(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

PCT

(43) Date de la publication internationale
12 septembre 2008 (12.09.2008)

(10) Numéro de publication internationale
WO 2008/107338 A1

(51) Classification internationale des brevets :
G06F 17/24 (2006.01)

(21) Numéro de la demande internationale :
PCT/EP2008/052274

(22) Date de dépôt international :
25 février 2008 (25.02.2008)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0753440 23 février 2007 (23.02.2007) FR

(71) Déposant et
(72) Inventeur : **MAIM, Enrico** [FR/FR]; 8, rue Dareau, F-75014 Paris (FR).

(74) Mandataire : **WARCOIN, AHNER, TEXIER, LE FORESTIER, CALLON DE LAMARCK, COLLIN, TETAZ-Cabinet Regimbeau**; 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) États désignés *(sauf indication contraire, pour tout titre de protection nationale disponible)* : AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Suite sur la page suivante]

WO 2008/107338 A1

(54) **Title:** METHODS FOR THE EXTRACTION, COMBINATION, SYNTHESIS AND VISUALISATION OF MULTI-DIMENSIONAL DATA FROM DIFFERENT SOURCES

(54) **Titre :** PROCEDES D'EXTRACTION, DE COMBINAISON, DE SYNTHESE ET DE VISUALISATION DE DONNEES MULTIDIMENSIONNELLES PROVENANT DE DIFFERENTES SOURCES

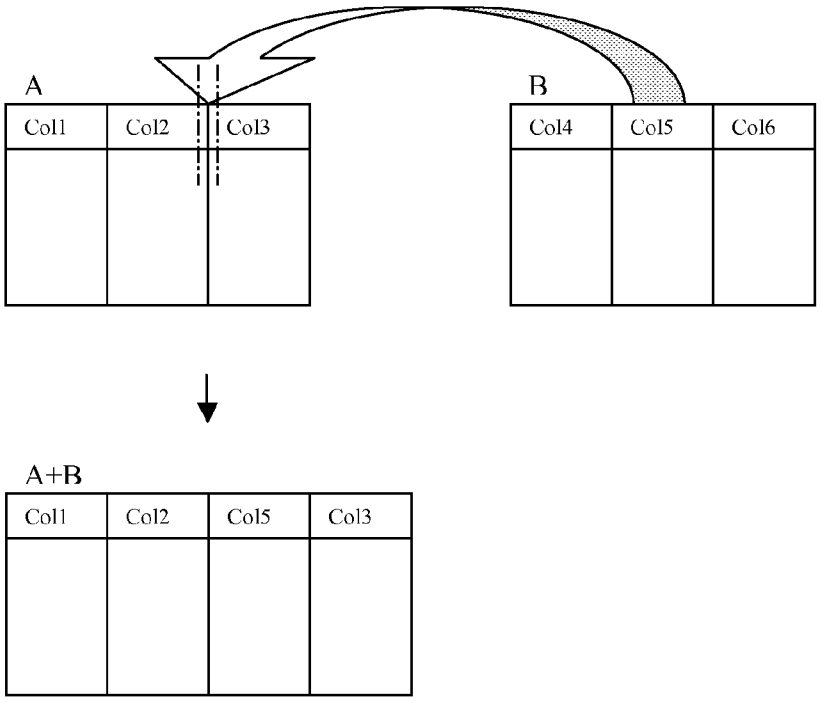


Figure 9

(57) Abstract: The invention relates to a method for the automatic combination of multi-dimensional data depending on manipulations at the level of their dimensions in a computer environment that comprises computer equipment capable of accessing multi-dimensional data sources. The method comprises the following steps: a) providing a first multi-dimensional data source; b) providing at least a second multi-dimensional data source, at least said second multi-dimensional data source being such that each data includes key values and non-key values in the different dimensions; c) identifying actions carried out using a user interface on representations of some dimensions of the sources; d) depending on said actions, combining the data sources by using the key values of each data of the second source(s) in order to establish associations between the multi-dimensional data of the first source and the multi-dimensional data of the second source(s) and to obtain combined multi-dimensional data, the combination being carried out as follows: i) adding to the multi-dimensional data of the first source at least a portion of the non-key values of the corresponding data of the second source(s); and ii) arranging the added non-key values with the pre-existing values for the same combined multi-dimensional data. The invention also relates to associated methods for the combination and enrichment of data and for the manipulation of the visualisation of resources.

[Suite sur la page suivante]

(84) États désignés *(sauf indication contraire, pour tout titre de protection régionale disponible)* : ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— *avec rapport de recherche internationale*

— *avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues*

(57) Abrégé : L'invention concerne un procédé de combinaison automatique de données multidimensionnelles en fonction de manipulations au niveau de leurs dimensions dans un environnement informatique comprenant un équipement informatique capable d'accéder à des sources de données multidimensionnelles. Le procédé comprend les étapes suivantes : a) prévoir une première source de données multidimensionnelles; b) prévoir au moins une deuxième source de données multidimensionnelles, au moins cette deuxième source étant telle que chacune de ses données comporte dans les différentes dimensions des valeurs clés et des valeurs non-clés; c) identifier des actions effectuées à l'aide d'une interface utilisateur sur des représentations de certaines dimensions des sources; d) en fonction desdites actions, combiner les sources de données en utilisant les valeurs clés de chacune des données de la ou des deuxièmes sources pour réaliser des associations entre les données multidimensionnelles de la première source et les données multidimensionnelles de la ou des deuxièmes sources et ainsi obtenir des données multidimensionnelles combinées, ladite combinaison étant effectuée: i) en ajoutant aux données multidimensionnelles de la première source au moins une partie des valeurs non-clés des données correspondantes de la ou des deuxièmes sources, et ii) en agençant les valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée. L'invention propose également des procédés associés de combinaison et d'enrichissement de données et de manipulation de la visualisation de ressources.

Titre : « Procédés d'extraction, de combinaison, de synthèse et de visualisation de données multidimensionnelles provenant de différentes sources »

La présente invention concerne des procédés de combinaison et de visualisation de données à partir de sources de données, et en particulier de nouveau types de services d'extraction, de combinaison et de visualisation aptes à combiner des informations complémentaires ou concurrentes disponibles notamment sur le Web, ou encore dans l'entreprise, et de naviguer facilement dans les combinaisons effectuées y compris lorsqu'elles représentent des volumes importants de données.

Dans l'état actuel de la technique, les outils dits de « Mashup » permettent combiner des données extraites de sites Web ou autres sources de données multidimensionnelles. Par exemple, des données extraites d'un site Web donnant des adresses d'hôtels peuvent être combinées avec des données extraites d'un site donnant des horaires d'avions, avec des données extraites d'un site de météo, etc.

Une difficulté de ces outils est qu'ils ont pour conséquence de démultiplier les données à considérer par l'utilisateur. Par exemple, si à chaque destination d'un vol d'avion correspondent 10 hôtels en moyenne, la combinaison des vols intéressant l'utilisateur avec des hôtels va augmenter la quantité des informations auxquelles il fera face d'un facteur 10.

Il existe donc un besoin pour des outils permettant de combiner les données en limitant la quantité d'information globale à prendre en compte, et d'explorer plus facilement ces données, et également d'enrichir des données avec d'autres, pertinentes par rapport aux attentes de l'utilisateur, le tout avec une bonne ergonomie, et en particulier avec des actions réalisées simplement notamment à l'aide d'un dispositif de pointage tel qu'une souris.

Par ailleurs, les outils de « mashup » peuvent être utilisés par combiner des sources d'informations par nature concurrentes. Par exemple, si l'on considère des extracteurs fournissant, à partir d'un site Web de vente de livres, un ensemble de données multidimensionnelles (typiquement présentées en lignes) composées des dimensions (typiquement les colonnes) telles que « Auteur principal », « Titre » et « Prix », une jointure peut être effectuée sur les colonnes « Auteur principal » et « Titre », pour comparer les prix d'ouvrages fournis par différents vendeurs sur différents sites...

Une jointure sur la colonne « Prix » peut également être effectuée pour comparer des livres différents ayant le même prix. Des extracteurs peuvent bien sûr être programmés pour combiner correctement de telles sources de données malgré leur recours à des vocabulaires différents.

Mais dans cette profusion de choix de sources de données et de manières de les combiner qui s'offre à l'utilisateur, il n'existe pas de moyens génériques pour notamment :

• permettre à un utilisateur de faire correspondre entre elles (ou de supprimer une telle correspondance), interactivement, n'importe quelles dimensions de sources de données multidimensionnelles, même si elles sont nommées différemment dans les différentes sources, afin de pouvoir ensuite combiner ces sources de manière automatisée ;

• tirer parti d'indicateurs (clés), le cas échéant déjà associées aux données elles-mêmes (dans l'exemple susmentionné seules les colonnes « Auteur principal » et « Titre » devraient typiquement servir de clés et pouvoir ainsi servir de colonnes de jointure, bien que l'utilisateur ait aussi fait correspondre la colonne « Prix ») pour automatiquement combiner les données de la façon la plus appropriée, et ceci pourrait être déterminé de manière automatique ;

• à partir du comptage des correspondances de dimensions déjà établies ou supprimées par d'autres utilisateurs, automatiquement suggérer à l'utilisateur non seulement des sources de données à combiner, mais également des correspondances de dimensions entre ces sources ;

• combiner les données en tenant compte de leur chronologie et notamment dans l'ordre de leur première apparition (par opposition à l'ordre de leur réception), et naviguer dans le temps pour explorer des combinaisons de données interdépendantes qui ne sont pas toutes forcément valides en même temps.

Ainsi et plus spécifiquement, des objets de l'invention sont :

• de permettre à l'utilisateur de simplement laisser des sources multidimensionnelles se combiner automatiquement selon des indicateurs (clés) associés aux données à combiner, après avoir simplement validé des suggestions automatiques de correspondances entre dimensions et/ou de sa propre initiative fait correspondre des dimensions entre elles.

• de tenir automatiquement à jour les résultats des combinaisons ;

• de retenir les temps de première apparition de chaque information pour que les informations de différentes sources ne soient combinées que dans leurs périodes de validité respectives et pour que l'utilisateur puisse naviguer dans le temps pour reconstituer, par combinaison de sources différentes, des "débats" tels que des ensembles d'informations contradictoires pour une actualité ,ou tout simplement revoir des offres concurrentes pour un produit.

Ainsi dans l'exemple précédent de combinaison de données extraites de sites de vente de livres, l'utilisateur pourrait faire correspondre les colonnes « Prix » des sources respectives pour faire afficher dans une même colonne le prix le plus récent de chaque livre ; puis en naviguant dans le temps, c'est-à-dire en se déplaçant dans le passé, il verrait défiler le prix de chaque autre vendeur. En effet, chaque offre a forcément été la plus récente une fois (au moins au moment précis de sa première apparition), et dans la mesure où elle n'a pas été retirée depuis ce moment (elle est encore valide), elle sera affichée. Enfin, l'utilisateur pourrait faire afficher les offres de différents vendeurs en même temps ou encore mettre ses propres livres en vente en utilisant le même procédé.

En outre, une difficulté des outils existants de visualisation de données en tableaux réside dans le fait que la disposition de la structure arborescente (branchements de gauche à droite) est préétablie et l'agencement des colonnes dans cette présentation peut être différent de celui du tableau de données de départ.

Ceci est désavantageux en ce que cela peut nuire à la clarté de lecture de l'ensemble des données ainsi réduites ou déployées, typiquement lorsque l'utilisateur est habitué à l'agencement de colonnes du document de départ.

Il existe ainsi un besoin pour accéder à des données sous forme de tableau dans lequel des réductions/déploiements de sous-ensembles des données puissent être effectuées directement dans le tableau d'origine, en conservant sa présentation et l'agencement de ses colonnes, de manière à améliorer la lisibilité par l'utilisateur des manipulations effectuées sur le tableau de données.

Pour atteindre au moins l'un des objectifs ci-dessus on propose selon un premier aspect de l'invention un procédé de combinaison automatique de données multidimensionnelles en fonction de manipulations au niveau de leurs dimensions dans un environnement informatique comprenant un équipement informatique capable d'accéder à des sources de données multidimensionnelles, caractérisé en ce qu'il comprend les étapes suivantes :

(a) prévoir une première source de données multidimensionnelles ;

(b) prévoir au moins une deuxième source de données multidimensionnelles, au moins cette deuxième source étant telle que chacune de ses données comporte dans les différentes dimensions des valeurs clés et des valeurs non-clés ;

(c) identifier des actions effectuées à l'aide d'une interface utilisateur sur des représentations de certaines dimensions des sources ;

(d) en fonction desdites actions, combiner les sources de données en utilisant les valeurs clés de chacune des données de la ou des deuxièmes sources pour réaliser des associations entre les données multidimensionnelles de la première source et les données multidimensionnelles de la ou des deuxièmes sources et ainsi obtenir des données multidimensionnelles combinées, ladite combinaison étant effectuée (i) en ajoutant aux données multidimensionnelles de la première source au moins une partie des valeurs non-clés des données correspondantes de la ou des deuxièmes sources, et (ii) en agençant les valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée.

L'invention propose selon un deuxième aspect un procédé de combinaison automatique de données multidimensionnelles provenant d'une pluralité de sources de données, caractérisé en ce qu'il comprend une succession de mises en œuvre en cascade du procédé tel que défini ci-dessus, les données combinées d'une mise en œuvre donnée dudit procédé constituant une source de données pour une mise en œuvre suivante dudit procédé.

Certains aspects préférés mais non limitatifs de ce procédé sont les suivants :

* l'agencement des valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée comprend la sélection d'une valeur non-vide parmi les valeurs provenant des différentes sources.

* l'agencement des valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée comprend la sélection d'une valeur parmi les valeurs provenant des différentes sources selon un processus décisionnel donné.

* à chaque donnée est associée une période de validité, et le processus décisionnel comprend la sélection d'une valeur appartenant à une donnée valide à une date donnée.

* à chaque donnée est associée une date de première apparition, et le processus décisionnel comprend la sélection d'une valeur provenant d'une donnée la plus récemment apparue à une date donnée éventuellement ajustable.

* le procédé comprend, pour les données d'une source auxquelles n'est associée aucune date de première apparition, une étape de création d'une date de première d'apparition égale à la date à laquelle la donnée a été mise en jeu dans une combinaison de données pour la première fois.

* au moins l'une des sources de données multidimensionnelles comprend au moins deux sources de données amont et des informations définissant une combinaison précédemment effectuée selon la revendication 1.

* lesdites actions sont des manipulations sur interface graphique d'une représentation d'au moins une dimension de la deuxième source pour la mettre en correspondance avec une représentation d'au moins une dimension de la première source ou pour l'insérer entre deux dimensions de la première source, les dimensions correspondant aux représentations manipulées déterminant soit lesdites associations entre données, soit lesdits agencements de valeurs non clés de la deuxième source avec des valeurs de la première source, selon que la dimension de la deuxième source correspondant à la représentation manipulée contient ou non des valeurs clés.

On propose selon un troisième aspect un procédé de combinaison de données multidimensionnelles, comprenant les étapes suivantes :

- prévoir un accès à une pluralité de sources de données,

- mémoriser des informations de correspondance entre sources de données à partir d'opérations de combinaison effectuées par le procédé de la revendication 1,

- lors d'un accès à une source de données ayant déjà fait l'objet de combinaison avec d'autres sources de données, signaler l'existence desdites autres sources de données.

Certains aspects préférés mais non limitatifs de ce procédé sont les suivants :

* les informations de correspondance comprennent également des informations de correspondance entre les dimensions desdites sources et le procédé comprend en outre, lors d'un accès à une source de données ayant déjà fait l'objet de combinaisons avec d'autres sources de données, signaler également les correspondances entre dimensions.

* le procédé comprend en outre l'exécution par défaut d'opérations de combinaison par le procédé de la revendication 1 entre la source de données à laquelle il est accédé et lesdites autres sources de données.

* les informations de correspondance sont mémorisées pour une pluralité d'utilisateurs, et l'étape de signalisation s'effectue en fonction de règles de prépondérance parmi les informations de correspondance.

Un quatrième aspect de l'invention vise un procédé de combinaison de données multidimensionnelles, comprenant les étapes suivantes :

- prévoir un accès à une pluralité de sources de données,

- mémoriser des informations de correspondance entre sources de données à partir d'opérations de combinaison effectuées par le procédé de la revendication 1,

- lors d'un accès à une source de données ayant déjà fait l'objet de combinaison avec d'autres sources de données, déterminer l'existence de correspondances en chaîne entre sources de données, et, en fonction de caractéristiques des informations de correspondance, exécuter sélectivement une opération de combinaison effectuée par le procédé de la revendication 1 entre la source de données à laquelle il est accédé et une source de données liées à celle-ci par une chaîne d'au moins deux correspondances.

Un cinquième aspect de l'invention vise un procédé d'enrichissement de données multidimensionnelles par combinaison automatique en fonction de manipulations au niveau de leurs dimensions dans un environnement informatique comprenant un équipement informatique capable d'accéder à des sources de données multidimensionnelles, caractérisé en ce qu'il comprend, après avoir appliqué à une source précédente de données une fonction de sélection pour obtenir une sélection précédente de données, les étapes suivantes :

- lors d'un accès à une source courante de données pour obtenir une sélection courante de données, déterminer l'existence d'au moins une mise en correspondance de dimensions entre les données des deux sources,

- si une telle existence a été déterminée, appliquer le procédé de la revendication 1 sur une première paire de première et deuxième sources de données constituées respectivement de la sélection courante et de la source précédente, et sur une seconde paire de première et deuxième sources de données constituées respectivement de la source courante, de laquelle est retirée la sélection courante, et de la sélection précédente.

Certains aspects préférés mais non limitatifs de ce procédé sont les suivants :

* la mise en correspondance de dimensions entre les données des deux sources est effectuée au cours de la mise en œuvre du procédé.

* l'accès aux sources s'effectue à l'aide d'un navigateur Web en ce que la mise en œuvre du procédé est effectuée par interception des requêtes vers des serveurs et extraction de données à partir de ces serveurs.

* ladite mise en correspondance est effectuée par affichage des sélections de données et glisser-déposer de valeurs dans les dimensions devant être mises en correspondance.

* le procédé comprend en outre une étape de synthèse pour afficher lesdites sélections dans leur environnement graphique et associer aux valeurs des moyens permettant le glisser-déposer.

* le procédé est mis en œuvre de façon répétée lors de l'accès à une succession de sources de données et, lors de l'accès à une source de données courante pour laquelle il n'existe aucune mise en correspondance de dimensions avec la source précédente, on recherche l'existence d'une mise en correspondance de dimensions entre la source courante et une source antérieure, et on applique le procédé d'enrichissement sur des paires de sources constituées par ladite source antérieure et toute source consultée plus récemment avec laquelle il existe une mise en correspondance de dimensions, puis sur une paire de sources constituée par ladite source antérieure ainsi enrichie et la source courante.

* le procédé est mis en œuvre lors de l'accès à une succession de sources de données SN-2, SN-1 et SN, et comprend les étapes suivantes :

si une mise en correspondance entre des dimensions des sources SN-2 et SN-1 d'une part, et SN-1, SN d'autre part existe, mettre en œuvre le procédé entre les sources SN et SN-1 en utilisant comme source SN-1 le résultat du procédé selon la revendication 15 mis en œuvre sur les sources SN-1 et SN-2,

si aucune mise en correspondance entre les sources SN-2 et SN-1 n'existe, déterminer s'il existe une mise en correspondance entre les dimensions des sources SN-2 et SN et, dans l'affirmative, mettre en œuvre le procédé selon la revendication 15 d'une part sur les sources SN et SN-1 et d'autre part sur les sources SN et SN-2, et

si aucune mise en correspondance entre les sources SN-1 et SN n'existe, déterminer s'il existe une mise en correspondance entre les dimensions des sources SN-2 et SN et entre les sources SN-2 et SN-1 et, dans l'affirmative, mettre en œuvre le procédé selon la revendication 15 d'une part sur les sources SN-1 et SN-2 et d'autre part sur les sources SN et SN-2.

et ainsi de suite pour des sources SN-3, SN-4, etc.

Selon un sixième aspect de l'invention, on propose un procédé pour manipuler la visualisation d'une ressource contenant des informations structurées sous forme de tableau à au moins deux dimensions, telle qu'obtenue notamment à partir du procédé tel que défini dans ce qui précède, où une dimension du tableau est constituée de colonnes représentant des types de données, et une autre dimension du tableau est constituée de lignes représentant des jeux de données associées ayant les types respectifs, et en ce que le procédé comprend :

(a) l'affichage sous forme d'une ligne unique d'un groupe de jeux ayant tous même valeur dans une colonne donnée,

(b) l'affichage, dans la colonne donnée, de ladite valeur, et

(c) l'affichage, en association avec au moins une autre colonne, d'un indicateur signalant qu'il existe au moins deux valeurs dans cette autre colonne pour le groupe de jeux de données.

Certains aspects préférés mais non limitatifs de ce procédé sont les suivants :

* le procédé comprend une étape consistant, en fonction d'actions exercées à l'aide d'une interface utilisateur d'entrée de façon associée à ladite autre colonne au niveau de ladite ligne unique, à provoquer l'affichage des différentes valeurs prise par le groupe de jeux de données dans cette autre colonne.

* ledit affichage est effectué valeur par valeur.

* ledit affichage est effectué dans un menu surgissant (pop-up).

* le procédé comprend en outre, en réponse à une action à l'aide d'une interface utilisateur d'entrée sur un indicateur, le déploiement de la ligne en question en un sous-tableau.

* ledit sous-tableau comprend autant de lignes qu'il existe, dans ledit groupe de jeux * chaque ligne dudit sous-tableau contient une valeur différente dans ladite autre colonne et représente un sous-groupe de jeux de données ayant tous cette valeur dans ladite autre colonne.

* le procédé comprend la répétition des étapes (b) à (d) pour au moins l'une des lignes constituant ledit sous-tableau, répétition appliquée à au moins une autre colonne encore dans laquelle il existe au moins deux valeurs pour le sous-groupe de jeux de données correspondant à ladite ligne au moins.

* il existe un type supplémentaire virtuel « ligne », le tableau étant initialement présenté sous forme d'une ligne unique regroupant toutes les lignes, un indicateur étant affiché en association avec chaque colonne dans laquelle il existe au moins deux valeurs.

* chaque ligne dudit sous-tableau représente, pour les colonnes à partir des indicateurs desquelles le sous-tableau a été formé, une combinaison spécifique de valeurs différente.

* le procédé comprend l'affichage d'un indicateur associé à ladite autre colonne après déploiement et sur lequel une action à l'aide d'une interface utilisateur d'entrée provoque le regroupement (réduction) dudit sous-tableau en ladite ligne unique.

* l'indicateur associé à ladite autre colonne comprend un symbole susceptible d'être orienté vers le bas ou vers le haut.

* le procédé comprend l'affichage dans ladite autre colonne, d'une des valeurs prises par le groupe de jeux de données dans cette colonne.

* le procédé comprend l'affichage dans ladite autre colonne, d'une combinaison des valeurs prises par le groupe de jeux de données dans cette colonne.

* le procédé comprend l'affichage dans ladite autre colonne, d'une propriété telle que la cardinalité de l'ensemble des valeurs prises par le groupe de jeux de données dans cette colonne.

* le procédé comprend une étape de détermination d'une clé de sélection de jeux de valeurs, déterminant un groupe de jeux de valeurs auxquels un changement de valeur dans une colonne au niveau d'une ligne affichée s'appliquera collectivement.

* ladite clé de sélection est constituée par les valeurs affichées dans la ou les colonnes à partir des indicateurs desquelles un sous tableau a été formé.

* ladite clé de sélection est constituée par les valeurs affichées dans toutes les colonnes, dont la valeur avant changement pour la colonne dans laquelle le changement est effectué.

* ladite clé de sélection est constituée par la valeur avant changement dans la colonne en question.

* le procédé comprend une étape d'ajout à la clé de sélection d'une valeur affichée dans une colonne par l'indicateur de laquelle aucun sous-tableau n'a été formé, par une action spécifique à l'aide d'une interface utilisateur d'entrée.

* le procédé comprend une étape de suppression de la clé de sélection d'une valeur affichée dans une colonne par l'indicateur de laquelle aucun sous-tableau n'a été formé, par une action spécifique à l'aide d'une interface utilisateur d'entrée.

* dans le cas où la ressource est construite dynamiquement à partir d'une source de données, le procédé est apte à afficher en association avec la ressource un indicateur visant à permettre un accès direct à ladite source de données.

Enfin l'invention propose un procédé de présentation de données, comprenant les étapes suivantes :

(a) définir un modèle de présentation de données à partir d'une ressource dont sont issues une partie des données dont la visualisation est manipulée dans le procédé tel que défini selon le sixième aspect,

(b) synthétiser au moins une partie desdites données manipulées dans ledit modèle de présentation selon des règles prédéterminées, et

(c) lorsqu'à une desdites données est associé un indicateur, afficher un équivalent dudit indicateur dans la présentation de données.

Les figures illustrent respectivement ces différentes approches d'utilisation. On suppose que l'utilisateur combine les sources de données[1] "Vendeur2" et "Vendeur3" à la source de données "Vendeur1"

Vendeur1

Auteur principal	Titre	Vendeur	Prix	Valide depuis	Valide jusqu'à
Auteur1	Titre1	Vendeur1	25	21 mars 2007 08:15	null
Auteur2	Titre2	Vendeur1	24	22 mars 2007 10:05	null
Auteur2	Titre3	Vendeur1	20	15 mars 2007 11:10	null
Auteur3	Titre4	Vendeur1	15	27 février 2007 11:50	22 mars 2007 11:49

"Vendeur2" et "Vendeur3" ayant les mêmes colonnes que "Vendeur1" mais avec respectivement les colonnes supplémentaires « Nombre de pages » et « Rating » (évaluation du livre en question).[2]

Vendeur2

Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Valide depuis	Valide jusqu'à
Auteur1	Titre1	Vendeur2		350	1 mars 2007 12:00	22 mars 12:00
Auteur2	Titre2	Vendeur2	25	430	23 mars 2007 14:15	null
Auteur3	Titre4	Vendeur2	12	62	23 mars 2007 09:00	

Vendeur3

Auteur principal	Titre	Vendeur	Prix	Rating	Valide depuis	Valide jusqu'à
Auteur1	Titre1	Vendeur3	25	***	22 mars 2007 10:00	null
Auteur2	Titre2	Vendeur3	26	**	23 mars 2007 14:14	null

A noter que les valeurs des colonnes temporelles « Valide depuis » et « Valide jusqu'à » proviennent d'une détection automatique de la première (resp. dernière) apparition des données dans les sources de données respectives, que les valeurs "null" indiquent une valeur non connue, et que les valeurs verticalement répétées ne sont pas présentées [3] (dans la figure suivante "Auteur2" est répétée dans deux lignes mais mentionnée dans une seule).

Dans la première approche (figure 1) l'utilisateur se positionne à "Maintenant" sur l'axe du temps et voit les offres *les plus récentes* pour les valeurs qu'il a déployées (qui dans la figure sont préfixées du symbole "^")[4] et qui sont *encore valides*. Il y voit notamment le livre {Auteur1, Titre1} proposé par Vendeur3 avec un « Rating » de "***" (dans la première ligne affichée dans la figure) et le livre {Auteur2, Titre2} proposé par Vendeur2 (dans la deuxième ligne affichée dans la figure)[5]. L'utilisateur remarque que, bien que la source de

[1] Dans la suite on utilise les termes « sources de données multidimensionnelles » et « sources de données » de manière interchangeable, et le terme « source de données » est parfois utilisé au lieu de « tableau de source de données ».
[2] Ainsi sur la base d'une clé composée des colonnes « Auteur principal » et « Titre », les sources "Vendeur2" et "Vendeur3" contribuent des valeurs d'une part ensemble dans la colonne « Prix » et d'autre part respectivement dans les colonnes « Nombre de pages » et « Rating ». On considère ici que les sources présentées sont en fait des « sélections » et que l'on applique le procédé d'enrichissement de sélections décrit plus loin.
[3] Chaque valeur manquante représente une répétition de la valeur au-dessus. Ceci arrive lorsqu'à partir d'une valeur déjà déployée d'une colonne d'une ligne, une valeur d'une autre colonne de la même ligne est déployée, ceci est décrit plus loin.
[4] Le concept de cellule « déployée » ou « réduite » est décrit plus loin. Pour simplifier, dans ces figures le symbole "ˇ" n'est pas affiché pour indiqué qu'une cellule est réduite, mais ceci sera le cas plus loin.
[5] Comme présenté dans la figure il est maintenant 15:20 et il peut optionnellement voir que cette offre est « Valide depuis » 14:15.

données Vendeur2 ne fournisse pas de « Rating », il y a une valeur "**" dans cette colonne dans la deuxième ligne. En amenant le curseur de la souris sur la valeur en question, il voit qu'elle a été obtenue par combinaison avec une offre antérieure (et qui est toujours valide) de Vendeur3. A noter enfin que la source Vendeur2 contribuant la colonne « Nombre de pages », toutes les lignes provenant de Vendeur2 présentent des valeurs dans cette colonne.

Dans la figure 2, l'utilisateur s'est positionné dans le passé (précisément au temps 22 mars 2007 10:10) par un glisser-déposer du curseur temporel[6] vers la gauche. Voici ci-dessous ce qu'il voit.

Première ligne de la figure 2 :

Bien que la source Vendeur3 ne fournisse pas de colonne « Nombre de pages », l'utilisateur y voit maintenant la valeur "350". Lorsqu'il y amène de curseur de la souris il lui est affiché que cette valeur est obtenue par combinaison avec la source Vendeur2 [7].

Deuxième ligne de la figure 2 :

L'utilisateur voit maintenant pour {Auteur2, Titre2} la ligne de Vendeur1 avec un « Prix » de "24", qui était l'offre *la plus récente* au temps 22 mars 2007 10:10 pour les valeurs déployées et qui est *encore valide maintenant*.

L'utilisateur qui s'intéresse à cette offre dont le prix est plus avantageux ("24" au lieu de "25"), peut vouloir y voir affiché (comme dans la figure 3) le « Nombre de page » ("430") et le « Rating » ("**") qui ne sont pas mentionnés et qu'il avait remarqué pour {Auteur2, Titre2} lorsqu'il était positionné au temps "Maintenant". Pour ce faire il va activer une option qui permette de compléter les données manquantes (cette option ne fonctionne que si la valeur de la cellule est à "null") par des données qui ont été insérés après le temps positionné et sont valides au temps « Maintenant ».

Troisième ligne de la figure 2 :

L'offre de Vendeur1 pour {Auteur2, Titre3} y est toujours[8]. Par contre, la quatrième ligne de la figure 1, qui présentait une offre pour {Auteur3, Titre4}, n'est pas présentée puisqu'elle était apparue après le 22 mars 2007 10:10 et il n'y a pas dans les sources de données d'autre offre pour ce livre qui soit apparue avant cette date et soit encore valide au temps "Maintenant".

Dans la figure 4, l'ouverture de la colonne optionnelle « Valide jusqu'à » a permis de présenter en plus les lignes (en l'occurrence il n'y en a qu'une seule) qui ne sont plus valides maintenant, bien qu'étant les plus récentes[9] au temps indiqué par la position du curseur temporel (au 22 mars 2007 10:10). Ainsi, en remplacement de la quatrième ligne de la première figure qui avait disparu de l'affichage dans les deux figures précédentes, pour {Auteur3, Titre4} est maintenant présentée une ligne qui était bien valide au temps 22 mars 2007 10:10 mais qui ne l'est plus depuis le 22 mars 2007 à 11:49 (ce qui explique donc qu'elle n'est pas montrée lorsque cette colonne n'est pas affichée).

L'offre de Vendeur1 pour le livre {Auteur2, Titre2} est plus avantageuse que celle de Vendeur2 et, bien que moins récente, elle est encore valide maintenant. Mais y a-t-il des offres encore plus avantageuses (et

---

[6] (qui, dans la figure précédente, était positionné sur "Maintenant")

[7] La ligne correspondante, qui est moins récente (« Valide depuis » : "1 mars 2007 12:00"), est encore valide au temps positionné : "|22 mars 2007 10:10". A noter que la colonne « Valide jusqu'à » n'étant pas ouverte, comme cette valeur "350" appartient à une donnée qui n'est plus valide Maintenant (sa date de fin de validité est le 22 Mars 12:00), une mise en œuvre différente pourrait ne pas la présenter.

[8] Etant toujours valide depuis le 15 mars 2007 11:10 elle était donc valide au temps positionné ("|22 mars 2007 10:10").

[9] (pour les valeurs déployées)

encore valides) ? L'utilisateur peut faire afficher toutes les offres valides maintenant des sources en présence, comme le présente la figure 5, en déployant toutes les valeurs de la colonne « Valide depuis ».[10]

L'utilisateur peut aussi demander à afficher les écarts de prix dans le temps pour chaque livre, comme le présente la figure 6, ou encore n'importe quelle autre fonction d'agrégation (telle que Min, Max, Moyenne, etc.) appliquée sur une cellule réduite, comme on va le décrire plus loin.

On suppose maintenant que l'utilisateur combine ces sources avec en plus la source de données ci-dessous "Vendeur4" comprenant les colonnes « Nombre de pages Min », « Nombre de pages Max », « Rating », « Vendeur » et « Prix », les trois premières étant la clé avec une condition sur la clé exprimant que le nombre de pages doit être entre les valeurs données dans les deux premières colonnes.

Vendeur4

Nombre de pages Min	Nombre de pages Max	Rating	Vendeur	Prix	Valide depuis	Valide jusqu'à
300	400	***	Vendeur4	23	22 mars 2007 10:01	null
400	500	**	Vendeur4	22	23 mars 2007 14:14	null

Le résultat est celui obtenu dans la figure 7 où Auteur2 Titre2 est donné par Vendeur4 au prix de 22 et si l'utilisateur se repositionne au temps "22 mars 2007 10:10" il obtient le résultat de la figure 8 où Auteur1 Titre1 est donné par Vendeur4 au Prix de 23.

Jusqu'ici on a fait l'hypothèse que les colonnes respectives des tableaux des différentes sources de données (Vendeur1, Vendeur2, Vendeur3 et Vendeur4) étaient déjà mises en correspondance, c'est-à-dire que par exemple « Auteur principal » du tableau "Vendeur1" correspond à « Auteur principal » du tableau "Vendeur2" (ces colonnes ayant pu être libellées différemment dans les différentes sources !).

On va maintenant décrire le procédé de mise en correspondance – et ensuite le procédé de suggestion ou application automatique de mises en correspondance – des sources de données multidimensionnelles et de leurs dimensions (colonnes) respectives.

L'intérêt principal du procédé est d'unifier les vocabulaires des sources combinées.

Le principe d'une interface utilisateur de *mise en correspondance de colonnes* est schématisé dans les figures 9 et 10. La figure 9 montre que le tableau B étant combiné avec le tableau A, et que la colonne Col5 de B étant glissée-déposée entre les colonnes Col2 et Col3 de A, les valeurs correspondantes de Col5 de B se trouvent affichées dans le tableau résultant A+B au sein d'une nouvelle colonne Col5 placée entre Col2 et Col3[11]. La figure 10 montre que le tableau B étant combiné avec le tableau A, et la colonne Col5 de B étant glissée-déposée sur la colonne Col2 de A, ces deux dernières *sont mises en correspondance* et ainsi, en résultat de combinaison les valeurs appropriées de Col5 de B se trouvent affichées dans le tableau résultant A+B au sein de la même colonne Col2 (appelée « Col2(Col5) » dans la figure). Les figures représentent au moyen de traits traitillés les régions permettant de distinguer (lors de la détection d'événements « on-drop ») ces deux cas de glisser déposer.

---

[10] La figure montre également un ensemble de curseurs temporels, un par date de début de validité (date de première apparition) présente dans le tableau. Placer (manuellement) plusieurs curseurs temporels signifierait (pour l'utilisateur) faire afficher l'union des lignes du tableau correspondant auxdits curseurs placés. Le fait de déployer toute la colonne « Valide depuis » (ceci pourra se faire en cliquant sur un symbole "ˇ" non montré dans la figure, mais utilisé plus loin) revient en effet à faire afficher l'union des lignes représentant les offres respectivement les plus récentes par rapport à ces curseurs temporels.

[11] Bien entendu, on fait ici l'hypothèse qu'une colonne clé a été mise en correspondance au préalable pour permettre la jointure entre A et B.

En mettant en correspondance la colonne Col5 de B et la colonne Col2 de A, l'utilisateur indique au système que ces colonnes contiennent des valeurs qui peuvent être combinées.[12]

Les mises en correspondance, par des utilisateurs, de tableaux de sources de données, et de colonnes entre tableaux, sont *comptées* et ceci permet de déterminer quelles mises en correspondance suggérer (ou appliquer par défaut) aux utilisateurs automatiquement[13].

Des poids sont associés aux mises en correspondance lors de leur comptage de manière à ce que les règles de prépondérance utilisées privilégient les mises en correspondance faites par des utilisateurs « proches », par exemple les utilisateurs travaillant dans le même domaine. Et, bien entendu, les mises en correspondance faites par l'utilisateur lui-même lui sont proposées en premier.[14]

Ainsi de nouvelles sources de données peuvent être combinées automatiquement par défaut, pourvu qu'elles aient déjà été combinées précédemment. Par exemple, un utilisateur crée lui-même une source de données "Vendeur5" (par exemple à partir d'une source déjà existante, en l'occurrence à partir de "Vendeur1") et présente l'offre de vente d'un livre "Auteur1" "Titre1" (par exemple un livre d'occasion qu'il voudrait revendre). Un autre utilisateur qui accède à "Vendeur1" prend connaissance de l'offre de "Vendeur5" par le simple fait qu'un nombre relativement grand d'autres utilisateurs ont déjà combiné "Vendeur5" avec "Vendeur1" et mis leurs colonnes respectives en correspondance. Comme déjà décrit, si l'offre de "Vendeur5" est la plus récente, ledit autre utilisateur verra l'offre de "Vendeur5" au lieu des offres des autres vendeurs ; sinon, il pourra la voir en se déplaçant dans le passé (en déplaçant le curseur temporel vers la gauche). Dans cette approche de combinaisons par défaut, un moyen graphique sera offert à l'utilisateur pour faire disparaître de l'affichage des valeurs provenant d'une source combinée par défaut, c'est-à-dire pour refuser la combinaison en question, ou défaire une mise en correspondance de colonnes effectuée par défaut, et ces refus sont comptabilisés dans les comptages mentionnés ci-dessus pour influencer la détermination des suggestions ultérieures.

De manière plus fine, les données présentées elles-mêmes peuvent être prises en compte dans les comptages. Reprenons l'exemple ci-dessus avec "Vendeur5" et précisons-le. L'utilisateur qui accède à "Vendeur1" va prendre connaissance de l'offre de "Vendeur5" non pas dans tous les cas, mais seulement dans le cas où "Auteur1" "Titre1" lui est présenté (dans le tableau "Vendeur1"), car c'est précisément lorsque "Auteur1" "Titre1" leur était présenté qu'un nombre relativement grand d'autres utilisateurs avaient

---

[12] Dans les exemples précédents il est entendu que la colonne « Auteur principal » du tableau "Vendeur1" et la colonne « Auteur principal » du tableau "Vendeur2", et ainsi de suite, avaient été mises en correspondance (à moins que bien sûr un dispositif induise les mises en correspondance par reconnaissance automatique de similarité des colonnes en question).

[13] La conception de la mise en œuvre de ce procédé est triviale pour le comptage et la suggestion de combinaisons. En ce qui concerne la suggestion de mises en correspondance de colonnes, la mise en œuvre peut consister à, pour chaque paire parmi les colonnes se trouvant dans la combinaison courante de tableaux, considérer l'ensemble des utilisateurs ayant combiné les tableaux en question (où se trouvent les colonnes formant cette paire) et qui ont gardé cette combinaison sous forme de version mémorisée (i.e. sous forme de « vues »), et compter le nombre de fois que cette paire est mise en correspondance dans ladite version mémorisée (prendre par utilisateur la moyenne pour toutes les versions mémorisées où cette combinaison est gardée) ainsi qu'éventuellement le nombre de fois qu'une suggestion de ladite mise en correspondance a été refusée par un utilisateur. Dans le cas où le nombre résultant du comptage est grand, et que le nombre correspondant des refus de suggestion est négligeable, suggérer ladite mise en correspondance lors d'une nouvelle combinaison (ou effectuer les mises en correspondance de colonne par défaut lors des combinaisons), les seuils "grand" et "négligeable" dépendant de la popularité des tableaux en question. A noter toutefois que le fait de gérer des vues implique d'éviter les cycles (une première vue faisant référence à une deuxième vue qui fait directement ou indirectement référence à la première) ; il suffit pour cela que le système refuse d'enregistrer sous le même nom une vue qui est référencée dans une autre vue et qui introduit un cycle du fait d'une référence qu'elle contient. Par suggérer on entend suggérer ou appliquer par défaut.

[14] En outre des droits d'accès peuvent être associés aux combinaisons et mises en correspondance, de manière à ce que par exemple les combinaisons faites par un utilisateur puissent lui être réservées à lui seul.

combiné "Vendeur5" avec "Vendeur1" (et non pas lorsqu'ils visualisaient des données sur n'importe quels autres livres). Ainsi, lesdits comptages peuvent en plus prendre en compte les données[15] visualisées par l'utilisateur lors des combinaisons.

Voici un exemple plus complet : Un extracteur fournit une source de données "Yamazuki" à partir du site Web du grand constructeur de motos Yamazuki qui présente toutes les motos de cette marque, avec toutes leurs caractéristiques.

Yamazuki

Type de moto	Caratéristiques...	Valide depuis	Valide jusqu'à
RS750	...	20 mars 2007 10:00	null
...			

Un particulier publie une source de donnée "Je vends" contenant une ligne présentant le type de moto (comme valeur clé), les détails, le prix et le lieu de vente d'une moto Yamazuki récente qu'il met en vente.

Je vends

Type de moto	Détails...	Prix	Lieu	Valide depuis	Valide jusqu'à
RS750	...	5000	Fontainebleau	23 mars 2007 17:00	null

Ensuite, lui-même et/ou d'autre(s) utilisateur(s) combinent cette source "Je vends" avec la source "Yamazuki", en mettant en correspondance la colonne qui identifie le type exact de la moto mis en vente.

Yamazuki + Je vends

Type de moto	Caratéristiques...	Détails...	Prix	Lieu	Valide depuis	Valide jusqu'à
RS750	...	...	5000	Fontainebleau	23 mars 2007 17:00	null
...						

Lorsqu'un utilisateur final va visiter le site de Yamazuki, et qu'il visualise les données sur le type de moto qui se trouve être celui de la moto que le particulier a mis en vente, l'offre de ce dernier va lui être spontanément présentée seulement si où le nombre de fois que "Je vends" a été combinée avec "Yamazuki" est relativement important.

Dans le cas contraire, c'est-à-dire même s'il y a trop de sources à combiner avec la source Yamazuki pour ce type de moto, en concurrence avec la source "Je vends", l'offre du particulier pourra être présenté par défaut *si l'utilisateur final s'intéresse dans la même session au lieu "Fontainebleau" qui se trouve être le lieu de vente de cette moto*. En effet la concurrence de données à combiner avec la source Yamazuki (pour la moto RS750) sera alors réduite.

Le scénario est le suivant : L'utilisateur final accède dans la même session non seulement au site "Yamazuki" mais aussi à un site "Châteaux" dans lequel l'utilisateur sélectionne[16] la ligne Fontainebleau. Dans ce cas, dans la mesure où la source "Je vends" est automatiquement combiné par défaut à ces deux sites, l'offre de la moto du particulier est présentée :

Yamazuki + Châteaux + Je vends

Type de moto	Caratéristiques...	Lieu	Détails...	Prix	Valide depuis	Valide jusqu'à

[15] Plus précisément les valeurs déployées de ces données, ce concept étant décrit plus loin.

RS750	...	Fontainebleau	...	5000	23 mars 2007 17:00 null

Le procédé d'enrichissement de sélections obtenues respectivement à partir des sources de données accédées dans une même session, permettant de mettre en oeuvre l'exemple ci-dessus, est décrit plus loin.

De manière encore plus fine, le contenu même des données présentées peut être pris en compte dans les comptages. Considérons l'exemple simple[17] suivant où les valeurs d'une colonne particulière sont prises en compte dans les comptages. Un utilisateur accède sur le Web à un moteur de recherche et lui fournit un mot-clé "mouche" représentant son domaine d'intérêt particulier. Un extracteur[18] présente, sous forme de tableau, le résultat retourné par le moteur de recherche[19] comme suit :

Moteur de recherche

Mot-clé	URL	Domaine	Valide depuis	Valide jusqu'à
mouche	...	Pêche à la mouche	23 mars 2007 17:00	null

...

On suppose ici que le moteur de recherche fournit, dans une colonne « Domaine », le domaine (en l'occurrence " Pêche à la mouche") correspondant au mot-clé ("mouche") donné. Dans le cas où un nombre relativement grand d'utilisateurs avaient, en visualisant précisément la valeur " Pêche à la mouche", combiné avec ce site "Moteur de recherche" la source "Vendeur1" (on suppose ici que "Vendeur1" est un vendeur de livre spécialisé dans le domaine "Pêche à la mouche"), cette dernière y sera automatiquement combiné :

Moteur de recherche + Vendeur1

Mot-clé	URL	Domaine	Auteur principal	Titre	Vendeur	Prix	Valide depuis	Valide jusqu'à
mouche	...	Pêche à la mouche	Auteur1	Titre1	Vendeur1	25	23 mars 2007 17:00	null

...

A chaque source de données[20] est associé le degré de finesse des informations à prendre en compte lors des comptages.

On va maintenant voir un autre exemple et introduire un procédé de suggestion qui ne reflète pas seulement un cas précédent de mise en correspondance, mais un enchaînement implicite de plusieurs cas précédents de mises en correspondance.

Dans le tableau "Mes articles" ci-dessous, un utilisateur associe un article ("Titre10", "Auteur10") à un livre ("Auteur1","Titre1") qu'il considère comme étant très « populaire » dans le domaine de l'article.

Mes articles

Article Titre	Article Premier Auteur	Revue	URL	Date parution	Livre Auteur principal	Livre Titre	Valide depuis	Valide jusqu'à

[16] (filtre sur)
[17] (et qu'on peut sophistiquer à souhait, voir plus loin l'exemple avec les expressions régulières)
[18] On décrit plus loin la mise en œuvre d'extracteurs de données à partir de sites Web.
[19] (qui devient ainsi une source de données au sens de la présente invention)
[20] (ou chaque extracteur)

Titre10	Auteur10	Revue10	Url10	Juin 2006	Auteur1	Titre1	23 mars 2007 16:00	null

Il met ensuite en correspondance les colonnes « Livre Auteur Principal » et « Livre Titre » (qui identifient ledit livre très populaire dans "Mes articles") avec les colonnes « Auteur Principal » et « Titre » de la source de données "Vendeur1".[21]

Vendeur1 + Mes articles

Auteur principal (Livre Auteur principal)	Titre (Livre Titre)	Article Titre	Article Premier Auteur	Revue	URL	Date parution	Valide depuis	Valide jusqu'à
Auteur1	Titre1	Titre10	Auteur10	Revue10	Url10	Juin 2006	23 mars 2007 16:00	

Ainsi, comme déjà décrit, lorsque plus tard l'utilisateur accède à la source "Vendeur1" et s'intéresse à ce même livre, sa combinaison avec "Mes articles" lui est rappelée automatiquement et l'article "Titre10" "Auteur10" lui est présenté.

Mais même lorsque l'utilisateur accède à une autre source (disons "Vendeur2") pour laquelle la combinaison avec "Vendeur1" aurait été automatiquement suggérée, sa source "Mes articles" peut[22] lui être suggérée.

En effet, ceci est justifié par le fait que "Mes articles" lui aurait de toutes façons été suggérée pour être combinée *indirectement* via[23] "Vendeur1" (et l'utilisateur aurait pu simplement faire disparaître les lignes et minimiser ("hide") toutes les colonnes provenant de "Vendeur1" pour se retrouver exactement dans le même cas).

Ainsi, une « chaîne de correspondances » existant entre "Vendeur2" et "Mes articles", et la correspondance de "Vendeur1" à "Mes articles" étant privilégiée (de poids fort) car établie par l'utilisateur lui-même, cette dernière source sera automatiquement combinée par défaut. La source "Mes articles" est ainsi rappelée à l'utilisateur même s'il ne se rappelle plus ni de son nom, ni même du nom de la source "Vendeur1" à laquelle il l'avait associé (combiné).

Bien évidemment, en fonction des règles de prépondérance utilisées, la combinaison de "Mes articles" avec "Vendeur1" ou "Vendeur2" sera aussi suggérée aux autres utilisateurs, dans la mesure où les sources en question leurs sont accessibles.[24]

Les fonctionnalités illustrées jusqu'ici sont rendues possibles grâce à un mécanisme de combinaison de sources de données multidimensionnelles, d'extracteurs (et synthétiseurs) de telles données à partir notamment de sites Web, et de gestion de la structure récursive de leurs attributs[25] multivalués (au sens que les valeurs d'un attribut multivalué peuvent elles-mêmes avoir un attribut multivalué), que nous allons maintenant décrire et illustrer davantage.

---

[21] A noter que l'on suppose ici que l'utilisateur a en plus « minimisé » (hide) les colonnes « Vendeur » et « Prix ».
[22] (en fonction des règles de prépondérence)
[23] Une chaîne d'indirection plus longue est ainsi aussi possible.
[24] En outre, on n'a pas considéré dans ce dernier exemple des degrés de finesses différents comme on l'avait fait dans les exemples précédents, ce qu'on aurait bien sûr pu faire.
[25] (dimensions, colonnes)

On va utiliser le concept de tag à deux parties (colonne et valeur), que l'on va appeler « **col:val** » (ou simplement **valeur**) ; le concept de relation entre des col:val, que l'on va appeler « **ligne** »[26] ; et le concept d'ensemble de lignes (tel qu'une table de BD relationnelle, chaque ligne étant alors une ligne de la table) que l'on va appeler « **tableau** ».

Exemple d'un tableau de deux lignes simples de trois valeurs :

1. Auteur principal:A, Titre:C, Prix:10
2. Titre:C, Prix:15, Editeur:D

Même tableau présenté sous forme tabulaire :

Auteur principal	Titre	Prix	Editeur
A	C	10	null
null	C	15	D

On va aussi utiliser le concept de « **ligne de modification** » qui spécifie une modification de lignes, par d'une part un ensemble de col:val donnés comme « **Clé** » et d'autre part un ensemble de col:val « **Non clé** » donnés comme valeurs de substitution lors des combinaisons.

Exemple d'un tableau de deux lignes de modification

1. Clé – Auteur principal:A, Titre:B ; Non clé – Titre:C, Prix:10
2. Clé – Titre:C ; Non clé – Prix:15, Editeur:D

ou sous forme tabulaire :

Clé		Non clé	(modification)	
Auteur principal	Titre	Titre	Prix	Editeur
A	B	C	10	null
null	C	null	15	D

Un « **tableau de modifications** » est un tableau de lignes de modification et un « **tableau simple** » est un tableau de lignes simples.

Tout tableau de modifications peut être vu comme un tableau de lignes simples. Ceci se fait en voyant chaque ligne de modification comme une ligne composée des col:val non clé qui y sont donnés et complétée, pour les col:val non inclus dans l'ensemble des col:val non clé qui y sont donnés, par des col:val clé qui y sont donnés. Par exemple le premier tableau présenté ci-avant représente un tableau de lignes simples tiré du tableau de modifications qui le suit.[27]

Pour une source d'informations apte à présenter des lignes simples, connaissant sa clé primaire (c'est-à-dire les colonnes identifiant ses lignes de manière unique), il est possible d'en extraire des lignes de modification.[28]

Exemple :

---

[26] (tel qu'une ligne de table de BD relationnelle, chaque col:val étant alors une valeur val dans une colonne col d'une ligne de la table – à noter que chaque ligne comprend implicitement un col:val ayant la valeur « null » pour chaque colonne non mentionnée)

[27] Bien entendu, un cas particulier de ligne de modification est le cas de ligne ne comprenant aucune valeur clé, et un cas particulier d'un tableau de modification est le cas où il n'y a aucune colonne clé.

[28] On verra plus tard que ceci est mis en œuvre dans les extracteurs.

Soit un tableau de lignes simples représentant une liste de livres, pour lequel on sait par ailleurs que les deux premières colonnes représentent la clé primaire :

Auteur principal	Titre	Prix
A	B	10
C	D	25

En voici la vue sous forme de tableau de modifications :

Clé		Non clé
Auteur principal	Titre	Prix
A	B	10
C	D	25

Un tableau simple (appelé premier tableau) peut être *combiné* avec un tableau de modifications (appelé deuxième tableau) en utilisant les valeurs clés de chacune des lignes du deuxième tableau pour réaliser, en fonction d'actions effectuées à l'aide d'une interface utilisateur sur des représentations des colonnes du premier tableau et du deuxième tableau, des associations entre les lignes du premier tableau et les lignes du deuxième tableau, en combinant les valeurs du premier tableau avec au moins une partie, déterminée également en fonction desdites actions, des valeurs non-clés du deuxième tableau, et en agençant les valeurs non-clés combinées avec des valeurs préexistantes également en fonction desdites actions. Lesdites actions sont des manipulations (telles que des glisser-déposer, comme déjà décrit) d'une représentation d'au moins une colonne du deuxième tableau pour la *mettre en correspondance* avec une représentation d'au moins une colonne du premier tableau (ou pour l'insérer entre deux colonnes du premier tableau), les colonnes correspondant aux représentations manipulées déterminant soit lesdites associations entre lignes, soit lesdits agencements de valeurs non clés du deuxième tableau avec des valeurs du premier tableau, selon que la colonne du deuxième tableau correspondant à la représentation manipulée contient ou non des valeurs clés.

Des conditions peuvent être associées à des colonnes clé et mémorisées sous forme de méta-données. Dans l'exemple du tableau "Vendeur4" donné précédemment (et reproduit ci-dessous[29]) comprenant les colonnes « Nombre de pages Min », « Nombre de pages Max », « Rating », « Vendeur » et « Prix », aux deux premières colonnes était associée une condition exprimant que le nombre de pages doit être entre les valeurs données dans ces deux premières colonnes. L'utilisateur qui cherche à mettre en correspondance une colonne (qui serait par exemple libéllée « #pages ») d'un premier tableau avec une colonne du tableau "Vendeur4" est alors invité à la mettre en correspondance avec le couple de colonnes « Nombre de pages Min », « Nombre de pages Max » au lieu d'une seule colonne.[30]

Clé			Non Clé			
Nombre de pages Min	Nombre de pages Max	Rating	Vendeur	Prix	Valide depuis	Valide jusqu'à
300	400	***	Vendeur4	23	22 mars 2007 10:01	null
400	500	**	Vendeur4	22	23 mars 2007 14:14	null

[29] (dans une table à clé, une colonne temporelle implicite « Valide depuis » existe en plus pour chaque colonne, comme décrit plus bas)

[30] Ces conditions sont alors vérifiées lors de la jointure des tables à clé décrite ci-après.

En plus des conditions, les méta-données peuvent contenir des actions. La ligne ci-dessous indique que si *condition :* la valeur en question correspond à l'expression "*Everest*8?844*", alors *action :* la transformer selon l'expression "*Everest*8?844<red>[Everest: 8 844,43 m]</red>]*".

Clé	Non clé
A trouver	A remplacer
*Everest*8?844*	*Everest*8?844<red>[Everest: 8 844,43 m]</red>]*

Enfin, les méta-données peuvent comprendre des indications et des conditions globales sur les sources de données à combinées.

Etant donné deux sources d'informations, la première source apte à fournir un tableau de lignes simples (ou un tableau de modifications vu comme un tableau de lignes simples), la deuxième source apte à fournir un tableau de modifications, une correspondance étant établie entre au moins une colonne du deuxième tableau (i.e. le tableau fournit par la deuxième source) et au moins une colonne du premier tableau (i.e. le tableau fournit par la première source), les lignes desdits tableaux pourront être combinées dans les cas où toutes les valeurs clé du deuxième tableau sont ainsi mises en correspondance avec des colonnes du premier tableau (même en l'absence de valeurs, c'est-à-dire même si elles ont une valeur nulle dans le premier tableau) et, si des valeurs manquent dans le premier tableau pour ces colonnes mises en correspondance, dans le cas où des colonnes clé sont données pour l'ensemble du premier tableau[31], toutes les col:val clés du premier tableau ont été mises en correspondance. Cette vérification peut être effectuée par un préprocesseur, avant la mise en œuvre du procédé de combinaison des tableaux proprement dite décrit ci-dessous.

A chaque valeur est associée un **début de validité** (i.e. temps de première apparition ou début de croyance de cette valeur). A chaque ligne est associée une période de validité : le début de validité de la ligne est égal au plus grand temps de début de validité associée à une valeur de la ligne, et la **fin de validité** de la ligne est son temps de dernière apparition[32] (ou temps de fin de croyance de cette donnée). Une fin de validité nulle signifie que la donnée est toujours valide (i.e. valeur toujours publiée par sa source ou valeur toujours crue).

Dans un tableau les lignes sont filtrées par rapport au curseur temporel positionné par l'utilisateur (comme illustré par les exemples donnés au début): seules les lignes ayant un temps de début de validité inférieur et un temps de fin de validité supérieur ou égal au temps indiqué par le curseur sont retenues (le temps du curseur indique le temps de croyance et seules les données crues au temps positionné sont considérés).

La mise en œuvre de la combinaison de tableaux utilise la notion de « table à clé ». Tout tableau peut être transformé en une table relationnelle, appellée « **table à clé** », où toutes les colonnes susceptibles d'être utilisées comme clé figurent comme colonne supplémentaire et contiennent la valeur de la clé pour la ligne en question.

La mise en œuvre du *procédé de combinaison d'un deuxième tableau avec un premier tableau* consiste à ajouter au premier tableau le résultat d'une jointure relationnelle (opérateur connu en soi) entre des tables à clé correspondant aux premier et deuxième tableaux (respectivement appelés première et deuxième tables à clé). Cette jointure s'effectue sur les valeurs clé se trouvant dans les colonnes de deuxième table à clé[33] mises en correspondance par l'utilisateur[34] en tenant compte des conditions[35] et/ou actions le cas

[31] On verra que l'information des colonnes clés peut être ainsi associée par un extracteur de données.
[32] (qui généralement nécessite d'être confirmée après une période d'incertitude)
[33] (par « colonne de table à clé » on entend les colonnes correspondantes dans le tableau correspondant)
[34] (ou par acceptation de suggestions automatiques de mises en correspondance de colonnes)

échéant associées se trouvant dans les méta-données (comme décrit plus haut) ; en fournissant lesdites valeurs clé avec pour chacune la plus grande valeur de début de validité[36] et, pour les autres valeurs des colonnes mises en correspondance, en fournissant les valeurs existantes ayant le début de validité le plus grand[37], les valeurs de début de validité associées auxdites valeurs fournies étant celles qu'elles avaient avant combinaison ; les lignes desdites tables à clé étant filtrées par rapport au temps positionné (curseur temporel, comme décrit plus haut).

Le procédé de déploiement / réduction décrit plus loin permet de ne présenter à l'utilisateur que les données les plus récentes (en réduisant la colonne « Début de validité », comme illustré plus haut dans les exemples).

Lorsqu'un résultat de combinaison de tableaux est combiné à son tour (en cascade), seules les lignes les plus récentes par rapport au temps positionné (curseur temporel) sont prises en considération.

On va maintenant tenir compte du fait que l'utilisateur qui accède à une source de données n'en visualise en général pas toutes les données en même temps, mais juste une sélection, et décrire un procédé *d'enrichissement* de sélections obtenues respectivement à partir des sources de données accédées dans une même session. Ce procédé est mis en œuvre en utilisant le procédé de combinaison de tableaux décrit ci-dessus et peut être vu comme en étant un perfectionnement.

Considérons que l'utilisateur accède tout d'abord à la deuxième source et ensuite à la première.[38] Appelons « première sélection » la sélection de données à partir de la première source et « deuxième sélection » la sélection de données à partir de la deuxième source.

Après la présentation à l'utilisateur d'une sélection de données à partir d'une deuxième source[39], lors de la présentation à l'utilisateur, dans la même session, d'une sélection de données à partir d'une première source[40], une mise en correspondance d'au moins une dimension ayant été faite[41] entre les deux sources, la première sélection va être « enrichie des deuxièmes et premières sources », à savoir : elle sera enrichie de la « combinaison » de la deuxième source avec elle et elle sera en plus enrichie en lui ajoutant la « combinaison » de la deuxième sélection avec la première source prise entièrement sauf la première sélection (puisque cette dernière a déjà été combinée avec la deuxième source prise en entier) ; par lesdites « combinaisons » on entend le procédé de combinaison déjà décrit plus haut.

Ensuite, lorsque l'utilisateur accède à une troisième source (et qu'une troisième sélection lui en est donc présentée), si une mise en correspondance d'au moins une dimension a été faite[42] avec la source de la sélection précédente (en l'occurrence la première), le même procédé est directement appliqué[43] (à noter

---

[35] (par exemple sous forme de « Where » ou « Having » en SQL)

[36] (entre la valeur de début de validité de ladite valeur clé dans la première table à clé et la valeur de début de validité de ladite valeur clé dans la deuxième table à clé)

[37] Autrement dit, entre les première et deuxième tables, pour chaque colonne non clé, lorsque la valeur existe des deux côtés c'est la valeur la plus récente qui est prise. Les informations sont ainsi complétées dans le temps. Ainsi, pour obtenir les valeurs « Prix » (comme présenté dans les exemples précédents), en SQL la clause SELECT comprendra une partie CASE WHEN Table2.[Prix_ValideDepuis] > Table1.[Prix_ValideDepuis] THEN ISNULL(Table2.Prix, Table1.Prix) ELSE ISNULL(Table1.Prix, Table2.Prix) END AS Prix.

[38] Ceci afin de pouvoir réutiliser la description du procédé de combinaison d'une deuxième source avec une première faite jusqu'ici.

[39] (dite « deuxième sélection »)

[40] (dite « première sélection »)

[41] (explicitement faite, ou appliquée par défaut, ou encore suggérée puis acceptée ; même remarque sous-entendue à chaque fois qu'on parle de mise en correspondance)

[42] (et que la ou les colonnes mises en correspondance sont susceptibles de couvrir les valeurs clés de la source précédente, du moins de la sélection précédente ; même remarque sous-entendue à chaque fois que ci-après on parle de mise en correspondance)

[43] (pour enrichir la troisième sélection à partir des première et troisième sources)

que si[44] la sélection précédente, en l'occurrence la première, n'avait pas été enrichie avec celle avant la précédente, en l'occurrence la deuxième, et qu'une mise en correspondance d'au moins une dimension avait été faite avec cette dernière, le même procédé aurait aussi dû être appliqué pour enrichir la troisième sélection avec les deuxième et troisième sources), sinon, si une mise en correspondance d'au moins une dimension a été faite avec celle avant la précédente, en l'occurrence la deuxième, le même procédé est appliqué pour enrichir la troisième sélection à partir des deuxième et troisième sources, sinon comme dans cet exemple aucune autre source n'a été accédée dans la session courante, la troisième sélection n'est pas enrichie.

Ainsi, en généralisant au cas de n sélections :

Lorsque l'utilisateur accède à une source courante, qu'une sélection courante lui en est présentée et que l'utilisateur avait dans la même session accédé à une source précédente et qu'une sélection lui en avait également été présentée,
SI une mise en correspondance d'au moins une dimension a été faite avec la source précédente, *la sélection courante est enrichie à partir des sources précédente et courante* (voir ci-après la définition de ces termes) et optionnellement, si l'utilisateur a dans la même session accédé à une source avant la précédente et qu'une sélection lui en a été présentée, dans le cas où la sélection précédente n'avait pas elle-même été enrichie avec celle avant la précédente, si une mise en correspondance d'au moins une dimension a été faite entre cette dernière et la source courante, la sélection courante est enrichie avec les sources avant la précédente et courante, et ainsi de suite jusqu'au début de la session, SINON
SI l'utilisateur a dans la même session accédé à une source avant la précédente et qu'une sélection lui en a été présentée, si une mise en correspondance d'au moins une dimension a été faite entre ladite source avant la précédente et la source courante, *la sélection courante est enrichie à partir de ladite source avant la précédente et la source courante* et optionnellement, si l'utilisateur a dans la même session accédé à une source avant 'la source avant la précédente' et qu'une sélection lui en a été présentée, dans le cas où la sélection avant la précédente n'avait pas elle-même été enrichie avec celle avant 'celle avant la précédente', si une mise en correspondance d'au moins une dimension a été faite entre cette dernière et la source courante, la sélection courante est enrichie avec les sources avant la précédente et courante, et ainsi de suite jusqu'au début de la session,
SINON une source encore antérieure est considérée et ainsi de suite jusqu'à qu'il y ait eu une mise en correspondance d'au moins une dimension entre une source précédente et la source courante ou qu'il n'y ait plus eu d'autre source précédemment accédée par l'utilisateur dans la même session.

Ledit *enrichissement de la sélection courante à partir d'une source précédente et de la source courante* consiste à ajouter à la sélection courante

- la combinaison de ladite source précédente avec la sélection courante et
- la combinaison de la sélection précédente avec la source courante de laquelle a été retirée la sélection courante.

A noter que le procédé ci-dessus est aussi applicable dans le cas d'une sélection vide[45]. Ceci est par exemple le cas lorsque, au sein d'une succession d'accès à des sites Web fournissant des données (vus comme sources de données via un extracteur, comme décrit plus loin), l'utilisateur visite la page d'accueil d'un site Web au lieu d'accéder directement à une page du site (qui elle en fournira par la suite une sélection de données).

Par session on entend une succession d'accès par l'utilisateur à des sources de données dont la combinaison est potentiellement pertinente. Typiquement, on va considérer que les accès rapprochés dans le temps

[44] (ce qui n'est pas le cas dans cet exemple)
[45] La combinaison d'une source avec une sélection vide résultant tout simplement en un tableau vide.

forment une même session. Une mise en œuvre particulière consiste à partitionner les sources accédées selon les dimensions qu'elles ont en correspondance[46].

On va maintenant décrire un procédé d'extraction/synthèse de données qui permettra d'effectuer les combinaisons (enrichissements de sélections) et déploiements directement au niveau des pages Web. La figure 11 présente schématiquement à gauche une page de résultats d'un site de vente de livres, groupés par auteurs et à droite le tableau résultant de son extraction[47].

L'utilisateur qui crée un extracteur lui associe des meta-données dans lesquelles il peut notamment indiquer quelles sont les colonnes clé[48] du tableau extrait. Il peut en indiquer plusieurs options. Ainsi pour l'exemple de la figure 11 il peut indiquer option1 : la colonne « ISBN », et option2 : le couple de colonnes « Auteur » et « Titre ». Lors de chaque combinaison le système choisira alors la première option (dans l'ordre des options données) qui fasse partie des colonne(s) mise(s) en correspondance. Par exemple, si l'utilisateur final met « Auteur » et « Titre » en correspondance lors d'une combinaison, c'est la deuxième option qui sera choisie.

Un extracteur fournit un tableau (simple ou de modifications) à partir des données en provenance d'une page web. Il doit donc indiquer d'une part la requête (url, paramètres GET ou POST) et d'autre part comment extraire les données de la page. Il peut également gérer la pagination et télécharger automatiquement plusieurs pages de résultats.

Le procédé de création d'un extracteur, à partir d'une page Web contenant un ensemble de données multidimensionnelles, est semi-automatique. Tout d'abord, l'utilisateur sélectionne dans la page Web un ou plusieurs objets correspondant chacun à une ligne du tableau, et indique quel objet de la page correspond à quelle ligne du tableau à générer. Le système compare les chemins de ces objets et construit classiquement un chemin générique (Xpath) couvrant au moins tous les objets indiqués par l'utilisateur.[49] Le système peut ainsi déterminer les valeurs pour chaque objet, et présenter le tableau ainsi obtenu à l'utilisateur.

Le synthétiseur est l'inverse de l'extracteur, il est créé automatiquement au moment de la création de l'extracteur correspondant, et permet d'afficher les données d'un tableau dans le style de présentation de la page Web, des zones graphiques étant placés à l'emplacement des objets contenant les valeurs du tableau pour permettre de les déployer ou réduire et de les glisser-déposer pour mettre en correspondance des colonnes de différents tableaux correspondant à différentes pages Web (i.e. à différents sites combinés comme nous le décrivons plus loin). Il est créé comme suit : L'utilisateur choisit un objet modèle correspondant à une ligne du tableau[50]. Tous les objets correspondant à d'autres lignes du tableau sont retirés de la page et tous les objets référencés par des objets correspondant à des lignes du tableau mais pas par l'objet modèle sont supprimés. Les valeurs contenues dans l'objet modèle sont modifiées pour

[46] Avantageusement on considérera ensemble (dans une même session) les sources accédées par l'utilisateur pour lesquelles des mises en correspondance de colonnes ont été suggérés à l'utilisateur et que l'utilisateur n'a pas rejeté.
[47] On voit que la colonne « auteur » répète les noms d'auteurs autant de fois que nécessaire, on verra plus loin comment pallier à ce problème grâce au procédé de déploiement réduction.
[48] (ou les colonnes clé « par défaut » dans le cas où des valeurs de ces colonnes peuvent être nulles)
[49] Dans une mise en œuvre préférée, tous les objets correspondant au chemin ainsi construit sont mis en évidence et l'utilisateur peut affiner le chemin en indiquant des objets additionnels ou en désélectionnant des objets mis en évidence. Le système affine alors le Xpath pour respecter ces contraintes. Lorsque l'utilisateur est satisfait de la sélection d'objets, il précise pour l'un de ces objets (l' « objet modèle ») tous les attributs qui correspondront aux colonnes du tableau. Pour chaque attribut, un objet dans la page, un nom de colonne et, si nécessaire, l'attribut HTML à extraire (par exemple, pour les liens, il a le choix entre la valeur de l'attribut href ou le texte du lien). Le système établit, pour chaque attribut, une paire (nom de colonne ; chemin Xpath), le chemin étant relatif à l'objet modèle, et enregistre cette information dans l'extracteur.
[50] (celui ayant servi comme modèle au moment de la création de l'extracteur, comme décrit dans la note précédente)

correspondre à la première ligne du tableau, et une copie de l'objet est insérée à la suite avec les valeurs de chaque autre ligne à afficher.[51]

On va maintenant brièvement introduire le procédé de déploiement/réduction offerte par l'affichage sous forme de tableau, au moyen d'un exemple (très schématique) de tableau présentant une liste de livres ayant les colonnes « Photo », « Auteur », « ISBN », « Titre » et « Langue ».

Dans la figure 12, seule la colonne « auteur » est *déployée*, ce qui fait qu'autant de lignes sont affichées qu'il y a d'auteurs différents. Comme ici tous les auteurs ont plus d'un livre, des boutons « déployer » (triangle dirigé vers le bas) permettent d'afficher la liste des livres écrits par un auteur donné.

L'affichage présenté dans la figure 13 est obtenu par clic sur le bouton « déployer » associé à la cellule Titre de la première ligne, ce bouton signifiant ici « déployer la liste des titres de l'auteur A1 ». Comme la colonne Auteur était déjà déployée, il est clair que les cellules de cette colonne, dans les lignes ainsi déployées, auront toutes la valeur A1. Afin d'alléger l'interface, A1 n'est donc indiqué qu'à la première des lignes déployées, les autres lignes laissant cette cellule vide.

Si maintenant l'utilisateur clique sur le bouton « réduire la liste des auteurs » (triangle dirigé vers le haut), tout le tableau se réduit à une seule ligne comme indiqué sur la figure 14. Des boutons « déployer » indiquent qu'il existe plus d'un auteur, plus d'un livre, plus d'une langue, etc. (figure 14).

Ensuite, à la figure 15, l'utilisateur a déployé la liste des langues. On voit qu'il n'y a pas de bouton « déployer » au niveau de A2 vu que dans cet exemple il n'y a pas d'autre auteur que A2 ayant un livre en anglais.

La même fonctionnalité est disponible lorsqu'un synthétiseur est utilisé : Les boutons « déployer » et « réduire » sont placés à l'emplacement de l'objet contenant la valeur de la cellule. Cependant certains synthétiseurs sont plus adaptés à un ordre de déploiement particulier. Par exemple un synthétiseur affichant les auteurs en évidence est plus adapté à un déploiement de la colonne Auteur en premier. On va maintenant décrire la sélection d'un synthétiseur approprié pour un ordre de déploiement de colonnes donné.

---

[51] Plus précisément, appelons « objet synthétisé » le plus petit objet contenant l'objet modèle ainsi que tous les objets correspondant à un attribut de la ligne modèle (appelons ces objets « objets attributs »), et soit o1, o2, …, oN la séquence d'objets dont chacun est parent du suivant, le premier est égal à l'objet synthétisé et le dernier égal à l'objet modèle. Une copie de l'objet synthétisé est effectuée, puis (dans le document lui-même) ses objets attributs sont modifiés pour correspondre à la première ligne affichée du tableau. Pour chaque ligne du tableau, est déterminé, dans l'objet synthétisé, le plus grand I (avec 1≤I≤N) tel que oI contient tous les objets attributs correspondant à des cellules non vides de la ligne courante. Une copie de oI (et donc également de oJ pour tous les J>I) est créée, ses objets attributs sont modifiés pour refléter la ligne courante, et elle est insérée à la suite (comme frère) de la dernière copie de oI à avoir été placée dans le document.
A noter que L'utilisateur peut demander à modifier un synthétiseur. Le même procédé ci-dessus est alors appliqué en se basant sur un tableau à une ligne contenant les noms des colonnes au lieu de valeurs, avec des marques spéciales permettant de les distinguer de texte normal (par exemple, « ${auteur} » dans la colonne auteur, et ainsi de suite). L'objet modèle est repéré par des marques spéciales (par exemple <model-object>…</model-object>). L'utilisateur peut modifier le document résultant à sa guise, par exemple à l'aide d'un éditeur de texte, et le renvoie au système. Pour afficher la page synthétisée, le procédé ci-dessus utilise désormais cette nouvelle structure (à condition qu'il y ait exactement une zone délimitée par les marqueurs d'objet modèle). A noter cependant qu'il est autorisé de supprimer ou dupliquer des marqueurs d'attributs. Il peut supprimer l'affichage d'un attribut qu'il juge peu important, et un exemple de duplication est de placer un attribut une fois à l'intérieur de l'objet modèle et une fois à l'extérieur, afin d'avoir une entête utilisant cet attribut, tout en affichant la valeur de l'attribut à chaque ligne de la liste affichée. Une autre application est de mettre la même valeur « url » comme texte et adresse d'un lien hypertexte (i.e. $url).

Pour un synthétiseur donné, à chaque colonne (affichée au moins une fois) peut être associé le plus petit objet oI (et donc le plus grand I, avec 1≤I≤N) contenant tous les marqueurs d'attributs correspondant à cette colonne. Ceci permet d'ordonner les colonnes selon l'importance leur étant attribuée par le synthétiseur (une petite valeur de I indique une importance plus élevée). On peut ainsi estimer dans quelle mesure un synthétiseur est approprié pour un ordre de déploiement de colonnes, en comparant l'ordre de déploiement avec l'ordre d'importance de ces colonnes selon le synthétiseur. Lorsque le système donne la liste des synthétiseurs pour une source donnée, cette liste pourra être triée selon ce critère, en fonction de déploiements déjà effectués par l'utilisateur, afin de permettre la sélection du synthétiseur.

Par ailleirs, le procédé *d'enrichissement* de sélections obtenues respectivement à partir des sources de données accédées dans une même session peut être appliqué à des sélections visualisées dans des pages Web jouant le rôle de sources de données (via extracteurs). Les mises en correspondance de colonnes peuvent être faites par l'utilisateur directement sur des présentations de ces sources de données (via synthétiseurs) sous forme de pages Web.

Deux sources étant ainsi affichées sous forme de pages Web positionnées de manière adjacente, l'utilisateur va simplement glisser-déposer une valeur d'une page Web sur une valeur d'une autre page Web pour mettre en correspondance les colonnes auxquelles ces valeurs appartiennent.

On va maintenant décrire le procédé de déploiement / réduction en détail. Les procédés décrits dans cette partie supposent l'existence d'un dispositif fournissant une interface similaire à un serveur de base de données et donnant accès au tableau affiché dans l'interface. Dans la suite on appelle ce dispositif la *source de données*. Un exemple typique est un calque extracteur.

La source de données stocke un « tableau » qui est une structure de données ayant un certain nombre de « colonnes » et de « lignes », et dont chaque ligne possède un certain contenu pour chaque colonne. En général, les lignes représentent des entités d'information et les colonnes des propriétés de ces entités, et il arrive fréquemment que pour une colonne, une même valeur se retrouve dans plusieurs lignes, par exemple dans le cas où une propriété d'une même entité peut par nature avoir plusieurs valeurs (on dit alors que la propriété en question est « multivaluée »). Lorsqu'il n'y a pas d'ambiguïtés, « le tableau » se réfère au tableau fourni par la source de données.

Accessoirement, pour permettre de réduire la quantité d'informations présentée, l'interface offrira un moyen d'appliquer un certain nombre de filtres sur les lignes (ou, autrement dit, d'effectuer une recherche dans le tableau). Lorsqu'un filtre sélectionne des lignes ayant une valeur spécifique dans une certaine colonne, on dit que la valeur de cette colonne est « spécifiée ». De manière plus générale, il peut aussi être possible d'imposer des *contraintes* (« Spécifier » une valeur de colonne devient alors un cas particulier de contrainte). Par exemple un filtre peut sélectionner les lignes contenant dans une colonne un mot donné. [52]

Lorsque la valeur d'une colonne est contrainte, une ligne n'ayant pas une valeur conforme à la contrainte pour cette colonne ne sera pas montrée dans le tableau affiché.

Le procédé de l'invention permet, en présence de champs multivalués, de remplacer les lignes ayant les mêmes valeurs dans un ensemble de colonnes (les colonnes «déployées ») par une seule ligne « réduite ».

Ayant un certain nombre de colonnes réduites et d'autres non, l'affichage se présente essentiellement comme suit: il contient une ligne (ci-après appelées « lignes affichées » par opposition aux lignes du tableau) pour chaque combinaison de valeurs dans les colonnes déployées qui existe dans le tableau. Pour une ligne affichée et une colonne réduite données, s'il existe une seule valeur possible selon le tableau, cette valeur

[52] Afin de ne pas alourdir le texte on décrit la présentation des données en supposant qu'aucun filtre n'est appliqué. En effet, il suffit de les ajouter à ceux produits par les algorithmes lors de chaque accès au tableau.

est montrée. Dans le cas contraire, le nombre de valeurs existantes (ou l'une de ces valeurs[53], ou encore n'importe quelle autre fonction de ces valeurs) est indiqué, et un bouton permet de déployer ces valeurs.

Plus précisément, les colonnes du tableau représentant des types de données et les lignes des jeux de données (tuples) ayant les types respectifs, pour certaines lignes l'affichage de certaines colonnes peut se faire en mode « déployé » et les lignes affichées comprennent alors les caractéristiques suivantes :

1. Chaque ligne affichée représente le sous-ensemble des tuples (du tableau) qui ont les valeurs présentées dans les colonnes déployées respectives[54].

2. Dans chaque ligne affichée, dans chaque colonne non déployée un bouton « déployer » est affiché si dans le tableau il existe au moins un autre tuple qui a

    a. une valeur différente pour cette colonne

    b. et les mêmes valeurs pour les colonnes déployées respectives.

3. Lorsque l'utilisateur positionne le curseur de la souris sur le bouton « déployer » (effectue un « mouse-over »), la liste des valeurs existant pour la colonne en question, parmi les tuples ayant les mêmes valeurs pour les colonnes déployées, est affichée dans une « fenêtre surgissante » (« pop-up »).

4. L'utilisateur peut alors cliquer sur une des valeurs affichées dans ledit pop-up, ce qui aura pour effet de changer les valeurs indiquées dans la ligne contenant ce bouton déployer, pour montrer un tuple ayant la valeur choisie dans cette colonne. On parle de « rotation » de ligne pour décrire cette opération.

5. Lorsque l'utilisateur clique sur le bouton « déployer », lesdits tuples (ayant une valeur différente pour la colonne en question et les mêmes valeurs pour les colonnes déployées – comme indiqué en point 2.) sont affichées (à noter que dans ces lignes déployées il y a maintenant une colonne déployée supplémentaire) et alors

    a. un bouton « réduire » remplace le bouton « déployer »

    b. comme indiqué en point 2., pour chaque ligne ainsi déployée, un bouton « déployer » est affiché dans chaque colonne pour laquelle au moins un tuple ayant une valeur différente existe dans le tableau.

6. En cascade, il est ensuite possible de déployer des cellules des lignes déployées au point précédent, et ainsi de suite jusqu'à obtenir un tableau ne comportant plus de bouton « déployer » et dont les lignes affichées correspondent exactement aux lignes du tableau.

7. On peut imaginer qu'il existe en plus une colonne virtuelle déployée appelée « ligne » (regroupant toutes les lignes) et que donc initialement une seule ligne représentant toutes les lignes du tableau peut être affichée, des boutons « déployer » figurant dans chaque colonne pour laquelle dans le tableau au moins un autre tuple ayant une valeur différente pour cette colonne existe.

8. Cliquer sur le bouton « réduire » apparu au point 5.a. a pour effet d'annuler le déploiement ayant causé son apparition, ainsi que tous les déploiements ayant eu lieu dans les lignes déployées correspondantes. Ce bouton sera typiquement placé au même endroit que le bouton ayant servi à déployer une cellule, de sorte que cliquer deux fois sur un bouton de déploiement permet de voir brièvement les lignes correspondant à une ligne affichée.

---

[53] (comme illustré dans les exemples présentés ci-dessous)

[54] Les sous-ensembles correspondant aux lignes affichées sont deux à deux disjoints.

9. Au lieu de ne placer qu'un bouton « réduire » pour un déploiement de colonnes, au niveau de la première ligne, un tel bouton peut être placé à chacune des lignes ainsi déployées, afin de simuler l'effet d'une réduction suivie d'une (ou plusieurs) rotation (on parle de « rotation globale »), de sorte que les valeurs montrées dans la ligne après réduction soient les mêmes que celle de la ligne se trouvant au niveau du bouton « réduire » utilisé, avant la réduction.

10. Pour chaque colonne, l'utilisateur peut choisir une fonction d'agrégation pour représenter les cellules réduites.

11. Des moyens peuvent éventuellement être offerts pour sélectionner plusieurs boutons « déployer » ou « réduire » et les actionner en un seul clic.

12. Aux entêtes des colonnes sont associés des éléments d'interface utilisateur permettant de réordonner les colonnes les unes par rapport aux autres et de supprimer des colonnes.

13. L'utilisateur a également un moyen de fournir un « filtre » sur les lignes à afficher dans le tableau. Par exemple pour ne montrer que les lignes qui ont une certaine valeur dans une certaine colonne, ou qui ont deux colonnes données égales, ou même correspondant à une expression SQL arbitraire (i.e. toute expression valide comme paramètre de WHERE ou de HAVING, en prenant bien sûr les précautions d'usage pour empêcher des accès interdits) fournie par l'utilisateur.

14. fonction « cacher les autres valeurs » sur les cellules des colonnes déployées, similaire à la fonction de réduction de lignes, mais qui se comporte comme si elle restait déployée. Ensuite un moyen de changer (par « rotation ») ou sortir de cet état.

On entend par « sous-tableau » l'ensemble des cellules nouvellement affichées comme décrit au point 5 ci-dessus.

A un tableau de lignes affichées sont associées les informations suivantes (constituant une structure de type « état de tableau de lignes affichées »)

- Un ensemble c1,c2,... de colonnes à montrer,
- une association colonne->valeur[55] f1->v1,f2->v2,... pour chaque valeur contrainte[56],
- un ensemble de colonnes déployées d1->t1,d2->t2,...., les valeurs ti, optionnelles, indiquant dans la liste des lignes affichées, laquelle est placée en premier (on parle de « rotation globale », cf. point 9 ci-avant où cette notion a été introduite)
- pour chaque ligne, un ensemble ordonné de rotations, chacune représentée par une paire colonne->valeur ri->wi, correspondant aux valeurs sélectionnées par l'utilisateur comme dit au point 4 ci-dessus. L'ordre correspond à l'ordre chronologique des rotations effectuées, et ne contient que les associations dont le résultat est encore visible : si une rotation annule l'effet d'une autre, cette dernière et celles qui la précédent sont retirées de la liste.
- un ensemble de sous-tableaux, également décrits par une structure de type « état de tableau de lignes affichées » (comme décrit ci-après).

Les données affichées par un tableau correspondent alors à une requête SQL SELECT c1,c2,... WHERE f1=v1 AND f2=v2 AND ... GROUP BY d1,d2,... (les lignes comprenant des rotations devant être altérées comme décrit plus bas), suivie de requêtes similaires pour chacun des sous-tableaux.

Si une colonne ci utilise une fonction d'agrégation a et que ci n'est pas déployé, alors la requête contient a(ci) au lieu de ci.

---

[55] (une valeur ou une expression représentant une contrainte)

[56] Dans la suite on présente les requêtes pour le cas de valeurs *spécifiées* (f1=v1...), mais les mécanismes décrits fonctionneront aussi dans le cas de valeurs *contraintes*.

Sauf contre indication, les ensembles de sous-tableaux et rotations d'une telle structure est initialement vide.

Une nouvelle rotation d'une colonne r à une valeur w dans une ligne affichée L est traitée comme suit : On ajoute la paire r->w en fin de la liste de rotations, pour obtenir r1->w1, r2->w2, ..., rn->wn, r->w. Ensuite on ajoute cette séquence aux colonnes spécifiées, ainsi que l'association des colonnes déployées d1, d2, ... aux valeurs qu'elles prennent dans la ligne L. Si au moins une ligne est trouvée, ses valeurs sont affichées pour L. Dans le cas contraire, la première association (r1->w1) est retirée de la table T, et le processus recommence, jusqu'à ce qu'au moins une ligne soit trouvée.

Lorsque dans un tableau T, une cellule d'une ligne affichée L et d'une colonne réduite C est déployée, un nouveau tableau T' est inséré dans la liste de sous-tableaux de T, avec les paramètres suivants:

Les colonnes à montrer sont c1,c2,..., les mêmes que T ; les valeurs spécifiées du sous-tableau sont f1->v1,f2->v2,...,d1->L(d1),d2->L(d2),..., celles de T plus les valeurs de L pour les colonnes déployées ; les colonnes déployées de T' sont d1,d2,...,C, les mêmes que celles du tableau contenant L plus la colonne C. Les rotations indiquées par T pour la ligne L sont retirées de T, et placées dans T', pour la même ligne, à l'exception, le cas échéant, d'une rotation de la colonne C pour la ligne L qui, elle, est enregistrée comme paramètre de rotation globale pour C.

T' représente donc toutes les lignes du tableau correspondant à L.

Plusieurs variantes peuvent être utilisées dans l'interface pour représenter un sous-tableau T' :

1. Il peut être montré par-dessus T, à la manière d'un « pop-up »,
2. être inséré à la place de la ligne L (décalant les lignes affichées de T suivant L), en le distinguant des autres par un cadre,[57]
3. être inséré à la place de la ligne L, en le distinguant des autres par un changement de couleurs.

La réduction d'un tableau T', sous-tableau d'un tableau T et déployant la cellule à la colonne c d'une ligne L, est effectuée en :

1. retirant T' de la liste de sous-tableaux de T,
2. ajoutant à la liste de rotations de la ligne que T' remplaçait l'ensemble des rotations se trouvant dans T' et ses sous-tableaux et correspondant à la ligne se trouvant au niveau du bouton de réduction utilisé. Ces rotations sont ordonnées en respectant l'ordre donné dans les sous-tableaux, et en plaçant celles se trouvant dans les sous-tableaux les plus profonds en premier ;
3. Si le bouton de réduction utilisé n'était pas celui de la première ligne mais celui en face d'une valeur v, une rotation c->v est ajoutée en fin de la liste de rotation de la ligne L.

Etant donné que l'ensemble des données se trouvant dans le tableau est typiquement trop grand pour être intégralement téléchargé depuis la source de données, seules les données effectivement affichées à l'utilisateur sont demandées. Lorsque l'utilisateur effectue un déploiement (ou changement des règles de filtrage), le système interroge la source de données pour n'obtenir que les informations non encore visibles par l'utilisateur.

Lors de la réduction d'un ensemble de lignes, aucune information ne doit être demandée à la source de données car toutes les informations nécessaires sont déjà présentes : afin de déterminer quelle cellule doit

[57] Il n'est pas nécessaire que T' contienne toutes les colonnes : en effet, les valeurs des colonnes spécifiées sont déjà visibles dans le tableau parent, et peuvent être omises dans le sous-tableau, afin d'alléger l'affichage. Cependant pour éviter, dans les variantes 1 et 2, d'avoir des « trous » dans le sous-tableau causés par ces colonnes omises (afin de préserver l'alignement avec le tableau racine : le tableau qui n'est sous-tableau d'aucun autre), on pourrait éventuellement imposer que les colonnes montrées dans les sous-tableaux soient un intervalle contigu des colonnes montrées dans le tableau racine.

contenir un bouton de déploiement il suffit de parcourir la colonne correspondante dans le sous-tableau à réduire, et vérifier s'il contient des cellules ayant un bouton de déploiement ou si deux cellules ont une valeur différente.

Alternativement, au moment de déployer un ensemble de lignes, l'état de la ligne avant déploiement est enregistré afin de pouvoir le restituer au moment de la réduction. L'opération de réduction rotative va au plus changer les valeurs se trouvant dans la ligne, pas l'existence ou non de boutons de déploiement.

Lors du déploiement d'un ensemble de lignes, une requête est envoyée à la source de données, contenant les propriétés du sous-tableau sur le point d'être créé (colonnes, valeurs spécifiées et informations de rotation).

La source de données détermine le contenu du sous-tableau (les valeurs à montrer, et l'ensemble des cellules devant contenir un bouton de déploiement), et le renvoie à l'utilisateur.

Les données reçues remplacent la ligne contenant le bouton ayant été cliqué par l'utilisateur.

Afin d'éviter de détruire des données lors d'une réduction de lignes (exigeant de les demander à nouveau au serveur si l'utilisateur déploie les colonnes à nouveau), il est possible de conserver l'ensemble des sous-tableaux ayant été créés, et de simplement les rendre visible à nouveau lorsque l'utilisateur effectue un déploiement pour afficher un sous-tableau ayant déjà été construit.

Pour ce faire, à chaque bouton de déploiement peut-être associée une référence au sous-tableau à rendre visible lorsque ce bouton est activé. Lorsqu'un bouton est pressé pour la première fois, un sous-tableau est créé comme décrit plus haut, et une référence vers ce dernier est enregistrée dans le bouton de déploiement. Lorsque le sous-tableau est réduit, il est simplement rendu invisible, et la ligne réduite est rendue visible. Si l'utilisateur utilise le bouton de déploiement à nouveau, la référence au sous-tableau est trouvée, et ce dernier est simplement rendu visible à nouveau.

A noter que, lors du déploiement d'une ligne, même si la première ligne du sous-tableau contient les mêmes valeurs que dans la ligne réduite, les boutons de déploiement sont des objets différents, et ont donc leur propre référence à un sous-tableau correspondant. Ceci permet de conserver en parallèle les différents ordres de déploiement des cellules.

Dans les figures 16 à 33 on considère la source de données suivante fournissant 5 lignes de 3 colonnes :

Organisation	Employé	Projet
O1	E1	P1
O1	E1	P2
O1	E2	P1
O2	E3	P1
O2	E4	P1

Tout d'abord, la figure 16 présente le cas où toutes les colonnes sont réduites. A noter que dans cet exemple l'interface utilisateur présente dans chaque cellule réduite une valeur (plutôt que par exemple combien de valeurs différentes cette cellule représente) et que le jeu de valeurs (O1, E1, P1) montrées dans les différentes colonnes correspond à une ligne existant réellement dans la source de données.

La figure 17 présente le sous-tableau T' présenté suite au déploiement de la colonne Organisation (l'utilisateur ayant cliqué sur le bouton associé à O1) dans la ligne unique qui était présentée dans l'exemple précédent. Remarquer que ce bouton est alors remplacé par un bouton inverse qui permet de réduire O1 à

nouveau (et revenir donc à la situation de la figure précédente). Dans la colonne Organisation, toutes les valeurs existantes (à savoir : O1 et O2) sont alors présentées, avec pour chacune une valeur associée présentée dans chaque autre colonne, les valeurs présentées dans chaque ligne formant ensemble un tuple existant dans la source de données[58].

La figure 18 présente le sous-tableau T'' qui apparaît suite au clic sur le bouton associé à E1 dans l'exemple précédent (pour déployer les employés de l'organisation O1). On remarque qu'il n y a pas eu besoin de répéter O1 dans la deuxième ligne, ce qui permet une présentation plus agréable à lire[59].

La figure 19 met en évidence le sous-tableau T''' qui apparaît suite au clic sur le bouton associé au projet P1 de la première ligne dans l'exemple précédent (pour déployer les projets de E1 de O1). E1 est implicite dans la deuxième ligne.[60]

La figure 20 met en évidence le sous-tableau qui apparaît suite au clic sur le bouton associé à E3 dans l'exemple précédent. On constate que l'on voit maintenant les 5 lignes du tableau de la source de données et qu'il n'y a ainsi plus de cellule à déployer.

En repartant de l'exemple présenté dans la figure 16, la figure 21 présente l'état du tableau affiché, suite au clic sur le bouton associé à P1 (déployer les projets). Dans la colonne Projet, toutes les valeurs existantes (P1 et P2) sont alors présentées, avec pour chacune une valeur associée présentée dans chaque autre colonne.

La figure 22 présente le sous-tableau T'' qui apparaît suite au clic sur le bouton associé à E1 de la première ligne de l'exemple précédent (ce clic ayant pour but de déployer les employés participant au projet P1). Le tableau est ainsi directement déployé au complet et, comme le deuxième tuple de la source de donnée est ainsi déjà présenté, le bouton déployer de O1 à la première ligne n'a plus de raison d'être.

En partant de la figure 21 on peut aussi choisir de déceler les organisations de la première ligne avant les employés, c'est ce que présente la figure 23.

Ensuite, dans la figure 24, E1 de la première ligne est déployée, et dans la figure 25, E3 est déployée pour compléter l'affichage des lignes de la source de données.

Enfin, en partant du premier exemple, on peut aussi commencer par déployer les employés, comme le montre la figure 26. Comme pour toutes les lignes présentées il n'y a plus d'organisation à déceler, il n'y a plus de bouton déployer dans la colonne organisation, par contre il reste un projet à montrer d'où le bouton déployer sur P1 à la première ligne.

L'utilisateur peut ensuite arriver directement au tableau entièrement déployé en cliquant sur le bouton associé à P1 dans la première ligne, comme le montre la figure 27.

---

[58] En variante, au lieu (ou en plus) de présenter une valeur dans chaque colonne on peut aussi présenter une combinaison ou agrégation ou cardinalité des valeurs existantes ou d'ailleurs n'importe quelle autre information pertinente, ou même rien du tout.

[59] (cette valeur est implicite dans la première colonne de la deuxième ligne)

[60] L'interface présente ainsi un arbre (une structure hiérarchique) dont la racine est O1, E1 et E2 forment deux branches, et ou P1 et P2 sont les deux feuilles de E1.

REVENDICATIONS

1. Procédé de combinaison automatique de données multidimensionnelles en fonction de manipulations au niveau de leurs dimensions dans un environnement informatique comprenant un équipement informatique capable d'accéder à des sources de données multidimensionnelles, caractérisé en ce qu'il comprend les étapes suivantes :

(a) prévoir une première source de données multidimensionnelles ;

(b) prévoir au moins une deuxième source de données multidimensionnelles, au moins cette deuxième source étant telle que chacune de ses données comporte dans les différentes dimensions des valeurs clés et des valeurs non-clés ;

(c) identifier des actions effectuées à l'aide d'une interface utilisateur sur des représentations de certaines dimensions des sources ;

(d) en fonction desdites actions, combiner les sources de données en utilisant les valeurs clés de chacune des données de la ou des deuxièmes sources pour réaliser des associations entre les données multidimensionnelles de la première source et les données multidimensionnelles de la ou des deuxièmes sources et ainsi obtenir des données multidimensionnelles combinées, ladite combinaison étant effectuée (i) en ajoutant aux données multidimensionnelles de la première source au moins une partie des valeurs non-clés des données correspondantes de la ou des deuxièmes sources, et (ii) en agençant les valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée.

2. Procédé de combinaison automatique de données multidimensionnelles provenant d'une pluralité de sources de données, caractérisé en ce qu'il comprend une succession de mises en œuvre en cascade du procédé de la revendication 1, les données combinées d'une mise en œuvre donnée dudit procédé constituant une source de données pour une mise en œuvre suivante dudit procédé.

3. Procédé selon la revendication 2, dans lequel l'agencement des valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée comprend la sélection d'une valeur non-vide parmi les valeurs provenant des différentes sources.

4. Procédé selon la revendication 2, dans lequel l'agencement des valeurs non-clés ajoutées avec des valeurs préexistantes pour la même donnée multidimensionnelle combinée comprend la sélection d'une valeur parmi les valeurs provenant des différentes sources selon un processus décisionnel donné.

5. Procédé selon la revendication 4, dans lequel à chaque donnée est associée une période de validité, et dans lequel le processus décisionnel comprend la sélection d'une valeur appartenant à une donnée valide à une date donnée.

6. Procédé selon la revendication 4, dans lequel à chaque donnée est associée une date de première apparition, et dans lequel le processus décisionnel comprend la sélection d'une valeur provenant d'une donnée la plus récemment apparue à une date donnée éventuellement ajustable.

7 Procédé selon la revendication 6, comprenant, pour les données d'une source auxquelles n'est associée aucune date de première apparition, une étape de création d'une date de première d'apparition égale à la date à laquelle la donnée a été mise en jeu dans une combinaison de données pour la première fois.

8. Procédé selon la revendication 1, dans lequel au moins l'une des sources de données multidimensionnelles comprend au moins deux sources de données amont et des informations définissant une combinaison précédemment effectuée selon la revendication 1.

9. Procédé selon la revendication 1, dans lequel lesdites actions sont des manipulations sur interface graphique d'une représentation d'au moins une dimension de la deuxième source pour la mettre en correspondance avec une représentation d'au moins une dimension de la première source ou pour l'insérer entre deux dimensions de la première source, les dimensions correspondant aux représentations manipulées déterminant soit lesdites associations entre données, soit lesdits agencements de valeurs non clés de la deuxième source avec des valeurs de la première source, selon que la dimension de la deuxième source correspondant à la représentation manipulée contient ou non des valeurs clés.

10. Procédé de combinaison de données multidimensionnelles, comprenant les étapes suivantes :

- prévoir un accès à une pluralité de sources de données,

- mémoriser des informations de correspondance entre sources de données à partir d'opérations de combinaison effectuées par le procédé de la revendication 1,

- lors d'un accès à une source de données ayant déjà fait l'objet de combinaison avec d'autres sources de données, signaler l'existence desdites autres sources de données.

11. Procédé selon la revendication 10, dans lequel les informations de correspondance comprennent également des informations de correspondance entre les dimensions desdites sources et comprenant en outre, lors d'un accès à une source de données ayant déjà fait l'objet de combinaisons avec d'autres sources de données, signaler également les correspondances entre dimensions.

12. Procédé selon la revendication 10 ou 11, comprenant en outre l'exécution par défaut d'opérations de combinaison par le procédé de la revendication 1 entre la source de données à laquelle il est accédé et lesdites autres sources de données.

13. Procédé selon l'une des revendications 10 à 12, dans lequel les informations de correspondance sont mémorisées pour une pluralité d'utilisateurs, et l'étape de signalisation s'effectue en fonction de règles de prépondérance parmi les informations de correspondance.

14. Procédé de combinaison de données multidimensionnelles, comprenant les étapes suivantes :

- prévoir un accès à une pluralité de sources de données,

- mémoriser des informations de correspondance entre sources de données à partir d'opérations de combinaison effectuées par le procédé de la revendication 1,

- lors d'un accès à une source de données ayant déjà fait l'objet de combinaison avec d'autres sources de données, déterminer l'existence de correspondances en chaîne entre sources de données, et, en fonction de caractéristiques des informations de correspondance, exécuter sélectivement une opération de combinaison effectuée par le procédé de la revendication 1 entre la source de données à laquelle il est accédé et une source de données liées à celle-ci par une chaîne d'au moins deux correspondances.

15. Procédé d'enrichissement de données multidimensionnelles par combinaison automatique en fonction de manipulations au niveau de leurs dimensions dans un environnement informatique comprenant un équipement informatique capable d'accéder à des sources de données multidimensionnelles, caractérisé en ce qu'il comprend, après avoir appliqué à une source précédente de données une fonction de sélection pour obtenir une sélection précédente de données, les étapes suivantes :

- lors d'un accès à une source courante de données pour obtenir une sélection courante de données, déterminer l'existence d'au moins une mise en correspondance de dimensions entre les données des deux sources,

- si une telle existence a été déterminée, appliquer le procédé de la revendication 1 sur une première paire de première et deuxième sources de données constituées respectivement de la sélection courante et de la source précédente, et sur une seconde paire de première et deuxième sources de données constituées respectivement de la source courante, de laquelle est retirée la sélection courante, et de la sélection précédente.

16. Procédé selon la revendication 15, dans lequel la mise en correspondance de dimensions entre les données des deux sources est effectuée au cours de la mise en œuvre du procédé.

17. Procédé selon l'une des revendications 15 et 16, dans lequel l'accès aux sources s'effectue à l'aide d'un navigateur Web en ce que la mise en œuvre du procédé est effectuée par interception des requêtes vers des serveurs et extraction de données à partir de ces serveurs.

18. Procédé selon les revendications 16 et 17 prises en combinaison, dans lequel ladite mise en correspondance est effectuée par affichage des sélections de données et glisser-déposer de valeurs dans les dimensions devant être mises en correspondance.

19. Procédé selon la revendication 18, comprenant en outre une étape de synthèse pour afficher lesdites sélections dans leur environnement graphique et associer aux valeurs des moyens permettant le glisser-déposer.

20. Procédé selon la revendication 15, mis en œuvre de façon répétée lors de l'accès à une succession de sources de données, dans lequel, lors de l'accès à une source de données courante pour laquelle il n'existe aucune mise en correspondance de dimensions avec la source précédente, on recherche l'existence d'une mise en correspondance de dimensions entre la source courante et une source antérieure, et on applique le procédé d'enrichissement sur des paires de sources constituées par ladite source antérieure et toute source consultée plus récemment avec laquelle il existe une mise en correspondance de dimensions, puis sur une paire de sources constituée par ladite source antérieure ainsi enrichie et la source courante.

21. Procédé selon la revendication 20, mis en œuvre lors de l'accès à une succession de sources de données SN-2, SN-1 et SN, comprenant les étapes suivantes :

si une mise en correspondance entre des dimensions des sources SN-2 et SN-1 d'une part, et SN-1, SN d'autre part existe, mettre en œuvre le procédé entre les sources SN et SN-1 en utilisant comme source SN-1 le résultat du procédé selon la revendication 15 mis en œuvre sur les sources SN-1 et SN-2,

si aucune mise en correspondance entre les sources SN-2 et SN-1 n'existe, déterminer s'il existe une mise en correspondance entre les dimensions des sources SN-2 et SN et, dans l'affirmative, mettre en œuvre le procédé selon la revendication 15 d'une part sur les sources SN et SN-1 et d'autre part sur les sources SN et SN-2, et

si aucune mise en correspondance entre les sources SN-1 et SN n'existe, déterminer s'il existe une mise en correspondance entre les dimensions des sources SN-2 et SN et entre les sources SN-2 et SN-1 et, dans l'affirmative, mettre en œuvre le procédé selon la revendication 15 d'une part sur les sources SN-1 et SN-2 et d'autre part sur les sources SN et SN-2.

et ainsi de suite pour des sources SN-3, SN-4, etc.

22. Procédé pour manipuler la visualisation d'une ressource contenant des informations structurées sous forme de tableau à au moins deux dimensions, telle qu'obtenue notamment à partir du procédé selon l'une des revendications 1 à 9 ou 12 à 21, où une dimension du tableau est constituée de colonnes représentant des types de données, et une autre dimension du tableau est constituée de lignes représentant des jeux de données associées ayant les types respectifs, et en ce que le procédé comprend :

(a) l'affichage sous forme d'une ligne unique d'un groupe de jeux ayant tous même valeur dans une colonne donnée,

(b) l'affichage, dans la colonne donnée, de ladite valeur, et

(c) l'affichage, en association avec au moins une autre colonne, d'un indicateur signalant qu'il existe au moins deux valeurs dans cette autre colonne pour le groupe de jeux de données.

23. Procédé selon la revendication 22, caractérisé en ce qu'il comprend une étape consistant, en fonction d'actions exercées à l'aide d'une interface utilisateur d'entrée de façon associée à ladite autre

colonne au niveau de ladite ligne unique, à provoquer l'affichage des différentes valeurs prise par le groupe de jeux de données dans cette autre colonne.

24. Procédé selon la revendication 23, caractérisé en ce que ledit affichage est effectué valeur par valeur.

25. Procédé selon la revendication 24, caractérisé en ce que ledit affichage est effectué dans un menu surgissant (pop-up).

26. Procédé selon l'une des revendications 22 à 25, caractérisé en ce qu'il comprend en outre, en réponse à une action à l'aide d'une interface utilisateur d'entrée sur un indicateur, le déploiement de la ligne en question en un sous-tableau.

27. Procédé selon la revendication 26, caractérisé en ce que ledit sous-tableau comprend autant de lignes qu'il existe, dans ledit groupe de jeux de données, de valeurs différentes dans ladite autre colonne.

28. Procédé selon la revendication 26 ou 27, caractérisé en ce que chaque ligne dudit sous-tableau contient une valeur différente dans ladite autre colonne et représente un sous-groupe de jeux de données ayant tous cette valeur dans ladite autre colonne.

29. Procédé selon l'une des revendications 26 à 28, caractérisé en ce qu'il comprend la répétition des étapes (b) à (d) pour au moins l'une des lignes constituant ledit sous-tableau, répétition appliquée à au moins une autre colonne encore dans laquelle il existe au moins deux valeurs pour le sous-groupe de jeux de données correspondant à ladite ligne au moins.

30. Procédé selon l'une des revendications 26 à 29, caractérisé en ce qu'il existe un type supplémentaire virtuel « ligne », le tableau étant initialement présenté sous forme d'une ligne unique regroupant toutes les lignes, un indicateur étant affiché en association avec chaque colonne dans laquelle il existe au moins deux valeurs.

31. Procédé selon l'une des revendications 26 à 30, caractérisé en ce que chaque ligne dudit sous-tableau représente, pour les colonnes à partir des indicateurs desquelles le sous-tableau a été formé, une combinaison spécifique de valeurs différente.

32. Procédé selon l'une des revendications 26 à 31, caractérisé en ce qu'il comprend l'affichage d'un indicateur associé à ladite autre colonne après déploiement et sur lequel une action à l'aide d'une interface utilisateur d'entrée provoque le regroupement (réduction) dudit sous-tableau en ladite ligne unique.

33. Procédé selon l'une des revendications 26 à 32, caractérisé en ce que l'indicateur associé à ladite autre colonne comprend un symbole susceptible d'être orienté vers le bas ou vers le haut.

34. Procédé selon l'une des revendications 26 à 33, caractérisé en ce qu'il comprend l'affichage dans ladite autre colonne, d'une des valeurs prises par le groupe de jeux de données dans cette colonne.

35. Procédé selon l'une des revendications 26 à 34, caractérisé en ce qu'il comprend l'affichage dans ladite autre colonne, d'une combinaison des valeurs prises par le groupe de jeux de données dans cette colonne.

36. Procédé selon l'une des revendications 26 à 35, caractérisé en ce qu'il comprend l'affichage dans ladite autre colonne, d'une propriété telle que la cardinalité de l'ensemble des valeurs prises par le groupe de jeux de données dans cette colonne.

37. Procédé selon l'une des revendications 26 à 36, caractérisé en ce qu'il comprend une étape de détermination d'une clé de sélection de jeux de valeurs, déterminant un groupe de jeux de valeurs auxquels un changement de valeur dans une colonne au niveau d'une ligne affichée s'appliquera collectivement.

38. Procédé selon la revendication 37, caractérisé en ce que ladite clé de sélection est constituée par les valeurs affichées dans la ou les colonnes à partir des indicateurs desquelles un sous tableau a été formé.

39. Procédé selon la revendication 37, caractérisé en ce que ladite clé de sélection est constituée par les valeurs affichées dans toutes les colonnes, dont la valeur avant changement pour la colonne dans laquelle le changement est effectué.

40. Procédé selon la revendication 37, caractérisé en ce que ladite clé de sélection est constituée par la valeur avant changement dans la colonne en question.

41. Procédé selon la revendication 37, caractérisé en ce qu'il comprend une étape d'ajout à la clé de sélection d'une valeur affichée dans une colonne par l'indicateur de laquelle aucun sous-tableau n'a été formé, par une action spécifique à l'aide d'une interface utilisateur d'entrée.

42. Procédé selon la revendication 37, caractérisé en ce qu'il comprend une étape de suppression de la clé de sélection d'une valeur affichée dans une colonne par l'indicateur de laquelle aucun sous-tableau n'a été formé, par une action spécifique à l'aide d'une interface utilisateur d'entrée.

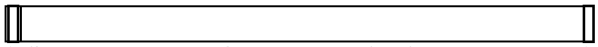
43. Procédé selon l'une des revendications 22 à 42, caractérisé en ce que, dans le cas où la ressource est construite dynamiquement à partir d'une source de données, il est apte à afficher en association avec la ressource un indicateur visant à permettre un accès direct à ladite source de données.

44. Procédé de présentation de données, comprenant les étapes suivantes :

(a) définir un modèle de présentation de données à partir d'une ressource dont sont issues une partie des données dont la visualisation est manipulée dans le procédé selon l'une des revendications 22 à 32,

(b) synthétiser au moins une partie desdites données manipulées dans ledit modèle de présentation selon des règles prédéterminées, et

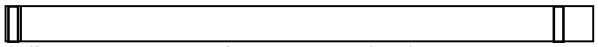
(c) lorsqu'à une desdites données est associé un indicateur, afficher un équivalent dudit indicateur dans la présentation de données.



|Glisser ce curseur pour ajouter un temps dans le passé |Maintenant : 23 mars 2007 15:20

Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur3	25	null	***	22 mars 2007 10:00
^Auteur2	^Titre2	Vendeur2	25	430	**	23 mars 2007 14:15
	^Titre3	Vendeur1	20	null	Null	15 mars 2007 11:10
^Auteur3	^Titre4	Vendeur2	12	62	Null	23 mars 2007 09:00

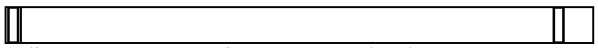
Figure 1



|Glisser ce curseur pour ajouter un temps dans le passé |22 mars 2007 10:10

Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur3	25	350	***	22 mars 2007 10:00
^Auteur2	^Titre2	Vendeur1	24	Null	null	22 mars 2007 10:05
^Auteur2	^Titre3	Vendeur1	20	Null	null	15 mars 2007 11:10

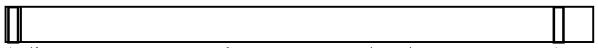
Figure 2



|Glisser ce curseur pour ajouter un temps dans le passé |22 mars 2007 10:10

Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur3	25	350	***	22 mars 2007 10:00
^Auteur2	^Titre2	Vendeur1	24	430	**	22 mars 2007 10:05
^Auteur2	^Titre3	Vendeur1	20	null	null	15 mars 2007 11:10

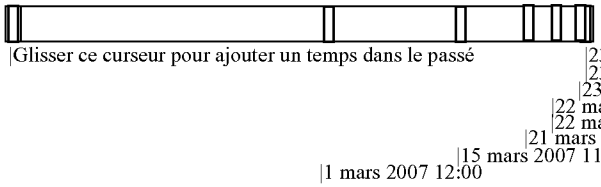
Figure 3



|Glisser ce curseur pour ajouter un temps dans le passé |22 mars 2007 10:10

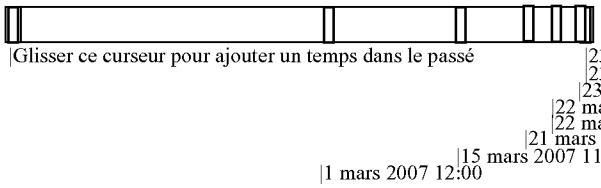
Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis	Valide jusqu'à
^Auteur1	^Titre1	Vendeur3	25	350	***	22 mars 2007 10:00	null
^Auteur2	^Titre2	Vendeur1	24	null	null	22 mars 2007 10:05	null
	^Titre3	Vendeur1	20	null	null	15 mars 2007 11:10	null
^Auteur3	^Titre4	Vendeur1	15	null	null	27 février 2007 11:50	22 mars 2007 11:49

Figure 4



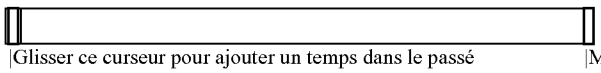
Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur2	26	350	null	^1 mars 2007 12:00
		Vendeur1	25	350	null	^21 mars 2007 08:15
		Vendeur3	25	350	***	^22 mars 2007 10:00
^Auteur2	^Titre2	Vendeur1	24	null	null	^22 mars 2007 10:05
		Vendeur3	26	null	**	^23 mars 2007 14:14
		Vendeur2	25	430	**	^23 mars 2007 14:15
	^Titre3	Vendeur1	20	null	null	^15 mars 2007 11:10
^Auteur3	^Titre4	Vendeur2	12	62	null	^23 mars 2007 09:00

Figure 5



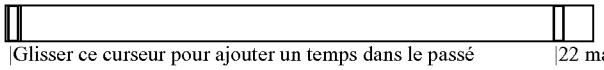
Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur2	□ = *1*	350	null	1 mars 2007 12:00
^Auteur2	^Titre2	Vendeur1	□ = *2*	Null	null	22 mars 2007 10:05
^Auteur2	^Titre3	Vendeur1	20	Null	null	15 mars 2007 11:10
^Auteur3	^Titre4	Vendeur2	□ = *3*	62	null	23 mars 2007 09:00

Figure 6



Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur3	25	null	***	22 mars 2007 10:00
^Auteur2	^Titre2	Vendeur4	22	430	**	23 mars 2007 14:14
	^Titre3	Vendeur1	20	null	null	15 mars 2007 11:10
^Auteur3	^Titre4	Vendeur2	12	62	null	23 mars 2007 09:00

Figure 7



Auteur principal	Titre	Vendeur	Prix	Nombre de pages	Rating	Valide depuis
^Auteur1	^Titre1	Vendeur4	23	350	***	22 mars 2007 10:01
^Auteur2	^Titre2	Vendeur1	24	null	null	22 mars 2007 10:05
^Auteur2	^Titre3	Vendeur1	20	null	null	15 mars 2007 11:10

Figure 8

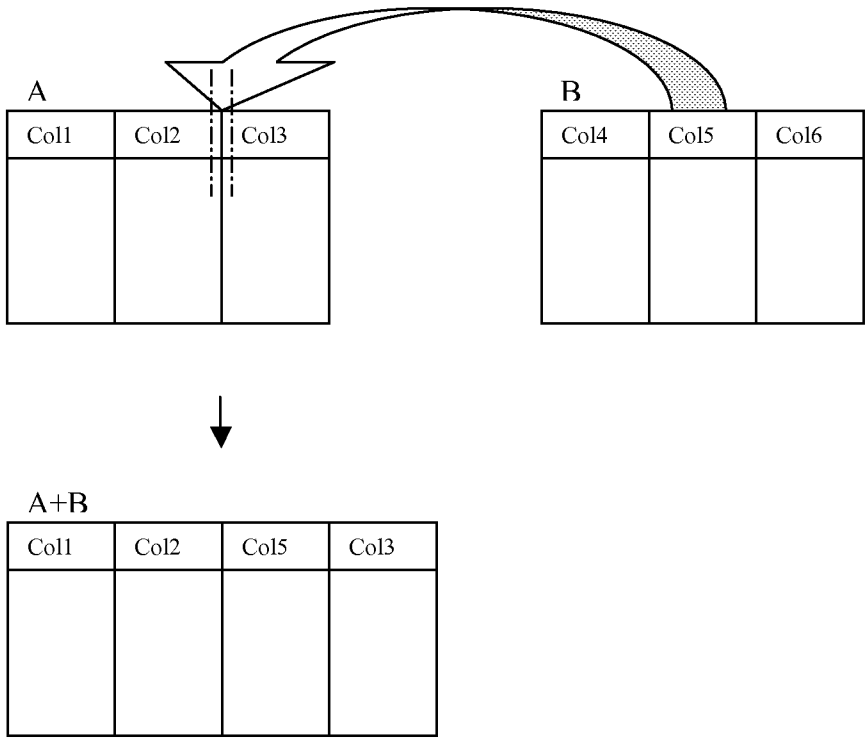


Figure 9

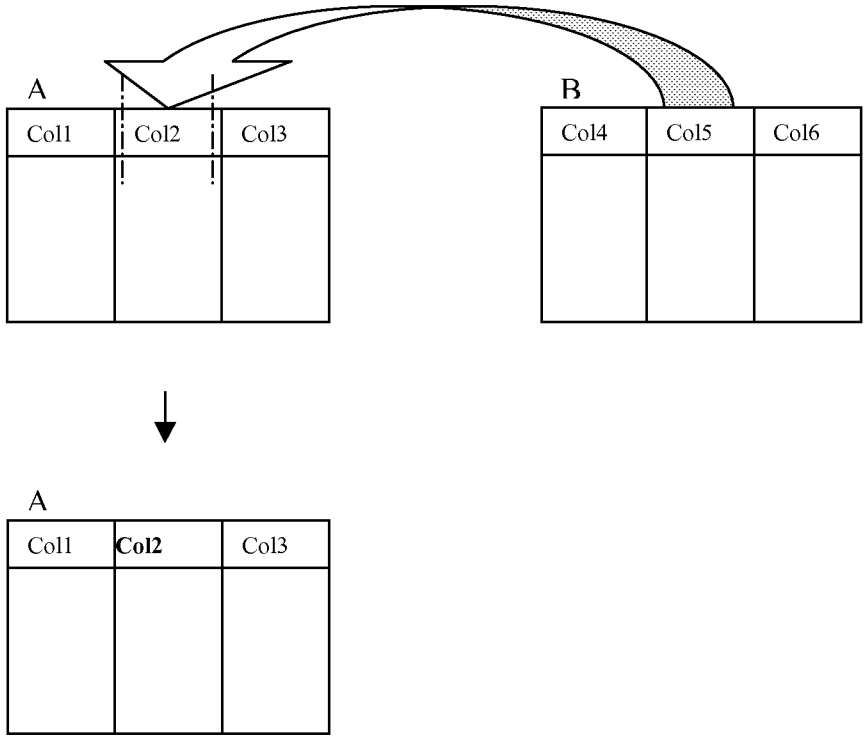


Figure 10

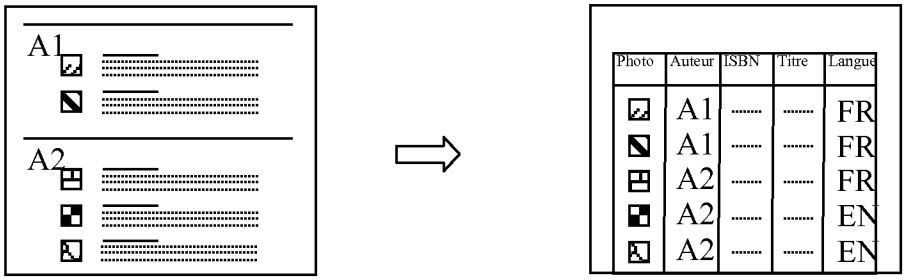


Figure 11

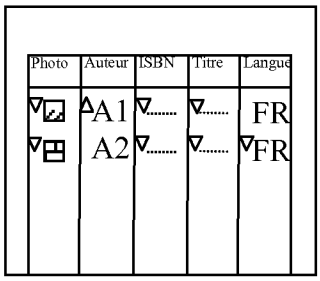


Figure 12

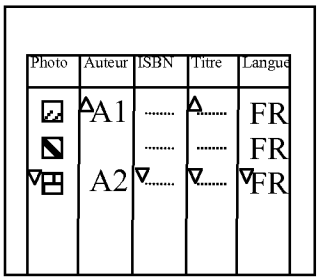


Figure 13

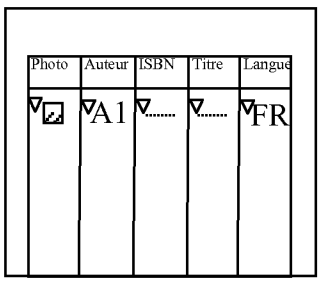


Figure 14

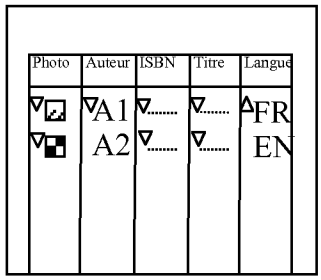


Figure 15

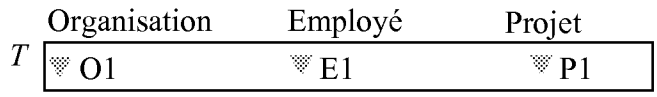


Figure 16

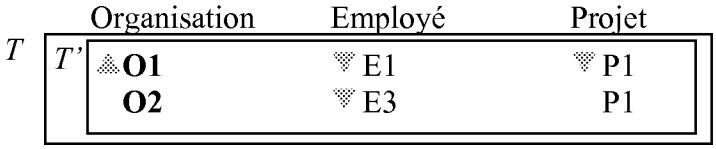


Figure 17

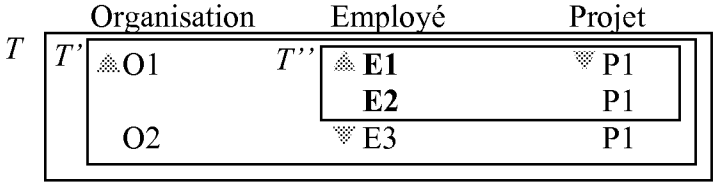


Figure 18

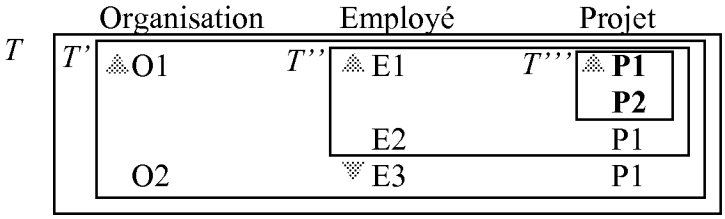


Figure 19

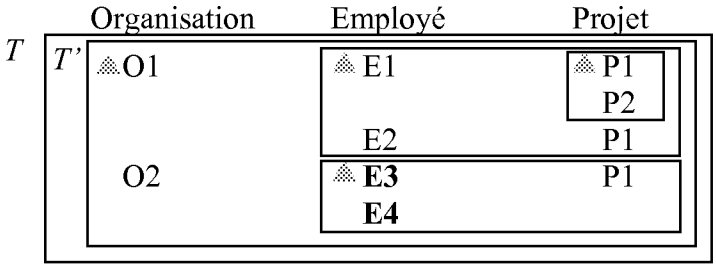


Figure 20

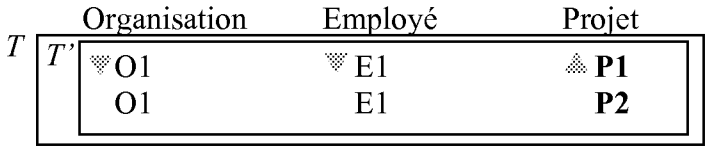


Figure 21

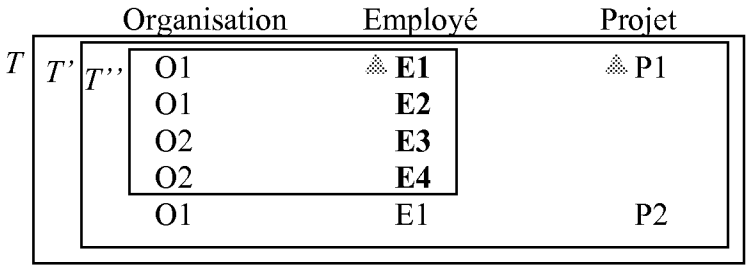


Figure 22

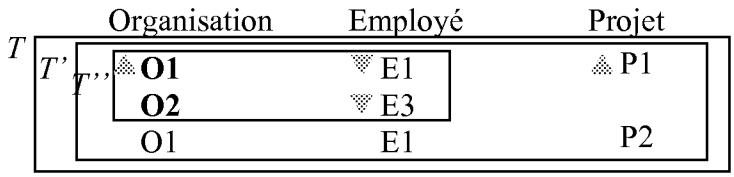


Figure 23

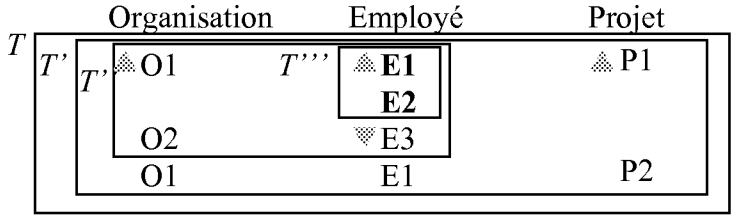


Figure 24

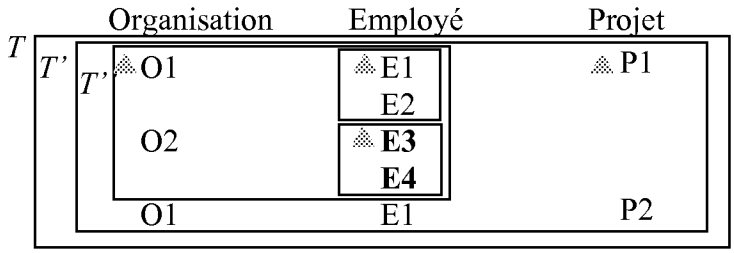


Figure 25

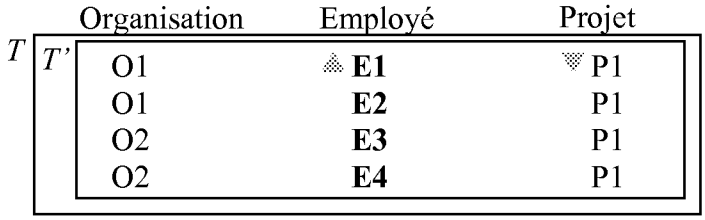


Figure 26

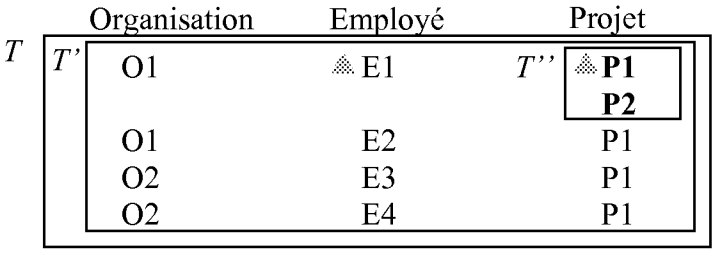


Figure 27

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/052274

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/24

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ALLEN G. TAYLOR: "SQL For Dummies." 14 August 2006 (2006-08-14), JOHN WILEY & SONS , XP002486106 ISBN: 0-470-04652-X	1-21
Y	pages 206-223 .	22-44
Y	WO 92/06437 A (HEWLETT PACKARD CO [US]) 16 April 1992 (1992-04-16) page 3, line 22 - page 4, line 26 page 13, line 24 - page 14, line 29 page 15, lines 24-29 figures 1,4A,4B	22-44

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 juillet 2008	21/07/2008

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Lopez, Patrice

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/052274

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2003/188259 A1 (AUREGLIA JEAN-JACQUES [FR] ET AL) 2 October 2003 (2003-10-02) abstract paragraphs [0014] - [0021] paragraphs [0024] - [0033] paragraphs [0120] - [0130] paragraphs [0197] - [0213]	22-44
Y	EP 0 400 620 A (SARATOGA SOFTWARE INC [US] MICROSOFT CORP [US]) 5 December 1990 (1990-12-05) abstract column 3, lines 12-43 column 9, line 49 - column 10, line 9 column 10, line 54 - column 11, line 23 column 11, line 49 - column 12, line 20	22-44
A	ANONYMOUS: "Dabble DB' s Features" INTERNET ARTICLE, [Online] 19 February 2007 (2007-02-19), XP002460295 Retrieved from the Internet: URL:http://web.archive.org/web/20070219132847/http://www.dabbledb.com/explore/features/> [retrieved on 2007-11-16] the whole document	1-44
A	-& ANONYMOUS: "Compact and Pivot" THE DABBLE BLOG, [Online] 4 August 2006 (2006-08-04), XP002460297 Retrieved from the Internet: URL:http://blog.dabbledb.com/2006/08/compact-and-piv.html> [retrieved on 2007-11-12] the whole document	1-44
A	-& ANONYMOUS: "A 60-second demo showing off the compact view" THE DABBLE DB BLOG, [Online] 4 August 2006 (2006-08-04), XP002460296 Retrieved from the Internet: URL:http://dabbledb.com/blog/clips/compactview.mov> [retrieved on 2007-11-16] the whole document	1-44
A	GLICK J.: "Autocomplete Addressing" MACHV/MOZILLA: MAIL & NEWS, [Online] 10 September 2001 (2001-09-10), XP002460298 Retrieved from the Internet: URL:http://www.mozilla.org/mailnews/specs/autocomplete/> [retrieved on 2007-11-16] page 6	1-44

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2008/052274

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9206437 A	16-04-1992	DE 69126295 D1	03-07-1997
		DE 69126295 T2	04-09-1997
		EP 0503060 A1	16-09-1992
		JP 5503388 T	03-06-1993
		US 5337405 A	09-08-1994
US 2003188259 A1	02-10-2003	NONE	
EP 0400620 A	05-12-1990	DE 69031149 D1	04-09-1997
		DE 69031149 T2	15-01-1998
		JP 2007386 C	11-01-1996
		JP 3163660 A	15-07-1991
		JP 7043707 B	15-05-1995

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/EP2008/052274

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
INV. G06F17/24

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
G06F

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela estréalisable, termes de recherche utilisés)
EPO-Internal

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	ALLEN G. TAYLOR: "SQL For Dummies" 14 août 2006 (2006-08-14), JOHN WILEY & SONS , XP002486106 ISBN: 0-470-04652-X	1-21
Y	pages 206-223	22-44
Y	WO 92/06437 A (HEWLETT PACKARD CO [US]) 16 avril 1992 (1992-04-16) page 3, ligne 22 - page 4, ligne 26 page 13, ligne 24 - page 14, ligne 29 page 15, ligne 24-29 figures 1,4A,4B	22-44

-/--

[X] Voir la suite du cadre C pour la fin de la liste des documents

[X] Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'inven tion revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'inven tion revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale
4 juillet 2008	21/07/2008

Nom et adresse postale de l'administration chargée de la recherche internationale	Fonctionnaire autorisé
Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Lopez, Patrice

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/EP2008/052274

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	US 2003/188259 A1 (AUREGLIA JEAN-JACQUES [FR] ET AL) 2 octobre 2003 (2003-10-02) abrégé alinéas [0014] - [0021] alinéas [0024] - [0033] alinéas [0120] - [0130] alinéas [0197] - [0213] ----	22-44
Y	EP 0 400 620 A (SARATOGA SOFTWARE INC [US] MICROSOFT CORP [US]) 5 décembre 1990 (1990-12-05) abrégé colonne 3, ligne 12-43 colonne 9, ligne 49 - colonne 10, ligne 9 colonne 10, ligne 54 - colonne 11, ligne 23 colonne 11, ligne 49 - colonne 12, ligne 20 ----	22-44
A	ANONYMOUS: "Dabble DB' s Features" INTERNET ARTICLE, [Online] 19 février 2007 (2007-02-19), XP002460295 Extrait de l'Internet: URL:http://web.archive.org/web/20070219132847/http://www.dabbledb.com/explore/features/> [extrait le 2007-11-16] le document en entier	1-44
A	-& ANONYMOUS: "Compact and Pivot" THE DABBLE BLOG, [Online] 4 août 2006 (2006-08-04), XP002460297 Extrait de l'Internet: URL:http://blog.dabbledb.com/2006/08/compact-and-piv.html> [extrait le 2007-11-12] le document en entier	1-44
A	-& ANONYMOUS: "A 60-second demo showing off the compact view" THE DABBLE DB BLOG, [Online] 4 août 2006 (2006-08-04), XP002460296 Extrait de l'Internet: URL:http://dabbledb.com/blog/clips/compactview.mov> [extrait le 2007-11-16] le document en entier ----	1-44
A	GLICK J.: "Autocomplete Addressing" MACHV/MOZILLA: MAIL & NEWS, [Online] 10 septembre 2001 (2001-09-10), XP002460298 Extrait de l'Internet: URL:http://www.mozilla.org/mailnews/specs/autocomplete/> [extrait le 2007-11-16] page 6 ----	1-44

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°
PCT/EP2008/052274

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9206437 A	16-04-1992	DE 69126295 D1	03-07-1997
		DE 69126295 T2	04-09-1997
		EP 0503060 A1	16-09-1992
		JP 5503388 T	03-06-1993
		US 5337405 A	09-08-1994
US 2003188259 A1	02-10-2003	AUCUN	
EP 0400620 A	05-12-1990	DE 69031149 D1	04-09-1997
		DE 69031149 T2	15-01-1998
		JP 2007386 C	11-01-1996
		JP 3163660 A	15-07-1991
		JP 7043707 B	15-05-1995