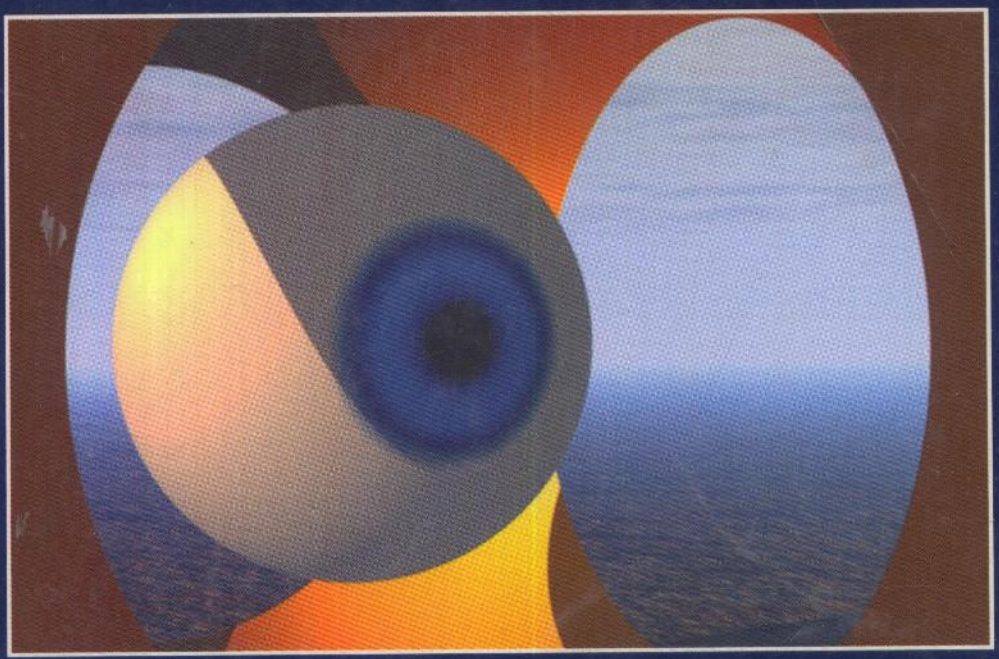

La recherche d'informations

du texte intégral au thésaurus

Philippe Lefèvre

Hermes

La recherche d'informations

du texte intégral au thésaurus

Philippe Lefèvre

Hermes Science publications

Table des matières

Comment y voir clair dans l'effervescence qui règne aujourd'hui parmi les outils d'accès à l'information : logiciels de gestion documentaire, moteurs de recherche, agents de toutes sortes, portails... Sur quelle base comparer leurs fonctionnalités très diverses ? Quels sont par exemple les rapports entre la fonction d'indexation-recherche, le classement des informations, le filtrage et le *push* ? Autant de questions auxquelles cet ouvrage apporte des réponses claires.

La recherche d'informations – du texte intégral au thésaurus propose en effet une vision de synthèse réalisant le lien entre plusieurs mondes : la documentation traditionnelle, la GEIDE et l'ingénierie linguistique, enfin le nouveau monde de l'internet.

Les difficultés à trouver les informations, y compris sur l'internet, sont surtout dues aux caractéristiques du langage. Partant de ce constat, ce livre fait l'inventaire des problèmes liés à la recherche d'informations ; puis il décrit les techniques appliquées pour les résoudre, des plus anciennes aux plus récentes.

L'indexation est à la base de tous les traitements de l'information textuelle ; cette fonction est étudiée en détail, ce qui amène à distinguer treize modes d'indexation différents. Une telle typologie permet de caractériser sur un plan fonctionnel les principaux moteurs de recherche du marché.

Enfin, l'internet est abordé en montrant ses différences avec les systèmes documentaires traditionnels, puis une analyse systématique est utilisée pour présenter les outils de recherche sur ce nouveau média : moteurs, métamoteurs... ainsi que les approches les plus récentes comme les portails.

L'auteur

Philippe Lefèvre, ingénieur ESE, est chercheur à la division Recherche et Développement d'EDF. Après avoir travaillé sur la reconnaissance de documents ou lecture optique, il s'intéresse depuis quelques années à la recherche d'informations et au traitement automatique du langage naturel.

hermes science publications

9 782746 201736