

La Volée 1974Concordance des classifications et étude comparative du recouplement des collections : les bibliothèques de la Faculté des Sciences de l'Université de Genève

Cynthia Dufaux

cynthia.dufaux@gmail.com

Jan Krause,

jan.krause@medecine.unige.ch

Bibliothèque de la faculté de Médecine, Genève

Aline Maurer

alinemau@gmail.com

Résumé

Cette étude évalue la collection de monographies de la Faculté des Sciences de l'Université de Genève. Sur cette base, une étude plus approfondie peut être menée dans le but de mettre en place une politique d'acquisition et une gestion des collections communes des sept bibliothèques de la faculté. Un tableau de concordance entre les sept classifications a été établi afin d'obtenir un référentiel commun et d'étudier les sujets couverts par plusieurs bibliothèques. Des scripts Python ont été développés pour obtenir des statistiques de recouplement des collections. L'algorithme UPGMA de GNU-R a été appliqué pour créer un dendrogramme, essai original de comparaison des collections. Les résultats mettent en évidence les recoulements et redondances entre les sept collections et un manque de consistance dans leur gestion.

En parallèle aux classifications en usage dans chaque bibliothèque, une classification commune basée sur le tableau de concordance faciliterait la mise en place d'une politique documentaire globale et cohérente. Affiner cette analyse requiert une étude des collections complètes, notamment en ce qui concerne les niveaux des documents et leurs usages.

Abstract

This analysis lends an initial assessment of the monograph collection of the Faculty of Science at the University of Geneva. These results can be used as a basis for a more detailed study aiming to help the establishment of a common management of the libraries' collections in the future. In order to allow the comparative study of the subjects covered by the libraries, a table of correspondence was constructed to provide insight and a common reference point for the classifications used in the seven locations. Statistical data concerning overlaps in the different collections were gathered using Python scripts, and a dendrogram was created using the GNU-R's UPGMA algorithm. The results showed some management inadequacies including cross-checking and redundancy across the seven collections.

In parallel with the currently used classifications in each library, a common classification based on the correspondence table can be implemented to facilitate the set-up of a global and coherent collection development and management policy. To take this study further, we can consider the entire collections and take into account the documents' levels and their usage.

Mots-clés

Etude des collections, Classification, Politique documentaire, Bibliothèque scientifique universitaire

1. Introduction

Cette étude a pour objet d'établir une concordance entre les classifications utilisées pour les monographies dans les bibliothèques de la Faculté des Sciences de l'Université de Genève, et d'étudier le recouplement documentaire d'une partie des collections de ces bibliothèques. De telles informations sont nécessaires pour débuter une étude détaillée des collections de ces bibliothèques, et finalement mettre en place une politique d'acquisition commune aux bibliothèques de la Faculté.

La Faculté des Sciences de l'Université de Genève comporte six bibliothèques, correspondant à des sections ou départements, soit les bibliothèques d'anthropologie, de mathématiques, de l'Observatoire, de physique, de Sciences II (biologie, chimie et sciences pharmaceutiques) et des Sciences de la Terre. Le Centre universitaire d'informatique (CUI) a été également intégré à cette étude car il est lié à plusieurs facultés dont celle des Sciences. Toutes ces bibliothèques sont situées à moins d'un quart d'heure de marche les unes des autres, sauf celle de l'Observatoire qui se trouve à 20 kilomètres.

En terme de collection (monographies et périodiques reliés), la bibliothèque de Sciences II est la plus importante avec 68'700 volumes, viennent ensuite les mathématiques avec 43'339 volumes, les Sciences de la Terre avec 43'000 volumes, la physique avec 34'360 volumes, l'anthropologie avec 20'600 volumes, l'Observatoire avec 20'095 volumes et le CUI avec 12'500 volumes.

Jusqu'à présent, ces sept bibliothèques fonctionnaient de manière relativement autonome. Elles sont sous la direction de leur section ou département respectif. Dans le but d'améliorer la qualité des services des différentes bibliothèques, plusieurs projets, tel un site Web commun, ont été lancés. De même, l'idée d'une politique d'acquisition commune à toutes les bibliothèques des Sciences est née.

Suite à l'arrivée d'une nouvelle Directrice de l'information scientifique à l'Université, un projet identique a été initié pour toute l'institution. Cette étude a débuté avant la création de ce poste, mais les résultats seront évidemment utiles dès le début du processus de création d'une politique d'acquisition, qu'elle soit faculaire ou institutionnelle.

2. Méthodologie

Cette étude a débuté par les visites des sept bibliothèques afin d'obtenir leurs classifications et de rencontrer leurs responsables pour avoir des renseignements supplémentaires concernant ces classifications (contexte, historique, évolution, etc.).

Un tableau de concordance a ensuite été établi entre la CDU (Classification décimale universelle) (UDC Consortium, 2004) et les classifications utilisées (voir tableau 1 pour un extrait de ce tableau). La CDU a été choisie comme référence car celle-ci est utilisée dans deux des bibliothèques concernées et possède une version récente datant de 2004 en français. Les concordances ont été définies manuellement par Aline Maurer et Cynthia Dufaux en comparant les diverses classifications à la CDU. Cela leur a été facilité par leurs connaissances en sciences, toutes deux étant de formation scientifique (respectivement biologie/botanique et sciences de la Terre).

Les classifications utilisées par les bibliothèques diffèrent dans leur degré de détails. Celle d'anthropologie est par exemple très générale, alors que celle du CUI, avec trois niveaux, est très détaillée. Le niveau hiérarchique de chaque indice CDU a été choisi au plus proche de celui de la classification locale, et est parfois plus élevé lorsque la CDU n'était pas assez détaillée. Pour la classification du CUI, tous les niveaux n'ont pas été pris en compte car ils étaient trop détaillés par rapport à la CDU.

Ce tableau a permis ensuite de déterminer les sujets communs à plusieurs bibliothèques. Seuls les indices CDU ayant un équivalent dans trois bibliothèques ou plus ont été retenus pour être étudiés, le nombre de documents concernés étant déjà conséquent.

A noter que dans le texte, le terme « domaine » est utilisé pour les thèmes représentés par deux chiffres dans la CDU, à l'exception de l'informatique (trois chiffres pour des raisons historiques). Les domaines sont par exemple les mathématiques, la physique, la biologie, etc. Le terme « sujet » désigne les subdivisions de ces domaines.

Afin de mieux comparer les sujets identiques dans plusieurs bibliothèques, une deuxième visite sur place a permis de déterminer de façon plus précise lesquels étaient traités par les documents présents dans les indices retenus.

Un programme informatique, comprenant deux scripts, a également été créé pour obtenir des statistiques concernant le nombre de documents se trouvant dans plusieurs bibliothèques, et donc les taux de recouplement entre les bibliothèques. Le premier script va chercher les notices bibliographiques dans le catalogue du Réseau des bibliothèques genevoises (Réseau des bibliothèques genevoises, 2007) pour un indice donné. La recherche est faite par classification, c'est-à-dire par un sigle décrivant la bibliothèque suivie de l'indice (ex : « ge-usa O » pour l'indice O en anthropologie). Cette classification est saisie lors du catalogage d'un document en zone 980_2 du format MARC21. Le deuxième script compare le résultat du premier avec l'indice en question pour éliminer les notices ne correspondant pas à la recherche (erreurs et doublons). Ce script produit également un tableau avec un indice recherché par ligne, et en colonne les bibliothèques possédant les documents représentés par chaque indice.

Les résultats obtenus permettent d'analyser la cohérence de l'échantillon étudié des collections de la Faculté des Sciences, et de faire des suggestions pour la mise en place d'une future politique d'acquisition commune aux sept bibliothèques.

3. Résultats

3.1. Tableau de concordance des classifications

Le tableau 1 présente un extrait du tableau de concordance (celui-ci comportant environ 110 pages, il n'est pas inclus dans cet article mais est disponible sur demande).

CDU		Anthropologie	Observatoire	CUI	Mathématiques	Physique	Science de la Terre	Sciences II
510	Considérations fondamentales et générales sur les mathématiques. Fondements. Logique, etc.				02 Logique et fondements			
510.2	Fondements des mathématiques							
510.22	Théorie des ensembles. Approche théorique des ensembles. Théorie des types d'ordre, des nombres ordinaux et cardinaux				04 Théorie des ensembles	146 Set theory		
510.6	Logique mathématique			F.4 Mathematical logic and formal languages	02 Logique et fondements	151 Logic, lattice theory		
511	Théorie des nombres				10 Théorie des nombres et théorie algébrique des nombres	153 Number theory		
512	Algèbre		32.2 Algèbre (linéaire, matrices) - Géométrie		09 Algèbres, manuels	157 Normed algebras		
510	Considérations fondamentales et générales sur les mathématiques. Fondements. Logique, etc.					158 Algebra (without specialized books on normed Algebras)		

Un indice CDU est attribué à chaque indice de chaque classification. Ainsi le tableau se compose, dans la colonne de gauche, des indices CDU retenus et des intitulés correspondants, puis à droite, sur deux colonnes par bibliothèque, des indices locaux et leur intitulé. Le degré de spécificité des indices CDU retenus est très variable, étant donné que les classifications étudiées ont des degrés de précision différents. Par contre, les indices généraux CDU qui n'avaient pas de correspondance, mais dans lesquels des sous-indices étaient utilisés dans le tableau, ont été rajoutés. Le tableau se compose ainsi de 1031 indices CDU, dont 151 ne correspondent pas à des indices appartenant aux autres classifications mais sont les indices généraux des indices spécifiques retenus.

Pour chaque indice CDU retenu, il est donc possible de savoir combien de bibliothèques possèdent des ouvrages y correspondant.

La figure 1 montre que 85% des indices se trouvent uniquement dans une bibliothèque. Le pourcentage d'indices présents dans deux bibliothèques ou plus peut donc paraître assez faible, pourtant le nombre de documents concernés est important étant donné l'étendue des collections, et concerne tout de même environ 150 indices et sous-indices.

Deux domaines et sujets seulement se trouvent dans six bibliothèques. Il s'agit de l'informatique et des probabilités et statistiques. Viennent ensuite les mathématiques présentes dans cinq bibliothèques.

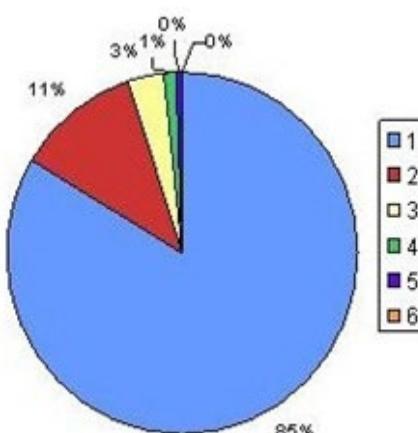


Fig. 1 : Pourcentage des indices présents dans 1, 2, 3, 4, 5 ou 6 bibliothèques

Les domaines et sujets présents dans quatre bibliothèques sont les suivants :

- Langages de programmation
- Histoire des sciences
- Mathématique numérique, analyse numérique, programmation (informatique), science des ordinateurs
- Astronomie, astrophysique, recherche spatiale, géodésie
- Physique
- Mécanique
- Mécanique des fluides, hydraulique
- Vibration, acoustique
- Optique
- Chaleur, thermodynamique
- Electricité, magnétisme, électromagnétisme
- Chimie, sciences minéralogiques, cristallographie
- Climatologie
- Biologie

Il s'agit de domaines généraux (biologie, chimie, physique), et de sujets (mathématique numérique, mécanique, etc.) appartenant principalement à la physique.

Les domaines et sujets apparaissant dans trois bibliothèques sont essentiellement sujets concernant l'informatique, les mathématiques, la physique, la chimie et les Sciences de la Terre :

- Langages de programmation
- Intelligence artificielle
- Traitement des images
- Bibliothéconomie, lecture
- Mathématiques. Sciences naturelles
- Généralités sur les sciences pures : Philosophie. Psychologie
- Sciences environnementales. Ressources naturelles. Conservation des ressources naturelles. Protection de l'environnement
- Logique mathématique
- Algèbre
- Géométrie
- Analyse mathématique
- Théorie de l'information: aspects mathématiques
- Théorie quantique

- Langages de programmation
- Intelligence artificielle
- Traitement des images
- Bibliothéconomie, lecture
- Mathématiques. Sciences naturelles
- Généralités sur les sciences pures : Philosophie. Psychologie
- Sciences environnementales. Ressources naturelles. Conservation des ressources naturelles. Protection de l'environnement
- Logique mathématique
- Algèbre
- Géométrie
- Analyse mathématique
- Théorie de l'information: aspects mathématiques
- Théorie quantique

En général, tous les indices présents dans trois bibliothèques ou plus représentent des sujets assez vastes et généraux ou des domaines. Les indices très spécifiques n'existent souvent que dans une seule bibliothèque.

Le graphique suivant (figure 2) a été élaboré en regroupant tous les indices spécifiques sous leur indice général (comportant au maximum trois chiffres) et indique donc combien de domaines ou sujets généraux sont présents dans plusieurs bibliothèques, même si à l'intérieur de ceux-ci, les sujets plus spécifiques traités ne sont pas les mêmes. Cela permet de déterminer si les bibliothèques contiennent des ouvrages concernant les mêmes types de domaines ou sujets généraux.

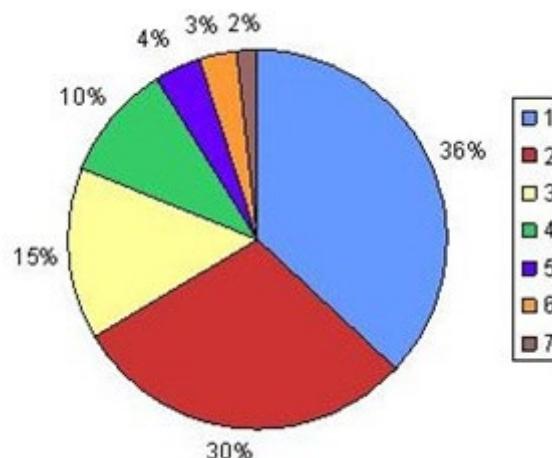
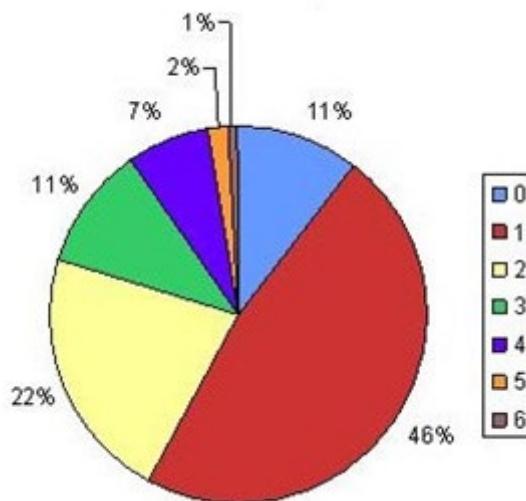


Fig. 2 : Pourcentages des indices spécifiques cumulés présents dans 1, 2, 3, 4, 5, 6 ou 7 bibliothèques

64% des domaines et sujets généraux représentés par trois chiffres dans la CDU sont donc traités dans plusieurs bibliothèques. Cela signifie que leurs collections ne sont pas limitées à leurs domaines respectifs.

Le troisième graphique (figure 3) permet de relativiser le précédent. En effet, celui-ci prend en compte uniquement les indices sans chiffre après la virgule. On remarque que ces indices généraux ne sont présents que dans un moindre pourcentage par rapport au graphique précédent. Seul 43% de ces domaines et sujets généraux existent dans deux à six bibliothèques. 11% ne sont présents dans aucune bibliothèque, ce sont des indices qui ont été rajoutés pour compléter des indices spécifiques existant dans le tableau.



Finalement, ces taux de recouplement montrent que certains domaines et sujets sont redondants. Il est sûr que certaines applications d'un domaine peuvent en concerter un autre, ou qu'une science a besoin d'éléments d'une autre, ce qui implique que certains ouvrages soient classés dans une autre partie de la CDU que celle qui concerne en premier lieu la bibliothèque.

3.2. Données statistiques

Comme la recherche des données statistiques a été effectuée sur la base locale genevoise de RERO, les bibliothèques touchées sont celles de la Faculté des Sciences et toutes celles faisant parties du Réseau des bibliothèques genevoises. L'analyse porte principalement sur les sept bibliothèques étudiées, mais quelques remarques seront également faites en relations avec les autres localisations genevoises.

Le tableau obtenu à l'aide du deuxième script a été travaillé de différentes manières, notamment en additionnant les résultats de tous les indices d'une bibliothèque pour avoir une comparaison des totaux par bibliothèques, en nombre de notices et en pourcentage (tableau 2 et 3). Il a également été nécessaire de produire une matrice triangulaire contenant la moyenne des totaux par couple de bibliothèques. En effet, la comparaison d'une bibliothèque A avec une B, ou de B avec A, devrait théoriquement donner le même résultat. Mais principalement à cause de différences de cotations et du fait que notre étude ne concerne qu'une partie des collections, ce n'est pas le cas. La moyenne des deux résultats a été calculée afin de n'avoir qu'un seul chiffre par couple de bibliothèques. Il faut comprendre par différences

de cotations le fait que, même si il a été décidé de faire correspondre deux indices, un même livre ne sera pas forcément coté dans la bibliothèque selon le tableau de concordance: la bibliothèque d'anthropologie mettra un livre de mathématiques dans l'indice correspondant aux mathématiques générales, alors que la bibliothèque de mathématiques mettra peut-être ce même livre dans un indice plus précis. Cela explique par exemple que le nombre de notices en commun obtenu en cherchant, dans les livres de la bibliothèque de mathématiques, ceux que possède la bibliothèque d'informatique, est égal à 145, alors que le contraire donne 62.

Seules les collections formées par les indices retenus grâce au tableau de concordance des classifications ont été étudiées et non les collections dans leur ensemble, il faut lire ces résultats en conséquence.

Dans les deux tableaux suivants, on trouve horizontalement un échantillon de notices d'une bibliothèque dans lequel on a cherché la présence des zones MARC 980_2 des bibliothèques représentées verticalement. Par exemple, 2101 documents ont été extraits pour la bibliothèque du CUI, et 62 de ces documents se trouvent également dans la bibliothèque de mathématiques (n'importe où dans la collection).

Il est important de remarquer que le nombre de notices prises en compte pour la bibliothèque des Sciences de la Terre est très faible car peu de documents possèdent une cote, ce qui rend les résultats peu comparables avec les autres bibliothèques.

Le tableau 1 présente un extrait du tableau de concordance (celui-ci comportant environ 110 pages, il n'est pas inclus dans cet article mais est disponible sur demande).

3.3. Analyse par bibliothèque

Le tableau 2 montre, pour chaque ligne, le nombre de notices de la bibliothèque en question retrouvées dans les bibliothèques en colonne. Le nombre total de notices pris en compte pour chacune des bibliothèques est en gris. Par exemple, dans les livres étudiés à la bibliothèque de mathématiques, 145 se trouvent aussi au CUI, mais ceux-ci ne sont pas forcément dans les indices étudiés au CUI.

Tableau 2 : Nombre de notices en commun entre les sept bibliothèques de la Faculté des Sciences

		Collections dans leur entier						
		CUI	Maths	Science II	Sc. Terre	Observatoire	Anthropologie	Physique
Collections étudiées dans les statistiques	CUI	2101	62	12	3	18	7	67
	Maths	145	6821	47	37	123	30	369
	Science II	14	53	1884	36	45	10	166
	Sc. Terre	1	0	4	128	4	0	2
	Observatoire	37	110	34	13	1027	7	159
	Anthropologie	14	37	19	20	8	872	10
	Physique	79	244	69	11	107	6	3112
		Nombre total de notices extraites par bibliothèques						
		>200						
		>100						
		>50						

Sur ce tableau, les bibliothèques partageant le plus grand nombre de notices avec d'autres localisations sont les bibliothèques de physique et de mathématiques. Entre elles deux, elles possèdent plus de 200 documents en commun, et passablement avec les cinq autres (surtout la bibliothèque de l'Observatoire et du CUI). Viennent ensuite le CUI, la bibliothèque de l'Observatoire et de Sciences II. Les bibliothèques des Sciences de la Terre et d'anthropologie sont celles qui ne partagent que peu de notices : moins de 50. Sur le tableau 3, le nombre de notices est pondéré par le nombre de notices total trouvées par bibliothèque, pour donner un pourcentage de notices en commun.

Tableau 3 : Pourcentage de notices en commun entre les sept bibliothèques de la Faculté des Sciences

		Collections dans leur entier						
		CUI	Maths	Science II	Sc. Terre	Observatoire	Anthropologie	Physique
Collections étudiées dans les statistiques	CUI	100.00	2.95	0.57	0.14	0.86	0.33	3.19
	Maths	2.13	100.00	0.69	0.54	1.80	0.44	5.41
	Science II	0.34	1.81	100.00	1.92	2.38	0.51	8.72
	Sc. Terre	0.78	0.00	3.13	100.00	3.13	0.00	1.56
	Observatoire	3.60	10.71	3.31	1.27	100.00	0.68	15.48
	Anthropologie	1.61	4.24	2.18	2.29	0.92	100.00	1.15
	Physique	2.54	7.84	2.22	0.35	3.44	0.19	100.00
Légende :		>1%			Nombre total de notices extraites par bibliothèques			
		>3%						
		>5%						

Ici, en lisant le tableau en colonne, on remarque que les bibliothèques de physique et de mathématiques sont celles qui possèdent le plus grand pourcentage de notices se trouvant dans les collections des indices retenus des autres bibliothèques.

Horizontalement, on voit par contre que la bibliothèque de mathématiques partage un pourcentage plutôt faible de sa collection avec les autres, physique à part. Par exemple, plus de 10% de la collection représentée par les indices retenus à l'Observatoire se trouve en mathématiques, alors que seulement 1.8% de la collection représentée par les indices retenus en mathématiques se retrouve dans la bibliothèque de l'Observatoire. Ceci est dû au nombre de livres, et au fait que ce sont essentiellement des indices généraux qui ont été retenus. D'une part, il y a relativement peu de livres qui concernent le domaine des mathématiques à l'Observatoire et ils sont classés dans des indices généraux (les indices détaillés concernent logiquement l'astronomie). En contrepartie, il y a évidemment un grand nombre de livres de mathématiques dans la bibliothèque de mathématiques et ils sont épargnés dans des indices détaillés ne correspondant pas aux indices retenus.

Toujours horizontalement, c'est la bibliothèque de l'Observatoire qui partage le plus grand nombre de notices avec les autres bibliothèques, en particulier avec les bibliothèques de mathématiques et physique, un peu moins avec le CUI et Science II. Viennent ensuite les bibliothèques de physique, Sciences II et mathématiques avec plus de 5% de notices en commun avec une autre bibliothèque, puis les bibliothèques d'anthropologie et du CUI avec plus de 3%. La collection de la bibliothèque des Sciences de la Terre, comme mentionné ci-dessus est difficile à commenter puisque le nombre de documents étudiés est faible. On peut tout de même remarquer environ 3% de notices en commun avec les bibliothèques de l'Observatoire et de Sciences II.

3.4. Analyse par dendrogramme

A partir de la matrice triangulaire, un dendrogramme a été produit à l'aide du programme R (R-Project, 2007), suivant l'algorithme (UPGMA Christian de Duve Institute of cellular pathology, 1997 ; Wikipédia, l'encyclopédie libre, 2007). Cet algorithme produit un arbre phylogénétique (dendrogramme) à partir d'une matrice de distances. Ceci est normalement utilisé pour étudier des séquences génomiques ou l'évolution des espèces à partir de leur génome. Le résultat illustre les distances entre bibliothèques, avec les deux bibliothèques les plus proches à l'extrémité de l'arbre (la branche la plus lointaine de la racine), puis la bibliothèque la plus proche de ce couple, etc., jusqu'à la dernière bibliothèque. Ce dendrogramme est un essai d'une méthode originale de comparaison des collections et il n'est pas aisément faire l'analyse.

Pour créer le dendrogramme de la figure 4, il a fallu prendre « 1-le pourcentage de notices en commun » (ramené à un nombre entre 0 et 1) du tableau 2, puis appliquer l'algorithme UPGMA. Il faut se souvenir que la matrice triangulaire à la base de ces données résulte d'un calcul de moyenne entre deux résultats. Ceci peut donc parfois tirer le pourcentage vers le haut ou vers le bas. Pour analyser ce dendrogramme, il faut prendre en compte la distance horizontale et verticale entre les bibliothèques. Les deux bibliothèques les plus proches en terme des collections étudiées sont celles de l'Observatoire et de physique (à l'extrémité du dendrogramme). Ensuite, la bibliothèque la plus proche de ce couple est celle de mathématiques, puis celle de Sciences II, etc. Il faut remarquer que chaque bibliothèque est comparée avec le groupe formé par celles qui la précèdent dans l'arbre. Du point de vue de la distance verticale, on voit que le groupe Observatoire-physique est relativement éloigné de la bibliothèque de mathématiques, qui est elle-même éloignée de celle de Sciences II. Par contre, Sciences II, le CUI, les bibliothèques des Sciences de la Terre et d'anthropologie sont relativement proches.

On peut donc distinguer trois parties dans le dendrogramme: le couple du bas (observatoire et physique), le groupe du haut (anthropologie, Sciences de la Terre, CUI et Sciences II), et la bibliothèque de mathématiques au milieu. Il faut toujours garder à l'esprit que les résultats concernant la bibliothèque des Sciences de la Terre ne sont pas représentatifs. On peut donc dire que les collections étudiées dans le premier couple sont relativement proches. La collection de la bibliothèque de mathématiques est la plus proche de ce couple, mais relativement éloignée vu la longueur de la barre verticale les séparant. La bibliothèque de Sciences II est la plus proche de ces trois mais avec une distance verticale également relativement grande. Les quatre dernières bibliothèques sont peu distantes verticalement.

Cluster Dendrogram

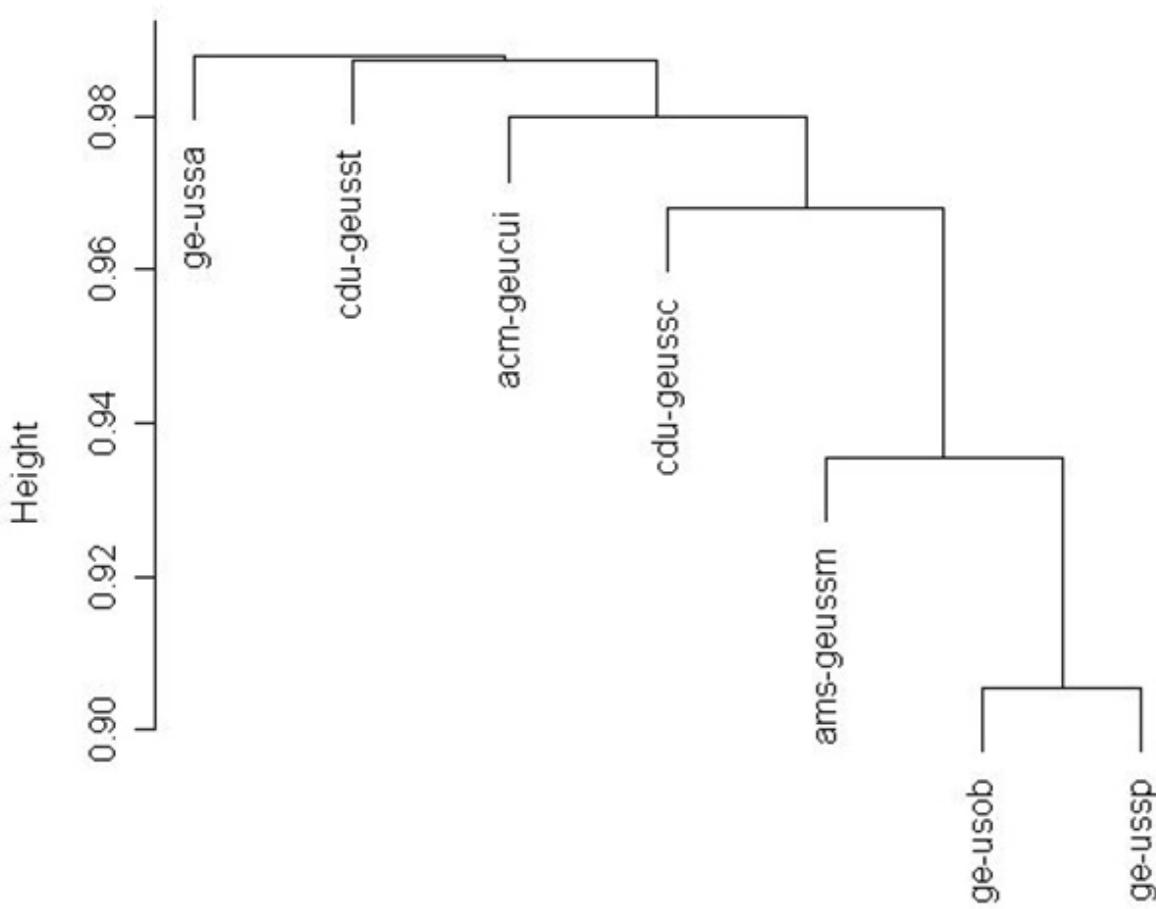


Fig. 4 : Dendrogramme obtenu par l'algorithme UPGMA

Légende: ge-ussa = anthropologie, cdu-geusst = Sciences de la Terre, acm-geucui = CUI, cdu-geussc = Sciences II, ams-geussm = mathématiques, ge-usob = Observatoire, ge-ussp = physique

3.5. Comparaison avec les autres bibliothèques genevoises

L'analyse des notices en commun entre les bibliothèques de la Faculté des Sciences et les autres bibliothèques genevoises montre que les bibliothèques des Sciences économiques et sociales (SES), et de la Faculté de psychologie et des sciences de l'éducation (FPSE) possèdent plus de 100 notices en communs avec la bibliothèque de mathématiques.

Entre la bibliothèque d'histoire des sciences et celle de physique, il y a plus de 50 notices en commun. Cela montre à quel point la bibliothèque de physique possède une grande collection de livres sur l'histoire des sciences et de la physique.

Plus de 50 notices en commun se trouvent également:

- entre les bibliothèques d'anthropologie, de l'Observatoire et de physique et celle de SES. Il est difficile de dire quels sont les sujets couverts par ces documents, mais on

peut imaginer qu'avec les bibliothèques de l'Observatoire et de physique, ça peut être des livres de statistiques et de mathématiques. Avec l'anthropologie, cela doit être un mélange de différents sujets : la sociologie et la démographie, la géographie et les statistiques, les mathématiques et l'informatique.

- entre la bibliothèque de philosophie et celles de mathématiques et physique. D'une part ces deux bibliothèques de la Faculté des Sciences possèdent un bon nombre de livres concernant la philosophie des sciences, et d'autre part, les logiques philosophique et mathématique sont liées.
- entre la bibliothèque de la Faculté de médecine (BFM) et Sciences II. Ces deux bibliothèques partagent probablement des documents de biologie, pharmacie et médecine.

4. Discussion

Il est évidemment difficile de dire que tel pourcentage de notices en commun entre deux bibliothèques est trop élevé ou pas, cela dépend du domaine, du type de bibliothèque et des usages de la bibliothèque.

La redondance des documents a des «pour» et des «contre». Les «pour» sont :

- L'accessibilité: la redondance entre bibliothèques rend les documents plus proches des usagers, d'autant plus que la grande majorité des monographies étudiées ici ne sont pas accessible en ligne à la Faculté des Sciences.
- La conservation: la redondance permet d'augmenter les chances de préservation à long terme. Ce n'est pas la mission principale des bibliothèques de la Faculté, bien que les documents anciens soient importants pour certains domaines.
- Usages simultanés: Les étudiants des différents départements ou sections de la Faculté ont une majorité de cours en commun au début de leur cursus et ont donc besoin des mêmes ouvrages.

Les «contre» sont :

- Coût élevé de la surface à Genève: plusieurs bibliothèques se trouvent dans des locaux inadaptés et manquent de place.
- Budget d'acquisition: il est en diminution systématique dans toutes les bibliothèques.
- Budget de fonctionnement: plus la collection est importante plus les coûts sont élevés.

Par ailleurs, pour pouvoir faire une bonne critique de ces résultats, il faudrait faire l'étude complète des collections, et notamment connaître le niveau des documents et leurs usages.

Le niveau des documents est intéressant car, comme le mentionne l'étude de Missingham & Walls (2003), il est justifiable que les ouvrages de référence et ceux destinés aux étudiants de premier cycle soient redondants, étant donné la nécessité d'un accès immédiat et multiple à ces documents.

Les usages et en particulier le taux de prêt permettent de savoir si une section de la collection est utile dans la bibliothèque étudiée bien que hors de son domaine. Cela permet également de se rendre compte si les usagers sont prêts à se rendre dans la bibliothèque de référence pour satisfaire leurs besoins en information.

Il faut également tenir compte l'environnement qui influence le comportement de l'usager:

- la distance entre les bibliothèques, qui peut être un obstacle au déplacement des usagers, même lorsqu'elle est faible ;
- les conditions de prêt qui limitent la disponibilité des documents (dans le temps et l'espace) ;
- etc.

4.1. 4.1 Discussion de l'analyse par bibliothèque

En analysant les pourcentages globaux de notices en commun pour chaque bibliothèque, il ressort nettement que les bibliothèques se recoupant le plus sont celles de physique, mathématiques et astronomie. Cette constatation est renforcée par le dendrogramme qui rassemble ces trois bibliothèques en un cluster. Cela paraît assez logique puisque l'astronomie est apparentée à la physique, et que ces deux domaines ont pour outils principaux les mathématiques.

Le peu de notices en commun entre les bibliothèques d'anthropologie et de Sciences II (contenant les collections de biologie) est en revanche plus étonnant. Les tableaux montrent peu de recouplement et les deux bibliothèques sont éloignées dans le dendrogramme. On aurait pu imaginer plus de similitudes car la bibliothèque d'anthropologie contient une collection importante de documents sur la biologie, la génétique et l'évolution, qui sont des sujets de recherche couverts par des chercheurs du département d'anthropologie.

4.2. 4.2 Discussion des résultats indice par indice

Une analyse de tous les indices présents dans trois bibliothèques ou plus a été faite à partir du tableau obtenu avec la deuxième partie du programme. Les détails ne sont pas présentés ici, mais en voici les points principaux :

- Une bibliothèque se distingue des autres : la bibliothèque de physique. Elle a souvent un taux de recouplement élevé avec les autres bibliothèques, en particulier avec celles de l'Observatoire, de mathématiques et dans une moindre mesure Sciences II.
- La bibliothèque de mathématiques ressort également du lot. Mais dans ce cas, c'est souvent les autres bibliothèques qui possèdent des livres de mathématiques, et il est donc normal que les mathématiques possèdent les ouvrages que les autres bibliothèques ont acquis dans son domaine. De plus, cette bibliothèque a un grand nombre de documents classés dans des indices qui a priori ne sont pas des mathématiques, mais il s'agit en réalité souvent d'ouvrages de mathématiques appliquées à un autre domaine et donc classés sous cet autre domaine, pour les distinguer des mathématiques non appliquées.
- La bibliothèque de Sciences II a, de manière générale et au contraire de ces deux premières bibliothèques, un taux de recouplement très faible avec les autres. Ses documents sont donc plus ciblés sur les sujets qu'elle couvre.

- Les taux de recouvrement relativement élevés qu'on peut trouver avec la bibliothèque du CUI, ne concernent pratiquement que des indices de sujets informatiques. Comme pour les mathématiques, ces taux de recouvrement sont donc à attribuer aux autres bibliothèques et non à celle du CUI qui reste bien limitée à son domaine.
- Les taux de recouvrement les plus élevés qu'on trouve avec la bibliothèque de l'Observatoire sont liés aux bibliothèques de mathématiques et physique, ce qui est logique vu la proximité de ces trois sciences.
- Pour la bibliothèque des Sciences de la Terre, il est difficile de tirer des conclusions pertinentes, car peu de documents de cette bibliothèque sont dotés d'une cote.
- Les résultats pour la bibliothèque d'anthropologie sont également difficiles à commenter car les indices utilisés couvrent plusieurs sujets différents. En général, les taux de recouvrement sont assez faibles. Cependant, les visites sur place ont montré qu'il existait un nombre relativement élevé de documents traitant d'autres sujets et domaines que ceux normalement couverts par cette bibliothèque. De plus, la majorité de ces documents « hors sujet » étaient pour la plupart très anciens (vieux de plusieurs dizaines d'années), ce qui diminue encore l'intérêt de leur présence dans cette bibliothèque.

4.3. Discussion de la précision des données (approximation d'erreur)

L'imprécision de la méthode de récolte des statistiques est due à trois sources d'erreurs :

1. les documents étudiés ne représentant pas la totalité de la collection des bibliothèques
2. les incohérences dans les méthodes de catalogage, notamment pour les documents en plusieurs volumes
3. la concordance établie entre les classifications dans cette étude, potentiellement différente des pratiques de cotation dans les bibliothèques

Etant donné l'asymétrie du tableau 2, une estimation de l'erreur a été calculée pour déterminer si les résultats sont tout de même significatifs. Pour chaque couple de bibliothèques (i, j), la formule $|n_{i,j} - n_{j,i}| / \max(n_{i,j}, n_{j,i})$ a été appliquée (tableaux 4 et 5).

Tableau 4: Approximation d'erreur pour toutes les bibliothèques: somme totale = 9.85, moyenne = 0.47

	CUI	Maths	Science II	Sc. Terre	Observatoire	Anthropologie	Physique
CUI							
Maths	0.57						
Science II	0.14	0.11					
Sc. Terre	0.67	1.00	0.89				
Observatoire	0.51	0.11	0.24	0.69			
Anthropologie	0.50	0.19	0.47	1.00	0.13		
Physique	0.15	0.34	0.58	0.82	0.33	0.40	
Somme	2.55	1.75	2.19	2.51	0.45	0.40	

**Tableau 5: Approximation d'erreur sans la bibliothèque des Sciences de la Terre:
somme totale = 4.78, moyenne = 0.32**

	CUI	Maths	Science II	Observatoire	Anthropologie	Physique
CUI						
Maths	0.57					
Science II	0.14	0.11				
Observatoire	0.51	0.11	0.24			
Anthropologie	0.50	0.19	0.47	0.13		
Physique	0.15	0.34	0.58	0.33	0.40	
Somme	1.88	0.75	1.30	0.45	0.40	

En comparant deux bibliothèques de deux façon différentes tel que dans le tableau 2, les résultats obtenus sont relativement proches car la variation entre les deux parties du tableau de cette figure est de 32% sans tenir compte de la bibliothèque des Sciences de la Terre. En général, la corrélation est bonne, ce qui montre que la méthode est assez représentative. Une mauvaise corrélation, par contre, ressort nettement chaque fois que la bibliothèque des Sciences de la Terre est concernée. Ceci était prévisible car le nombre de documents cotés et donc étudiés dans cette bibliothèque est faible et non représentatif.

4.4. Synthèse de l'analyse

Des tendances se détachent clairement de l'étude de ces statistiques et donnent une bonne idée des relations entre les échantillons des collections des sept bibliothèques étudiées. Elle souligne notamment la redondance, nécessaire ou non, de certains sujets et documents dans les diverses bibliothèques.

Selon le nombre de notices communes entre les sept bibliothèques, celles de mathématiques et physique partagent le plus de documents avec les autres localisations. Viennent ensuite les bibliothèques du CUI, de l'Observatoire et de Sciences II. En ce qui concerne le dendrogramme, l'ordre est un peu différent : les bibliothèques de l'Observatoire et de physique sont les plus proches, suivies de celles de mathématiques. Les bibliothèques du CUI et de Sciences II se rapprochent plus, elles, de l'anthropologie et des Sciences de la Terre.

L'étude des statistiques et les visites sur place ont fait ressortir certains points :

- Pour la bibliothèque de physique, le nombre de documents concernant l'histoire et la philosophie des sciences est important. De plus, le nombre de notices en commun avec certaines bibliothèques hors de la Faculté des Sciences comme celles du Musée d'histoire des sciences, de la Faculté des sciences économiques et sociales et du Département de philosophie est élevé. Il serait nécessaire de mener une réflexion sur la politique à suivre pour la gestion de cette partie de la collection.
- Pour la bibliothèque de physique en particulier, mais également pour celles de mathématiques, d'anthropologie, et dans une moindre mesure en astronomie, Sciences de la Terre et à Sciences II, les collections d'informatique contiennent beaucoup d'ouvrages sur des sujets redondants, souvent obsolètes. Sachant que la bibliothèque du CUI ne conserve en général pas les ouvrages trop anciens, les autres bibliothèques devraient d'autant plus désherber leurs collections dans ce domaine.
- Il y a beaucoup de recouplement entre les collections des bibliothèques de physique et de l'Observatoire, mais vu la distance entre les deux localisations, la proximité des domaines de recherche et le prêt exclu à l'Observatoire, ceci est en partie justifié. Il serait intéressant d'étudier l'usage des documents de ces deux bibliothèques pour savoir si ce recouplement correspond réellement à un besoin.

- Pour la bibliothèque de l'Observatoire, outre les remarques ci-dessus, on trouve un recouplement important avec la bibliothèque de mathématiques, que la distance peut à nouveau expliquer.
- La bibliothèque de mathématiques a beaucoup de recouplement avec les autres bibliothèques, mais il s'agit surtout de documents de mathématiques possédés par les autres. Par ailleurs, cette bibliothèque possède plus de 250 livres de physique et 90 livres d'histoire/philosophie des sciences. Ici aussi, l'étude de l'usage de documents pourrait apporter d'avantage d'éléments pour décider de la nécessité ou non d'un tel recouplement.
- En ce qui concerne la bibliothèque de Sciences II, la plus grande des sept, il y a peu de recouplement avec les autres bibliothèques. La seule remarque à faire concerne le domaine de la physique qui est très représenté, d'autant plus que la bibliothèque de physique se trouve à environ 200 mètres.
- La collection de la bibliothèque du CUI est ciblée car le recouplement avec les autres localisations concerne principalement des livres d'informatique. Les mathématiques y sont bien représentées, mais elles constituent un outil essentiel de l'informatique.
- Le tableau 2 montre que très peu des livres étudiés dans la bibliothèque des Sciences de la Terre se retrouvent dans les autres collections. Il y a par contre plus de livres en commun si on étudie les collections des autres bibliothèques par rapport à celle des Sciences de la Terre. Cela veut dire qu'il s'agit dans ce cas de livres catalogués mais non cotés dans cette localisation. Si tous les livres étaient cotés, plus de recouplement aurait peut-être été mis en évidence.
- Pour la bibliothèque d'anthropologie, les observations sur place ont montré la présence de beaucoup de livres hors des sujets étudiés au département. Ceci est dû à l'historique, au manque de désherbage et à l'évolution des sujets couverts par la bibliothèque. Seuls 2% des livres étudiés se retrouvent à la bibliothèque de Sciences II (comprenant la section de biologie), ce qui est très peu sachant qu'un des sujets d'étude du département d'anthropologie est la génétique des populations et l'évolution. Cela tient peut-être au fait qu'il s'agit d'un sujet très précis de la génétique.

Ces résultats mériteraient d'être complétés par d'autres études concernant les usages et des indicateurs détaillés sur les ouvrages eux-mêmes. Cependant ils posent une base utile pour une future étude complète des collections à la Faculté des Sciences, étape nécessaire à la mise en place d'une politique d'acquisition commune, et à une gestion des collections concertée entre les bibliothèques.

4.5. Les politiques d'acquisition à la Faculté des Sciences et à l'Université de Genève

Il n'y a pas de politique d'acquisition commune aux bibliothèques de la Faculté des Sciences, et pas non plus de politique d'acquisition formelle par bibliothèque. Chaque bibliothécaire responsable des acquisitions a bien sûr sa politique, mais elle n'est pas formalisée.

La situation de l'ensemble de l'Université de Genève ressemble à une plus large échelle à celle de la Faculté des Sciences: il n'y a pas véritablement de politique documentaire

commune à l'Université. Plus précisément, concernant la politique d'acquisition, aucune charte des collections n'existe. Ce manque est sur le point d'être comblé puisqu'une directrice de l'information scientifique (fonction qui n'existait pas auparavant) vient d'être engagée au niveau de l'Université. Celle-ci a mis en place plusieurs groupes de travail, dont un est chargé des collections. Ce groupe est actuellement en train d'établir une vision à long terme des collections (10 ans), et a commencé à faire une évaluation des collections pour réunir les indicateurs nécessaires à la mise en place d'un plan de développement des collections.

Etant donné qu'une politique d'acquisition au niveau de l'Université ⁽¹⁾ va bientôt voir le jour, celle de la Faculté des Sciences devra s'y conformer. Mais il faudra détailler celle-ci davantage pour chaque bibliothèque afin établir des plans de développement des collections ainsi que des protocoles de sélection propres à chaque domaine, et rendre ainsi la politique d'acquisition de l'Université applicable à chaque bibliothèque de la Faculté des Sciences.

Voici quelques remarques dont il serait intéressant de tenir compte au moment où la politique d'acquisition sera mise en place dans les sept bibliothèques.

Dans les bibliothèques de la Faculté des Sciences, les publics sont assez bien définis, suite à un travail de diplôme de la HEG sur les pratiques documentaires de leurs usagers (Bui et al., 2006). Par contre, les collections sont assez mal connues, totalement indépendantes les unes des autres (il n'y a pas de concertation entre les bibliothèques à ce sujet) et parfois, à cause de leur historique, surprenantes dans leur contenu. Une étude globale des sept collections paraît fondamentale si l'on désire avoir un jour un ensemble de collections cohérent, et pouvoir considérer les acquisitions en réseau ou centralisées pour la Faculté.

Ce travail s'inscrit donc dans un premier pas vers cette étude des collections. Pour cela, le tableau de concordance des classifications peut servir comme repère pour la segmentation des collections en catégories comparables d'une bibliothèque à l'autre. Même s'il semble peu probable qu'une classification unique soit utilisée dans toutes les bibliothèques pour classer les monographies, ce tableau pourrait être utilisé pour gérer de manière centrale les acquisitions, et les répartir en fonction d'une classification unique. Les sujets couverts par les bibliothèques étant pointus, ce serait, dans certains cas, une perte de précision que de vouloir tout classer selon la CDU. Par contre, une cote CDU attribuée à chaque document pourrait être utilisée pour la gestion globale des collections, et une cote selon la classification locale permettrait de le ranger en rayon.

Les résultats montrent que les collections se recoupent, tant du point de vue des sujets que des documents eux-mêmes. Il est visible que chaque collection a été gérée comme une entité indépendante, et les vestiges de l'histoire de la bibliothèque n'ont pas été remis en cause. Par exemple, quatre bibliothèques possèdent une collection non négligeable de documents sur l'histoire et la philosophie des sciences, sujets tout à fait d'actualité dans un environnement universitaire et scientifique. Mais on peut imaginer qu'il n'est pas nécessaire d'en avoir autant, épars dans quatre localisations. Une bibliothèque pourrait être désignée comme référence sur le sujet, et le reste des documents non conservés pourraient être donnés à la bibliothèque d'Histoire des Sciences.

Il en est de même pour l'informatique où toutes les bibliothèques regorgent de livres sur les logiciels et langages de programmation, dont certains tout à fait obsolètes. Pour ce qui est des statistiques et probabilités, toutes les sciences utilisent ces outils. Mais on pourrait ne laisser

dans chaque bibliothèque que les documents présentant les mathématiques appliquées au domaine couvert par la bibliothèque, ainsi qu'un ou deux ouvrages généraux de qualité, et rassembler l'excédant à la bibliothèque de mathématiques.

En général, définir les responsabilités des bibliothèques en matière d'achat et de conservation des différents domaines en présence paraît être la solution pour rationaliser la gestion des collections.

5. Conclusion

La redondance des sujets mise en lumière par cette étude est en partie nécessaire car elle concerne les catégories suivantes : les sujets inhérent à l'environnement scientifique (comme l'histoire ou la philosophie des sciences), les outils de base en sciences (comme les mathématiques ou l'informatique).

Diverses études se sont penchées sur la redondance de collections, notamment celle de Missingham & Walls (2003). Contrairement à celles-ci, notre étude se base sur sept bibliothèques spécialisées par domaines. De plus, nous n'avons analysé que les documents correspondant aux indices présents dans trois bibliothèques ou plus. Ainsi, notre étude concerne un sous ensemble des collections des bibliothèques de la Faculté des Sciences, choisi pour obtenir une estimation maximale de la redondance. Il est donc inapproprié de comparer nos résultats à ceux des études mentionnées précédemment. La suite logique de ce travail serait d'étudier les collections dans leur ensemble à l'aide d'un tableau d'indicateurs détaillé. Cela permettrait d'obtenir entre autres des chiffres comparables à la littérature mentionnée ci-dessus. Pour pouvoir tirer des conclusions définitives, il serait également important de prendre en compte les statistiques d'usages des collections.

Sur la base du tableau de concordance, une classification unique destinée à la gestion des collections pourrait être attribuée à chaque document, tout en gardant les classifications propres à chaque localisation pour le rangement en rayon. Cela permettrait de développer une vue d'ensemble des collections de la Faculté des Sciences et de mettre en place une gestion globale cohérente.

Si la gestion centralisée des sept bibliothèques, telle qu'envisagée à une période, semble difficile en tenant compte des remarques des chercheurs, on pourrait en revanche tout à fait imaginer une « centralisation des documents par domaines ». Selon ce modèle, chaque bibliothèque s'occuperait de ses domaines, ce qui est en partie déjà le cas du point de vue des acquisitions, mais pas de la gestion de la collection en général (désherbage par exemple). Pour les bibliothèques ayant des domaines se chevauchant, mathématiques et physique par exemple, chaque bibliothèque pourrait être désignée comme référence pour un certain nombre de sujets. Ainsi la répartition des acquisitions serait clarifiée et il serait plus facile de prendre des décisions en matière de désherbage. Six des sept bibliothèques se trouvant dans le même quartier, la contrainte de déplacement est légère, et le sentiment de réseau en serait renforcé.

6. Remerciements

Nous remercions Anne-Christine Robert, coordinatrice des bibliothèques de la Faculté des Sciences de l'Université de Genève, qui a commandité cette étude, et Bertrand Calenge pour la direction de ce travail dans le cadre de la formation du CESID.

NOTES

⁽¹⁾ Le poste de Direction est vacant depuis août et les projets subiront certainement quelques retards.

BIBLIOGRAPHIE

AMERICAN INSTITUTE OF PHYSICS (2007). PACS, physics and astronomy classification scheme [en ligne]. [Consulté le 21.05.2007]. <http://www.aip.org/pacs/>

AMERICAN MATHEMATICAL SOCIETY (2007). MathSciNet: mathematical reviews on the Web [en ligne]. [Consulté le 21.05.2007]. <http://ams.u-strasbg.fr/mathscinet/otherTools.html>

ASSOCIATION FOR COMPUTING MACHINERY (2007). The guide to computing literature [en ligne]. [Consulté le 25.07.2007].

<http://portal.acm.org/ccs.cfm?part=author&coll=GUIDE&dl=GUIDE&CFID=24852534&CFTOKEN=71835811>

BAKER, Sharon L., LANCASTER, F. Wilfrid (1991). The measurement and evaluation of library services. Second ed. Alington, Information Resources Press

BUI, Céline, LEHNER Susanne, MORESI Nadia (2006). étude des pratiques documentaires des usagers : quels services pour la bibliothèque de demain ? 153 p. Mémoire de diplôme HES, Haute école de gestion de Genève

CALENGE, Bertrand (1994). Les politiques d'acquisition : constituer une collection dans une bibliothèque. Paris, Editions du Cercle de la Librairie. Collection Bibliothèques

CALENGE, Bertrand (1999). Conduire une politique. Paris, Editions du Cercle de la Librairie. Collection Bibliothèques

CALENGE, Bertrand (2006). Politique documentaire et gestion des collections : cours donnée au CESID, volée 2005-2007

CHRISTIAN DE DUVE INSTITUTE OF CELLULAR PATHOLOGY (1997). Construction of a distance tree using clustering with the Unweighted Pair Group Method with Arithmatic Mean (UPGMA) [en ligne]. [Consulté le 04.06.2007]. <http://www.icp.ucl.ac.be/~opperd/private/upgma.html>

MISSINGHAM, Roxanne, WALLS, Robert (2003). Australian university libraries: collections overlap study. The Australian Library Journal [en ligne], [Consulté le 14.11.2007]. <http://www.alia.org.au/publishing/alj/52.3/full.text/missingham.walls.html>

R-PROJECT (2007). The R Project for Statistical Computing [en ligne]. [Consulté le 06.06.2007]. <http://www.r-project.org/>

RESEAU DES BIBLIOTHEQUES GENEVOISES (2007). Catalogue [en ligne]. [Consulté le 06.06.2007]. <http://opac.ge.ch/gateway?lng=fr-ch>

TABAH, Albert N (1995). Evaluation des collections. In CALENGE, Bertrand ... [et al.] Diriger une bibliothèque d'enseignement supérieur. Sainte-Foy, Presses universitaires du Québec. P. 285-296. Gestion de l'information. ISBN 2760508706

UDC CONSORTIUM, (2004). Classification décimale universelle. Edition moyenne internationale. Liège, Editions du CEFAL

WIKIPEDIA, L'ENCYCLOPEDIE LIBRE (2007). UPGMA [en ligne]. [Consulté le 04.06.2007]. <http://fr.wikipedia.org/wiki/UPGMA>