

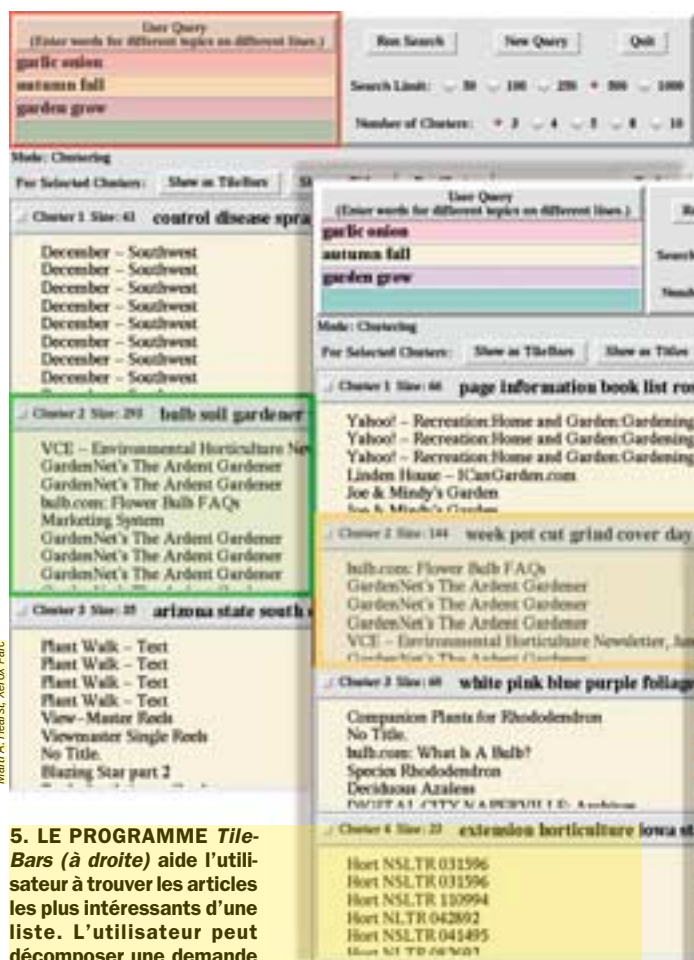
les documents par la totalité des catégories qui leur correspondent, ainsi que par un autre ensemble d'attributs (par exemple, l'origine, la date, le genre et l'auteur). Les informaticiens de la bibliothèque de l'Université Stanford élaborent actuellement une interface nommée *SenseMaker*, fondée sur ce principe.

Au Centre de recherches Xerox, nous avons mis au point une autre procédure pour grouper les pages fournies par un moteur de recherche. Cette procédure nommée *Scatter/Gather* construit une table des matières qui

se modifie à mesure que l'utilisateur comprend mieux la nature des documents disponibles et découvre ceux qui sont les plus intéressants.

Imaginons qu'effectuant une recherche à l'aide du programme *Excite*, Alice reçoive les 500 premières pages proposées par le réseau. Le système *Scatter/Gather* examine alors ces pages et les répartit en trois à cinq groupes selon leurs ressemblances (voir la figure 4). Alice peut alors examiner ces groupes et choisir ceux qui lui semblent plus intéressants.

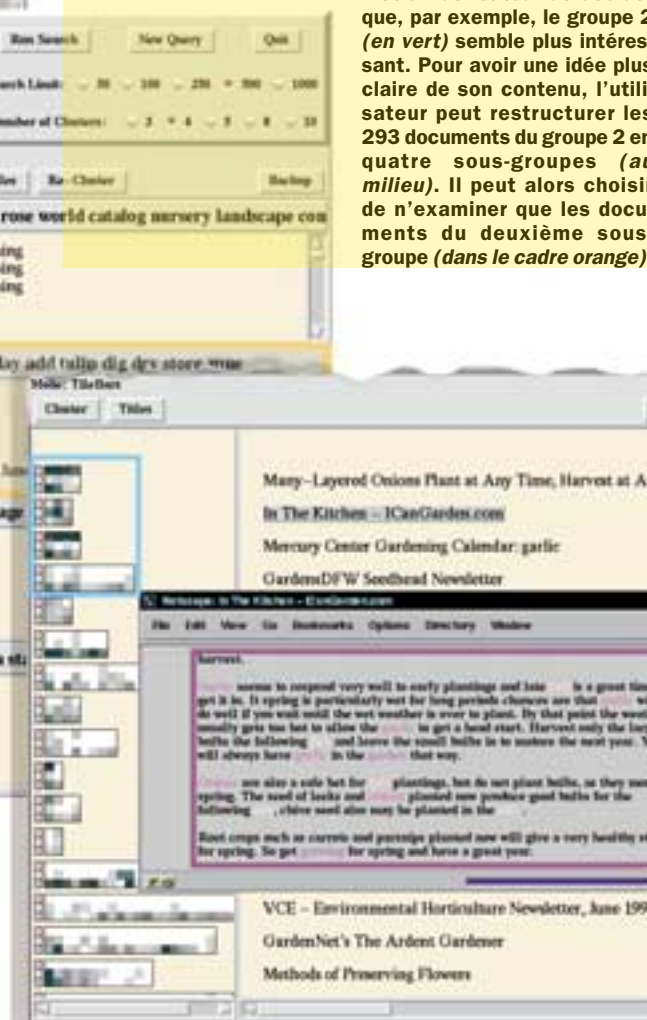
Bien que le comportement des utilisateurs soit variable, les expériences indiquent que les arborescences facilitent l'accès aux documents intéressants. Si Alice juge que le groupe des 293 textes résumés par «bulbes», «sol» et «jardinier» est prometteur, par exemple, elle peut utiliser *Scatter/Gather* pour le parcourir et le diviser en sous-groupes plus petits. Après quelques subdivisions, elle ramène son ensemble de 500 pages, pour la plupart sans intérêt, à un groupe de quelques dizaines de pages qui, elles, sont intéressantes.



5. LE PROGRAMME *TileBars* (à droite) aide l'utilisateur à trouver les articles les plus intéressants d'une liste. L'utilisateur peut décomposer une demande en plusieurs thèmes (en rouge), chacun contenant plusieurs mots clés. Le programme crée alors une fenêtre (en bleu) où il affiche les occurrences trouvées. La taille de la fenêtre est proportionnelle à la longueur du document. Chaque fenêtre est un quadrillage : les colonnes correspondent aux passages dans plusieurs paragraphes de texte, et les lignes aux thèmes de recherche demandés par l'utilisateur. Le programme *TileBars* affiche ainsi immédiatement, pour chaque article, les passages qui contiennent un sujet particulier, indiquant en outre le nombre d'occurrences de chaque sujet (plus les carrés sont sombres, plus les occurrences sont nombreuses). Ici, la ligne supérieure de chaque rectangle indique le nombre d'occurrences des mots «ail» ou «oignon» dans le premier thème ; la ligne du milieu correspond au nombre d'occurrences du mot «automne», etc. Cette interface guide les recherches. Les deux articles supérieurs, où les trois sujets figurent dans plusieurs pas-

4. L'INTERFACE *Scatter/Gather* groupe les pages de premier niveau identifiées par le moteur de recherche *Excite*, et il étiquette chaque groupe par des mots qui semblent caractéristiques (ces mots, puisés dans les documents réels, sont probablement différents de ceux que *Yahoo* aurait choisis).

L'examen des étiquettes permet à l'utilisateur de décider que, par exemple, le groupe 2 (en vert) semble plus intéressant. Pour avoir une idée plus claire de son contenu, l'utilisateur peut restructurer les 293 documents du groupe 2 en quatre sous-groupes (au milieu). Il peut alors choisir de n'examiner que les documents du deuxième sous-groupe (dans le cadre orange).



sages, seront probablement plus intéressants que le document inférieur, qui ne contient aucune mention des mots «ail» ou «oignon». En cliquant sur un carreau d'une fenêtre, on affiche une fenêtre (en mauve) où s'affiche le passage correspondant à l'intérieur du texte, avec les mots clés mis en relief par des couleurs appropriées.