amis et leurs amis directs. Lorsque vous vous poserez une question concernant la confidentialité, vous pourrez prendre exemple sur l'un ou l'autre des réseaux sociaux de votre choix pour trancher. Bien entendu l'amitié est une relation symétrique. Il existe une possibilité de ne pas être sollicité par quelqu'un en particulier, et de blacklister une personne. D'une certaine façon cette liste noire se rapproche d'une relation d'inimitié avec un ancien ami.

Vous devrez pouvoir répondre aux requêtes suivantes :

1. Quels sont les personnes qui ne portent le nom d'aucun de leurs parents ?
2. Y a-t-il des parents ou des enfants qui ne se reconnaissent pas en tant que tel ? (problème de cohérence de déclarations ?)
3. Quelles sont les personnes qui ne sont pas amis avec leurs enfants (il y a une nuance entre famille et amis) ?

4. Quelles sont les personnes qui ne sont pas amis avec leurs parents ?
5. Quelles sont les personnes qui ne sont pas amis avec les enfants de leurs parents ?
6. Quelles sont les personnes en couple qui font des déclarations contradictoires ?
7. Quelles sont les personnes en couple et qui le disent de manière cohérente, mais dont un de leurs amis se déclare également en couple avec l'un d'eux ? Est-ce un ami commun du couple ?
8. Quelles sont les personnes qui n'ont pour amis que des gens qui partagent la même opinion politique ?
9. Quels sont les utilisateurs qui ont des amis de toutes les confessions ?
10. Quels sont ceux qui sont le plus populaire ? (le plus d'amis)
11. Quels sont ceux qui sont le plus haïs ? (mis en liste noire)
12. Quels sont ceux qui ont le plus grand réseau d'ennemis ? (des ennemis qui sont amis entre eux)

2 Les discussions

Les utilisateurs peuvent écrire des messages, cela se fait sur leur propre "mur" dont la confidentialité est définie dans les paramètres. Ils peuvent aussi écrire sur le mur de leurs amis qui l'autorisent. Un message initie un fil de discussion sur lequel chaque utilisateur qui en a le droit peut ajouter son commentaire.

1. Quels sont les utilisateurs les plus actifs ? (en nombre de messages écrits)
2. Quels sont ceux qui en moyenne suscitent le plus de réponses ? (pas de monologue)
3. Le calcul de l'activité peut être faussé par une activité de spams. Proposez une ou plusieurs définitions de notions qui favorisent le décompte de messages suscitant des discussions partagées. Écrivez les requêtes correspondant au calcul de cette activité.
4. Un utilisateur peut poster un message sur le mur d'un ami et recevoir des réponses d'inconnus. Quelles sont les personnes avec qui un utilisateur a discuté sans qu'ils soient amis ? Quelles sont celles qui sont devenues amis après une telle discussion ?
5. Une personne témoin d'une discussion entre deux autres personnes souhaite connaître tous les messages échangés entre ces deux personnes, s'il en a le droit. Écrivez cette requête.

3 La publicité

Votre site servira d'intermédiaire pour des publicités. Elles seront composées d'un court message, d'un objectif d'affichage en nombre de diffusion, d'un nombre de clics obtenu, du bénéfice attendu par affichage et par clic, et de dates pour la campagne. Elles seront ciblées en fonction d'une combinaison de certains critères comme la catégorie d'âge, le sexe, la situation familiale, le lieu de résidence etc.

Le site est chargé de placer ces publicités quelque part sur la page de connexion des utilisateurs, de la façon la plus efficace possible. Pour cela on souhaite pouvoir répondre à un certain nombre de questions :

1. Quels sont les utilisateurs qui correspondent à la cible définie pour une publicité ?
2. Quels sont ceux à qui la publicité n'a pas déjà été proposée dans les deux dernières semaines ?
3. Parfois le publicitaire souhaiterait raffiner ses critères de cible. Vous pouvez l'aider en trouvant, parmi ceux qui ont cliqué sur une publicité, quel est le critère majoritaire qui apparaît mais n'a pas été retenu dans la définition de la campagne.

4 Interface graphique

Vous devez programmer une interface graphique en **Java** pour votre projet. Il s'agit de dialoguer avec votre base de données pour implanter en particulier les fonctions suivantes :

- connexion à un compte utilisateur ;
- affichage du mur de l'utilisateur avec les derniers messages publiés et leur fil de discussion, ainsi que la possibilité d'accéder à l'historique de tous les messages du mur ;
- affichage des publicités éventuelles ;
- écriture d'un commentaire sur le mur d'un ami (il peut s'agir d'une réponse à un commentaire existant ou d'un nouveau message) ;
- création, modification et suppression d'un compte ;
- déclaration et suppression d'un ami ou d'un ennemi ;
- statistiques sur l'activité du site ;
- toute autre fonctionnalité qui vous paraît intéressante.

Les étudiants Isifar ne programmeront pas d'interface graphique mais devront donner un exemple de ces opérations sous la forme de requêtes SQL.

5 Les jeux de données

Nous vous fournissons un jeu partiel de fichiers de données que vous devrez compléter. Ces fichiers comportent un grand nombre de données et ne doivent être chargés dans la base que lors de la dernière phase de validation. Si beaucoup de grosses tables sont chargées dans la base de données, le serveur risque de s'écrouler. Il vous faudra donc constituer de petites tables de données pour tester les requêtes en cours d'élaboration.

5.1 Les fichiers de données fournis

Afin de tester l'efficacité de vos requêtes, nous vous fournissons des fichiers comportant un grand nombre de données. Ces fichiers se trouvent sur la page web du cours et sont accompagnés d'un fichier `readme.txt` qui explique ce que contient chaque fichier de données. Attention, ces fichiers ne sont absolument pas un modèle pour la base de données. Il vous faudra constituer vos tables à partir des fichiers de données, mais les tables ne suivront pas le format imposé par les fichiers. Afin de construire vos tables à partir des fichiers, vous pouvez par exemple :

- modifier les fichiers de données, afin qu'ils aient la forme adéquate pour être copiés dans une table ;
- ne sélectionner qu'un sous-ensemble de colonnes du fichier à insérer dans une table. Pour cela il faut copier intégralement le fichier dans une table intermédiaire `t`, puis sélectionner les colonnes de `t` à copier dans la table `maTable` que l'on veut créer en utilisant la commande sql :

```
CREATE TABLE maTable AS SELECT nom_colonne1, nom_colonne2... FROM t;
```
- ou toute autre méthode appropriée.

L'efficacité des méthodes utilisées pour créer vos tables à partir des fichiers fournis sera prise en compte dans l'évaluation du projet. Vous pourrez consulter la documentation de la commande `\copy` pour en connaître les options.

5.2 Les relations de parenté

Les fichiers fournis ne contiennent pas de relations de parenté (parents-enfants et réciproquement) : vous devez en ajouter entre 30 et 100. Vous ferez en sorte que les liens que vous créez permettent aux requêtes précédentes de fournir des résultats intéressants (notamment, les parents/enfants n'auront pas forcément le même nom de famille, il y aura des déclarations contradictoires, etc.). Pour cela, la génération peut être manuelle.

5.3 Génération de grands fichiers de données

En plus des fichiers fournis, il vous reste à créer les fichiers de données portant sur les murs de commentaires et les publicités. Ces fichiers doivent être conséquents et leur ordre de grandeur vous est donné ci-

dessous. Pour créer ces fichiers, vous pouvez utiliser votre langage de programmation préféré et son générateur aléatoire.

5.3.1 Les relations d'amitié et d'inimitié

Il faut engendrer un fichier de données sur les relations d'amitié, ainsi qu'un fichier sur les relations d'inimitié. Concernant le fichier d'amis, vous n'oublierez pas de spécifier qui peut écrire sur le mur de qui. Le fichier d'amis devra comporter de l'ordre de 10 000 données et celui d'ennemis de l'ordre de 1 000 données. Attention, la relation d'amitié doit être symétrique.

5.3.2 Le fichier de commentaires (pour les étudiants de L3 informatique)

A partir du fichier `commentaires.data`, vous créerez un fichier qui devra recenser, pour chaque commentaire, qui a écrit ce commentaire, sur quel mur, à quels date et horaire et à quel commentaire il répond si c'est une réponse. Le fichier devra être cohérent, intrinsèquement et avec le fichier d'amis.