

# Analyse du comportement des clients

## Partie 1

Un site web marchand propose de la documentation thématique pour ses clients (jardinage, informatique, ..). Le fichier de log du site nous permet d'identifier le type de documentation qu'a téléchargé chaque usager. Voici un exemple type d'enregistrement dans un fichier de log d'Apache :

**127.0.0.1 - frank [10/Oct/2000:13:55:36 -0700] "GET /apache\_pb.gif HTTP/1.0" 200 2326**

Dans cet exemple : le 10 octobre à 13h55 Franck a téléchargé l'image apache\_pb.gif

Après « nettoyage », le contenu utile de ce fichier est joint en annexe (log\_clients\_themes.txt). Dans ce fichier, chaque ligne représente un téléchargement (ordre d'enregistrement, nom de l'usager, thèmes). Par exemple : « **5 ; Pierre ; Bricolage** » signifie que le 5 eme enregistrement indique que Pierre a consulté une information sur le bricolage.

On vous demande d'écrire un programme en JAVA permettant construire la matrice MUT (usagers, thèmes) dans laquelle chaque élément représente le nombre d'accès de chaque usager à un thème donné. Cette matrice vous permettra de répondre aux questions qui suivent. Le programme que vous écrirez vous sera aussi utile dans les exercices suivants.

Votre programme doit extraire 3 fichiers texte

- La liste des thèmes
- La liste des usagers
- Une matrice de chiffre MUT (relations entre usagers et thèmes)