

Sciences numériques et technologie

Classe de seconde

Enseignement commun

Document de travail - Ne pas diffuser

Préambule

L'enseignement « sciences numériques et technologie » en classe de seconde a pour objet de permettre d'appréhender les principaux concepts des sciences numériques, mais également de permettre aux élèves, à partir d'un objet technologique, de comprendre le poids croissant du numérique et les enjeux qui en découlent. La numérisation généralisée des données, les nouvelles modalités de traitement ou de stockage et le développement récent d'algorithmes permettant de traiter de très grands volumes de données numériques constituent une réelle rupture dans la diffusion des technologies de l'information et de la communication. Cette révolution multiplie les impacts majeurs sur les pratiques humaines.

Par exemple, l'actuel mobile multifonction est un objet technologique qui permet, comme le téléphone du XX^e siècle, de téléphoner, mais qui sert également à bien d'autres activités : envoyer des messages, photographier, filmer, enregistrer, chercher et partager une information, écouter de la musique, regarder des vidéos, repérer où l'on se trouve, réserver des billets de train, vérifier son rythme cardiaque, programmer le chauffage de son appartement, etc. Ainsi, il est devenu une interface universelle d'accès à l'information et de commande d'autres objets.

Cette convergence d'activités encore récemment indépendantes est un phénomène généralisé lié au développement de la science informatique et des technologies associées, et notamment à leur intégration avec le domaine des télécommunications, à l'informatisation massive de domaines variés (communication, audiovisuel, transports, instrumentation scientifique médicale et technique, outillage numérique, objets connectés, etc.), et bien sûr à la création du réseau internet.

Malgré leur grande variété, ces avancées se fondent toutes sur l'universalité et la flexibilité d'un petit nombre de concepts en interaction :

- les **données**, qui représentent sous une forme numérique unifiée des **informations** très diverses : textes, images, sons, mesures physiques, sommes d'argent, etc. ;
- les **algorithmes**, qui spécifient de façon abstraite et précise des traitements à effectuer sur les données à partir d'opérations élémentaires ;
- les **langages**, qui permettent de traduire les algorithmes abstraits en **programmes** textuels ou graphiques de façon à ce qu'ils soient exécutables par les machines ;
- les **machines**, et leurs systèmes d'exploitation, qui permettent d'exécuter des programmes en enchaînant un grand nombre d'instructions simples, assurent la persistance des données par leur stockage et de gérer les communications. On y inclut les **objets connectés** et les **réseaux**.

À ces concepts s'ajoute un élément transversal : les interfaces qui permettent la communication avec les humains, la collecte des données et la commande des systèmes.

L'enseignement de « sciences numériques et technologie » aide à mieux comprendre les enjeux scientifiques et sociétaux de la science informatique et de ses applications, à adopter un usage réfléchi et raisonné des technologies numériques dans la vie quotidienne et à se préparer aux mutations présentes et à venir de tous les métiers.



CONSEIL SUPÉRIEUR DES PROGRAMMES

Cet enseignement s'inscrit dans le prolongement de l'enseignement d'algorithmique, d'informatique et de programmation dispensé au collège en mathématiques et en technologie.

Le programme est organisé autour de huit thèmes : internet ; le Web ; les réseaux sociaux ; les données structurées et leur traitement ; l'intelligence artificielle ; localisation, cartographie et mobilité ; informatique embarquée et objets connectés ; la photographie numérique.

Chacun des huit thèmes a vocation à être enseigné sur une durée d'environ quatre semaines. L'ordre dans lequel sont traités les huit thèmes est au libre choix des enseignants.

La présentation de chaque thème débute par des éléments de culture scientifique et technologique qui peuvent proposer des repères historiques, expliciter les concepts et décrire les impacts sur les pratiques humaines des technologies présentées. Elle se termine par un tableau qui circonscrit précisément les connaissances et les capacités attendues des élèves, puis d'une liste, non-exhaustive ni impérative, d'activités qui peuvent être menées avec les élèves.

Cet enseignement a vocation à multiplier les occasions de mise en activité des élèves, **sous des formes variées** (exposés, travaux en groupe, mini-projets, productions individuelles ou collectives, etc.) qui permettent de développer des compétences transversales :

- faire preuve d'autonomie, d'initiative et de créativité ;
- présenter un problème ou sa solution, développer une argumentation dans le cadre d'un débat ;
- coopérer au sein d'une équipe ;
- rechercher de l'information, apprendre à utiliser des sources de qualité, partager des ressources ;
- faire un usage responsable et critique des sciences et technologies numériques.

Un langage de programmation est nécessaire pour l'écriture des programmes : un langage simple d'usage, interprété, concis, libre et gratuit, multiplateforme, largement répandu, riche de bibliothèques adaptées aux thématiques étudiées et bénéficiant d'une vaste communauté d'auteurs dans le monde éducatif est nécessaire. Au moment de la conception de ce programme, le langage choisi est Python version 3 (ou supérieure).



■ Internet

Introduction

Grâce à sa souplesse et son universalité, internet est devenu le moyen de communication principal entre les hommes et avec les machines.

Repères historiques

Dès les années cinquante, les ordinateurs ont été mis en réseau pour échanger des informations, mais de façon très liée aux constructeurs d'ordinateurs ou aux opérateurs téléphoniques. Les réseaux généraux indépendants des constructeurs sont nés aux États-Unis avec ArpaNet (1970) et en France avec Cyclades (1971). Cet effort a culminé avec internet, né en 1983.

Le protocole TCP/IP

Internet est défini par le **protocole IP** (*Internet Protocol*), ensemble de normes qui permettent d'identifier et de nommer de façon uniforme tous les ordinateurs ou objets qui lui sont connectés. IP est accompagné de protocoles de transmission pour transférer l'information par **paquets**, le principal étant **TCP/IP** (*Transmission Control Protocol*). De nature logicielle, internet s'appuie sur une grande variété de réseaux physiques où IP est implémenté. Il uniformise l'accès à tous les ordinateurs, les téléphones et les objets connectés.

Les données et l'information

Internet manipule deux types d'information : les contenus envoyés et les adresses du destinataire et de l'émetteur. Ces deux types d'information sont regroupés dans des paquets de taille fixe, de façon uniforme et indépendante du type de données transportées : texte, images, sons, vidéos, etc. Les adresses sont numériques et hiérarchiques, comme 78.109.84.114, mais l'utilisateur connaît surtout des **adresses symboliques** normalisées, comme wikipedia.fr. Le système **DNS** (*Domain Name System*) transforme une adresse symbolique en adresse numérique. Il est réalisé par un grand nombre d'ordinateurs répartis sur le réseau et constamment mis à jour.

Les algorithmes et les programmes

Le principal algorithme d'internet est le **roulage** des paquets de leurs émetteurs vers leurs destinataires. Il est effectué par des machines appelées routeurs, qui échangent en permanence avec leurs voisins pour établir une carte locale de ce qu'ils voient du réseau. Chaque paquet transite par une série de routeurs, chacun l'envoyant à un autre routeur selon sa carte locale et la destination prévue. Les routeurs s'ajustent en permanence et de proche en proche quand on les ajoute au réseau ou quand un routeur voisin disparaît. Il n'y a plus besoin de carte globale, ce qui permet le roulage à grande échelle.

Lors du roulage, un paquet peut ne pas arriver pour deux raisons : une panne matérielle d'une ligne ou d'un routeur, ou sa destruction. Chaque paquet contient l'information d'un nombre maximal de routeurs à traverser : pour ne pas encombrer le réseau, il est détruit si ce nombre est atteint. C'est le protocole TCP qui fiabilise la communication en redemandant les paquets manquants. Il garantit que tout paquet finira par arriver, sauf panne matérielle incontournable. TCP réordonne aussi les paquets arrivés dans le désordre, et diminue la congestion du réseau en gérant au mieux les redemandes. Mais ni internet ni TCP ne possèdent de garantie temporelle d'arrivée des paquets, ce qui nuit à la qualité du *streaming* du son ou des vidéos et de la téléconférence. En effet, dans une vidéo on peut perdre une image isolée, mais pas le fil du temps.

D'autres protocoles s'appuient sur ceux d'internet, par exemple les protocoles du *Web* (HTTP et HTTPS) et le protocole NTP (*Network Time Protocol*) qui permet de synchroniser finement les heures des



ordinateurs et objets connectés.

Les machines

Réseau mondial, internet fonctionne à l'aide de routeurs, de lignes de transmissions à très hauts débits (fibres optiques) entre routeurs, de réseaux de téléphonie mobile, et de réseaux locaux. Ses protocoles étant logiciels, il peut s'appuyer sur **n'importe quel réseau physique** qui les implémente : 4G, Ethernet, ADSL, Wi-Fi, Bluetooth, etc. TCP/IP n'est pas implémenté dans l'infrastructure, mais dans chacun des ordinateurs connectés, et un serveur DNS est aussi un ordinateur connecté. Des mécanismes complexes assurent la continuité de la connexion, par exemple pour passer sans interruption de téléphonie 4G au Wi-Fi, ou son ubiquité, par exemple pour passer de façon invisible d'antenne à antenne avec un téléphone portable quand on voyage.

Dans les **réseaux pair-à-pair** s'appuyant sur internet et souvent utilisé pour le transport de vidéos, chaque ordinateur sert à la fois d'émetteur et de récepteur.

Impacts sur les pratiques humaines

Internet a fait progressivement disparaître beaucoup des moyens de communication précédents : télégramme, télex, le courrier postal pour une bonne partie, et bientôt le téléphone fixe grâce à VoIP (voix sur IP).

Son trafic prévu pour 2021 est de 3 300 milliards de milliards d'octets ($3,3 \times 10^{21}$ octets).

Internet a aussi ses problèmes : absence de garantie temporelle sur l'arrivée des paquets, et possibilité d'attaques par saturation en envoyant un très grand nombre de messages à un site donné, pour y provoquer un déni de service.

La neutralité du Net, présente dès l'origine du réseau, exprime l'idée que les routeurs doivent transmettre les paquets indépendamment du type de leur contenu : texte, vidéo, etc. Mais elle est constamment remise en cause par certains lobbies industriels.

Contenus	Capacités attendues
Protocole TCP/IP : paquets, routage des paquets	Distinguer le rôle des protocoles IP et TCP. Caractériser les principes du routage et ses limites. Distinguer la fiabilité de transmission et l'absence de garantie temporelle.
Adresses symboliques et serveurs DNS	Sur des exemples réels, retrouver une adresse IP à partir d'une adresse symbolique et inversement.
Réseaux pair-à-pair	Décrire l'intérêt des réseaux pair-à-pair ainsi que les usages illicites qu'on peut en faire.
Indépendance d'internet par rapport au réseau physique	Caractériser quelques types de réseaux physiques : obsolètes ou actuels, rapides ou lents, filaires ou non. Caractériser l'ordre de grandeur du trafic de données sur internet et son évolution.
Exemples d'activités	Illustrer le fonctionnement du routage et de TCP par des activités débranchées ou à l'aide de logiciels dédiés, en tenant compte de la destruction de paquets. Déterminer l'adresse IP d'un équipement et l'adresse du DNS sur un réseau. Analyser son réseau local pour observer ce qui y est connecté. Suivre le chemin d'un courriel en utilisant une commande du protocole IP.



■ Le Web

Introduction

Le *Web* (toile ou réseau) désigne un système donnant accès à un ensemble de données (page, image, son, vidéo) reliées par des liens hypertextes et accessibles sur le réseau internet.

Repères historiques

- 1965 : invention et programmation du concept d'hypertexte par Ted Nelson ;
- 1989 : naissance au CERN par Tim Berners Lee ;
- 1993 : mise dans le domaine public, disponibilité du premier navigateur Mosaic ;
- 1995 : mise à disposition de technologies pour le développement de site Web interactif (langage JavaScript) et dynamique (langage PHP) ;
- 2001 : standardisation des pages grâce au DOM (*Document Object Model*) ;
- 2010 : mise à disposition de technologies pour le développement d'applications sur mobiles.

Normalisation de la présentation de l'information

Sur le *Web*, les textes, photos, vidéos, graphiques, sons, programmes sont exprimés et assemblés dans divers formats normalisés par un consortium mondial (W3C : *World Wide Web Consortium*), ce qui permet une circulation standardisée de ces informations.

Les pages *Web* sont écrites dans le **langage de balises HTML** (*HypertextMarkupLanguage*). Leur style graphique est défini dans le **langage CSS** (*Cascading Style Sheets*).

Les pages ont une adresse unique, nommée **URL** (*Uniform Ressource Locator*). Elles sont accessibles via internet en utilisant le protocole HTTP (*Hypertext Transfer Protocol*) ou sa version sécurisée HTTPS qui crypte les échanges. L'affichage des pages est réalisé chez l'utilisateur par un programme appelé navigateur.

Un hypertexte est un texte augmenté de renvois automatiques à des textes, des images ou des sons. Initialement, un hypertexte se restreignait à la mémoire d'un seul ordinateur. Dans une page *Web*, ce renvoi se fait sur n'importe quelle machine du réseau internet, par le truchement de l'adresse de la page *Web* du texte (URL) auquel il fait référence. La toile d'araignée construite par les liens peut être représentée sous forme d'un graphe qui matérialise la structure du *Web*.

Moteurs de recherche

Les moteurs de recherche permettent de trouver des informations dans des pages dont on ne connaît pas l'adresse, voire dont on ignore l'existence. La méthode de recherche appelée référencement naturel se décompose en trois grandes activités, réalisées par les moteurs de recherche : (1) le parcours automatique du *Web* pour collecter les pages visitées (aspiration des pages *Web* effectuée par des robots) ; (2) l'analyse du contenu des pages et leur indexation sur les mots qu'elles contiennent (constitution d'un annuaire inversé qui associe à chaque terme les URL des pages où il apparaît) ; (3) la troisième activité, réalisée à chaque fois qu'un internaute fait une requête, construit une liste ordonnée des pages (classement) comportant les mots clés de la requête. Leur ordre dépend notamment de leur popularité (principe des liens), de leur pertinence (aux mots de la requête), et de l'ordre des termes de la requête.

Les concepteurs de site *Web* peuvent améliorer le référencement de leurs pages en choisissant bien les mots et en les plaçant à des endroits stratégiques dans les pages.

Interaction client/serveur

Le *Web* s'appuie sur le dialogue entre **clients** et **serveurs**. L'interaction est à l'initiative des clients (les applications qui se connectent au *Web*, dont les navigateurs), qui envoient des **requêtes HTTP** aux serveurs. Ces derniers renvoient leur résultat : des pages qu'ils ont stockées ou qu'ils créent dynamiquement en fonction de la requête formulée. Les pages reçues par les clients peuvent contenir des codes exécutables (souvent en JavaScript) qui permettent aux clients d'effectuer des traitements en accédant aux ressources de son ordinateur et en interagissant avec les serveurs.

Les applications peuvent être paramétrées pour autoriser ou interdire l'accès à des ressources locales aux programmes téléchargés par les pages.

Sécurité et confidentialité

En formulant des requêtes sur des sites *Web* dynamiques et en laissant des programmes s'exécuter sur sa machine, l'utilisateur prend des risques : il peut communiquer des informations personnelles à son insu à des serveurs qui en gardent une trace, à distance ou localement par des *cookies*, ou encore charger des pages contenant des programmes malveillants, par exemple permettant d'espionner en continu les actions de l'utilisateur. Par ailleurs, un navigateur peut garder un historique de toutes les interactions, et le laisser accessible aux sites connectés. L'utilisateur peut utiliser des services qui s'engagent à ne pas garder de traces de ses interactions, par exemple certains moteurs de recherche. Il peut aussi **paramétrer son navigateur** de façon à ce que celui-ci n'enregistre pas d'historique des interactions. De fausses pages peuvent encore être utilisées pour l'hameçonnage des utilisateurs. Un nom de lien pouvant cacher une adresse *Web* malveillante, il faut examiner cette adresse avant de l'activer par un clic.

Impacts sur les pratiques humaines

Dans l'histoire de la communication, le *Web* est une révolution : il a ouvert à tous la possibilité et le droit de publier ; il permet une coopération d'une nature nouvelle entre individus et entre organisations : commerce en ligne, création et distribution de logiciels libres multi-auteurs, création d'encyclopédies mises à jour en permanence, etc. ; il devient universel pour communiquer avec les objets connectés.

Le *Web* permet aussi de diffuser toutes sortes d'informations dont ni la qualité, ni la pertinence, ni la véracité ne sont garanties et dont la vérification des sources n'est pas toujours facile. Il conserve des informations, parfois personnelles, accessibles partout sur de longues durées sans qu'il soit facile de les effacer, ce qui pose la question du droit à l'oubli. Il permet une exploitation de ses données, dont les conséquences sociales sont encore difficiles à estimer : recommandation à des fins commerciales, bulles informationnelles, etc. En particulier, des moteurs de recherche permettent à certains sites d'acquérir de la visibilité sur la première page des résultats de recherche en achetant de la publicité qui apparaîtra parmi les liens promotionnels.

Contenus	Capacités attendues
Repères historiques	Définir les étapes du développement du <i>Web</i> .
Hypertexte	Maîtriser les renvois d'un texte à différents contenus.
Langages HTML et CSS	Distinguer ce qui relève du contenu d'une page et de son style de présentation. Étudier et modifier une page HTML simple.
URL	Décomposer l'URL d'une page. Reconnaître les pages sécurisées.
Requête HTTP	Décomposer le contenu d'une requête HTTP et identifier les paramètres passés.
Modèle client/serveur	Inspecter le code d'une page hébergée par un serveur et distinguer



	ce qui est exécuté par le client et par le serveur.
Moteurs de recherche : principes et usages	Mener une analyse critique des résultats fournis par un moteur de recherche.
Paramètres de sécurité d'un navigateur	Maîtriser les réglages les plus importants concernant la gestion des cookies, la sécurité et la confidentialité d'un navigateur.
Exemples d'activités	
Construire une page <i>Web</i> simple contenant des liens hypertextes, la mettre en ligne. Modifier une page <i>Web</i> existante, changer la mise en forme d'une page en modifiant son CSS. Insérer un lien dans une page <i>Web</i> . Comparer les paramétrages de différents navigateurs. Utiliser plusieurs moteurs de recherche, comparer les résultats et s'interroger sur la pertinence des classements. Réaliser à la main l'indexation de quelques textes sur quelques mots puis choisir les textes correspondant à une requête. Calculer la popularité d'une page à l'aide d'un graphe simple puis programmer l'algorithme. Paramétrer un navigateur de manière qu'il interdise l'exécution d'un programme sur le client. Comparer les politiques des moteurs de recherche quant à la conservation des informations sur les utilisateurs. Effacer l'historique du navigateur, consulter les cookies, paramétrer le navigateur afin qu'il ne garde pas de traces. Utiliser un outil de visualisation tel que Cookieviz pour mesurer l'impact des <i>cookies</i> et des traqueurs lors d'une navigation.	

■ Les réseaux sociaux

Introduction

Les réseaux sociaux sont des applications basées sur les technologies du *Web* qui offrent un service de mise en relation d'internautes pour aider à développer des communautés d'intérêts.

Repères historiques

- 1995 : Classmates est l'un des premiers réseaux sociaux qui permettent aux étudiants de rester en relation ;
- 2003 : apparition de Myspace, aujourd'hui en perte de vitesse, et de LinkedIn (racheté depuis par Microsoft), à vocation professionnelle ;
- 2004 : apparition de Facebook, d'abord réservé aux étudiants de l'université Harvard, puis ouvert au grand public en 2006 ;
- 2006 : apparition de Twitter, qui permet l'échange de courts messages, limités au départ à 140 puis à 280 caractères (on parle de microblogage) ;
- 2009 : lancement de la messagerie instantanée WhatsApp (rachetée depuis par Facebook) qui se substitue à l'utilisation des SMS et MMS chez beaucoup d'utilisateurs ;
- 2010 : arrivée d'Instagram (racheté depuis par Facebook), qui permet le partage de photos et de vidéos ;
- 2011 : début de Snapchat qui permet, sur plateformes mobiles, le partage de photos et de vidéos, avec une limitation de durée ;
- 2018 : on estime à 3,2 milliards le nombre d'utilisateurs actifs des réseaux sociaux.

En 2018, les réseaux sociaux utilisés en France sont états-uniens, toutefois il en existe bien d'autres : en Chine, par exemple, apparaît en 2009 l'application de microblogage Weibo avec plus de 350 millions

d'utilisateurs actifs en 2018 ; en 2012 naît l'application de messagerie Weixin (développée par Tencent) qui compte en 2018 plus d'un milliard de comptes utilisateurs.

Les données et l'information

Les différents réseaux sociaux permettent l'échange d'informations de natures différentes : textes, photos, vidéos. Certains limitent strictement la taille des informations, d'autres autorisent la publication, mais de façon limitée dans le temps. Certains permettent l'adjonction d'applications tierces (*plug-ins*) qui peuvent ajouter des fonctionnalités supplémentaires.

Toutes les applications de réseautage social utilisent d'importantes bases de données qui gèrent leurs utilisateurs, l'ensemble des données qu'ils partagent, mais aussi celles qu'ils consentent à fournir (sans toujours le savoir), y compris sur leur vie personnelle.

Les algorithmes et les programmes

De très nombreux algorithmes sont mis en œuvre par les applications de réseautage social.

Toutes les applications s'appuient sur des services de mise en relation avec des internautes membres du réseau, relations ou amis communs : des algorithmes opérant sur les graphes et sur les bases de données sont au cœur de ces services.

À l'aide d'algorithmes de recommandation, les réseaux sociaux suggèrent aux utilisateurs des amis, des contenus, des annonces promotionnelles. Ils permettent aussi aux plateformes sociales d'étudier les comportements de leurs utilisateurs à des fins commerciales, politiques ou d'amélioration du service.

Impacts sur les pratiques humaines

Le développement des réseaux sociaux introduit un nouveau type de liens sur le *Web*, qui ne relève pas de l'hypertexte : il s'agit de l'abonnement à des relations/des amis et de la possibilité de recommander de l'information en fonction du réseau ainsi constitué.

L'objectif annoncé des applications de réseautage social est de mettre les individus en relation les uns avec les autres. Quelle est la réalité ? L'expérience de **Milgram** (1967) semble indiquer la constitution de « **petits mondes** » où chacun est au plus à six liens de distance d'un autre. Peut-on éviter les phénomènes de communautés liées à des recommandations se renforçant les unes les autres pouvant aller jusqu'à un appauvrissement de la pensée critique ? Ces questions font référence au concept de *bonding* (renforcement de liens existants au sein d'un même groupe) versus *bridging* (construction de nouveaux liens non redondants).

Les affaires de fuite de données personnelles mettent en avant les questions liées aux modèles économiques des applications de réseautage social symbolisés par le slogan « *quand c'est gratuit, c'est vous le produit* ».

Les réseaux sociaux peuvent être le support d'un **harcèlement numérique**, par le biais de photographies partagées sans consentement ou impossibles à retirer, par la diffusion de fausses nouvelles, de dénégations ou de calomnies. Des pratiques, des outils et des services permettent de se protéger, lutter et dénoncer de tels agissements.

Contenus	Capacités attendues
Réseaux sociaux existants	Distinguer plusieurs réseaux sociaux selon leurs caractéristiques, y compris un ordre de grandeur de leurs nombres d'abonnés. Paramétrer des abonnements pour assurer la confidentialité de données personnelles.



Modèle économique des réseaux sociaux	Identifier les sources de revenus des entreprises de réseautage social.
Rayon, diamètre et centre d'un graphe	Déterminer ces caractéristiques sur des graphes simples.
Notion de « petit monde » Expérience de Milgram	Décrire comment l'information présentée par les réseaux sociaux est conditionnée par le choix préalable de ses amis.
Harcèlement numérique	Analyser l'article 222-33-2-2 du code pénal.
Exemples d'activités	
Construire ou utiliser une représentation du graphe des relations d'un utilisateur. S'appuyer sur la densité des liens pour identifier des groupes, des communautés. Sur des exemples de graphes simples, en informatique débranchée, étudier les notions de rayon, diamètre et centre d'un graphe, de manière à illustrer la notion de « petit monde ». Comparer les interfaces et fonctionnalités de différents réseaux sociaux. Dresser un comparatif des formats de données, des possibilités d'échange ou d'approbation (bouton <i>like</i>), de la persistance des données entre différents réseaux sociaux. Analyser les paramètres d'utilisation d'un réseau social. Analyser les autorisations données aux applications tierces. Discuter des garanties d'authenticité des comptes utilisateurs ou des images. Lire et expliquer les conditions générales d'utilisation d'un réseau social. Consulter le site nonauharcèlement.education.gouv.fr .	

■ Les données structurées et leur traitement

Introduction

Les données constituent la matière première de toute activité numérique. Afin de permettre leur réutilisation, il est nécessaire de les conserver de manière persistante. Les structurer correctement garantit que l'on puisse les exploiter facilement pour produire de l'information. Cependant, les données non structurées peuvent aussi être exploitées, par exemple par les moteurs de recherche.

Repères historiques

- 1930 : utilisation des cartes perforées, premier support de stockage de données ;
- 1956 : invention du disque dur permettant de stocker de plus grandes quantités de données, avec un accès de plus en plus rapide ;
- 1970 : invention du modèle relationnel (E. L. Codd) pour la structuration et l'indexation des bases de données ;
- 1979 : création du premier tableur, VisiCalc ;
- 2009 : *Open Government Initiative* du président Obama ;
- 2013 : charte du G8 pour l'ouverture des données publiques.

Les données et l'information

Une **donnée** est une valeur décrivant un objet, une personne, un événement digne d'intérêt pour celui qui choisit de la conserver. Par exemple, le numéro de téléphone d'un contact est une donnée. Plusieurs **descripteurs** peuvent être utiles pour décrire un même objet (par exemple des descripteurs permettant de caractériser un contact : nom, prénom, adresse et numéro de téléphone).

Une **collection** regroupe des objets partageant les mêmes descripteurs (par exemple, la collection des contacts d'un carnet d'adresses). La structure de table permet de présenter une collection : les objets en ligne, les descripteurs en colonne et les données à l'intersection. Les données sont alors dites structurées. Pour assurer la persistance des données, ces dernières sont stockées dans des fichiers. Le format CSV



(*Comma Separated Values*, les données avec des séparateurs) est un format de fichier simple permettant d'enregistrer une table. À tout fichier sont associées des **métadonnées** qui permettent d'en décrire le contenu. Ces métadonnées varient selon le type de fichier (date et coordonnées de géolocalisation d'une photographie, auteur et titre d'un fichier texte, etc.).

Les données comme les métadonnées peuvent être capturées et enregistrées par un dispositif matériel ou bien renseignées par un humain. Elles sont de différents types (numériques, textes, dates) et peuvent être traitées différemment (calcul, tri, affichage, etc.).

Certaines collections typiques sont utilisées dans des applications et des formats standardisés leur sont associés : par exemple le format ouvert vCard (extension .vfc) pour une collection de contacts.

Une **base de données** regroupe plusieurs collections de données reliées entre elles. Par exemple, la base de données d'une bibliothèque conserve les données sur les livres, les abonnés et les emprunts effectués.

Les algorithmes et les programmes

La recherche dans des **données structurées** a d'abord été effectuée selon une indexation préalable faite par l'homme. Des algorithmes ont ensuite permis d'automatiser l'indexation à partir de textes, d'images ou de sons.

Une table de données peut faire l'objet de différentes opérations : rechercher une information précise dans la collection, trier la collection sur une ou plusieurs propriétés, filtrer la collection selon un ou plusieurs tests sur les valeurs des descripteurs, effectuer des calculs, mettre en forme les informations produites pour une visualisation par les utilisateurs.

La recherche dans une base comportant plusieurs collections peut aussi croiser des collections différentes sur un descripteur commun ou comparable.

Les machines

Les fichiers de données sont stockés sur des supports de stockage : internes (disque dur ou SSD) ou externes (disque, clé USB), locaux ou distants (**cloud**). Ces supports pouvant subir des dommages entraînant des altérations ou des destructions des données, il est nécessaire de réaliser des sauvegardes.

Des recherches dans les fichiers se font à l'intérieur même des ordinateurs, soit sur la base de leurs métadonnées, soit sur la base d'une indexation (à la manière des moteurs de recherche sur le *Web*).

Les grandes bases de données sont souvent implémentées sur des serveurs dédiés (machines puissantes avec une importante capacité de stockage sur disques). Ces centres de données doivent être alimentés en électricité et maintenus à des températures suffisamment basses pour fonctionner correctement.

Impacts sur les pratiques humaines

L'évolution des capacités de stockage, de traitement et de diffusion des données fait qu'on assiste aujourd'hui à un phénomène de surabondance des données et au développement de nouveaux algorithmes capables de les exploiter.

L'exploitation de données massives (*Big Data*) est en plein essor dans des domaines aussi variés que les sciences, la santé ou encore l'économie. Les conséquences sociétales sont nombreuses tant en termes de démocratie, de surveillance de masse ou encore d'exploitation des données personnelles.

Certaines de ces données sont dites ouvertes (*OpenData*), leurs producteurs considérant qu'il s'agit d'un bien commun. Mais on assiste aussi au développement d'un marché de la donnée où des entreprises collectent et revendent des données sans transparence pour les usagers. D'où l'importance d'un cadre



juridique permettant de protéger les usagers, préoccupation à laquelle répond le règlement général sur la protection des données (RGPD).

Les centres de données (*datacenter*) stockent des serveurs mettant à disposition les données et des applications les exploitant. Leur fonctionnement nécessite des ressources (en eau pour le refroidissement des machines, en électricité pour leur fonctionnement, en métaux rares pour leur fabrication) et génère de la pollution (manipulation de substances dangereuses lors de la fabrication, de la destruction ou du recyclage). De ce fait les usages numériques doivent être pensés de façon à limiter la transformation des écosystèmes (notamment le réchauffement climatique) et à protéger la santé humaine.

Contenus	Capacités attendues
Données	Identifier les principaux formats et représentations de données.
Données structurées	Identifier les différents descripteurs d'un objet. Distinguer la valeur d'une donnée de son descripteur. Utiliser un site de données ouvertes, pour sélectionner et récupérer des données.
Traitement de données structurées	Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.
Métadonnées	Retrouver les métadonnées d'un fichier personnel.
Données dans le nuage (<i>cloud</i>)	Utiliser un support de stockage dans le nuage. Partager des fichiers, paramétrer des modes de synchronisation. Identifier les principales causes de la consommation énergétique des centres de données ainsi que leur ordre de grandeur.
Exemples d'activités	Consulter les métadonnées de fichiers correspondant à des informations différentes et repérer celles collectées par un dispositif et celles renseignées par l'utilisateur. Télécharger des données ouvertes (sous forme d'un fichier au format CSV avec les métadonnées associées), observer les différences de traitements possibles selon le logiciel choisi pour lire le fichier : programme Python, tableur, éditeur de textes ou encore outils spécialisés en ligne. Explorer les données d'un fichier CSV à l'aide d'opérations de tri et de filtre, effectuer des calculs sur ces données, réaliser une visualisation graphique des données. À partir de deux tables de données ayant en commun un descripteur, montrer l'intérêt des deux tables pour éviter les redondances et les anomalies d'insertion et de suppression, réaliser un croisement des données permettant d'obtenir une nouvelle information. Illustrer, par des exemples simples, la consommation énergétique induite par le traitement et le stockage des données.

■ L'intelligence artificielle

Introduction

L'intelligence artificielle ou IA, née en 1956, visait à simuler sur ordinateurs les facultés cognitives humaines et recouvrait des approches scientifiques relevant des sciences informatique et mathématique, et à la frontière des sciences cognitives. Aujourd'hui, on parle surtout d'IA pour désigner l'apprentissage machine, qui envahit un grand nombre de domaines de la vie courante (analyse d'image, reconnaissance de la parole, etc.). Il existe néanmoins d'autres applications de l'intelligence artificielle comme le traitement et la traduction automatique des langues naturelles, la démonstration automatique de théorèmes ou la représentation des connaissances que l'on utilise couramment.



Repères historiques

- 1950, article fondateur d'Alan Turing, *Computing Machinery and Intelligence* ;
- 1956, école d'été de Dartmouth, organisée par John McCarthy, Marvin Minsky, Nathaniel Rochester et Claude Shannon ;
- 1957, publication de l'algorithme Perceptron de Frank Rosenblatt ;
- 1959, programmation par Arthur Samuel des premiers algorithmes d'apprentissage par renforcement ; 1997, victoire de Deep Blue (IBM) aux échecs ;
- 2006, introduction de l'apprentissage par réseaux de neurones profonds, par Geoffrey Hinton, Joshua Bengio et Yann Le Cun ;
- 2014, mise en œuvre du logiciel de reconnaissance de visage DeepFace (Facebook) ;
- 2016, victoire de AlphaGo (Google DeepMind) au jeu de Go.

L'apprentissage machine a pris un essor considérable ces dernières années, avec des résultats spectaculaires grâce à l'apprentissage profond.

Les données et l'information

Une machine apprend si elle tire parti de sa propre expérience pour modifier son programme et améliorer son comportement. Les données jouent un rôle fondamental pour la plupart des systèmes d'apprentissage : c'est à partir d'elles qu'ils sont entraînés.

L'apprentissage machine utilise souvent des données massives et de toutes natures (mesures de capteurs d'une voiture, textes pour la traduction automatique de langues, sons pour la reconnaissance de la parole, images médicales, etc.).

Le choix des données, de leur structure et des propriétés qui vont les décrire a un impact sur la qualité de l'apprentissage. La qualité de ces données est essentielle pour celle des résultats.

Une interprétation trop rapide des données et un amalgame entre **corrélation** et **causalité** peuvent entraîner des conclusions incorrectes. Les techniques d'apprentissage ont comme danger la conservation et même l'amplification des **biais** des données.

Les algorithmes et les programmes

On distingue plusieurs types d'apprentissage.

Apprentissage non supervisé. Il vise à découvrir des régularités dans les données et permet une analyse exploratoire des données. On peut ainsi organiser une collection de données en regroupant des données dans des groupes (*clusters*), en fonction de leur proximité (principe de similarité). Des exemples d'algorithme non supervisé sont le *clustering* hiérarchique qui construit une hiérarchie de groupes représentée par un arbre, et la méthode des k-moyennes qui regroupe les données en k groupes en fonction de leur distance au centre du groupe.

Dans l'analyse des données, on peut aussi s'intéresser à la découverte d'**associations** dans une collection de données décrites par des descripteurs communs. C'était le cas initialement pour la recommandation en ligne, où l'on étudie les paniers d'achats des utilisateurs, pour en déduire des règles d'association. Ces règles indiquant des co-occurrences possibles et ne sont pas des causalités.

Apprentissage supervisé. Il part de données étiquetées et vise à construire une procédure de classification compatible avec l'étiquetage des données. Appartiennent à cette famille d'algorithmes les arbres de décision dont les branches décrivent les propriétés qui doivent être satisfaites pour que la classe indiquée dans les feuilles soit prédite, et la méthode des k plus proches voisins qui prédit la classe

d'un élément en considérant la classe majoritaire de ses k plus proches voisins. La surreprésentation d'une classe pour les données étiquetées peut biaiser les prédictions.

On distingue aussi l'**apprentissage par renforcement**, dans lequel le système décide d'effectuer une action en fonction de son état pour interagir avec son environnement, celui-ci lui renvoyant un renforcement sous la forme d'une récompense positive ou négative.

Les succès récents des méthodes d'apprentissage doivent beaucoup à l'augmentation de la puissance de calcul et de stockage des machines, ainsi qu'à l'explosion des données disponibles et à la capacité de les traiter à grande échelle.

Impacts sur les pratiques humaines

Les algorithmes d'apprentissage profond sont très performants, mais souffrent d'une opacité de leur fonctionnement et de l'imprédictibilité de leurs résultats.

L'IA a une présence de plus en plus importante dans de nombreuses activités humaines, comme le domaine de la santé ou de l'accompagnement des personnes, mais aussi dans les activités de commerce (recommandation de produits), les activités sociales (accompagnement des personnes, proposition de nouveaux amis sur les réseaux sociaux), de communication (traduction automatique de la langue, assistant numérique personnel) ou encore les transports (développement de la voiture intelligente) ou la sécurité. On assiste à une transformation progressive des métiers qu'il faut savoir anticiper.

Contenus	Capacités attendues
Différents champs d'application de l'IA	Identifier l'impact de l'intelligence artificielle dans différentes activités humaines.
Distinguer les types d'apprentissage	Décrire des exemples d'utilisation de différents types d'apprentissage.
Corrélation et causalité	Sur des exemples simples, montrer qu'une corrélation ne correspond pas toujours à une relation de causalité.
Biais des données et pertinence d'un résultat	Sur des exemples simples, montrer l'importance d'éviter tout biais des données fournies à un algorithme d'apprentissage.
Exemples d'activités	
<p>Dans une collection d'objets représentés par une table, repérer des répétitions de valeurs, pour découvrir des associations entre descripteurs. Montrer comment des descripteurs manquants peuvent suggérer de fausses causalités.</p> <p>En utilisant un jeu de données, établir une règle permettant de résoudre un problème donné. Utiliser un algorithme de <i>clustering</i> avec un autre jeu de données afin de mettre en défaut une règle trop simpliste. Aborder alors la différence entre causalité et corrélation.</p> <p>Avec l'algorithme des k plus proches voisins, montrer qu'on n'effectue pas nécessairement les mêmes prédictions selon la collection initiale de données étiquetées et la valeur du paramètre k choisi. Montrer l'importance d'utiliser des données non biaisées.</p> <p>Illustrer par un exemple simple ou une vidéo le principe de l'apprentissage par retour d'erreur du Perceptron ou d'un réseau de neurones simple.</p>	

■ Localisation, cartographie et mobilité

Introduction

La cartographie est essentielle pour beaucoup d'activités : agriculture, urbanisme, transports, loisirs, etc. Elle a été révolutionnée par l'arrivée des cartes numériques accessibles depuis les ordinateurs, tablettes et téléphones, bien plus souples à l'usage que les cartes papier.



Les **cartes numériques** rassemblent toutes les échelles et permettent de montrer différents aspects de la région visualisée sur une seule carte. Les algorithmes de recherche permettent de retrouver sur la carte les endroits en donnant simplement leur nom, et de calculer des itinéraires entre points selon des modes de transports variés.

Repères historiques

Les cartes ont été systématiquement numérisées à la fin du XX^e siècle.

Le principal instrument de localisation, **GPS** (*Global Positioning System*), a été conçu par l'armée américaine dans les années soixante. Le premier satellite GPS fut lancé en 1978. Il y en a actuellement une trentaine, de sorte qu'à tout moment quatre à six satellites au moins sont visibles depuis tout point de la Terre. Couplé aux cartes numériques, le système GPS permet de se situer. Il n'est pas toujours efficace en ville, et peut être complété par d'autres moyens de localisation comme la détection de bornes Wi-Fi proches. D'autres systèmes plus précis, dont **Galileo**, sont en cours de déploiement.

Les données et l'information

Les informations des cartes numériques proviennent de nombreuses sources : services géographiques des États, photos prises par des satellites, avions ou voitures, données fournies par les utilisateurs, etc. Ces informations sont de natures diverses : topographiques, géologiques, photographiques, liées aux transports, à l'activité industrielle ou touristique, etc. Des projets collaboratifs comme *OpenStreetMap* permettent à chaque utilisateur d'ajouter des informations à une carte en libre accès, qui deviennent alors visibles par tous les utilisateurs.

Un satellite GPS contient des horloges atomiques mesurant le temps à une très grande précision et envoyant régulièrement des messages contenant cette heure. Chaque message se propageant à la vitesse de la lumière, le récepteur peut calculer sa distance au satellite. On peut en déduire sa position en suivant plusieurs satellites, ce que fait automatiquement le récepteur GPS.

Les algorithmes et les programmes

Les algorithmes cartographiques concernent principalement l'affichage sélectif d'informations variées et le **calcul d'itinéraires**. L'affichage est paramétré par les informations à montrer, que l'on peut choisir par simples clics. Une difficulté est liée au mélange d'informations de types différents lors des changements d'échelle : les graphismes peuvent être très différents et beaucoup d'informations doivent être supprimées pour les grandes échelles, mais une route doit être représentée avec à peu près la même largeur, quelle que soit l'échelle.

Les récepteurs GPS fournissent la localisation sous une forme normalisée facilement décodable, par exemple selon le **protocole NMEA 0183** (*National Marine Electronics Association*), ou directement dans les métadonnées EXIF d'une photo. La localisation et les cartes se couplent dans le suivi permanent de la position sur la carte ou sur un itinéraire précalculé.

Les machines

Les machines utilisées pour la cartographie sont surtout les ordinateurs, tablettes et téléphones classiques équipés d'une application ad hoc. Les récepteurs GPS spécialisés restent importants pour la navigation maritime ou aérienne, mais ceux pour la randonnée pédestre sont en voie de disparition, supplantés par les téléphones.

L'heure fournie par le GPS sert aussi de base pour la synchronisation précise des horloges internes des



ordinateurs connectés à internet, ce qui est très important pour tous les échanges d'informations.

Impacts sur les pratiques humaines

Les cartes numériques, accessibles depuis un téléphone, remplacent progressivement les cartes sur papier. Leurs interfaces permettent d'accéder commodément à de nombreux types d'information. Couplé aux algorithmes de calculs d'itinéraires, le GPS est utilisé systématiquement pour les transports, l'agriculture, la randonnée, la navigation à voile, etc.

Le maintien à jour des cartes numériques est un problème difficile qui demande beaucoup de ressources au plan mondial. Les erreurs dans les cartes, inévitables à cause de l'énorme quantité d'informations à collecter, peuvent avoir des conséquences dramatiques.

Par ailleurs, de nombreuses applications ont accès à la localisation dans un téléphone, ce qui leur permet d'envoyer des publicités non désirées, de suivre vos itinéraires, ou de localiser une personne. Enfin, le GPS n'est pas toujours sûr, car facile à brouiller à l'aide d'appareils simples.

Contenus	Capacités attendues
GPS, Galileo	Décrire le principe de fonctionnement de la géolocalisation.
Cartes numériques	Identifier les différentes couches d'information de GeoPortail pour extraire différents types de données. Contribuer à OpenStreetMap de façon collaborative.
Protocole NMEA 0183	Décoder une trame NMEA pour trouver des coordonnées géographiques.
Calculs d'itinéraires	Utiliser un logiciel pour calculer un itinéraire. Représenter un calcul d'itinéraire comme un problème sur un graphe.
Confidentialité	Régler les paramètres de confidentialité d'un téléphone pour partager ou non sa position.
Exemples d'activités	
Expérimenter la sélection d'informations à afficher et l'impact sur le changement d'échelle de cartes (par exemple sur GeoPortail), ainsi que les ajouts d'informations par les utilisateurs dans OpenStreetMap. Mettre en évidence les problèmes liés à un changement d'échelle dans la représentation par exemple des routes ou de leur nom sur une carte numérique pour illustrer l'aspect discret du zoom. Calculer un itinéraire routier entre deux points à partir d'une carte numérique. Connecter un récepteur GPS sur un ordinateur afin de récupérer la trame NMEA, en extraire la localisation. Extraire la géolocalisation des métadonnées d'une photo. Situer sur une carte numérique la position récupérée.	

■ Informatique embarquée et objets connectés

Introduction

Embarquer l'informatique dans les objets a beaucoup d'avantages : simplifier leur fonctionnement, leur donner plus de possibilités d'usage et de sûreté, et leur permettre d'intégrer de nouvelles possibilités à matériel constant par simple modification de leur logiciel.

Après avoir transformé les chaînes de montage des automobiles et les avions dans les années quatre-vingt-dix, l'informatique intervient maintenant dans des domaines toujours plus nombreux :



automobile, réseau ferroviaire et transports urbains, domotique, robotique, loisirs, etc., conduisant à un nouvel internet des objets.

Pour les avions par exemple, l'informatique gère le vol en commandant finement des servomoteurs électriques, plus légers et plus fiables que les vérins hydrauliques, les réacteurs, la navigation et le pilotage automatique, et permet l'atterrissage automatique par temps de brouillard. Elle a eu un impact décisif sur l'amélioration de la sécurité aérienne.

Les objets informatisés avaient autrefois des **interfaces homme-machine (IHM)** dédiées, souvent dépendantes d'une liaison filaire directe. Mais les technologies du *Web* intégrées au téléphone portable permettent maintenant d'y rassembler les interfaces des objets du quotidien, ce qui en simplifie et uniformise l'usage. Les objets informatisés deviennent ainsi connectés.

Repères historiques

- 1967 : premier système embarqué de guidage lors de la mission lunaire Apollo ;
- 1971 : premier processeur produit par Intel ;
- 1984 : sortie de l'Airbus 320, premier avion équipé de commandes électriques informatisées ;
- 1998 : mise en service du métro informatisé sans conducteur Météor (ligne 14 à Paris) ;
- 1999 : introduction de l'expression « internet des objets » par Kevin Ashton ;
- 2007 : arrivée du *smartphone*.

On estime à 50 milliards le nombre d'objets connectés en 2020.

Les données et l'information

Dans les **systèmes informatiques embarqués**, l'information provient soit des IHM soit des capteurs, pour contrôler automatiquement ou manuellement le fonctionnement physique par des actionneurs et transmettre des informations aux utilisateurs. Le flux d'informations à travers les IHM permet ainsi une interaction continue entre l'homme et la machine.

Les algorithmes et les programmes

Le développement des logiciels embarqués est délicat, car il pose souvent des questions de temps-réel, c'est-à-dire de respect de temps de réponse imposé. Ceci conduit à des méthodes de programmation spécifiques.

Les machines

Les microprocesseurs sont beaucoup plus nombreux dans les objets que dans les ordinateurs et téléphones, mais ils sont souvent plus petits, moins chers et moins rapides. Les **capteurs** et **actionneurs** reposent sur des technologies physiques et électroniques variées, allant quelquefois vers l'électronique de puissance. Un problème essentiel est la réduction de la consommation électrique, surtout pour les appareils sur pile.

Impacts sur les pratiques humaines

L'impact de l'informatisation des objets devient considérable, surtout depuis que leurs interfaces s'unifient. Le but est de fabriquer des machines d'utilisation facile permettant des fonctionnalités améliorées, voire complètement nouvelles comme la voiture autonome. Celle-ci utilise à la fois des techniques de systèmes embarqués pour son fonctionnement et sa navigation et de l'intelligence artificielle pour l'analyse en temps-réel de l'environnement à l'aide de capteurs variés (caméras, radars, lidars, etc.).



Comme l'informatique embarquée interagit avec le monde physique en exposant quelquefois des vies humaines ou des équipements critiques (réseaux électriques par exemple), elle est soumise à de fortes contraintes de sûreté (absence d'erreurs) et de sécurité (résistance aux attaques). En avionique, ferroviaire ou autres applications critiques, des processus lourds de certification externe sont utilisés. Cependant dans beaucoup de systèmes embarqués moins critiques, la sécurité reste souvent un point faible, et les objets connectés sont de plus en plus utilisés comme robots pour lancer des attaques sur internet.

Contenus	Capacités attendues
Systèmes informatiques embarqués	Identifier des algorithmes de contrôle des comportements physiques à travers les données des capteurs, l'IHM et les actions des actionneurs dans des systèmes courants.
Interface homme-machine (IHM)	Réaliser une IHM simple d'un objet connecté.
Commande d'un actionneur, acquisition des données d'un capteur	Écrire des programmes simples d'acquisition de données ou de commande d'un actionneur.
Exemples d'activités	Identifier les évolutions apportées par les algorithmes au contrôle des freins et du moteur d'une automobile, ou à celui de l'assistance au pédalage d'un vélo électrique. Réaliser une IHM pouvant piloter deux ou trois actionneurs et acquérir les données d'un ou deux capteurs. Gérer des entrées/sorties à travers les ports utilisés par le système. Utiliser un tableau de correspondance entre caractères envoyés ou reçus et commandes physiques (exemple : le moteur A est piloté à 50 % de sa vitesse maximale lorsque le robot reçoit la chaîne de caractères « A50 »).

■ La photographie numérique

Introduction

Les technologies de la photographie argentique ont eu une évolution très lente, liée aux progrès en optique, mécanique et chimie. Ce n'est plus du tout le cas de l'évolution actuelle, davantage due aux algorithmes qu'à la physique : algorithmes de développement et d'amélioration de l'image brute, algorithmes d'aide à la prise de vue. Cet exemple est caractéristique des façons de procéder de la révolution informatique par rapport aux approches traditionnelles.

La photographie numérique présente un coût marginal très faible et une diffusion par internet facile et immédiate : chaque jour, des milliards de photos sont prises et partagées.

Repères historiques

- 1826 : naissance de la photographie argentique ;
- 1900 : photographie en couleurs. Après la dernière guerre mondiale, généralisation du format 24 x 36 et de la visée reflex ;
- 1969 : arrivée des premiers capteurs CCD (*Charge Coupled Device*) ;
- 1975 : apparition des premiers appareils numériques ;
- 2007 : arrivée du *smartphone*.



Les données et l'information

En entrée, le capteur est formé de **photosites** en matrice de petits carrés de quatre photosites, deux verts, un bleu et un rouge, correspondant à la répartition des cônes de la rétine. La **résolution du capteur** se mesure en millions de photosites.

En sortie, l'image est formée de **pixels** colorés homogènes, représentés par trois nombres RVB (rouge, vert, bleu). La **résolution de l'image** se compte en mégapixels ; elle n'est pas forcément égale à celle du capteur. La **profondeur de couleur** est en général de 8 bits par pixel et par couleur pour l'image finale.

Des **métadonnées** sont stockées dans les fichiers images sous format **EXIF** (*Exchangeable Image File Format*) : modèle de l'appareil, objectif, vitesse, diaphragme, distance de mise au point, auteur, copyright, localisation, etc.

Les couleurs peuvent être représentées dans différents systèmes : RVB, TSL (teinte, saturation, lumière), avec des formules empiriques de passage d'un modèle à l'autre. On distingue différents formats des fichiers images, compressés ou non, avec ou sans perte : RAW, BMP, TIFF, JPEG.

Les algorithmes et les programmes

Des algorithmes permettent de traiter toutes les lumières, d'effectuer une retouche facile, avec une qualité maintenant bien supérieure à l'argentique. Avec l'arrivée du téléphone mobile, des algorithmes de fusion d'images permettent de concilier une excellente qualité avec un capteur et un objectif minuscules.

De nombreux algorithmes sophistiqués sont utilisés dans les appareils de photographie numérique :

- Lors de la prise de vue : calcul de l'exposition, mise au point, stabilisation par le capteur et/ou l'objectif, le tout en automatique ou manuel assisté, *focus-peaking* (scintillement des contours nets), prise en rafales rapides d'images multiples avant et après appui sur le déclencheur.
- Lors du développement de l'image issue du capteur en une image pixellisée : gestion de la lumière et du contraste, balance des blancs, netteté, débouchage des ombres, correction automatique des distorsions ou des aberrations optiques.
- Après le développement : compression du fichier (TIFF sans perte, JPEG avec perte).
- En utilisant la fusion d'images : réduction du bruit et amélioration de la netteté, panoramas, HDR (*High Dynamic Range*), super-résolution par micro-décalages du capteur, *focus stacking* pour étendre la netteté avec plusieurs mises au point successives, réduction du bruit et amélioration de la netteté.

Certains appareils peuvent augmenter leurs fonctionnalités par téléchargement de nouveaux logiciels.

Les machines

Comme les algorithmes de prise de vue et de développement demandent beaucoup de calcul, les appareils embarquent plusieurs processeurs, généraux ou spécialisés.

Les algorithmes prennent le relais des capteurs physiques en calculant les pixels de l'image finale : ils compensent par exemple les distorsions des lentilles. Des algorithmes permettent également de commander la mise au point et l'exposition automatique, ainsi que de compenser le bougé de l'utilisateur (stabilisation).



Impacts sur les pratiques humaines

La gratuité et l'immédiateté de la réplication des images introduisent de nouveaux usages de la photographie : à la photographie archive (histoire de famille) s'ajoutent la photographie à partager et la photographie utilitaire, prothèse de la mémoire (photo d'un ticket de caisse, d'une présentation lors d'une réunion de travail, d'une place de parking, etc.). Les images s'intègrent à tous les dispositifs de communication et de partage, téléphones, Web et réseaux sociaux.

De nouveaux problèmes apparaissent, liés à la diffusion de photos qui ne disparaîtront jamais (notion de droit à l'oubli), au truchage difficile à détecter des images, au pistage des individus ou à l'obsolescence des supports. Est ainsi posée la question de l'archivage de photographies historiques, scientifiques ou culturelles.

Contenus	Capacités attendues
Photosites, pixels, résolution (du capteur, de l'image), profondeur de couleur	Distinguer les photosites du capteur et les pixels de l'image en comparant les résolutions du capteur et de l'image selon les réglages de l'appareil.
Métadonnées EXIF	Retrouver les métadonnées d'une photographie.
Traitement d'image	Traiter par programme une image pour la transformer en agissant sur les trois composantes de ses pixels.
Rôle des algorithmes dans les appareils photo numériques	Expliciter des algorithmes associés à la prise de vue. Identifier les étapes de la construction de l'image finale.
Exemples d'activités	
Programmer un algorithme de passage d'une image couleur à une image en niveaux de gris : par moyenne des pixels RVB ou par changement de modèle de représentation (du RVB au TSL, mise de la saturation à zéro, retour au RVB).	
Programmer un algorithme de passage au négatif d'une image.	
Programmer un algorithme d'extraction de contours par comparaison entre pixels voisins et utilisation d'un seuil.	
Utiliser un logiciel de retouche afin de modifier les courbes de luminosité, de contraste, de couleur d'une photographie.	