

Terminale STMG Sciences de gestion Systèmes d'information de gestion

L'information pour agir et décider

Le thème *L'organisation informatisée* permet de montrer comment l'information est mobilisée dans la conduite des activités d'une organisation. Le thème *L'information pour agir et décider* étudie les logiques de production et de mémorisation de l'information, ses finalités et ses limites.

La résolution d'un problème de gestion consiste à répondre à une question par la production d'informations, à partir de données et de règles de gestion. La réalisation de programmes, en automatisant les règles de gestion, aide les acteurs à agir et décider dans les organisations (première question). Quand les données à manipuler sont volumineuses mais structurées, leur stockage dans des bases de données relationnelles facilite leur exploitation et la production d'informations (deuxième question). Mais certaines informations détenues dans les systèmes d'information ne proviennent pas de bases de données relationnelles et ne peuvent y être stockées du fait d'un contenu plus composite : ce sont des documents produits et manipulés par les acteurs de l'organisation ou de ses partenaires qui nécessitent un traitement afin d'en valoriser le contenu (troisième question).

Ce thème se prête particulièrement à la mise en place d'expérimentations et de projets d'ambition volontairement limitée s'appuyant sur un environnement de développement d'applications - simple à prendre en main - ou encore d'autres outils numériques, comme un tableur, notamment pour l'approche de la simulation (première question), un système de gestion de base de données relationnelle (deuxième question) ou encore un système de gestion de contenu (troisième question).

Les outils numériques mobilisés, notamment l'environnement de développement et le ou les langages mobilisés, ne sont pas des objets d'enseignement et doivent servir à mettre en évidence les notions du programme, à les éprouver ou à les illustrer. Par la pratique de ces outils, l'élève en apprécie les potentialités, les effets et les limites.

Question : La résolution de tous les problèmes de gestion est-elle automatisable ?

Intentions

Les technologies de l'information et de la communication contribuent pour une large part à accompagner l'organisation et ses acteurs dans leurs activités, en automatisant certaines d'entre elles.

L'augmentation exponentielle des performances des outils numériques a permis de faire évoluer le périmètre des problèmes qu'ils peuvent résoudre aujourd'hui. Les programmes d'intelligence artificielle et les outils de simulation permettent ainsi d'assister l'utilisateur dans ses prises de décision face à des problèmes de plus en plus complexes sans toutefois pouvoir totalement se substituer à lui.

Répondre à la question en distinguant ce qui est automatisable et ce qui ne l'est pas est l'occasion de comprendre les limites d'un programme informatique à travers l'étude de ses constituants et de la manière dont ceux-ci sont organisés pour apporter une réponse à un problème de gestion. Cette

question permet aussi d'explorer les potentialités des outils de simulation et d'en aborder différents usages aussi bien comme outils d'aide à la décision que comme outils au service des apprentissages. Elle permet aux élèves de s'initier à la programmation et de développer à cette occasion des qualités nécessaires à la poursuite d'études : logique, rigueur, curiosité, créativité et ténacité.

Objets d'un traitement informatique : constantes, variables, paramètres, opérations associées (déclaration et affectation), types scalaires (entier, flottant, booléen, caractère), type chaîne de caractères, tableaux, listes

Lien avec la question de gestion

Un programme est l'abstraction d'un processus de résolution de problèmes qui agit sur des objets virtuels, les variables, les constantes et les paramètres, qui représentent des faits ou des éléments (physiques ou abstraits) du monde réel. Cette représentation plus ou moins simplifiée s'appuie sur les fonctionnalités d'un langage de programmation qui permettent de déterminer le type des objets ou des ensembles d'objets qu'il peut traiter.

Pré-requis, transversalité

La plupart des notions traitées sont abordées par le programme de mathématiques ou ont été mises en pratique à travers les apprentissages liés à l'acquisition des compétences du B2i et les situations proposées dans l'enseignement de la classe de première : exploitation d'un tableur, d'un PGI, d'un environnement numérique de travail. Il s'agit de s'appuyer sur ces acquis pour fixer les concepts associés.

Attendus et limites

En s'appuyant sur des programmes permettant de résoudre des problèmes de gestion simples, il s'agit de faire découvrir la représentation des objets réels liés au problème, à travers l'exploitation de ces programmes et l'étude des variables manipulées. On s'intéressera notamment, sur des exemples simples en liaison avec le programme de mathématiques (algorithmique), au mode de représentation des données manipulées (codage binaire) pour aborder notamment la notion de précision d'un programme et la notion de format numérique.

Objets d'un traitement informatique : types scalaires (entier, flottant, booléen, caractère)

Un **type de données** définit les valeurs que peut prendre une donnée (qu'il s'agisse d'une variable, d'une constante ou d'un paramètre), ainsi que les opérations qui peuvent lui être associées.

Il ne s'agit pas de dresser une liste exhaustive de tous les types existants mais de montrer à travers l'exploitation d'un langage de programmation l'importance de choisir le type de données adapté à l'objet qu'on souhaite représenter en se limitant aux types proposés.

Objets d'un traitement informatique : type chaîne de caractères

Une **chaîne de caractères** est une suite ordonnée de caractères encadrée par des délimiteurs différents selon les langages. Le type *chaîne de caractères* est associé à des opérateurs ou des primitives permettant de les manipuler : concaténation, extraction, recherche, etc.

Objets d'un traitement informatique : constantes, variables, paramètres, opérations associées (déclaration et affectation)

Une variable détermine un emplacement de la mémoire associé à un programme et destiné à mémoriser une donnée représentative des éléments réels manipulés. La détermination de cet emplacement peut se faire de façon dynamique (à la première utilisation d'une variable dans un

programme) ou par une déclaration préalable. L'opération d'affectation consiste à attribuer une nouvelle valeur à une variable.

Une constante correspond à une donnée dont la valeur est non modifiable dans un programme.

Un paramètre est une donnée manipulée par un programme qui provient d'un autre programme. Cette notion est indissociable de celle de fonction.

Objets d'un traitement informatique : tableaux, listes

Tableau et liste sont des structures de données qui permettent de stocker plusieurs données élémentaires ou éléments de même type. Même si leur implémentation peut varier d'un langage à un autre, tableau et liste peuvent se différencier par les caractéristiques suivantes :

- Un tableau est une structure de données dans laquelle chaque élément est accessible directement grâce à une information permettant de le repérer, souvent appelé indice. On parle de structure à accès direct ;
- Une liste est une structure de données dans lequel un élément ne peut être accédé qu'à travers le parcours de tous les éléments qui le précède : on parle de structure à accès séquentiel.

Logique d'un traitement informatique : algorithme, opérateurs arithmétiques et logiques, séquences, instructions conditionnelles (tests), répétitions (boucles), fonctions, formule

Lien avec la question de gestion

Un programme répond aux besoins d'un utilisateur en transformant des données, que celui-ci lui fournit, en résultats qu'il attend. Il se déroule en suivant des règles de comportement (son algorithme) par rapport à son environnement en automatisant ainsi la résolution de la catégorie de problèmes pour lequel il a été conçu.

Pré-requis, transversalités

Les notions liées aux opérateurs arithmétiques et logiques sont abordées dans le programme de mathématiques. Les réalisations de macro-instructions dans un tableur en classe de première peuvent servir de point de départ à l'analyse d'un algorithme.

Attendus et limites

À travers l'exploitation et l'analyse de programmes permettant de résoudre des problèmes de gestion simples, il s'agit de faire découvrir comment sont représentées et exécutées les règles de gestion permettant de transformer les objets d'un programme.

Logique d'un traitement informatique : algorithme

Un algorithme est une série finie et non-ambiguë d'opérations à réaliser pour résoudre un problème en transformant des données en résultats. Il est mis en œuvre par un ordinateur sous la forme d'un programme. Réalisation et analyse d'algorithmes à travers des scénarios permettant leur déroulement sont l'occasion d'insister sur leur caractère déterministe : les mêmes données donnent toujours les mêmes résultats.

Logique d'un traitement informatique : opérateurs arithmétiques et logiques, formules

Les opérateurs permettent à un programme de manipuler les données qui lui sont fournies. Pour chaque type abordé, il s'agit de montrer le lien entre le type d'une variable et les opérateurs qu'il peut supporter.

Logique d'un traitement informatique : séquences, instructions conditionnelles (tests), répétitions (boucles)

La réalisation d'un traitement informatique repose sur l'exécution successive (séquence) d'opérations qui peuvent être conditionnées par l'environnement (instructions conditionnelles) et s'appliquer sur un ensemble de situations ou de données de natures identiques (instructions répétitives).

Logique d'un traitement informatique : fonctions, formule

Une fonction est un ensemble d'opérations permettant la transformation d'un ensemble de données qui lui sont fournies (paramètres ou arguments) en un résultat attendu (image de la fonction) selon un contrat donné. Une fonction peut être prédéfinie (mise à disposition par le langage de programmation utilisé) ou réalisée par le développeur pour répondre à un besoin spécifique. La réalisation de fonctions répond à la fois à la volonté de réduction de la complexité d'un programme (analyse descendante) et de réutilisabilité.

Selon les langages utilisés, on pourra faire référence aux notions de procédure ou de méthode en établissant le lien avec la notion de fonction, en s'attachant à distinguer les arguments en entrée (nécessaires pour l'exécution et non modifiables) et les arguments en sortie (résultats rendus à la fin de l'exécution).

Une formule dans une feuille de calcul (ou dans un programme) est une expression mathématique qui associe des opérateurs et des données (constantes, variables, paramètres, fonctions, etc.) pour définir comment obtenir le résultat attendu de la cellule (ou de la variable d'affectation).

Langage, programme. Tests et mise au point

Lien avec la question de gestion

La phase de test et de mise au point d'un programme consiste à s'assurer à l'aide de jeux d'essai qu'il répond aux besoins exprimés en transformant les données fournies en résultats conformément aux règles de gestion du problème à résoudre.

Pré-requis, transversalités

La question « Pourquoi la qualité du système d'information est-elle un enjeu pour l'organisation ? » du thème L'organisation informatisée permet d'aborder la qualité des données et des informations. Le test d'un programme contribue à la qualité des informations produites par le système d'information.

Attendus et limites

En s'appuyant sur la réalisation et la mise au point de programmes permettant de résoudre des problèmes de gestion simples, il s'agit de faire découvrir les bases de leur construction et de construire des réflexes de rigueur et de contrôle à travers la mise en place de procédures de test permettant de s'assurer de quelques critères de qualité de base d'un programme : exactitude (justesse, fiabilité, précision) et lisibilité.

Langage, programme

L'apprentissage d'un langage de programmation n'étant pas une fin en soi, le choix du langage doit être subordonné aux intentions pédagogiques exprimées dans la colonne « contexte et finalités » du programme. Dans ce cadre, on pourra s'appuyer sur des langages tels que PHP, VB ou Javascript.

Tests et mise au point

Tests et mise au point d'un programme ont pour objectif de livrer à l'utilisateur un programme répondant au problème à traiter. On distinguera sur des exemples simples, par exemple l'étude et l'utilisation d'une fonction retournant le minimum de deux valeurs pour la construction d'une fonction retournant le minimum de trois valeurs, les notions de test unitaire (vérification de la conformité d'une fonction à son contrat) et de test d'intégration (vérification que toutes les parties développées séparément fonctionnent ensemble conformément au cahier des charges).

Simulation, scénario, modèle

Lien avec la question de gestion

La modélisation et la simulation d'un phénomène permettent d'en comprendre le fonctionnement et d'en prédire dans une certaine mesure les évolutions. Ce sont des outils d'aide à la prise de décision en permettant de tester des hypothèses à travers différents scénarios. Ce sont donc des outils au service de l'action.

Elles favorisent la compréhension et l'analyse d'un phénomène à travers des possibilités de manipulation de modèles simplifiés. Ce sont donc aussi des outils au service de l'apprentissage.

Pré-requis, transversalité

Les élèves ont manipulé des outils de simulation en sciences de gestion en première pour traiter la question « La recherche de l'amélioration de la performance comporte-t-elle des risques ? » du thème Temps et risque.

Attendus et limites

Simulation

La simulation représente le processus d'expérimentation de différents scénarios pour en comprendre ou en évaluer le fonctionnement ou pour en mesurer les conséquences. Elle s'appuie sur la conception d'un modèle du phénomène ou du système réel étudié.

La simulation s'appuie de plus en plus sur les outils numériques qui permettent de multiplier les scénarios et d'améliorer les modèles mathématiques en les rapprochant du phénomène modélisé. On parle alors de simulation numérique ou informatique, qui permet :

- De simuler des événements difficiles à reproduire (par exemple les conséquences d'un réchauffement climatique) ;
- Ou entraînant la destruction ou l'altération du système réel (par exemple les conséquences d'un accident sur les passagers d'un véhicule) ;
- Ou encore présentant des risques pour le système réel (par exemple les conséquences d'un acte médical sur un patient).

L'exploitation et l'analyse de différents outils de simulation liés au domaine de la gestion et de la finance doivent permettre de fixer les notions du programme. La mise en œuvre d'outils simples de simulation liée à une prise de décision, notamment à travers la mise en place de scénarios sur tableur, permet d'en mesurer les limites.

Scénario

Un scénario représente une combinaison d'hypothèses dont la simulation va permettre de mesurer les conséquences (que se passe-t-il si ...) ou d'observer le fonctionnement du problème modélisé pour le comprendre.

Modèle

Un modèle est une représentation simplifiée de la structure d'un objet ou d'un système, de son fonctionnement ou encore de la dynamique d'un phénomène. Il peut être graphique (carte routière, schéma d'un réseau, schéma relationnel, diagramme de séquences), analogique ou physique (maquette ou « modèle réduit »), linguistique (cas d'utilisation, système expert) ou encore mathématique (système d'équations, formules).

Un modèle est statique s'il est indépendant du temps (comme un schéma relationnel), dynamique dans le cas contraire (comme un algorithme, un programme ou une feuille de calcul).

Proposition d'exploitation pédagogique

La proposition ci-dessous est un exemple d'exploitation pédagogique permettant d'aborder progressivement l'ensemble des notions liées à la question.

Point de départ

Un outil de simulation écrit à l'aide d'un tableur associé à son langage de macro-commandes ou à l'aide d'un langage de programmation, accessible en ligne ou non (des thèmes possibles : simulateur de prêt, de placement ou de crédit, de calcul de rentabilité ou de *cashflow*, de calcul de pension, d'impôts, etc.).

Principe

Balayer une partie du cycle de vie d'une solution informatique.

- Première étape – côté testeur (approche « boîte noire »)

Prendre en main l'outil de simulation à l'aide d'un mode d'emploi, le tester sur des jeux d'essai afin d'en comprendre les finalités (à quoi il sert) et de contrôler la qualité des résultats obtenus.

- Deuxième étape – côté développeur (approche « boîte blanche »)

Analyser tout ou partie du code source de l'outil de simulation pour en comprendre le comportement (comment il permet d'obtenir les résultats attendus).

- Troisième étape – côté utilisateur

Exploiter l'outil afin de produire et d'interpréter les résultats de simulations dans le cadre de scénarios.

- Quatrième étape – côté développeur

Améliorer l'outil, en corriger certaines faiblesses ou l'adapter à la prise en compte d'une nouvelle règle de gestion puis contrôler la vraisemblance et la pertinence des résultats obtenus après ces modifications grâce à la réalisation d'un jeu d'essai.

Bibliographie

- Algorithmique (3ème édition) - Thomas Cormen, Charles Leiserson, Ronald Rivest, Clifford Stein – Éditions Dunod – 2010 - ISBN-10: 2100545264
- Algorithmique - Raisonner pour concevoir - Haro C.- Éditions ENI - 2009 - ISBN 978-2-7460-4844-7
- Informatique et sciences du numérique Spécialité ISN en terminale S - Gilles Dowek, Jean-Pierre Archambault, Emmanuel Baccelli, Claudio Cimelli, Albert Cohen, Christine

Eisenbeis, Thierry Viéville, Benjamin Wack – Éditions Eyrolles - version numérique ou pdf
<http://www.editions-eyrolles.com/Livre/9782212135435>

- PHP 5 avancé (6e édition) - Éric Daspet, Cyril Pierre De Geyer – Éditions Eyrolles – 2012 - ISBN10 : 2-212-13435-5
- Savoir tout faire pour le web - PHP et MySQL - Jonathan Buttigieg , Nicolas Nunge – Éditions Oracom – 2012 - ISBN10 : 2-36145-081-X
- Introduction à la programmation avec VBA, Applications avec Excel - Robert Chevallier – Éditions Pearson – 2012 - ISBN-10 2-7440-7608-2

Question : Comment peut-on produire de l'information à partir de données contenues dans une base ?

Intentions

Les technologies de l'information et de la communication contribuent pour une large part à accompagner l'organisation et ses acteurs, en mettant à leur disposition des informations pertinentes pour leurs activités.

L'augmentation exponentielle des performances des outils numériques permet de stocker de plus en plus de données significatives de l'activité d'une organisation, rendant de plus en plus difficile leur exploitation. Une base de données relationnelle est l'un des moyens de mettre à disposition des acteurs d'une organisation des données structurées correspondant aux informations dont ils ont besoin.

Répondre à la question en distinguant donnée et information est l'occasion de comprendre les limites d'une base de données relationnelle à travers l'étude de ses règles de représentation et du langage de requêtes permettant d'en extraire des informations pertinentes en réponse à un problème de gestion.

L'usage d'une base de données permet aux élèves de s'initier à la conception de requêtes, de percevoir la différence entre modèle, schéma et base de données et de développer à cette occasion des qualités de logique, d'analyse, de synthèse et de rigueur.

Dans ce cadre, langage et outil graphique de requête ne sont pas des objets d'enseignement : ils servent à concrétiser et à donner corps aux notions abordées en mobilisant une démarche d'analyse et de résolution de problèmes.

Modèle relationnel : attribut, domaine, relation, clé, schéma relationnel, dépendance fonctionnelle, contrainte de clé, contrainte d'intégrité de domaine, contrainte d'intégrité référentielle

Lien avec la question de gestion

La production d'informations à partir du contenu d'une base de données passe par une maîtrise du modèle de représentation choisi. Il s'agit d'abord de donner du sens à un ensemble de données structurées avant de pouvoir produire de nouvelles informations. On peut par exemple montrer comment, à partir d'une base de données permettant de gérer l'activité d'une organisation, il est possible de s'appuyer sur des extractions statistiques, obtenues par des requêtes fournies, permettant d'infléchir sa stratégie (statistiques de vente par produit, par catégorie de client, par zone géographique, en quantité et en chiffre d'affaires, etc.) en s'appuyant notamment sur les exemples étudiés en classe de première dans le cadre de la question « qu'est-ce qu'une organisation performante ? » du thème Évaluation et performance.

Pré-requis, transversalités

La compréhension des concepts de type et de modèle abordés à l'occasion de la réalisation de programmes et à l'utilisation d'outils de modélisation et de simulation peut servir de base à l'introduction du modèle relationnel comme cas particulier d'outil de modélisation (voir la question « la résolution de tous les problèmes de gestion est-elle automatisable ? » du présent thème).

Attendus et limites

En s'appuyant sur des bases de données représentant des réalités de gestion simples, il s'agit de faire découvrir la représentation des objets réels (comme un produit ou un client) ou abstraits (comme une catégorie ou une vente) et les liens entre eux (comme les produits vendus à un client ou les clients appartenant à une catégorie) et comment il est possible d'en déduire des

informations utiles à l'action à travers l'exploitation de requêtes (comme le chiffre d'affaires réalisé par catégorie de client).

Modèle relationnel : attribut, domaine, relation, clé, schéma relationnel, dépendance fonctionnelle

Le **modèle relationnel** définit le mode de structuration d'une base de données et ses règles de manipulation. Il sert de référence pour la représentation, sous la forme d'un schéma relationnel, de la réalité d'une organisation et de son environnement.

Un **schéma relationnel** est donc l'abstraction de liens entre des objets physiques (des clients, des fournisseurs, des produits, des assurés, etc.) ou abstraits (des commandes, des livraisons, des règlements, des assurances, des accidents, etc.) du monde réel.

Chaque objet ou lien entre des objets est représenté sous la forme d'une **relation** constituée d'attributs. *Par exemple, un produit est représenté par une référence, un libellé, une quantité disponible et un prix.* Chaque produit présent dans l'organisation fait l'objet d'un n-uplet dans cette relation, c'est-à-dire une valorisation des différents attributs représentant une assertion (*il y a 20 exemplaires disponibles du produit de référence 215, de libellé "tablette numérique" coûtant 450 €*).

Un **attribut** représente une caractéristique d'un objet physique (le nom d'un client) ou abstrait (le montant d'un règlement, le numéro d'une assurance) ou d'un lien entre des objets (la quantité commandée d'un produit) présent dans une relation. Il prend une valeur dans un ensemble de valeurs possibles appelé un **domaine** (extension du concept de type, vu en programmation). *Par exemple, le nom d'un client est une chaîne de caractères, le montant d'un règlement est une valeur monétaire, le type de règlement en ligne prend une valeur parmi un ensemble de valeurs (comme "carte bancaire", "chèque", "paypal", "carte prépayée") et le code d'un établissement scolaire prend une valeur toujours constituée de 7 chiffres suivis d'une lettre.*

La **clé primaire** est un attribut ou la combinaison minimale de plusieurs attributs permettant de repérer un n-uplet de façon unique dans une relation : *par exemple, connaissant la référence d'un produit, on peut en déduire sans ambiguïté possible son libellé, son prix et la quantité disponible en stock.*

Une **clé étrangère** est un attribut d'une relation qui exprime une règle de gestion en faisant référence à la clé primaire d'une autre relation. *Ainsi, exprimer qu'une commande est passée par un client consiste à placer dans la relation COMMANDE un attribut faisant référence à la clé primaire présente dans la relation CLIENT.*

Les notions de clé sont indissociables de la notion de **dépendance fonctionnelle** qui exprime un lien fort entre deux attributs : la connaissance d'une valeur du premier implique la connaissance d'une valeur unique du second. *Ainsi, lorsqu'on connaît la référence d'un client on connaît son nom mais il n'est pas possible d'en déduire le numéro d'une commande qu'il a passée (car il a pu en passer plusieurs). Si on connaît le numéro d'une commande, on peut en déduire sans ambiguïté la référence du client qui l'a passée, mais pas la référence d'un produit commandé (car il peut y en avoir plusieurs).*

La notion de dépendance fonctionnelle permet ainsi de regrouper tous les attributs qui caractérisent un objet (comme un client, une commande) ou un lien entre objets (comme une ligne de commande) du système d'information de façon la plus structurée possible afin d'éviter tout risque d'ambiguïté ou d'erreur. La qualité d'un schéma relationnel dépend de l'application rigoureuse de la notion de dépendance fonctionnelle et des trois formes normales qui en découlent. Il s'agit pour les élèves de comprendre l'utilité du rôle des trois premières formes normales à travers des exemples simples.

Modèle relationnel : contrainte de clé, contrainte d'intégrité de domaine, contrainte d'intégrité référentielle

Une base de données représente l'implantation d'un schéma relationnel sous formes de tables, dont la mise à jour et l'exploitation sont réalisées à l'aide d'un système de gestion de bases de données. Chaque table de la base correspond à une relation, chaque ligne (ou enregistrement) d'une table correspond à un n-uplet d'une relation, chaque colonne (ou champ) correspond à un attribut.

C'est donc une façon de revenir à une représentation numérique concrète d'une réalité sans oublier qu'il ne s'agit toujours que d'une représentation : comme une carte routière, une base de données reste une abstraction de la réalité, plus ou moins fidèle (comme il est possible de le vérifier après s'être perdu-e en suivant les informations proposées par un GPS).

Un système de gestion de bases de données relationnelles fournit les moyens de vérifier automatiquement la cohérence des données d'une base de données avec son schéma relationnel.

La **contrainte de clé** (ou de relation) permet de vérifier que deux lignes ne peuvent être identifiées par une même valeur de clé primaire (deux clients ne peuvent avoir le même identifiant, il ne peut pas y avoir deux fois le même produit dans une commande) et qu'un objet est toujours identifié (un client doit posséder un identifiant, il ne peut pas y avoir une ligne de commande ne concernant pas un produit).

La **contrainte d'intégrité de domaine** permet d'interdire à un attribut de prendre une valeur incorrecte : le montant d'un règlement doit être numérique, le mode de règlement en ligne ne peut pas être "espèces", etc.

La **contrainte d'intégrité référentielle** (ou de clé étrangère) permet de vérifier la cohérence des règles de gestion exprimées par une clé étrangère : lorsqu'une clé étrangère a une valeur, cette valeur fait obligatoirement référence à une valeur de la clé primaire en référence. *Une commande concerne toujours un client, un règlement fait toujours référence à une facture, une livraison à une commande, etc.*

Langage de requêtes : opérateurs logiques, opérateurs relationnels, regroupement et fonctions associées

Lien avec la question de gestion

La production d'informations à partir du contenu d'une base de données repose sur la maîtrise d'une démarche de conception de requêtes et sur la pratique d'un outil de requête. Il s'agit de pouvoir extraire les données nécessaires à la production de l'information souhaitée. On peut par exemple, à partir d'une base de données permettant de gérer l'activité d'une organisation, déterminer les requêtes nécessaires à des prises de décisions opérationnelles telles que la mise en place d'une campagne promotionnelle, la détermination d'une quantité minimale de réapprovisionnement ou encore la planification des visites d'un commercial.

Pré-requis, transversalités

L'exploitation d'un langage de requêtes, en mode texte ou en mode graphique, peut être mise en parallèle avec l'exploitation de moteurs de recherche et l'utilisation de métadonnées pour l'indexation de données non structurées (voir la question « La numérisation suffit-elle à valoriser l'information ? » du présent thème).

Attendus et limites

Il s'agit d'interpréter la structure d'une base de données relationnelle représentant une réalité de gestion simple au regard d'un besoin de gestion et dans le respect des règles du modèle

relationnel pour préparer l'extraction d'informations soit à l'aide de requêtes directes, soit à l'aide de requêtes intégrées à un programme.

Langage de requêtes : opérateurs logiques, opérateurs relationnels

Un langage de requêtes permet de manipuler les données d'une base de données. On s'intéresse en priorité au langage d'interrogation d'une base de données relationnelle, graphique ou littéral – SQL- afin de permettre de comprendre le principe de fonctionnement des principaux opérateurs relationnels (sélection, projection et jointure) et logiques (ET, OU, NON) à travers différentes extractions.

Les fonctionnalités des langages de définition de données seront mobilisées pour concrétiser des opérations d'adaptation de structure d'une base de données permettant de prendre en charge de nouvelles règles de gestion simples, comme l'ajout d'un champ, d'une table ou d'une clé étrangère. On évitera les exemples trop complexes comme la création d'une clé primaire ou étrangère multi-attributs.

Les fonctionnalités d'un langage de manipulation de données seront mises en œuvre pour des requêtes d'interrogation et de mise à jour des données, notamment pour montrer comment un programme peut modifier le contenu d'une base de données à partir de la saisie d'un formulaire grâce à une requête intégrée.

L'apprentissage du langage SQL n'est pas une fin en soi. L'essentiel est de construire une démarche d'analyse et de conception de requêtes reposant sur la compréhension des données manipulées et sur un questionnement permettant de passer du besoin d'information exprimé à la requête à formuler, sans oublier la phase de contrôle de cohérence des résultats obtenus.

On pourra s'appuyer au besoin sur une représentation schématique permettant de tirer parti de l'approche déclarative du langage, y compris en exploitant un outil graphique de requêtes permettant une traduction automatique en langage SQL.

Langages de requêtes : regroupement et fonctions associées

Les opérateurs de regroupement permettent d'obtenir des informations agrégées à partir du contenu d'une base de données, en regroupant ensemble des lignes ayant des caractéristiques communes avant de leur appliquer une fonction agrégative. *On peut par exemple déterminer les quantités vendues de chaque produit en regroupant les lignes de commande par produit et en cumulant pour chaque produit les quantités commandées.*

Proposition d'exploitation pédagogique

La proposition ci-dessous est un exemple d'exploitation pédagogique permettant d'aborder progressivement l'ensemble des notions liées à la question.

Point de départ

Une solution informatique de type application *web* associée à une base de données pouvant être interrogée à l'aide d'un langage de requêtes et dont le schéma relationnel est disponible ou peut être reconstitué à partir de l'étude de la base de données.

Principe

Balayer une partie du cycle de vie d'une solution informatique.

Première étape – côté utilisateur

Exploiter la solution informatique afin de mémoriser et de traiter des informations, de produire et d'interpréter les résultats correspondant à des opérations liées à une prise de décision en lien avec une activité d'un processus prise en charge par la solution.

Deuxième étape – côté développeur

Analyser tout ou partie du schéma relationnel de la base de données pour en comprendre le sens (comment elle permet de représenter la réalité de l'organisation).

Troisième étape – côté correspondant informatique¹

Exploiter la compréhension du schéma relationnel de la base de données pour en extraire de nouvelles informations à la demande des utilisateurs.

Quatrième étape – côté développeur

Adapter la base de données à la prise en compte de nouvelles règles de gestion, tester sa validité en contrôlant la vraisemblance et la pertinence des résultats obtenus après ces modifications. Dans le prolongement des activités réalisées en liaison avec la première question du thème, il est possible d'envisager la prise en charge de modifications de la solution informatique (modification d'une interface par exemple).

Bibliographie

- Variations pédagogiques autour de l'enseignement du modèle relationnel et des SGBDR en 1ère STG: <http://www.reseaucerta.org/didactique/pub.php?num=328>
- Apprendre SQL avec MySQL - Christian Soutou – Éditions Eyrolles - 2006 - ISBN10 : 2-212-11915-1
- Des bases de données à l'Internet - Philippe Mathieu - Éditions Vuibert - 2000 - ISBN10 : 2711786692
- Conception des bases de données relationnelles - En pratique - Jacky Akpka, Isabelle Comyn-Wattiau – Éditions Vuibert - 2001 - ISBN : 978-2-7117-8678-7

¹ Le correspondant informatique, appelé aussi correspondant fonctionnel, assiste, accompagne et conseille les utilisateurs dans l'utilisation de leur poste de travail et des applications liées à leur métier.

Question : La numérisation suffit-elle à valoriser l'information ?

Intentions

Cette question doit permettre de montrer que l'abondance et la diversité des informations produites par les technologies de l'information et de la communication ne suffisent pas à leur donner de la valeur. Les données deviennent des informations par un processus d'interprétation. Donner du sens aux données pour en faire une information passe par le classement, l'indexation et la création de liens sémantiques. Ces opérations sont d'autant plus importantes que les informations ne sont pas uniquement produites à partir des données structurées stockées dans les bases de données mais sont aussi contenues dans des sources plus composites et moins structurées que sont les documents.

Document : numérisation, structuration, indexation

Lien avec la question de gestion

Un document est un objet écrit, iconographique, sonore, audiovisuel qui contient des informations. La numérisation est un préalable à son exploitation. Elle consiste à stocker le document sur un support électronique, ce stockage est de plus en plus souvent natif car les documents sont de plus en plus souvent créés sous forme électronique. Un ensemble de données se transforme en document après avoir subi les étapes de numérisation, structuration et indexation. La structuration et l'indexation permettront l'organisation de ces informations en vue de leur valorisation.

Pré-requis, transversalités

L'étude, réalisée en classe de première STMG, a permis d'initier les élèves à la recherche de documents. Ils ont ainsi déjà mobilisé certaines capacités du programme de terminale. Ils ont pu constater que l'utilisation de moteurs de recherche ne conduit pas forcément vers les sources les plus pertinentes. Un travail d'étude sémantique, de recherche de sens, incombe à l'utilisateur.

On retrouve aussi les notions de donnée, d'information et de connaissance qui sont traitées dans le programme de sciences de gestion de première, thème information et intelligence collective, question « En quoi les technologies de l'information transforment-elles l'information en ressource ? ».

Enfin, cette question de gestion est à rapprocher de la question « pourquoi la qualité du système d'information est-elle un enjeu pour l'organisation ? » du thème L'organisation informatisée puisque la problématique de la valorisation de l'information est déterminante pour la qualité du système d'information.

Attendus et limites

Disposer de documents sous forme numérique permet de les réunir, de les classer, de les conserver puis de les diffuser et les exploiter de façon automatisée. L'étape finale du cycle de vie du document est l'archivage, éventuellement la destruction. Entre la numérisation et la destruction, des traitements mis en œuvre par un ensemble d'outils et de méthodes permettent l'exploitation des ressources documentaires.

On peut considérer qu'un document matérialise la transformation de l'information en ressource pour une personne ou pour une organisation. Son élaboration, son traitement et sa diffusion conditionnent largement son accessibilité et sa capacité à être perçu comme étant pertinent.

Étant donné les capacités attendues, il s'agit de distinguer les différentes opérations de numérisation, de structuration, d'indexation et de référencement pour un document donné. La structuration ne peut se faire que sur des documents éditables auxquels on applique un marquage descriptif (balises). Ce marquage pourra être observé, ou modifié pour des éléments simples que

ce soit pour des documents bureautiques ou des documents affichables dans un navigateur. En outre, les contraintes liées à ces opérations devront être mises en évidence lors des manipulations réalisées par les élèves : normes, contraintes réglementaires, équipement à mobiliser, accessibilité, format, etc.

La numérisation nécessite du matériel qui dépend de la nature du support sur lequel se trouve le document. Ses coûts sont liés à la manipulation des documents originaux, aux capacités de stockage utilisées par certains documents (images, photographies) et à la nécessité d'une main d'œuvre qualifiée pour effectuer ce travail.

Un document numérisé a la même valeur juridique qu'un écrit sur support papier à condition d'en assurer la traçabilité et la conservation dans des conditions de nature à en garantir l'intégrité. Dans la mesure où il doit être conservé plusieurs années, il faut veiller à ce que l'évolution technologique ne soit pas un obstacle à son exploitation future. La Norme NF Z42 013 – ISO 14641-1 propose un ensemble de spécifications techniques et organisationnelles concernant le stockage, la pérennité et l'intégrité des documents numériques. Le service interministériel des archives de France diffuse les normes professionnelles en matière de numérisation.

Les droits d'auteur ou d'éditeur sont à prendre en considération lors de la numérisation qui pose un problème particulier compte tenu de la facilité de diffusion et de reproduction des documents.

Langage de définition de documents

Lien avec la question de gestion

Les langages de définition des documents servent à structurer ou formater des documents. Ils permettent un travail automatisé de présentation des contenus qui en facilite l'exploitation. Ils participent donc au processus de valorisation de l'information numérique.

Pré-requis, transversalités

Voir supra.

Attendus et limites

On retiendra le XML pour illustrer la notion de langage de définition de document.

Compte-tenu des attendus, l'élève doit être capable de distinguer données (le contenu) et métadonnées (la description du contenu) au sein d'un document. Il s'agit également de montrer comment le contenu d'un document peut être exploité par différentes applications qui en proposent des présentations différentes, par exemple aux formats PDF ou HTML.

Structuration de contenu documentaire : hyperlien, métadonnées, syndication, référencement

Lien avec la question de gestion

Pour être valorisés, les contenus documentaires doivent passer par une étape de structuration : hiérarchisation par un système d'hyperliens, description par un système de métadonnées, syndication à l'aide d'agrégateurs de contenus, référencement sur des moteurs de recherche.

Pré-requis, transversalités

La préparation au brevet informatique et internet (B2i Lycée) a initié les élèves à la conception et l'exploitation de documents numériques. Cette initiation trouve un prolongement avec la notion étudiée ici.

Attendus et limites

Un hyperlien est un moyen d'atteindre un document quelle qu'en soit la forme, à partir d'un autre document. Ce système est au cœur de la navigation entre documents sur le web. La notion d'hyperlien largement connue des élèves, doit être étudiée comme un moyen d'accéder à des informations pertinentes, soit par le biais des moteurs de recherche, soit par la navigation entre documents.

Les métadonnées sont des données qui permettent de décrire un document. Elles peuvent être utilisées dans les balises des pages web pour leur référencement mais aussi de façon explicite sous la forme de tags ou de mots-clés associés à des documents.

La syndication permet de produire un flux d'information à partir des documents publiés sur un ou plusieurs sites. Il s'agit d'un outil de veille informationnelle qui permet de passer d'une publication en mode tiré (l'internaute tire l'information d'un site qu'il visite) à une publication en mode poussé (l'internaute reçoit automatiquement l'information depuis les sites de son choix).

Le référencement de documents consiste à le rendre accessible de façon pertinente sur le web, notamment en garantissant qu'ils seront trouvés à l'aide des moteurs de recherche par les internautes, en fonction de l'information qui s'y trouve.

Gestion de contenu documentaire : fonctions et outils

Lien avec la question de gestion

La gestion de contenu documentaire est le principal moyen de valoriser le patrimoine informationnel accumulé par les organisations. Elle concerne les informations non structurées, souvent textuelles, qui ne sont pas stockées dans une base de données et qui peuvent être de nature différente : documents textes, graphiques, feuilles de calcul, courriels, *blogs*, pages *web*...

L'objectif est de faciliter la recherche et le partage d'informations pour les capitaliser et les retrouver aisément au bénéfice d'un travail collectif. Ces dispositifs contribuent à l'émergence d'une intelligence collective au sein de l'organisation.

Pré-requis, transversalités

Voir supra.

Attendus et limites

Le processus de gestion de contenu numérique est pris en charge par des logiciels et/ou matériels informatiques. Chaque logiciel est adapté à la gestion d'un type de contenu particulier.

Les WCMS (web content management system) concernent la gestion des contenus web.

La GED (gestion électronique des documents) couvre les fonctions de collecte, stockage et diffusion de l'information. Elle favorise le classement et la disponibilité des documents numérisés.

Les outils de DAM (Digital Asset Management) permettent la gestion de contenus multimédias.

Les systèmes d'exploitation modernes intègrent des fonctionnalités et outils de gestion documentaire (recherche, indexation, classement, archivage). Avant même d'utiliser un outil spécialisé, il convient de maîtriser ces premiers outils. Il n'est pas pertinent de faire manipuler plusieurs outils de gestion documentaire mais d'en choisir au plus un, y compris intégré au système d'exploitation.

Propositions d'exploitation pédagogique

Cette question peut être traitée à travers des activités transversales telles que :

- Utiliser un outil de gestion documentaire alimenté tout au long de l'année sur un thème du programme appliqué à des cas d'entreprise, par exemple la notion de trace numérique ou la sécurité des échanges. Ces contenus pourront ensuite être mutualisés au sein de la classe participant ainsi à un processus d'intelligence collective dans l'optique de la préparation de la question relative à une problématique de gestion de la partie écrite de l'épreuve de spécialité ;
- À partir d'un corpus de documents numériques, décrire ces documents grâce à des métadonnées qui permettront de définir leur contenu et de faciliter les recherches. Ce corpus peut être constitué des études produites en classe de première par les élèves de la classe. Les systèmes d'exploitation sont équipés d'outils de consultation et de modification des métadonnées. Le même travail peut être effectué sur des images ou des fichiers sonores. Cet exercice ne présente pas de difficulté technique. Il s'agit d'apprendre à repérer les informations pertinentes qui permettent la valorisation d'un document ou d'un ensemble de documents ;
- Utiliser un agrégateur tel que Google Reader ou Netvibes pour gérer des contenus sur une notion du programme (trace numérique, réseaux sociaux, communautés de pratiques, protection des données, etc.). Les contenus sont consultés et enrichis régulièrement par les élèves qui en proposent une synthèse à leurs camarades. Ces synthèses sont mises à disposition de tous via un ENT ;
- À partir d'une base de données, générer le fichier XML correspondant, le réinjecter dans une feuille de calcul pour faire l'objet de traitements issus de règles de gestion à mettre en œuvre (lien avec la question « La résolution de tous les problèmes de gestion est-elle automatisable ? » du même thème).

Cependant certaines capacités peuvent être mobilisées lors de travaux pratiques plus spécifiques :

- Créer un document avec OpenOffice. Les fichiers qui composent un document OpenOffice sont encapsulés dans une archive que l'on peut ouvrir avec un logiciel de décompression tel que 7zip. On accède alors aux fichiers XML qui composent l'archive. Les fichiers styles.xml et content.xml montrent la séparation entre le contenu du document et sa mise en forme.

Bibliographie

- http://www.openoffice.org/fr/FAQ/general_fr/xml.html
- https://help.libreoffice.org/Common/XML_File_Formats/fr
- <http://windows.microsoft.com/fr-FR/windows7/Tag-pictures-so-theyre-easier-to-find>
- http://www.bnf.fr/fr/professionnels/gerer_les_droits/i.gerer_droits_typedoc/s.numerisation_dr_oits_typedoc_manuscrits.html?first_Rub=non
- http://www.journaldunet.com/solutions/0304/030424_juridique.shtml
- http://www.bnf.fr/fr/professionnels/gerer_les_droits/i.gerer_droits_typedoc/s.numerisation_dr_oits_typedoc_manuscrits.html?first_Rub=non
- [http://fr.wikipedia.org/wiki/Marqueur_\(m%C3%A9tadonn%C3%A9e\)](http://fr.wikipedia.org/wiki/Marqueur_(m%C3%A9tadonn%C3%A9e))