



Introduction à la sécurité
dans le domaine des Big Data.

**Protection des données personnelles
et
Introduction à la cryptographie**

Thierry Baritaud

14/02/2018

Plan de la présentation

- Sécurité des systèmes d'informations et des données
- Big Data et sécurité : nouveaux challenges, nouvelles menaces
- Authentification, protection des données personnelles et privacy
- Contexte normatif de la sécurité
- Introduction à la cryptographie pour la protection des données
 - Cryptographie à clé secrète, à clé publique
 - Chiffrement homomorphe dans un contexte Big Data



Evolutions du secteur Télécoms : des ruptures technologiques historiques à fort impact de sécurité

1

• Broadband everywhere



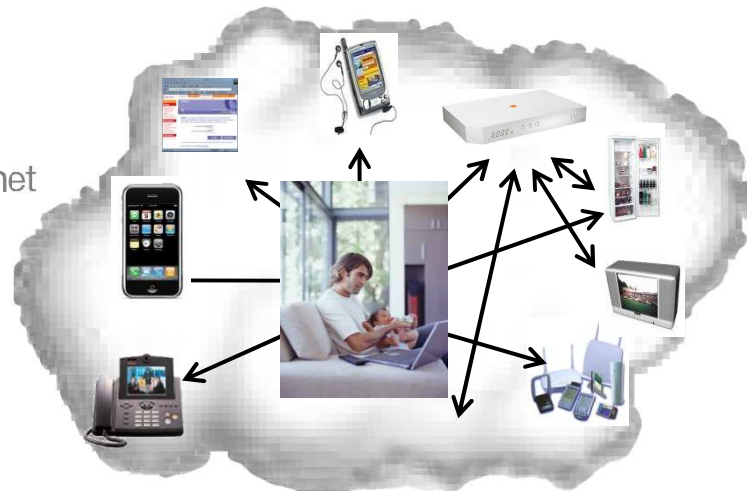
- Images
- X-DSL
- GigaEthernet

2

• Mobility everywhere



- Wifi, Bluetooth, NFC
- Wimax, 3G, 4G, 5G
- Electromagnetic



Security

3

• IT platforms on open networks



- E-Commerce
- Instant Messaging
- Web and Intermediation services
- Cloud computing, Big data
- Customer intimacy, Trust and Privacy

4

• Innovative multi-access terminals



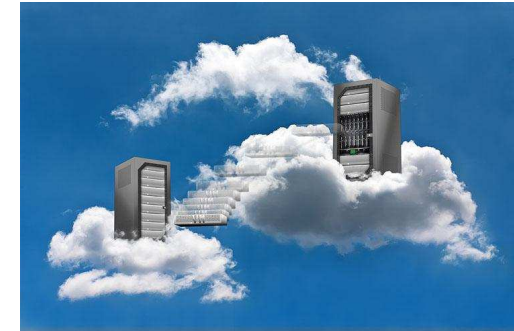
- Terminals
- Home Gateways
- Voice services
- M2M, internet of things



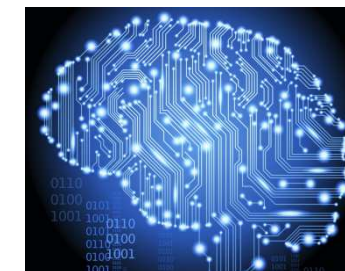
Quelle sécurité pour les nouvelles technologies ?



Cloud

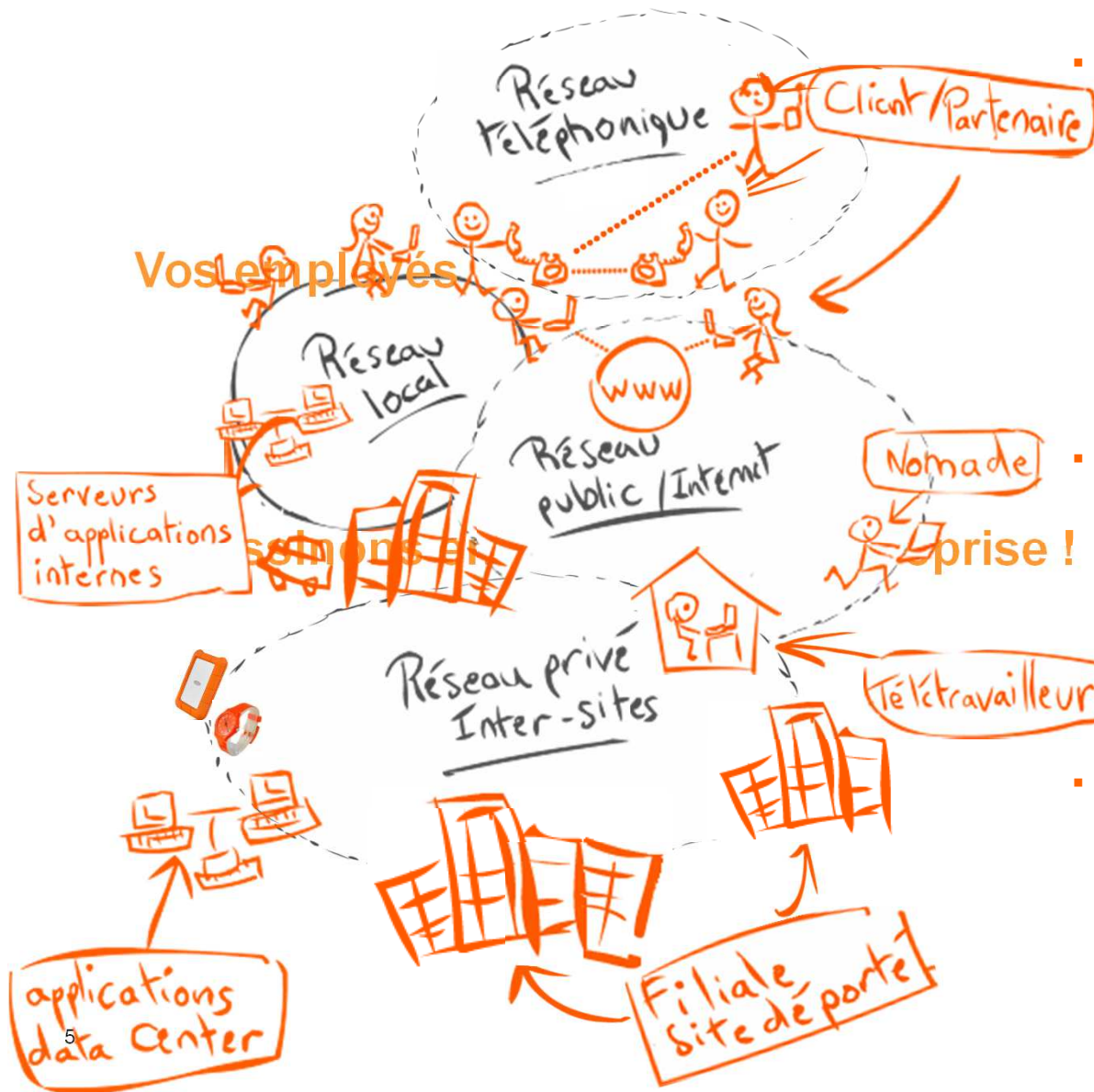


Systèmes industriels SCADA



Intelligence artificielle

L'entreprise se transforme et les risques augmentent: *maintenir le niveau de sécurité d'origine devient difficile*



▪ L'ouverture des entreprises engendre:

- De nouveaux accès géographiques
- De nouveaux réseaux d'accès
- Des accès à de nouvelles applications
- Une multiplicité des moyens et technologies d'accès

▪ Cette multiplicité des accès offre des opportunités pour :

- Les utilisateurs légitimes
- Les personnes malveillantes souhaitant s'introduire sur les SI / réseaux

▪ Objectifs:

- Simplifier les procédures d'accès
- Gérer les risques en maintenant un niveau de sécurité et de confiance fort

Vers une cybercriminalité facilitée...?



- Convergence vers le Tout IP : **fin de la confiance ?**
 - Fin des réseaux étanches, dépendances de technos dominées par d'autres pays/acteurs
- Nouveaux usages : **multiplication des accès légitimes ou malveillants**
 - Nomadisme, accès aux mêmes ressources via plusieurs médias
 - Accès permanent : ADSL, sans contact, bluetooth, Wifi, Wimax, UMTS, 4G...
- Echanges dématérialisés : **argent, biens de valeurs sur le réseau**
 - Services à valeur ajoutée: E-commerce/banking, E-monnaie, musique, videos, images
 - Stockage réseau/distant de données sensibles ou stratégiques, cloud, big data
- Connaissances des techniques d'attaques : **fraudeurs en position de force**
 - Explosion du nb de « hackers » maîtrisant les attaques sophistiquées
 - Multiplication des outils d'attaques sur le web, et sources d'informations sur les failles



Et une sécurisation difficile...

- Utilisateurs « humains » non sensibilisés donc vulnérables...
 - Démunis face à la complexité technique, phishing, virus, vol de données.
- Sentiment d'anonymat et d'impunité pour les attaquants
 - Législations difficilement applicables, entr'aide internationale variable
 - Volumétrie des infos à traiter, nombres d'infractions à sanctionner
- Une double hétérogénéité rend la sécurisation technique complexe
 - Hétérogénéité des réseaux/services : niveaux de sécurité internes variables
 - Hétérogénéité des solutions complémentaires de sécurisation du marché



Login : toto
Password : X&t*\$K7u



- En entreprise: la sécurité souvent vue comme un coût ou frein à l'innovation
 - Solutions souvent locales, dépendant des métiers, contextes, réglementations
 - Rarement une approche globale de gestion des risques de sécurité
 - Comment concilier sécurité, ergonomie, et développement des usages ?

Sécurité des SI et des réseaux : nécessité d'une approche globale par la gestion des risques



Système d'information :

Sécuriser les informations et les supports d'information

- les données et applications

- annuaires d'identité, serveurs de fichiers, bases de données
- en local, lors de leur transfert, hébergées à distance hors de l'entreprise
- applications logicielles, web et intranet

- les infrastructures

- les réseaux internes et les passerelles d'interconnexions
- les postes de travail et supports amovibles
- les serveurs applicatifs, et les systèmes industriels
- les équipements bureautiques (ex: imprimantes, mopieurs...)
- les interfaces et réseaux d'accès sans fils : Wifi, sans contact, 2G, 3G, 4G...

- les outils de communications

- téléphonie, messagerie...
- accès distants, solutions VPN, solutions de mobilité
- outils collaboratif audio, vidéo, partage de documents,

- ...

Quelles sont les zones à protéger



Actifs & patrimoine



Les moyens de production matériel et immatériels de l'entreprise et/ou ses produits

Acteurs

Employés

- Isolé
- Petit Site
- Grands Site
- Site spécialisé



Fournisseurs Partenaires Joint-Venture



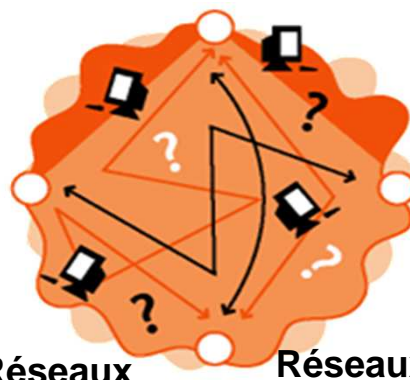
Clients

- BtoB
- BtoC
- Prospect



Infrastructure & Services

Réseaux Publiques & Internet



Réseaux
Locaux

Réseaux
inter-sites

Réseaux
téléphonique

Applications

Applications Communication

- Messagerie,
- Collaboratif,
- Téléphonie,
- Vidéo



Applications Intermédiation

- EDI,
- Portails,
- Transactionnel
- Transfert de données



Applications Métier

- ERP
- Ap. Spécialisées
- Reporting
- Bases de données



Processus d'entreprise

De nombreux défauts sont à l'origine des failles

- Défauts dans la conception des protocoles et des applications
 - Aspects sécurité non traités ou partiellement
- Défauts dans l'implémentation logicielle de protocoles et applications
 - Aspects sécurité non traités ou partiellement
 - Erreurs de programmation:
 - Estimation entre 5 à 15 erreurs sur 1000 lignes de code.
 - World of WarCraft : près de 5 M de lignes de codes ; Facebook : 60 M
 - Windows : 50 M ; Mac OS X Tiger : 80 M
- Défaut dans la configuration des systèmes et des réseaux
 - L'autorisation de certaines actions peut permettre des actions illicites
 - La combinaison

Prévention...

- Définir une architecture de sécurité
 - Référencement des besoins, analyse du risque
 - Implanter des équipements de sécurité : Firewalls, IDS.
 - Sécuriser les équipements (plate-formes, serveurs, postes de travail) et les applications.
 - Évaluation des nouveaux produits (logiciels et matériels) introduits dans le système
 - N'autoriser que ce qui est strictement nécessaire.

- Faire vivre la sécurité : politique de sécurité
 - Administrer les équipements de manière régulière
 - Ce ne sont pas des boîtes noires délaissées une fois qu'elles fonctionnent.
 - Avoir du personnel compétent et ayant du temps dédié à la sécurité !
 - Appliquer les mises à jour de sécurité sur les systèmes et les applications
 - Veille sur les attaques, failles de produits, vulnérabilités...

- Audit de sécurité
 - Évaluer régulièrement le niveau de sécurité d'un système et son adéquation avec les risques
 - Pour y déceler les failles et ajuster les configurations

Méthode de protection classique (2/2)

Détection...

- Se rendre compte que le bien à été endommagé
- Utilisation de systèmes de détection d'intrusion (IDS)
- Analyse des logs

Réaction...

- Avoir une politique de réaction sur incident (piratage, sauvegarde...) et l'appliquer.
 - Journaliser les attaques réalisées en cas de poursuites.
- Prendre en compte les avis des CERT : Computer Emergency Response Team
 - Cert-IST, Cert-Renater, Cert-A
 - centralisation des demandes d'assistance suite aux incidents de sécurité
 - traitement des alertes et réaction aux attaques informatiques
 - établissement et maintenance d'une base de donnée des vulnérabilités
 - prévention par diffusion d'informations sur les précautions à prendre
- Réponse aux attaques
 - Changer l'architecture de sécurité et mettre à jour la politique de sécurité
 - Réparer les dommages

Conseils concrets pour sécuriser un SI (1/3)

- **Adopter une politique de mot de passe rigoureuse**
 - Identifiant / mot de passe: 1ère des protections d'accès à un ordinateur ou fichier.
 - Mot de passe individuel, secret et difficile à deviner: minimum 8 caractères avec chiffres, lettres, caractères spéciaux. Il doit être renouvelé fréquemment (ex: 3 mois),
 - Le nouveau mot de passe doit être différent des 3 derniers utilisés.
 - Le mot de passe attribué par l'administrateur doit être modifié dès la 1ère connexion.
 - Les administrateurs doivent modifier les mots de passe qu'ils utilisent eux-mêmes...

- **Concevoir une procédure de création et de suppression des comptes utilisateurs**
 - Les comptes utilisateurs (ou administrateurs) doivent être nominatifs (i.e. non-génériques)
 - pour pouvoir tracer les actions faites et responsabiliser les utilisateurs.

- **Sécuriser les postes de travail**
 - Verrouillage automatique des postes en cas d'inactivité (après 10 minutes maximum) ;
 - Les utilisateurs doivent également être incités à verrouiller systématiquement leur poste dès qu'ils s'absentent de leur bureau.

Conseils concrets pour sécuriser un SI (2/3)

- Identifier précisément qui peut avoir accès aux données
 - Accès aux données perso: limité aux personnes légitimes et selon le profil d'habilitation.
 - Mise à jour des droits d'accès en cas de nouvelle affectation, et validation hiérarchique.
 - Vérification périodique des profils des applications et des droits d'accès
- Veiller à la confidentialité des données vis-à-vis des prestataires
 - Clause de confidentialité (imposée par la loi) dans les contrats de sous-traitance.
 - Présence d'un salarié en cas d'intervention d'un prestataire sur des bases de données
 - Chiffrement des données « sensibles » et données utilisateurs
- Sécuriser le réseau local
 - Sécuriser le SI ou réseau vis-à-vis des attaques extérieures.
 - Routeurs filtrants, pare-feu, sonde anti intrusions, anti virus et anti logiciels espions
 - Sécurisation des connexions entre sites distants, par VPN.
 - Chiffrement des réseaux sans fil, contrôle des adresses physiques des nomades autorisés
 - Authentification des postes nomades: IPsec, SSL/TLS ou encore HTTPS.
- Sécuriser l'accès physique aux locaux sensibles (ex: serveurs informatiques)
 - Vérification des habilitations, gardiennage, portes fermées (clé, digicode, badge nominatifs)
 - La DSI assure la protection des docs techniques, plans d'adressages réseau, contrats...

Conseils concrets pour sécuriser un SI (3/3)

- **Anticiper le risque de perte ou de divulgation des données**
 - Sauvegardes régulières et supports de sauvegarde stockés dans un lieu distinct sécurisé
 - Sécurisation particulière pour les supports nomades (ex: disque chiffré).
 - Destruction physique des matériels informatiques en fin de vie, retrait des données sensibles en cas de maintenance
- **Anticiper et formaliser une politique de sécurité du système d'information**
 - Document accessible à l'ensemble des salariés.
 - Evolutions en cas de modifications des systèmes et outils informatiques
 - Le paramètre « sécurité » doit être pris en compte en amont de tout projet lié au SI.
- **Sensibiliser les utilisateurs aux «risques Sécurité» et à la loi Informatique et libertés**
 - Principal risque en matière de sécurité : l'erreur humaine.
 - Formation, diffusion de notes de service, envoi périodique de fiches pratiques.
 - Formalisé dans une « charte informatique »: règles à respecter en sécurité informatique
 - bon usage des outils de communications
 - conditions de création de fichiers de données perso...
 - engagement de responsabilité à signer par chaque utilisateur.
- **Mettre en place des plans de gestion de crise**
 - avant toute nouvelle ouverture de service ou lancement de produit

La sécurité, c'est l'affaire de tous...

Protection des données

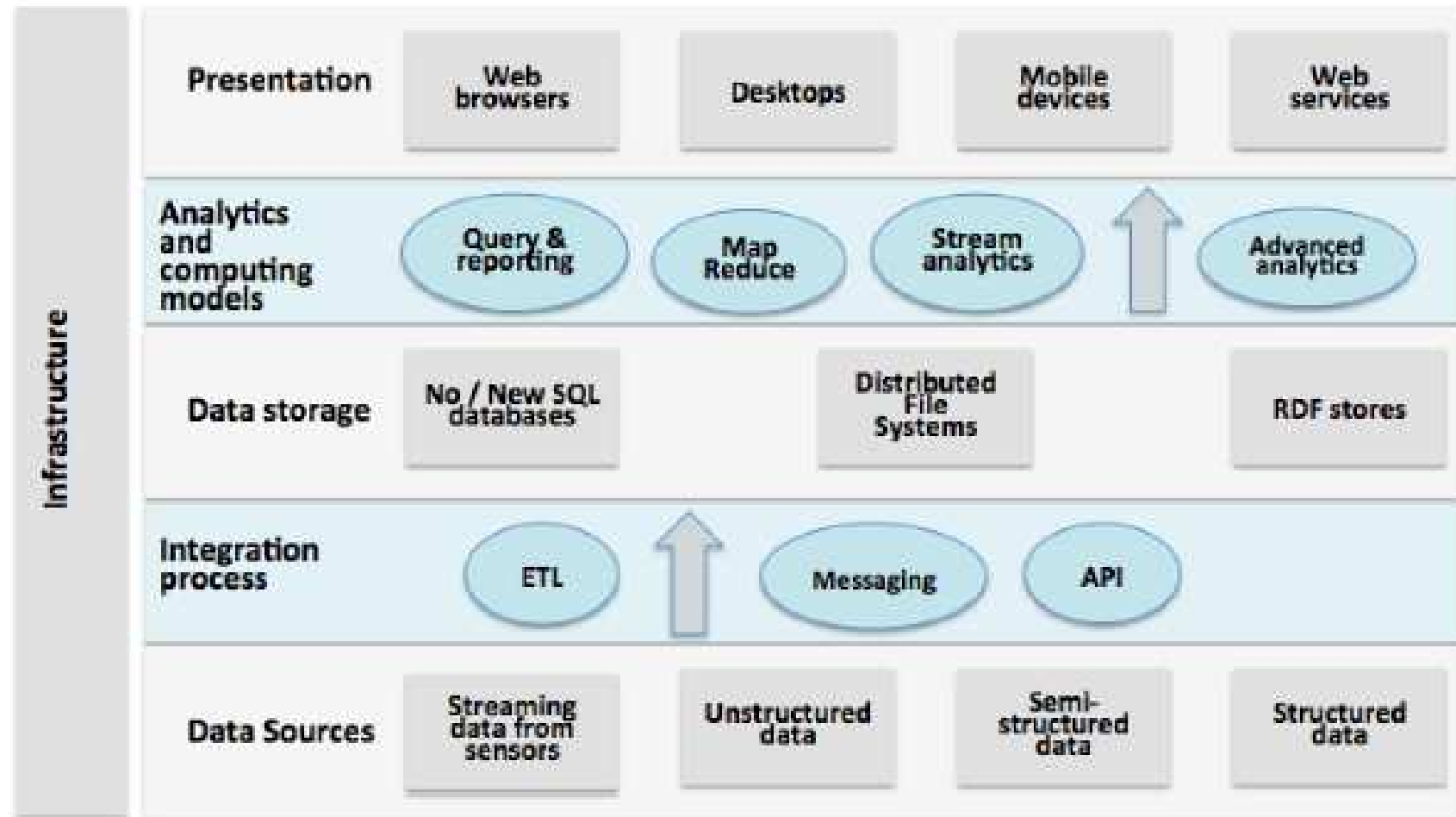


Sécurité des données

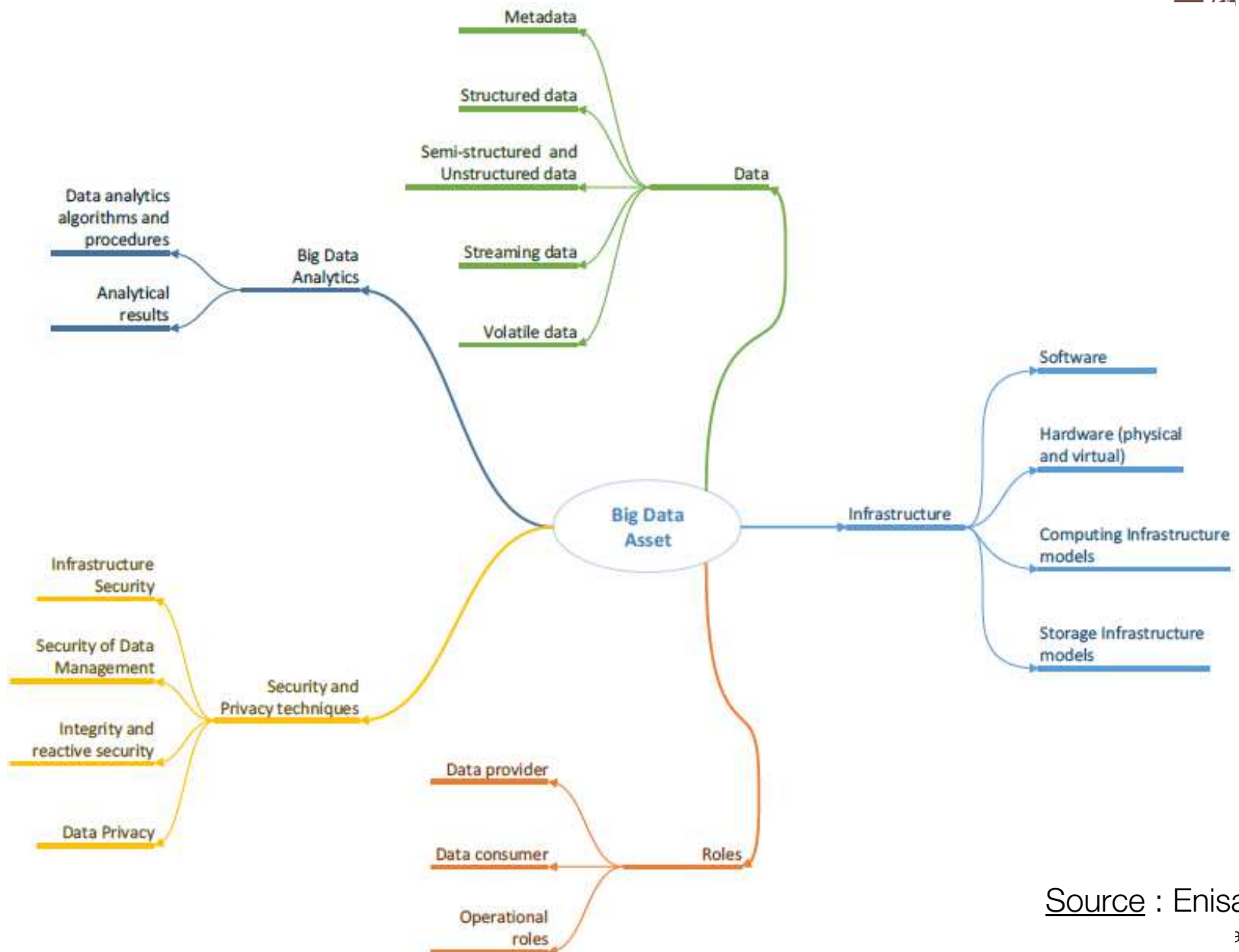
- Techniques classiques de protection des données
 - Confidentialité
 - Intégrité
 - Contrôle d'accès

- Mais le problème est plus complexe pour les données personnelles
 - Respect de la vie privée
 - Anonymat
 - La protection des données à caractère personnel est une **composante sociétale** du développement durable.

Architecture en couches des systèmes Big Data



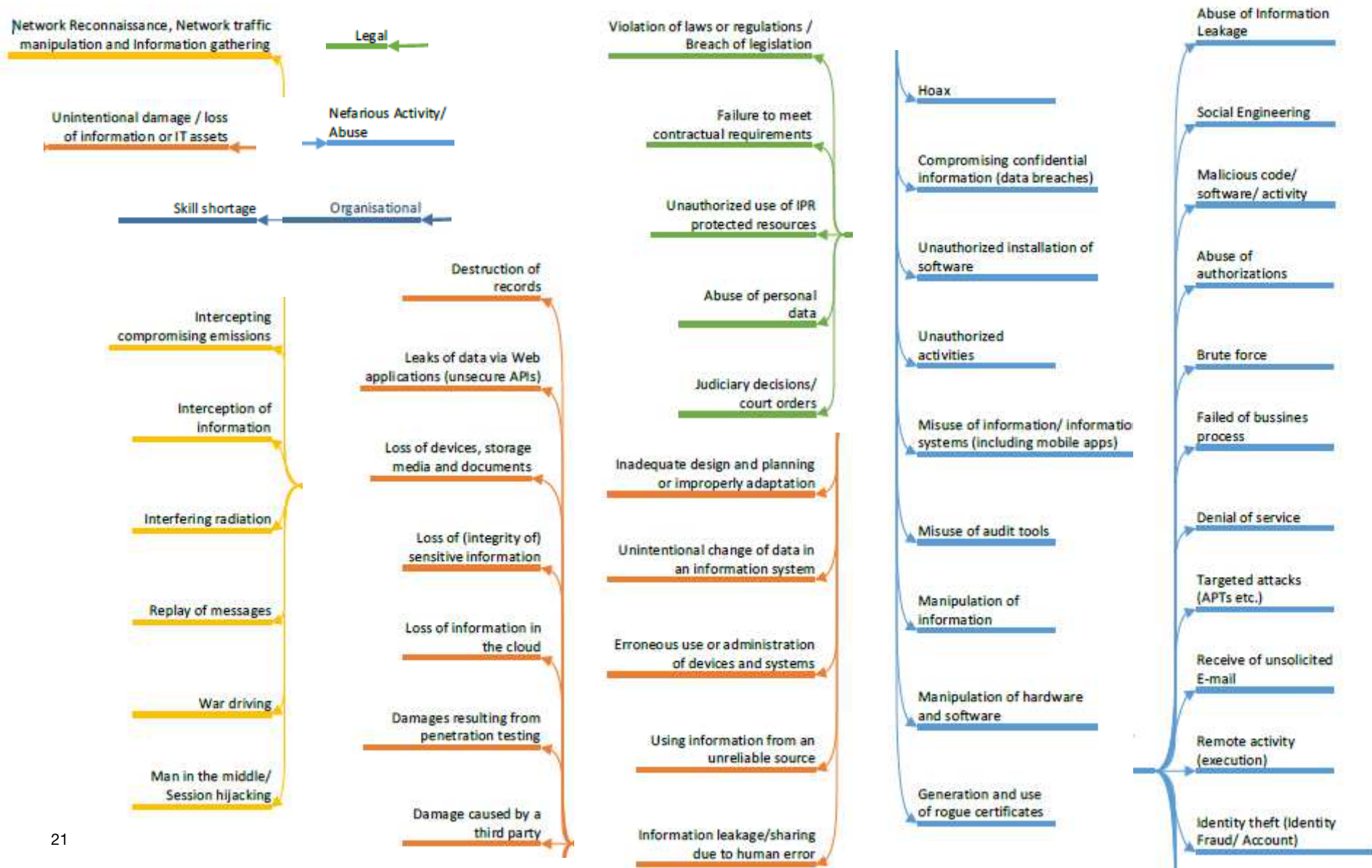
Environnements et actifs Big Data



Source : Enisa

aritaud

Exemples de menaces en environnement Big Data



Impact du Big Data sur la sécurité

Pourquoi le Big Data bouscule la sécurité... ?

- Le *big data* bouscule la sécurité en raison de son ambivalence :
 - Il se révèle à la fois profitable, pour les opérateurs économiques
 - connaître les besoins des consommateurs,
 - aider les acteurs de la santé, prédire des épidémies...
 - aider les gouvernements dans la lutte contre le terrorisme...
 - et potentiellement néfaste, principalement pour les individus.
 - risques de détournement de fichiers et d'interconnexion des données ayant trait à la vie privée des individus.
 - surveillance des personnes...
- Le Big Data bouscule les responsabilités de chaque acteur:
 - besoin de sécurisation du stockage d'une telle masse de données, dans un data center privé, dans le cloud, ou hébergé par un prestataire,...
 - prévoir dans les contrats des clauses de sécurité et des clauses précisant les modalités de la protection des données personnelles...

Une variété de situations critiques

- Les problèmes de sécurité varient en fonction de :
 - l'origine des données (publiques, privées ou mixtes),
 - la loyauté de leur recueil,
 - la présence ou non, directe ou indirecte, de données personnelles
 - l'objectif poursuivi (recherche scientifique, avantage concurrentiel...)
 - la transparence ou l'opacité des buts poursuivis,
 - les infrastructures (publiques, privées ou mixtes) de stockage et de calculs mises en œuvre
 - et le caractère ouvert ou fermé des traitements algorithmiques.

- Les attaques sont donc multiples:
 - attaques informatiques classiques, atteintes aux infrastructures
 - usages détournés des puissances de calculs
 - falsification, clonage ou manipulation des données et de l'information
 - atteinte à la dignité ou vie privée des individus

Sécurité en environnement Big Data :

Quelques questions qui prennent encore plus de sens...

- Quelle confiance porter aux bases de données ?
- Comment protéger les sources, processus et décisions contre le vol et la corruption ?
- Comment est assurée la confidentialité des informations, quels politiques et processus ont été mis en place vis à vis des employés ?
- Comment assurer le stockage sécurisé d'une telle masse de données
- Quels sont vos actes qui peuvent être exploités par nos adversaires ?
- A quels types appartiennent les informations qui sont collectées, et quels sont les défis juridiques et réglementaires ?
- Quelles sont les responsabilités des acteurs lors de l'hébergement de données dans un data center privé, dans le Cloud, ou hébergé par un prestataire externe

Quelques challenges de sécurité pour les nouveaux environnements: Big Data, cloud...

- Insuffisance des protections périmétriques:
 - virtualisation et ubiquité, constitutives des architectures massives, augmentent les surfaces d'attaques et les délocalisent.
- Origine des données
- Contrôle d'accès légitime aux données
- Communication et échanges sécurisés entre acteurs
- Stockage sécurisé étanche entre bases de données
- Chiffrement cryptographique adapté
- Supervision des activités et des connexions
- Sécurité des supports d'information
- Politiques de sécurité globale et gestion des droits
- Privacy
- Transferts de responsabilité de sécurisation des données

Un nouvel équilibre juridique à trouver

- Face aux multiples problèmes posés par le *big data*, le droit doit donc trouver un nouvel équilibre entre :
 - Les intérêts commerciaux de l'entreprise et ceux du consommateur
 - les intérêts de l'Etat et ceux des citoyens ;
 - la protection et la circulation des données personnelles.



Les principaux textes de l'UE

GDPR (General Data Protection Regulation)

Cadre unique de **protection des données personnelles**, applicable aux Etats membres, mais aussi aux ressortissants de l'Union en situation extra-territoriale, et à toutes les organisations traitant des données provenant d'organisations européennes. Il s'applique ainsi aux entreprises non-européennes qui proposent des biens et des services en Europe.



NIS (Network and Information Security directive)



Directive destinée à créer au sein de l'UE un **niveau commun de sécurité dans les réseaux et les systèmes d'information**, et mettant en œuvre :

- Des obligations pour chaque pays membre de créer une autorité nationale de cybersécurité, et de se définir une stratégie nationale de cybersécurité ;
- Une coopération technique entre Etats membres ;
- Le renforcement de la cybersécurité des « opérateurs de services essentiels », avec obligations de cybersécurité et de notification des incidents de sécurité.

eIDAS (Electronic IDentification Authentication and trust Services regulation)

Règlement sur l'**identification électronique et les services de confiance pour les transactions électroniques**. Il institue un cadre juridique uniforme pour les interactions électroniques sécurisées entre les citoyens, les entreprises et les autorités publiques.

Il prévoit des exigences sur la signature électronique, le cachet électronique, l'horodatage électronique, l'envoi en recommandé électronique, l'authentification de sites internet.

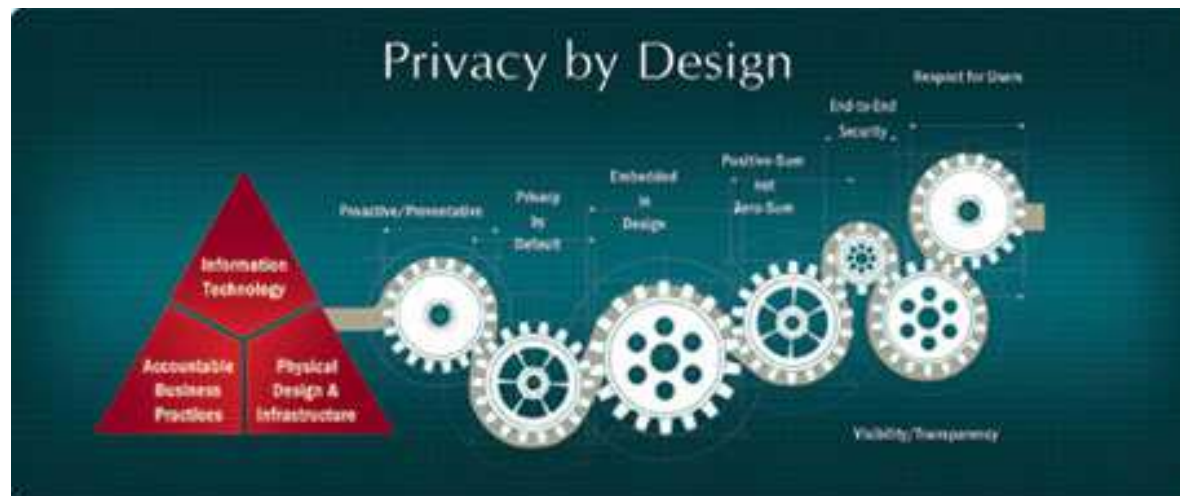


Privacy

- Notion complexe qui recouvre à la fois :
 - les droits des individus à la protection de leur vie privée
 - mais aussi les contrôles étatiques ou organisationnels permettant de les protéger.
- Beaucoup de mesures du domaine sont d'ordre légal ou réglementaire et ne se prêtent donc pas facilement à une approche normative.
- D'autres mesures sont techniques et font le plus souvent appel à des mécanismes d'identification/authentification.
- On commence à voir apparaître des notations (notamment américaines) portant sur la manière dont des organismes ou des prestataires de services gèrent les données privées qu'ils sont amenés à traiter.

Privacy by Design: 7 principes

- **anticiper** : être pro actif et non pas réactif, préventif à l'égard des risques et non curatif
- **se substituer au client ou la vie privée par défaut** : les données personnelles sont automatiquement protégées, même si l'utilisateur ne fait rien
- **intégrer la démarche** : dans la conception et l'architecture des systèmes d'information et des pratiques commerciales
- **positiver vie et sécurité** : des fonctionnalités complètes qui n'opposent pas vie privée et sécurité mais les rendent compatibles
- **exhaustivité** : une protection d'un bout à l'autre du cycle de vie des données, de la création à la destruction en passant par le stockage
- **visibilité et transparence** : assurer que tous les principes sont respectés, y compris par des contrôles indépendants.
- **respecter la vie privée du client**



Anonymisation traditionnelle

- Ne collecter les données qu'au niveau de finesse strictement nécessaire
- Répartir les données, dont le croisement risque de lever l'anonymat, dans des fichiers ou des systèmes informatiques distincts
- Dans la procédure de collecte ou de saisie des données, cloisonner la collecte et la saisie des données en les répartissant auprès de personnels ou organismes différents
- Ne pas fournir systématiquement un logiciel d'interrogation généraliste permettant de croiser n'importe quels critères
- Interdire certains croisements
- Dans les requêtes d'interrogation, ne pas fournir de résultat si le nombre est trop faible

Contexte normatif

Une norme c'est quoi ?

- *Document établi par consensus et approuvé par un organisme reconnu qui fournit, pour des usages communs et répétés, des lignes directrices ou des caractéristiques, pour des activités ou leurs résultats, garantissant un niveau d'ordre optimal, dans un contexte donné. (définition ISO)*

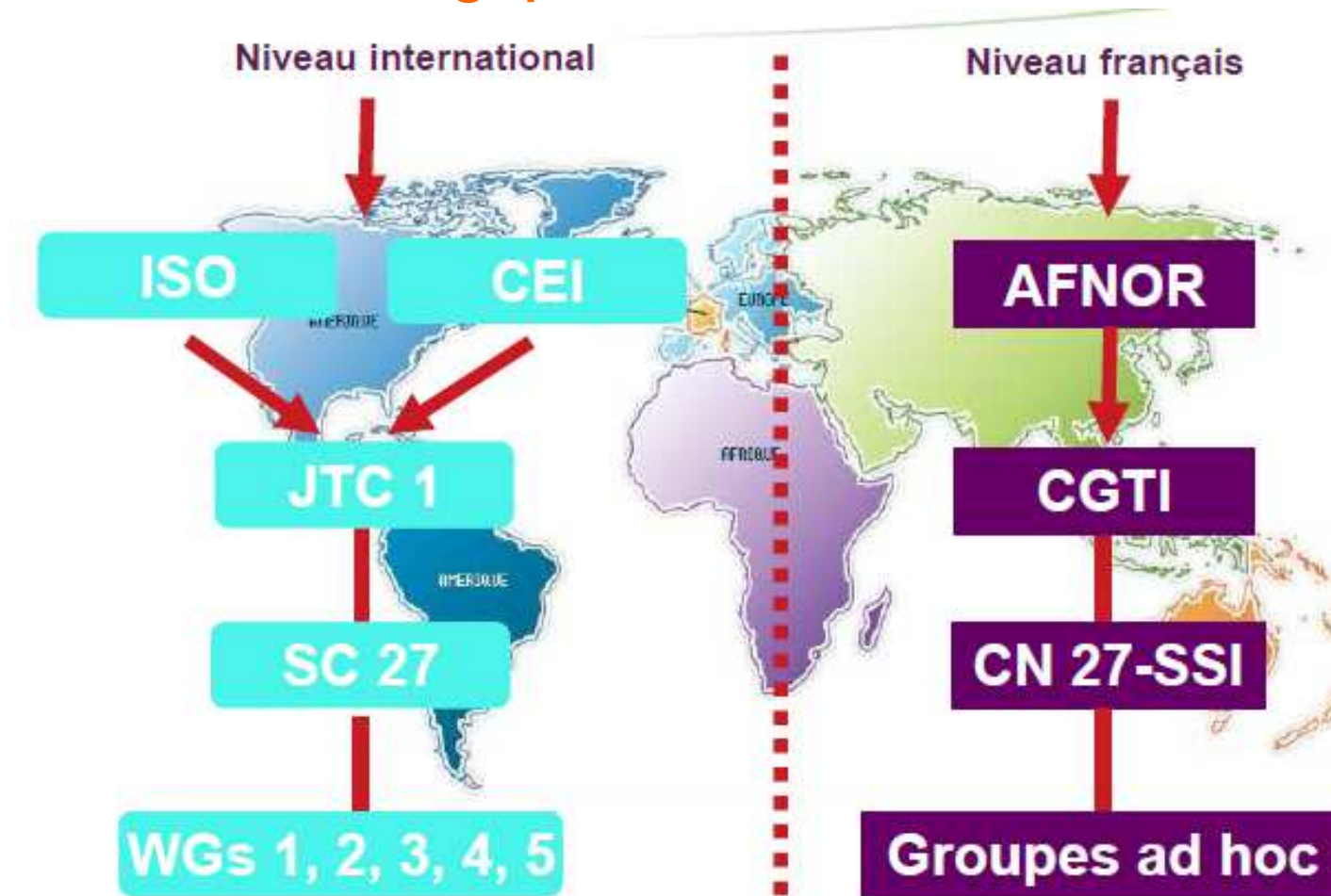
- Objectif :
 - Clarifier et harmoniser des pratiques et des services
 - Définir des niveaux qualité, sécurité, moindre impact environnemental
 - Assurer la compatibilité technique de matériels ou interopérabilité de systèmes
 - Définir des méthodes de caractérisation des produits
 - Faciliter les échanges commerciaux (nationaux, régionaux, internationaux)
 - Fournir des modes de preuves de conformité (réglementation, contrat, .)

- Respecter une norme garantit des niveaux d'exigence et d'interopérabilité : C'est un gage de confiance pour les utilisateurs.

Norme vs réglementation

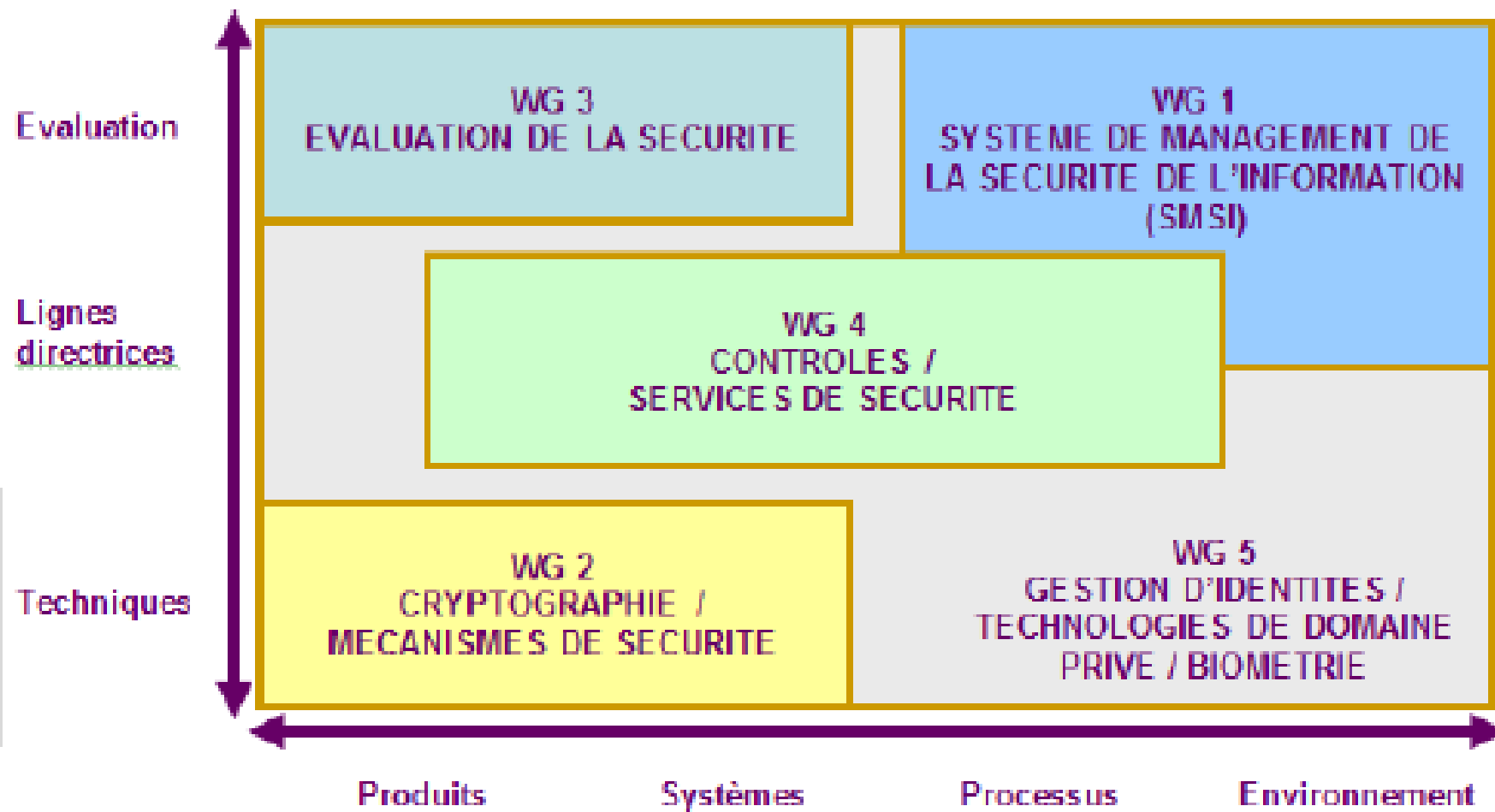
- La Réglementation relève des pouvoirs publics
 - Expression d'une loi, d'un décret ou d'un règlement. Son application est imposée
- Les normes ont un caractère volontaire
 - Reflet de l'engagement d'une organisation à vouloir satisfaire un niveau de qualité et de sécurité reconnu et approuvé
 - La norme est d'application volontaire contrairement à la réglementation
- Certaines normes peuvent être d'application obligatoire
 - Parce que citées dans un texte réglementaire
 - ou comme moyen unique de satisfaire aux exigences de ce texte
- Raisons pour rendre obligatoires certaines normes
 - Ordre public et sécurité publique,
 - Protection de la santé et vie des personnes
 - Exigences liées à la défense du consommateur
 - ...

Organisation de la normalisation de la sécurité des SI : une logique de miroirs...

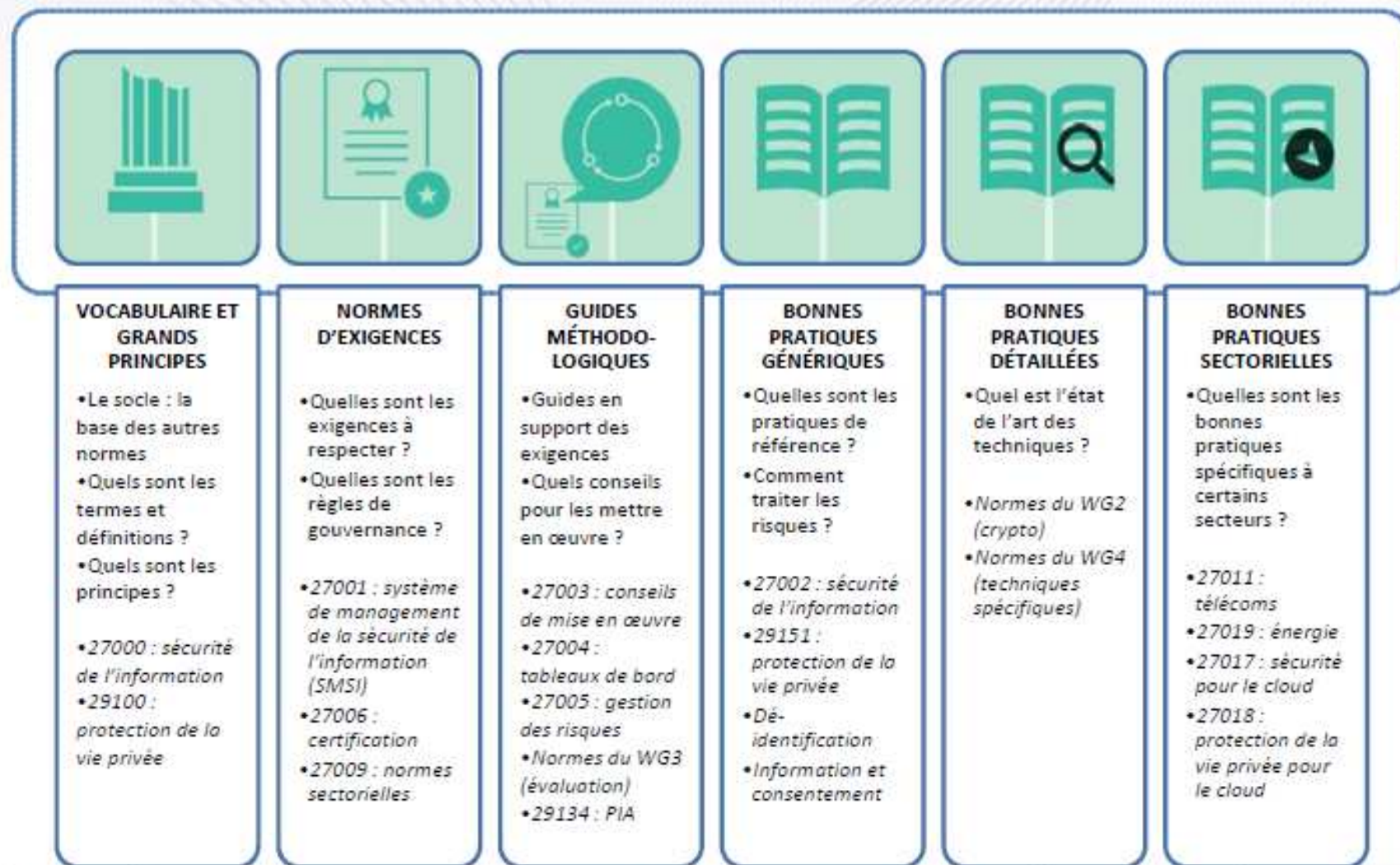


CEI = **Commission électrotechnique internationale** : organisation internationale de normalisation chargée des domaines de l'électricité, de l'électronique, de la compatibilité électromagnétique, de la nanotechnologie et des techniques connexes. Elle est complémentaire de l'Organisation internationale de normalisation (ISO), qui est chargée des autres domaines.

Organisation du comité d'études ISO-JTC1-SC27



Différents types de normes à l'ISO/JTC1/SC27



[images adaptées de Marco Galtarossa, anbileru adaleru, Thomas Helbig et Christopher Holm-Hansen sur <https://thenounproject.com>]

- Trois normes fondamentales
 - IS 27001 ISMS *requirements*,
 - IS 27002 *Code of practice for ISM*,
 - IS 27005 *Information security risk management*.

- Et d'autres normes sur différents thèmes comme :
 - IS 27000 ISMS *fundamentals and vocabulary*,
 - IS 27003 ISM *implementation guidance*,
 - IS 27006 *Requirements for bodies providing audit and certification of ISMS*,
 - IS 27009 *Sector-specific application of ISO/IEC 27001- Requirements*,
 - IS 27010 *ISM for inter-sector and inter-organizational communications*.
 - ...

ISO 27005 : Analyse de risques

- ISO/CEI 27005

Identification : Identification des biens, des menaces et des impacts.
Évaluation des mesures existantes et prévues.

Estimation : Méthodologies d'estimation du risque,
Évaluation des conséquences,
Évaluation de la probabilité des menaces et des
vulnérabilités.

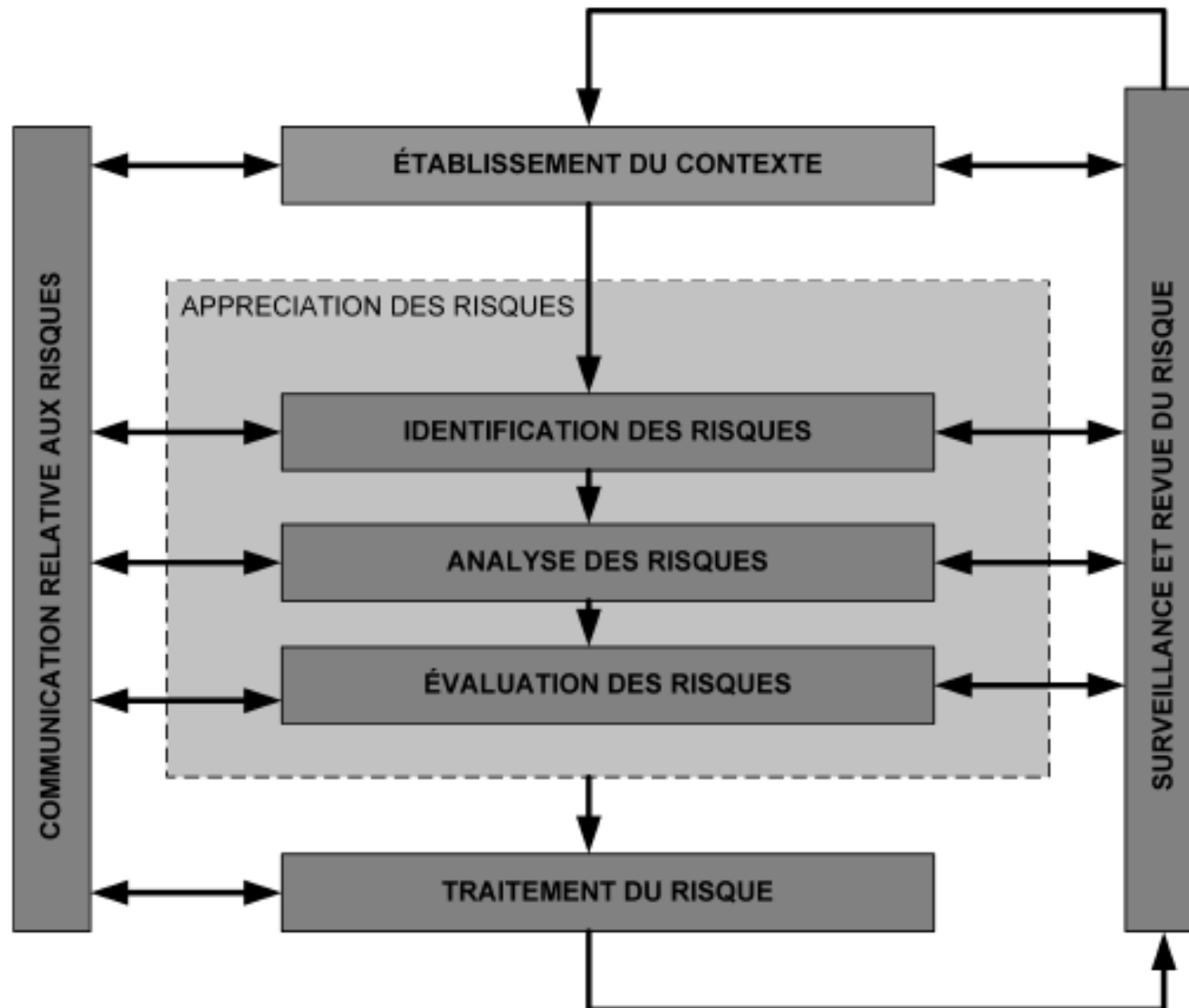
Options de traitement du risque
Évitement,
Réduction,
Transfert,
Acceptation.

Risque résiduel.

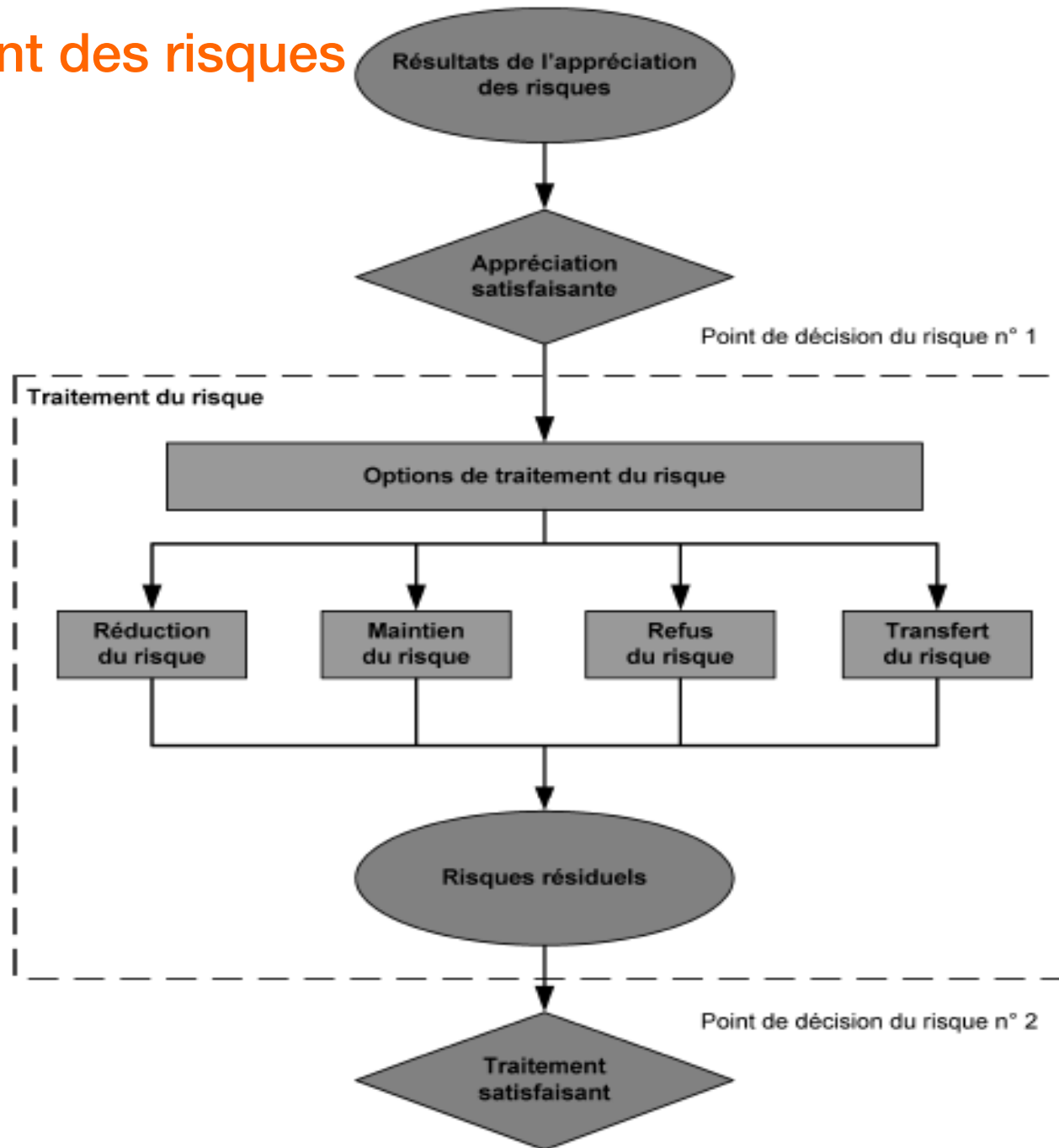
Communication.

Supervision et révision.

ISO 27005 : Processus de gestion des risques



Traitement des risques



PDCA : Plan, Do, Check, Act appliqué au SMSI : Système de Management de la Sécurité de l'Information

Processus SMSI	Processus de gestion des risques en sécurité de l'information
Planifier	Établissement du contexte Appréciation des risques Élaboration du plan de traitement des risques Acceptation des risques
Déployer	Mise en œuvre du plan de traitement des risques
Contrôler	Surveillance et revue continues des risques
Agir	Maintien et amélioration du processus de gestion des risques en sécurité de l'information

Vocabulaire relatif à l'identité

- Identification : « *J'annonce qui je suis* »
 - Login, Nom complet, numéro de compte

- Authentification : « Je prouve qui je suis »
 - Ce que je connais (mot de passe)
 - Ce que j'ai (une carte à puce, un "jeton")
 - Ce que je fais (signature)
 - Ce que je suis (empreinte digitale : données biométriques)
 - Où je suis (station de travail dédiée)

Principe
logique

Principe
physique

Login : toto
Password : X&t*\$K7u



- Autorisation : « Ce que j'ai le droit de faire et de ne pas faire »
 - Accéder au système
 - Entrer dans la salle
 - Lire/Écrire des données
 - Acheter...

Authentification : 3 niveaux

■ Authentification faible à une passe

- envoi simple de mot de passe
- possibilité de génération automatique de mots de passe
- la difficulté à mémoriser est source de failles (compromission par manque de précaution)
- compromission difficile à détecter.

Login : toto
Password : X&t*\$K7u

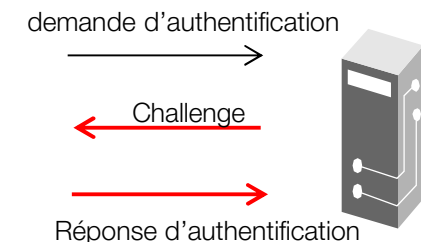
■ Authentification forte à une passe

- mot de passe dynamique qui change à chaque fois (avec le temps)
- nécessite une bonne synchronisation pour le contrôle



■ Authentification forte à deux passes

- une procédure challenge-réponse
- le vérifieur pose une question (défi) au prouveur,
- question différente à chaque fois et non prédictible
- la réponse du prouveur dépend de :
 - la question (le défi) et de l'élément secret qui caractérise le prouveur
- la réponse varie à chaque session, permettant d'éviter le rejeu



Authentification par ce que l'on sait

Classiquement le mot de passe ou la pass-phrase



Caractéristiques

- Dépend beaucoup de l'**implémentation**
- Possibilité de **génération automatique** de mots de passe
 - *alternative aux difficultés de mémorisation*
- Existence de variantes (demandes de réponses préétablies à des questions),
- **Difficulté à mémoriser** : source de failles (compromission par manque de précaution)
- Problèmes d'utilisation sur les réseaux (compromission et cascade, difficile à détecter).

Précautions à prendre

- Durée de **validité** (min. et max.), modalités de changement à mettre en œuvre
- **Choix trivial** (longueur, composition). Utiliser un mécanisme de composition simple, facile à mémoriser et difficile à compromettre,
- **Stockage** (éviter que l'on puisse accéder aux mots de passe ou les reconstituer),
- Traitement des **erreurs**
 - *éviter la reconstitution des mots de passe à partir des enregistrements d'erreurs.*

Authentification par ce que l'on a

- Nombreux dispositifs faciles à transporter
 - badges, carte Secur ID, Activcard, bouchons, cartes à mémoire, cartes à microprocesseur, dongle USB...
- Peut poser des problèmes de généralisation
 - nécessité d'une interface
- Divers protocoles sont envisageables
 - niveaux de sécurité différents
 - risque de n'authentifier que le dispositif détenu
- Compromission limitée et facile à détecter (sauf copie)
- Coût intermédiaire entre mot de passe et dispositif biométrique.

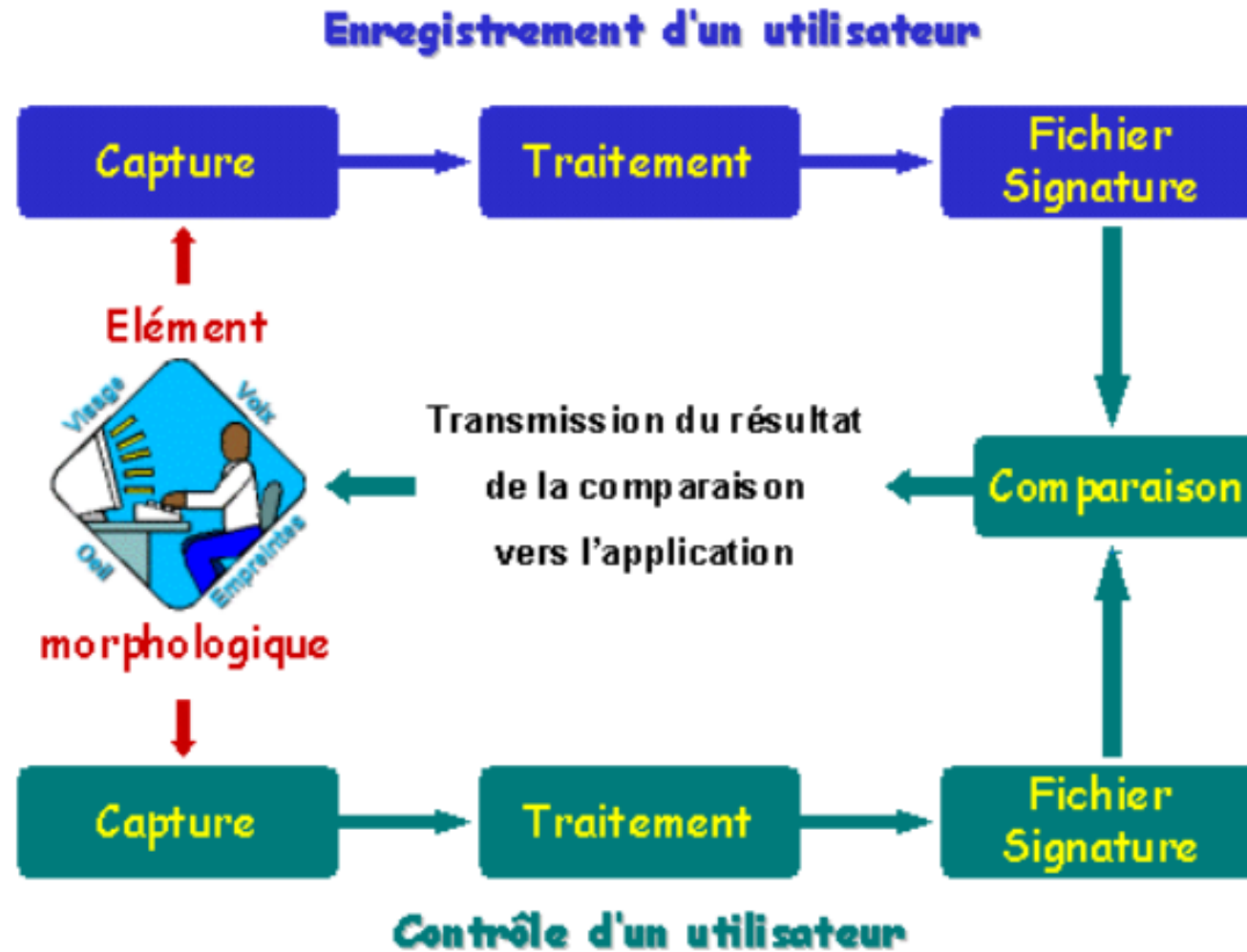
Authentification par ce que l'on est :

- Quasiment impossible à transmettre à un tiers
 - à la différence des cas précédents
- Difficile à imiter ou à contourner en raison de la complexité des caractéristiques biologiques retenues
- Variabilité des caractéristiques biologiques
 - peut conduire à recourir à des facteurs de tolérance limitant l'efficacité du mécanisme
- Problèmes d'acceptation de la part des utilisateurs
 - (empreintes rétiniennes, vie privée)
- Coût élevé pour une sécurité de haut niveau
 - mais la situation évolue fortement depuis quelques années.

Biométrie : 2 catégories

- Techniques d'analyse du comportement :
 - La dynamique de la signature (la vitesse de déplacement du stylo, les accélérations, la pression exercée, l'inclinaison).
 - La façon d'utiliser un clavier d'ordinateur (la pression exercée, la vitesse de frappe).
- Techniques d'analyse de la morphologie humaine : (empreintes digitales, forme de la main, traits du visage, dessin du réseau veineux de l'œil, la voix). Ces éléments ont l'avantage d'être stables dans la vie d'un individu et ne subissent pas autant les effets du stress par exemple, que l'on retrouve dans l'identification comportementale.

Biométrie : principe technique



Biométrie

Techniques	Avantages	Inconvénients
Empreintes digitales	Coût. Ergonomie moyenne. Facilité de mise en place. Taille du capteur.	Qualité optimale des appareils de mesure (fiabilité). Acceptabilité moyenne. Possibilité d'attaque. (rémance de l'empreinte...)
Forme de la main	Très ergonomique. Bonne acceptabilité.	Système encombrant. Coût. Perturbation possible par des blessures et l'authentification des membres d'une même famille.
Visage	Coût. Peu encombrant. Bonne acceptabilité.	Jumeaux. Psychologie, religion. Déguisement... Vulnérable aux attaques.
Rétine	Fiabilité. Pérennité.	Coût. Acceptabilité faible. Installation difficile.
Iris	Fiabilité.	Acceptabilité très faible. Contrainte d'éclairage.
Voix	Facile.	Vulnérable aux attaques.
Signature	Ergonomie.	Dépendance de l'état émotionnel de la personne. Fiabilité.
Frappe au clavier	Ergonomie.	Dépendant de l'état physique de la personne

Authentification mixte

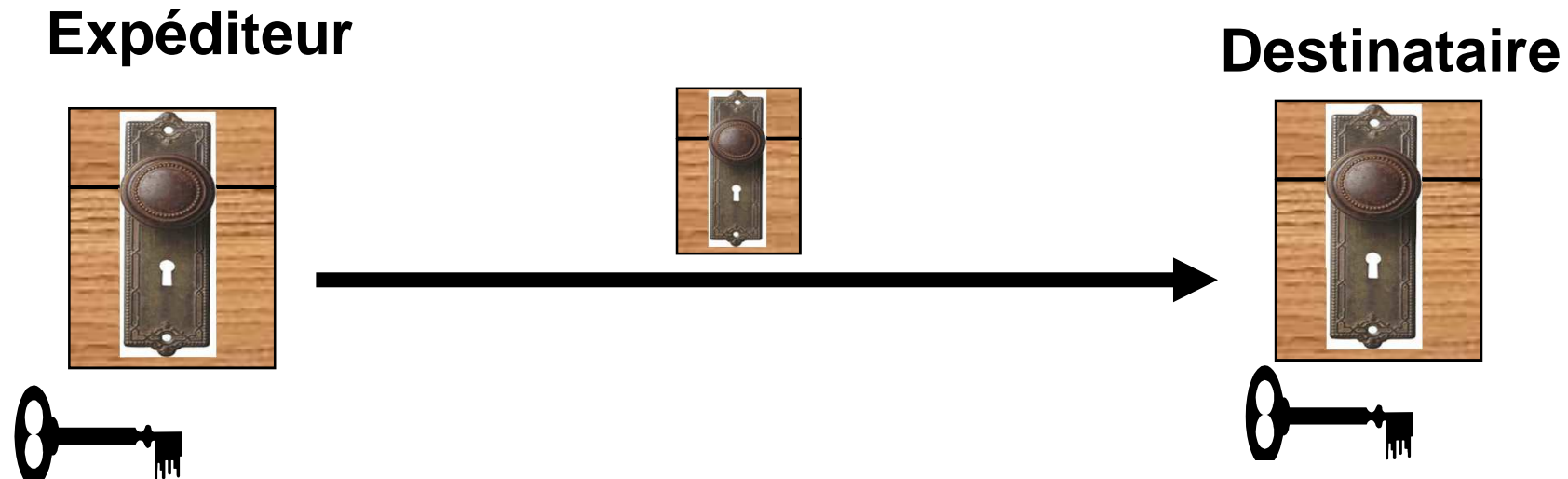
- Possibilité d'utiliser plusieurs mécanismes selon le niveau de sécurité requis
- Conditions d'exploitation ou des contraintes de disponibilité
 - Dongle USB (ou carte à puce) et mot de passe local
 - Lecteur de badge ou carte à puce et empreintes digitales,
 - Utilisation de dispositifs biométriques différents selon la sensibilité des informations ou l'utilisateur (sans compter l'aspect secours),
 - ...
- Attention à l'emplacement des dispositifs
 - éviter le contournement ou la cascade

Un peu de serrurerie...
et de cryptographie...

Où trouve-t-on de la cryptographie ?



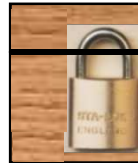
Serrurerie symétrique... (à clé secrète)



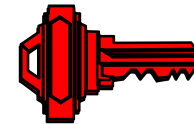
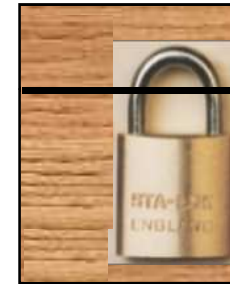
- La même clé est nécessaire pour fermer la boîte et pour l'ouvrir
- Seuls l'expéditeur et le destinataire peuvent ouvrir ou fermer la boîte
- Expéditeur et destinataire doivent maintenir leur clé secrète

Serrurerie asymétrique... (à clé publique)

Expéditeur

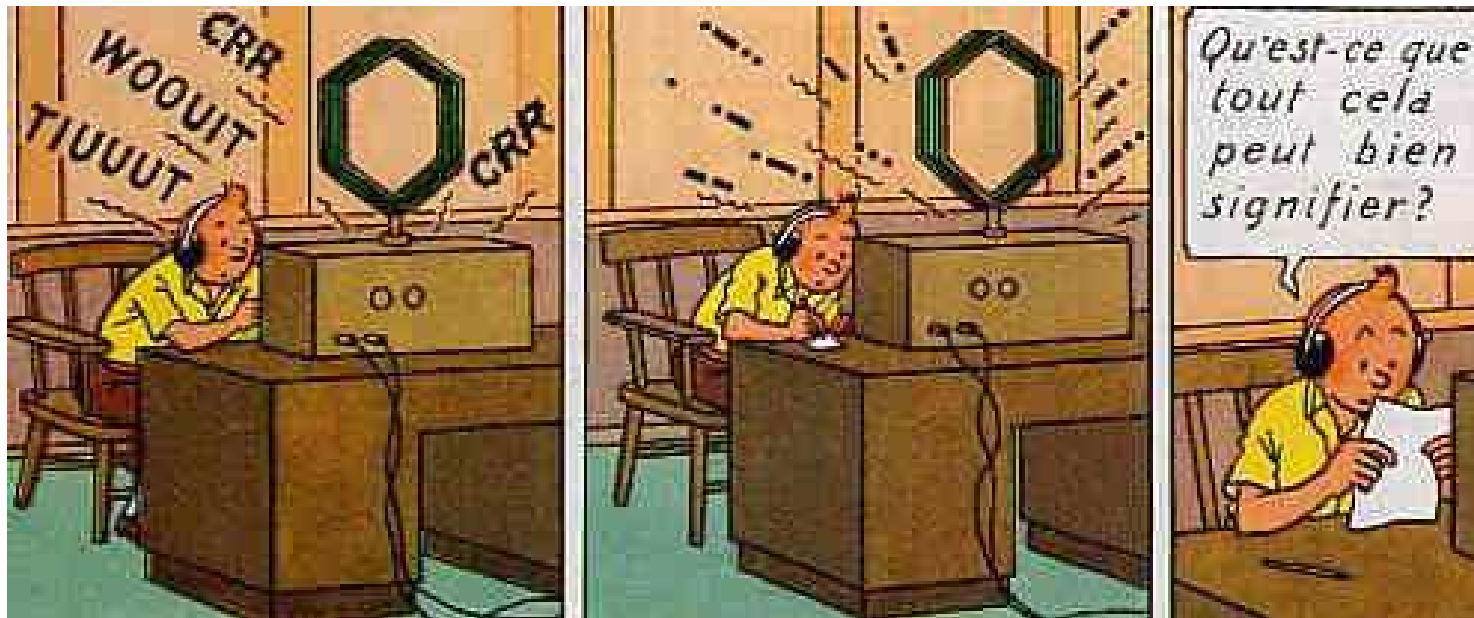


Destinataire



- Une fois que l'expéditeur a récupéré le cadenas (ouvert) appartenant au destinataire, il peut fermer le cadenas et lui envoyer une boîte fermée
- Aucune clé n'est nécessaire pour fermer la boîte.
- Seul le destinataire peut ouvrir la boîte
- Seul le destinataire doit maintenir sa clé privée

Cryptologie et cryptographie



- Cryptologie, étymologiquement la *science du secret* (*KRYPTOS* = *caché*), englobe :
 - la cryptographie — l'écriture secrète —
 - et la cryptanalyse — l'analyse de cette dernière.
 - s'attachant à protéger des messages
 - assurant confidentialité, authenticité et intégrité en s'aidant souvent de *secrets* ou clés.
- La cryptographie se scinde en deux parties nettement différenciées :
 - la cryptographie à clef secrète, encore appelée *symétrique* ou bien *classique* ;
 - la cryptographie à clef publique, dite également *asymétrique* ou *moderne*.

9

Ve SIÈCLE

Pendant la guerre du Péloponnèse, les Spartiates utilisent la **scytole**, un bâton de bois autour duquel s'enroule une bande de cuir sur laquelle est écrit le message. Celui-ci est illisible si le destinataire n'a pas un bâton du même diamètre. C'est le premier dispositif de cryptographie militaire connu.

1^{er} SIÈCLE

Jules César utilise une méthode de chiffrement auquel il légua son nom (le chiffre de César) et qui consiste à décaler les lettres de l'alphabet d'un nombre n de cases (si $n=3$, A devient D, etc.). C'est la plus ancienne méthode de substitution mono-alphabétique.

IX^e SIÈCLE

Le philosophe arabe **Al-Kindi** (801-873), écrit le premier ouvrage de cryptanalyse (ou déchiffrement) connu. Retrouvé en 1987 dans les archives ottomanes d'Istanbul, ce manuscrit présente une technique d'analyse fréquentielle des lettres du texte chiffré.

1586

Le diplomate et cryptographe français **Blaise de Vigenère** (1523-1596) présente une technique de chiffrement par substitution polyalphabétique : une même lettre peut, suivant sa position dans le message, être remplacée par des lettres différentes. Cette technique restera le nec plus ultra de la cryptographie jusqu'au XIX^e siècle.



1942

Le génial mathématicien britannique Alan Turing (1912-1954) casse le code de la version navale de la **machine Enigma** utilisée par l'armée allemande pour crypter ses messages.



1977

Adi Shamir, Ronald Rivest et Leonard Adleman inventent l'algorithme RSA, premier système de cryptographie associant une clé publique et une clé privée. Cette méthode est aujourd'hui au cœur des systèmes de paiement du commerce électronique.



2000

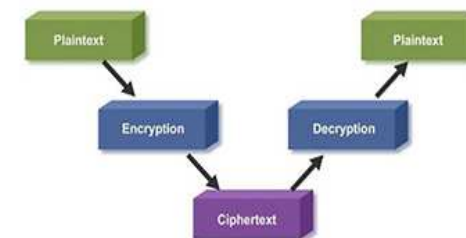
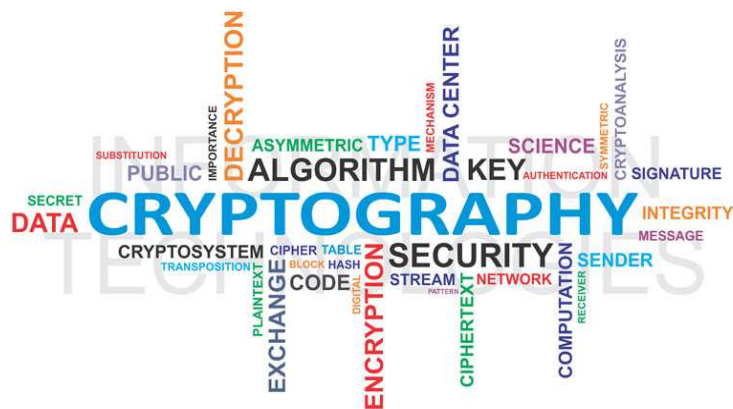
Les Belges Vincent Rijmen et Joan Daemen créent l'algorithme AES, qui est à ce jour le plus solide système de cryptographie symétrique : pour en venir à bout, il faut effectuer 10^{77} opérations. Un nombre proche de celui de toutes les particules élémentaires contenues dans l'univers.



* LES FOMES * / PROTEIL 39

Readable format.
Non-encrypted data.

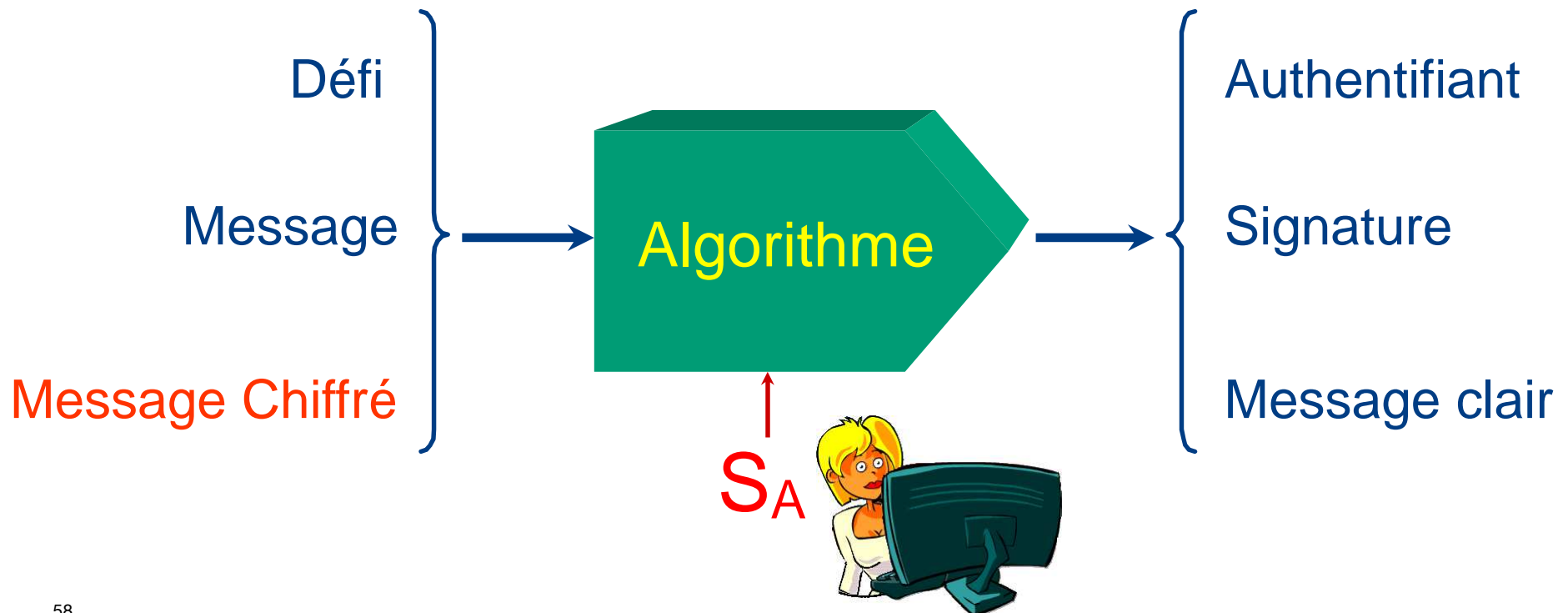
Non-readable format.
Encrypted data.



Cryptology
Study of cryptography
and cryptanalysis

Algorithmes cryptographiques

Un algorithme cryptographique va permettre à Alice de s'authentifier, de signer ou de déchiffrer un message, en utilisant son secret S_A , mais *sans le révéler*.



Cryptographie symétrique (à clé secrète)

- Repose sur des techniques simples (XOR, permutations de bits, additions...)
- Nécessité d'un partage/échange préalable du secret entre chaque utilisateur
- Une clé S par couple ou groupe d'utilisateurs

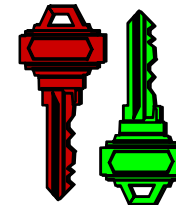


- La clé secrète S doit être maintenue secrète :
 - Chez chaque utilisateur
 - Lors de sa création et transfert/échange
- L'algorithme de déchiffrement est l'inverse de l'algorithme de chiffrement

59

Cryptographie asymétrique (à clé publique)

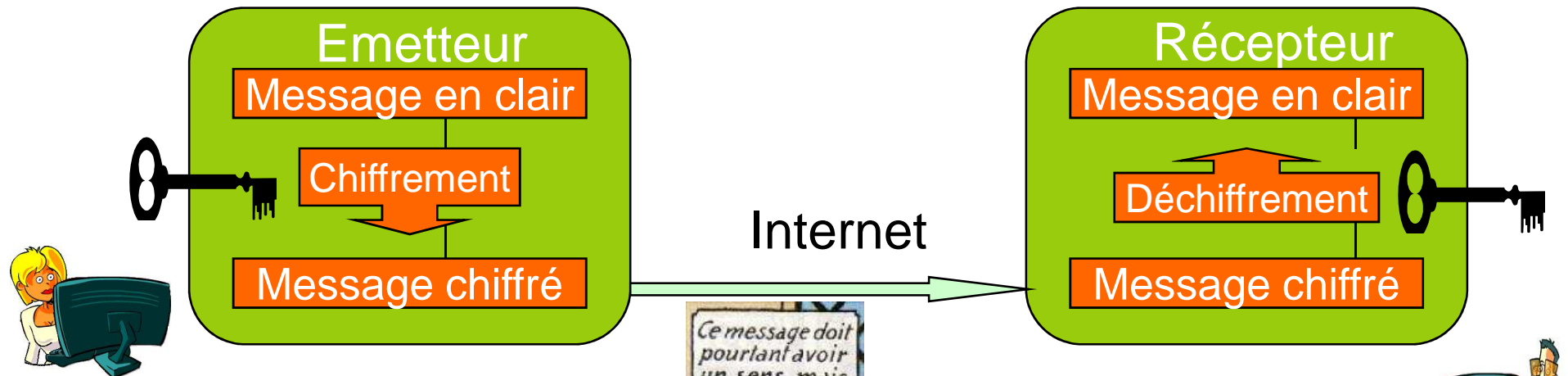
- Repose sur des technos mathématiques complexes
- Pas de partage au préalable d'un secret pour chaque couple d'utilisateurs
- Un couple de clés personnelles (S , P) pour chaque utilisateur



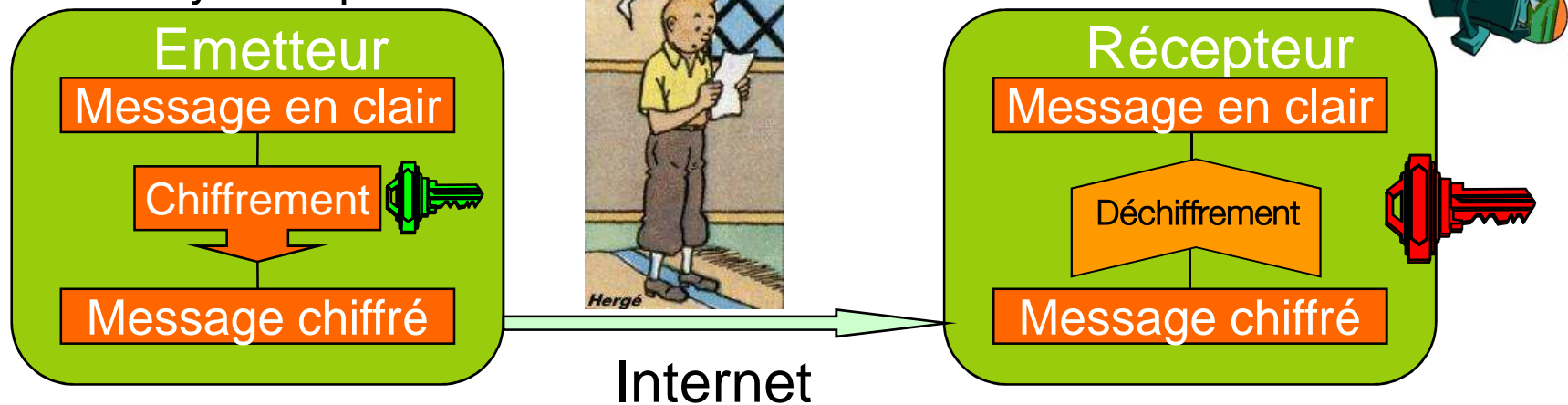
- P est dérivée de S par une fonction à sens unique
 - P est rendue publique
 - P définit la fonction de chiffrement utilisée par toute personne désirant établir une communication sécurisée avec celui qui l'a publiée
- S est gardée secrète/privée par son propriétaire
 - S définit la fonction déchiffrement de messages reçus

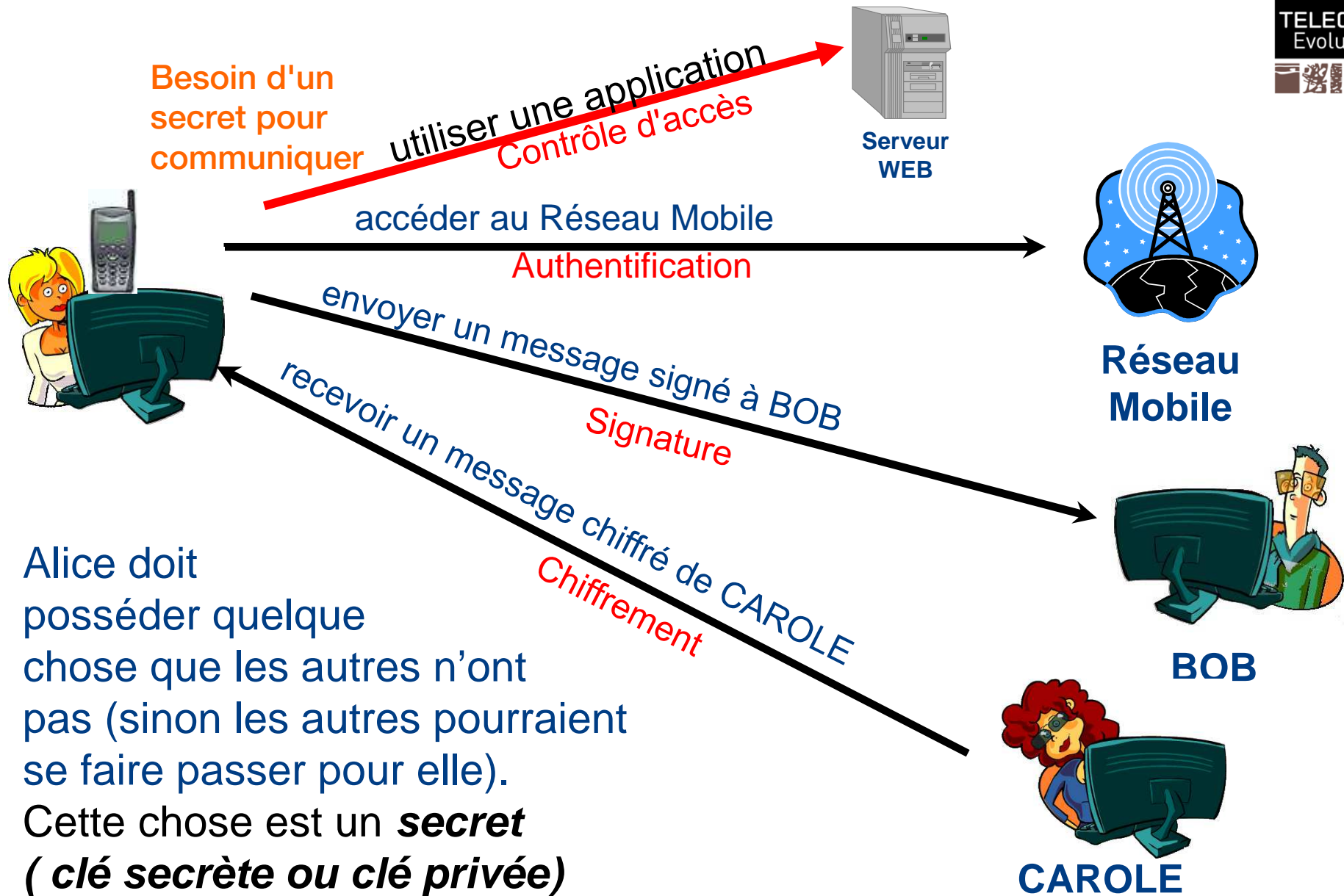
Exemple dans le cas du chiffrement : *confidentialité des données*

Chiffrement symétrique



Chiffrement asymétrique





Crypto symétrique

Alice partage son secret

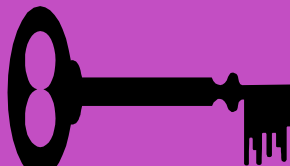


Le réseau Mobile
authentifie Alice



Réseau
Mobile

avec un
algorithme symétrique



Bob vérifie la
signature d'Alice



BOB

Carole chiffre un
message pour Alice



CAROLE

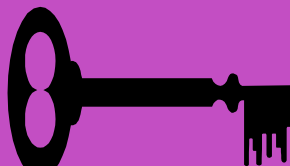
Crypto symétrique



Le réseau Mobile
authentifie Alice



avec un
algorithme symétrique



Bob vérifie la
signature d'Alice



BOB

Carole chiffre un
message pour Alice



CAROLE

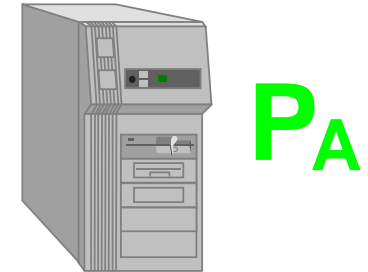
N'est pas adaptée aux systèmes « ouverts »

Crypto asymétrique

Alice donne sa clé publique

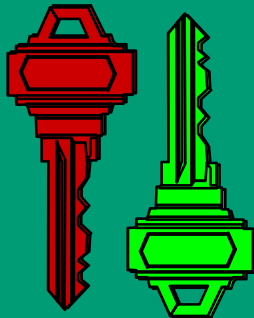


le serveur WEB
authentifie Alice



Serveur
WEB

avec un
algorithme asymétrique



Bob vérifie la
signature d'Alice



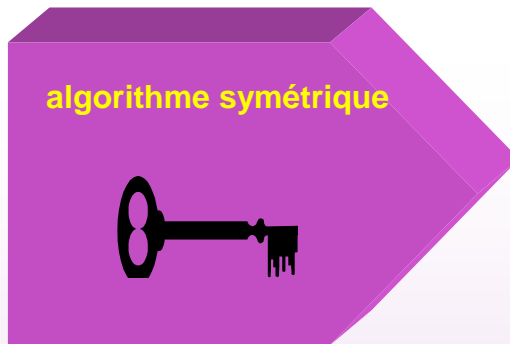
BOB

Carole chiffre un
message pour Alice



CAROLE

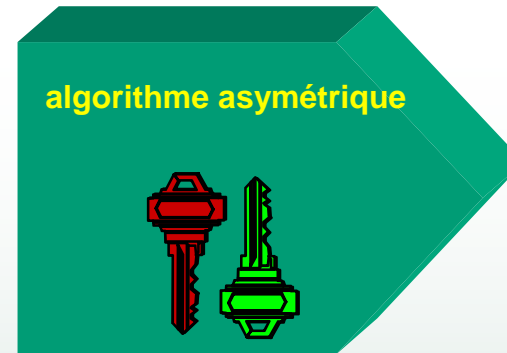
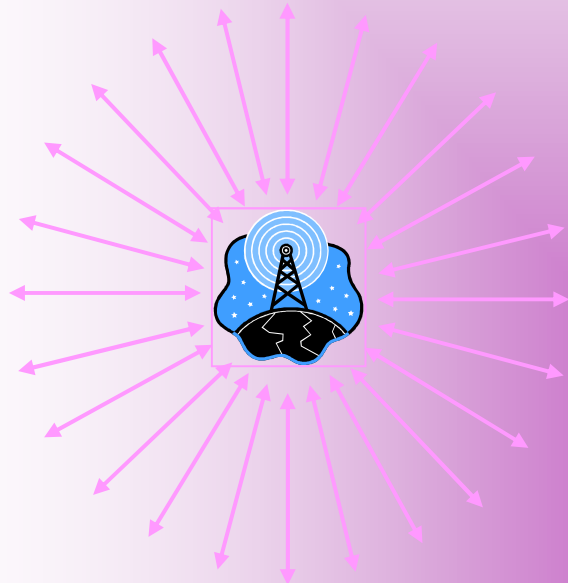
Algorithmes : symétrique - asymétrique



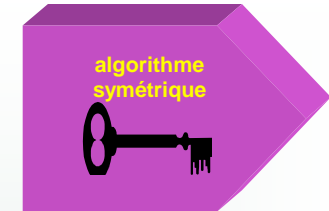
Systèmes « fermés »

Relation en « étoile »

Exemple : les réseaux mobiles



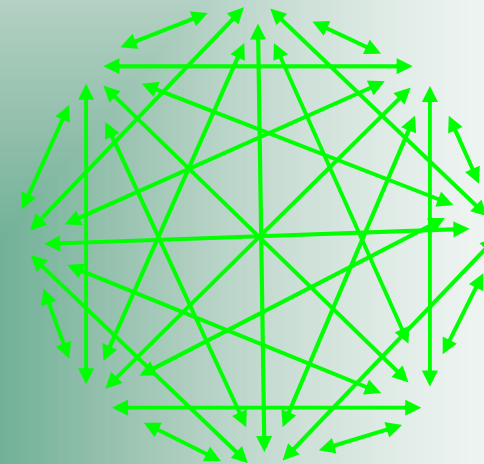
+



Systèmes « ouvert »

Relation « peer to peer »

Exemple : Internet



Cryptographie à clés secrètes (symétrique)

Algorithmes Symétriques ou à clé secrète

- Il ya bien longtemps !

- Chiffre de César
- Vigenère

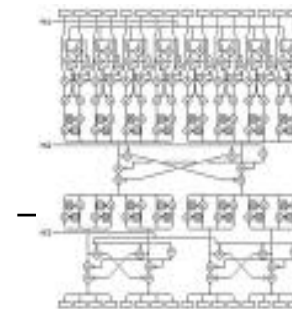
A→D
B→E
C→F
...
Z→C

- Électromécanique : enigma



- Avec l'électronique et l'informatique :

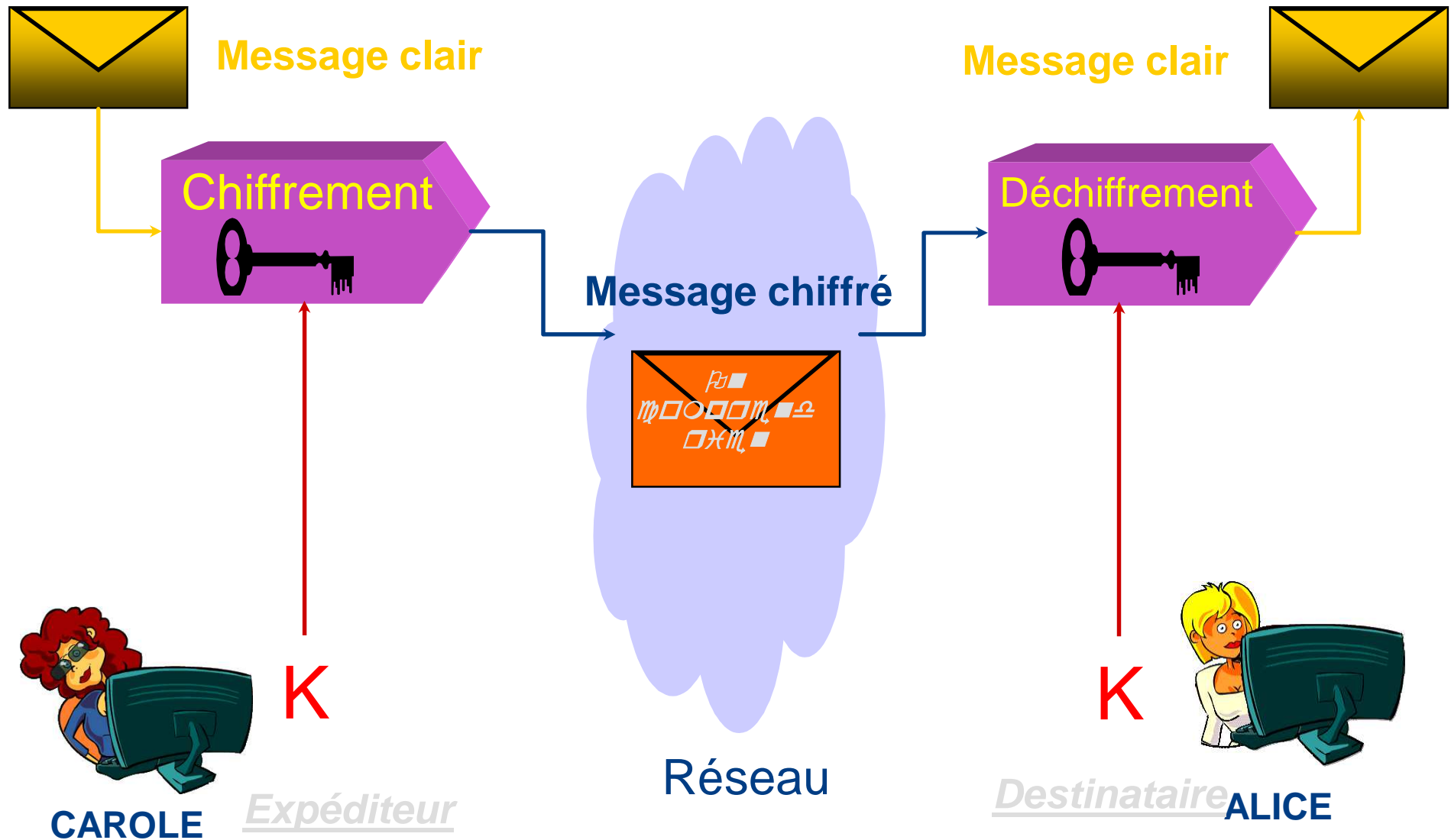
- Data Encryption Standart (DES) 1970
- IDEA, RC2, RC4 ... 1980 - 1990
- Advanced Encryption Standart : (AES – Rijndael) Oct 2000



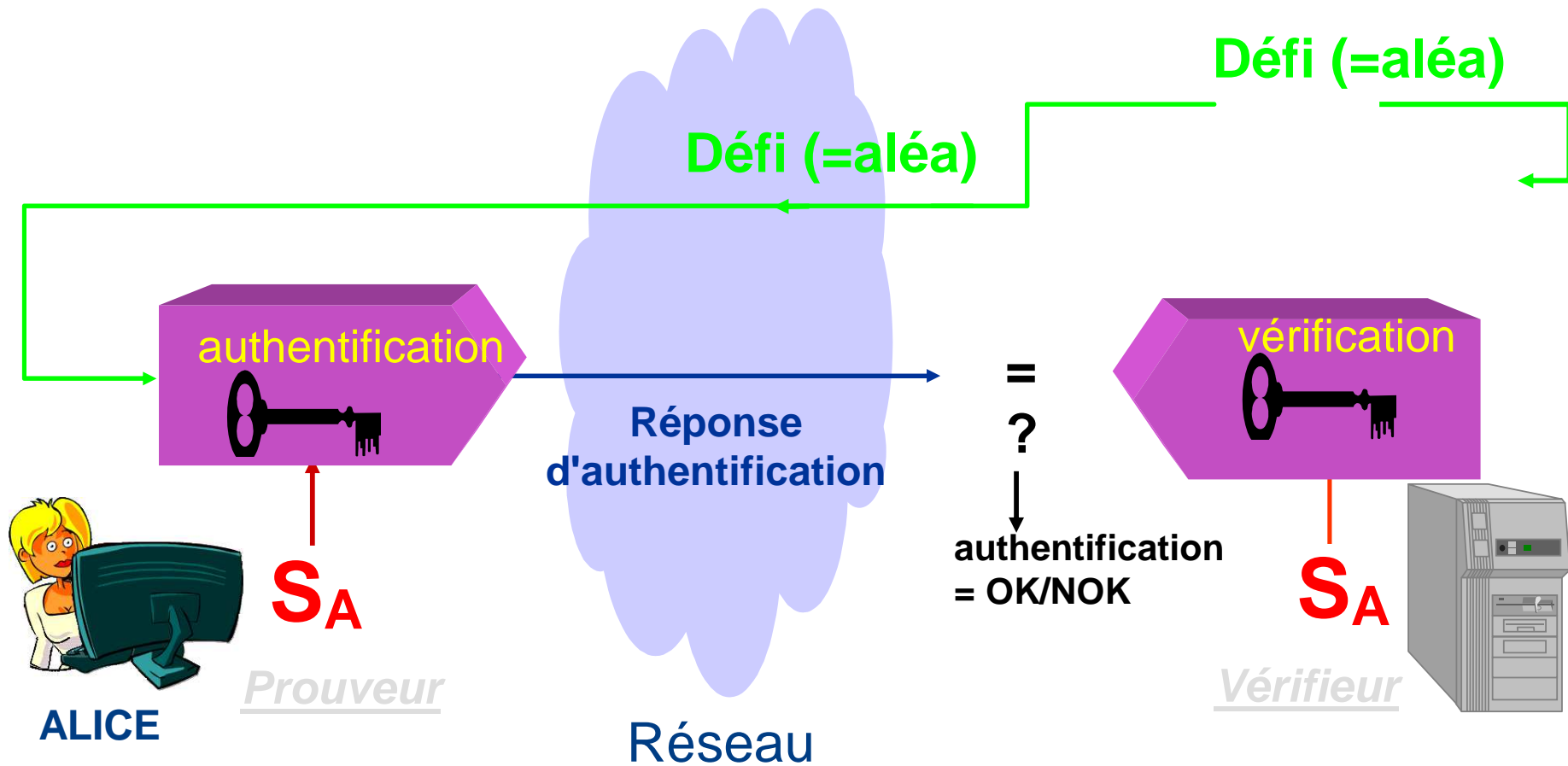
Cryptographie à clé secrète

- Repose sur des techniques simples (XOR, permutations de bits, additions...)
- Nécessité d'un partage/échange préalable du secret entre chaque utilisateur
- Une clé S par couple ou groupe d'utilisateurs
- La clé secrète S doit être maintenue secrète :
 - Chez chaque utilisateur
 - Lors de sa création et transfert/échange
- L'algorithme de déchiffrement est l'inverse de l'algorithme de chiffrement

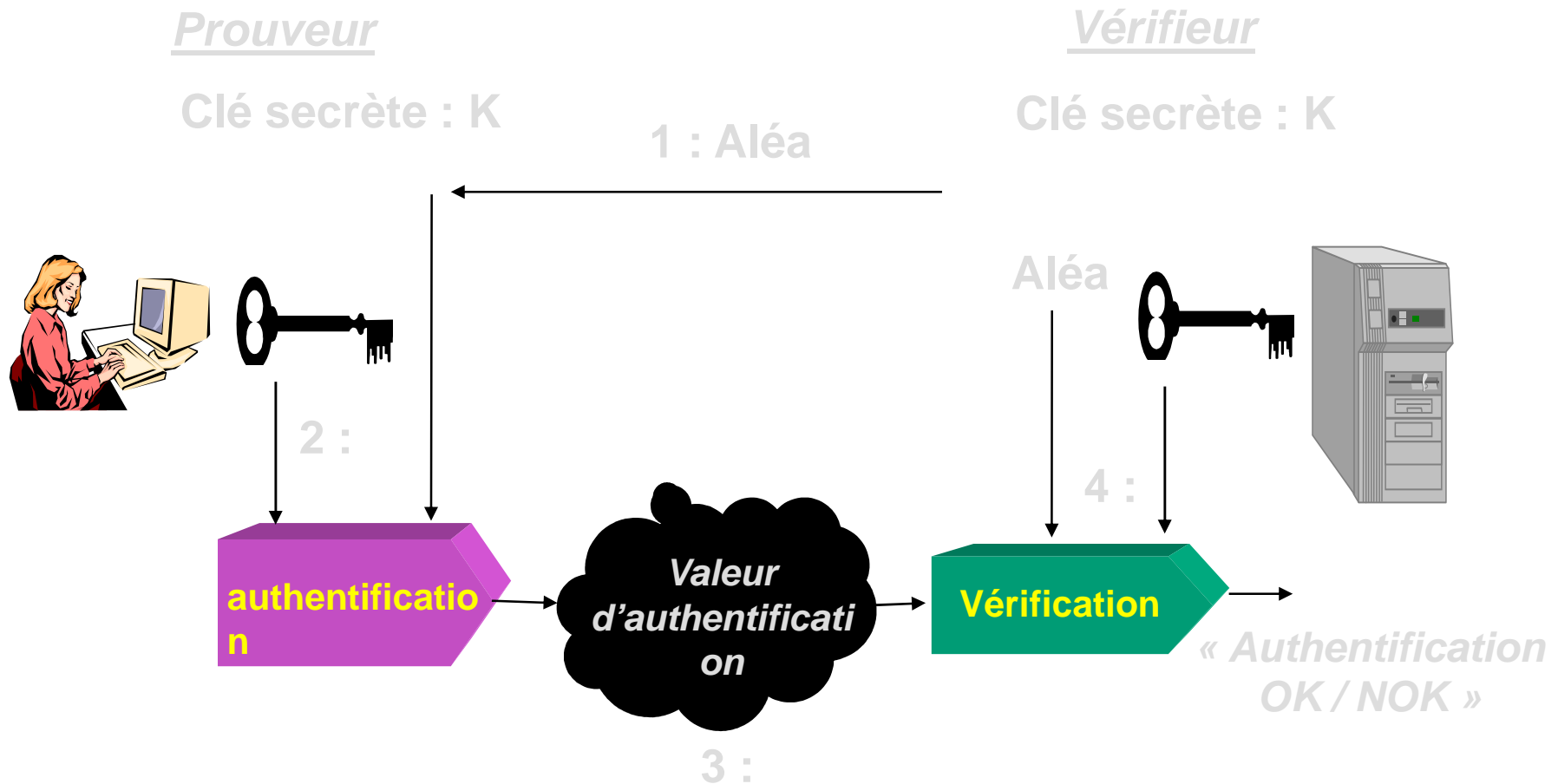
Chiffrement symétrique (à clé secrète)



L'authentification (forte) à clé secrète



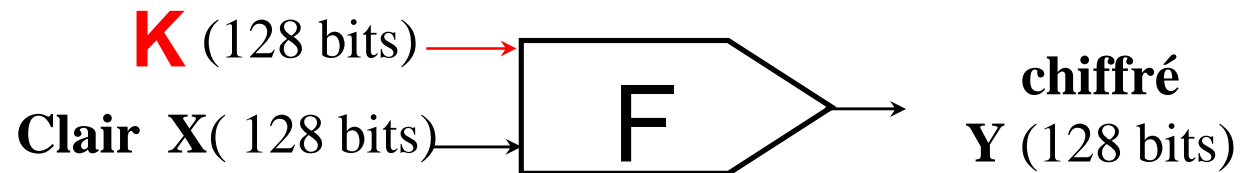
Crypto. symétrique (à clé secrète) : application à l'authentification



Algorithmes symétriques par blocs vs par flot

Algorithmes par blocs (ex : AES <http://csrc.nist.gov/CryptoToolkit/aes/>)

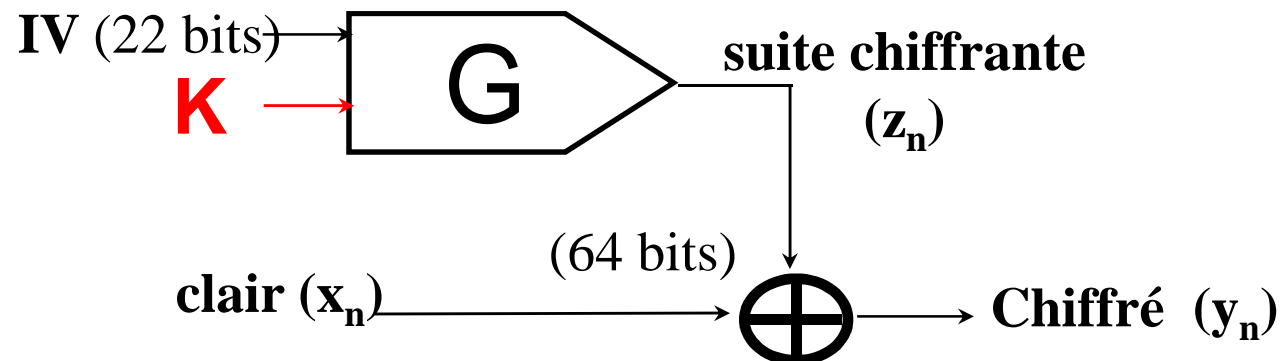
A la clé K est associée une fonction F_K



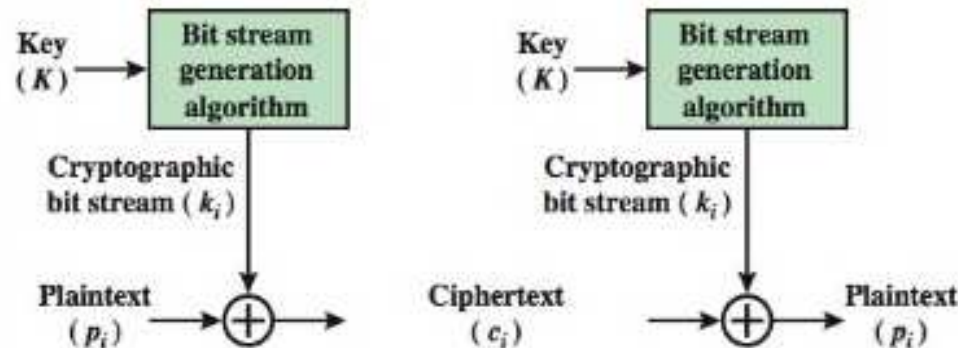
*Autres exemples .
DES, 3-DES*

Algorithmes par flot (ex : GSM A5)

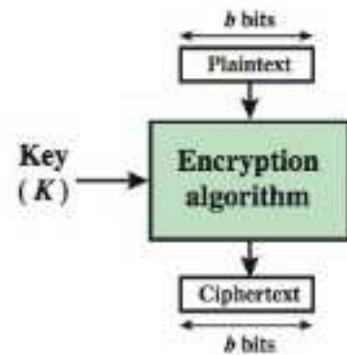
A la clé K est associée une suite chiffrante (z_n)



Block cipher vs stream cipher



(a) Stream Cipher Using Algorithmic Bit Stream Generator



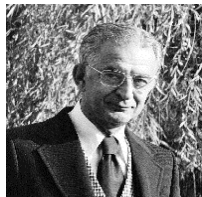
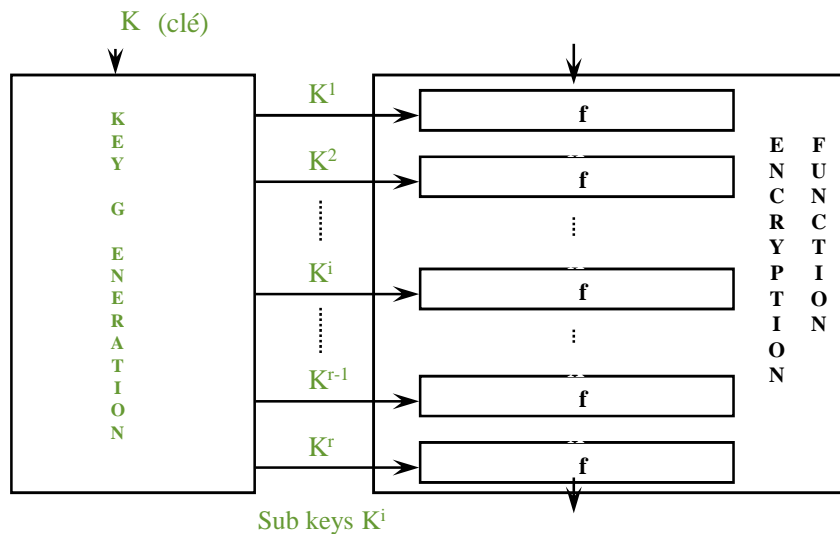
(b) Block Cipher

Block Cipher	Stream Cipher
Processing or encoding of the plain text is done as a fixed length block one by one. A block for example could be 64 or 128 bits in size.	Processing or encoding of plain text is done bit by bit. The block size here is simply one bit.
The same key is used to encrypt each of the blocks	A different key is used to encrypt each of the bits.
A Pad added to short length blocks	Bits are processed one by one in as in a chain
Uses Symmetric Encryption and is NOT used in asymmetric encryption	High speed and low hardware complexity
Confusion factor: The key to the cipher text relationship could be really very complicated.	Key is often combined with an initialization vector
Diffusion Factor: output depends on the input in a very complex method.	Long period with no repetition
Most block ciphers are based on Feistel cipher in structure	Statistically random
Looks more like an extremely large substitution and Using the idea of a product cipher	Depends on a large key and Large liner complexity
More secure in most cases	Equally secure if properly designed
Usually more complex and slower in operation	Usually very simple and much faster
Examples of Block Cipher are: Lucifer / DES, IDEA, RC5, Blowfish etc.	Examples of Stream Cipher are: FISH, RC4, ISAAC, SEAL, SNOW etc.

Block cipher algorithms examples

Algorithm	key length	block length	from	in
DES 3DES	56 112	64 64	IBM	still in >50% products
RC2 RC6	40-256 128-256	64 128	Rivest RSA Labs	S/MIME
IDEA	128	64	Massey Lai	PGP
MISTY KASUMI	128 128	64 64	Matsui 3GPP	UMTS
AES (Rijndael) <i>new standard</i>	128-256	128	Daemen Rijmen	replaces DES

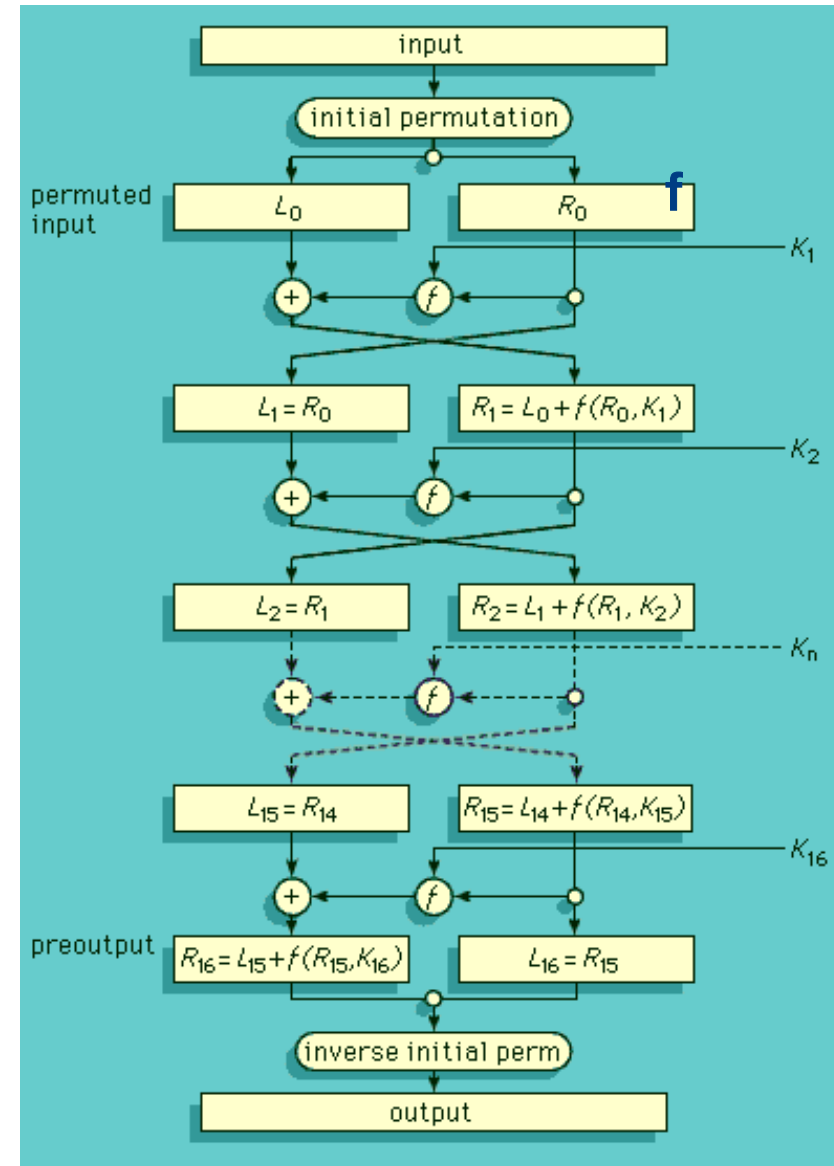
DES (Data Encryption Standard)



Horst Feistel



Don Coppersmith



AES: Advanced Encryption Standard

<http://csrc.nist.gov/CryptoToolkit/aes/>



Vincent Rijmen and Joan Daemen

1. The key that is provided as input is expanded into an array of forty-four 32-bit words.

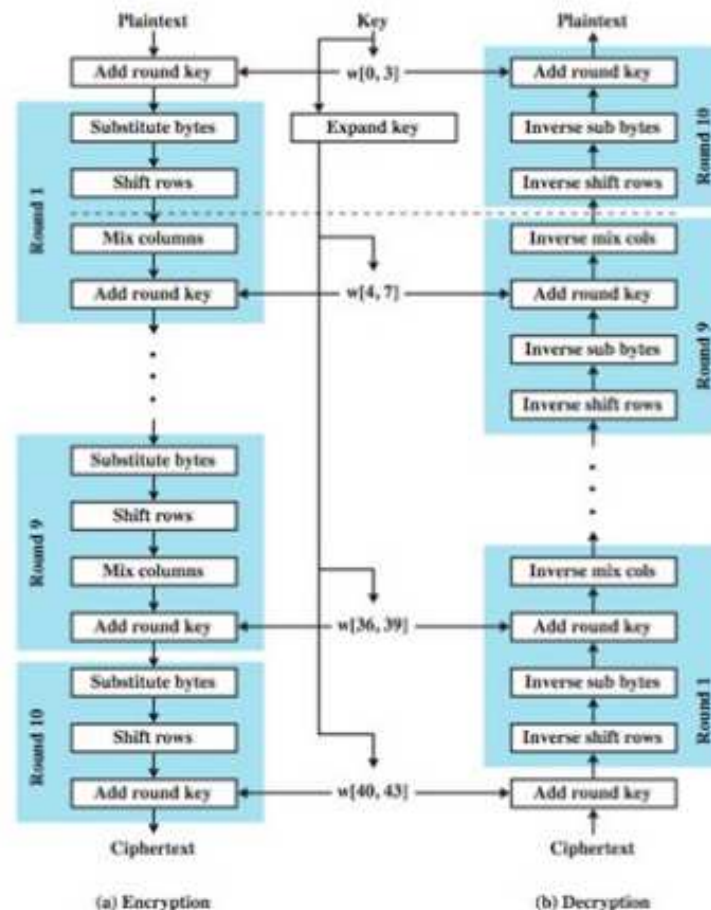
2. Four different stages are used, one of permutation and three of substitution

Substitute bytes: Uses a table, referred to as an S-box, to perform a byte-by-byte substitution of the block

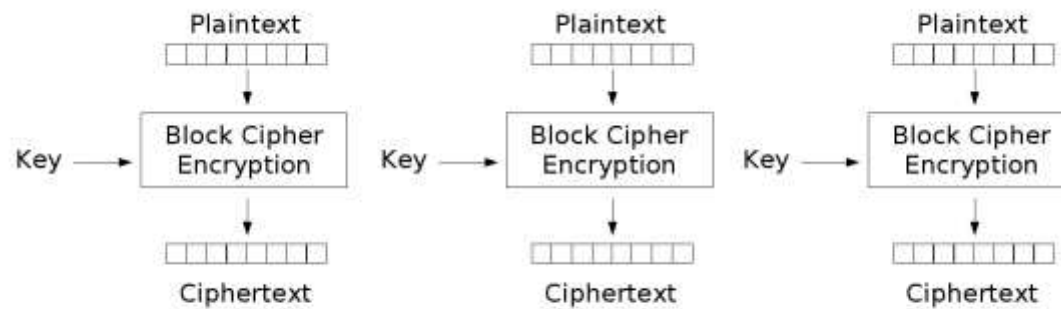
Shift rows: A simple permutation that is performed row by row

Mix columns: A substitution that alters each byte in a column

Add round key: A simple bitwise XOR of the current block with a portion of the expanded key

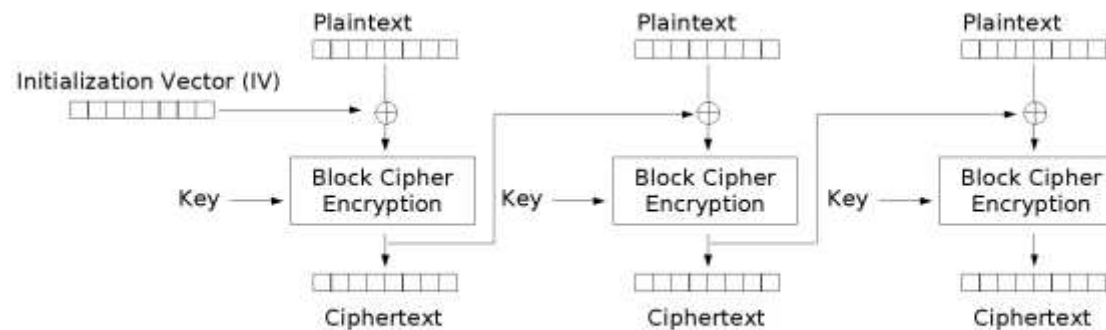


Implementation methods impact security



Electronic Codebook (ECB) mode encryption

- In ECB mode, each plaintext block gives a single block of ciphertext. It is therefore possible to recognize a pattern of the plaintext watching the ciphertext (e.g. same plaintext blocks give same ciphertext blocks).
- An active attacker can substitute a ciphertext block with another block of ciphertext produced with the same key (a block authentication problem).
- These problems are solved by the CBC method :



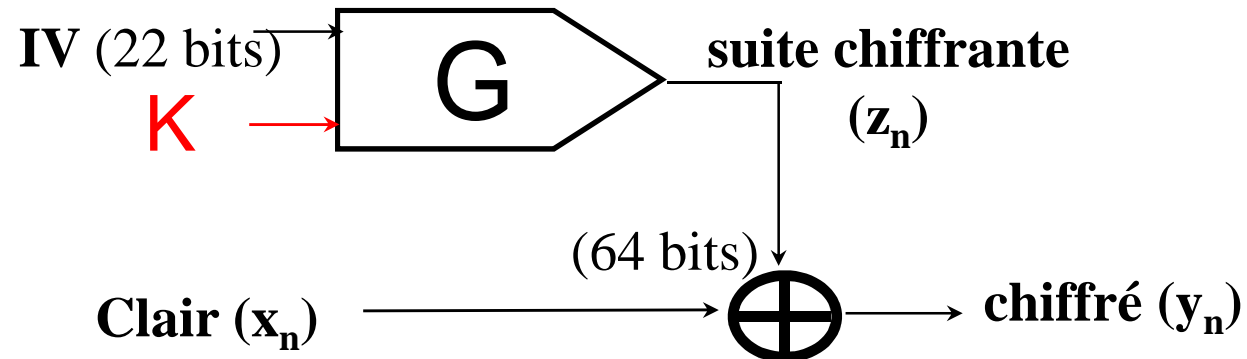
Cipher Block Chaining (CBC) mode encryption

Exemples d'algorithmes par blocs

Algorithme	Taille de clé	Taille de bloc	Origine	Utilisé dans
DES 3DES	56 112	64 64	IBM	>50% produits
RC2 RC6	40-256 128-256	64 128	Rivest RSA Labs	S/MIME
IDEA	128	64	Massey Lai	PGP
MISTY KASUMI	128 128	64 64	Matsui 3GPP	UMTS
AES (Rijndael) <i>nouveau standard</i>	128-256	128	Daemen Rijmen	À terme : remplacera DES

Exemples d'algorithmes par flot

A la clé K est associée une suite chiffrante (z_n)



Algorithme	Taille des clés	Origine	Utilisé dans
RC4	40-256	Rivest RSA-Labs	SSL
A5/1	64	ETSI	GSM
GEA2	64	ETSI	GPRS
SEAL	128	IBM	

Les attaques cryptographiques sur algos à clés secrètes (blocs ou flux)

Plus la clé est longue, plus le système est sûr :

- « 64 bits » signifie: 18446744073709551616 clés possibles de longueur 64 bits
- « 256 bits » signifie: 15792089237316195423570985008687907853269984665640564039457584007913129639936 clés

Toujours considérer que les algos sont connus de tous. Le seul élément secret doit être la clé secrète

- Recherche exhaustive : force brute (Balayage de l'ensemble des clefs)
 - Sans faiblesse cryptanalytique connue, une attaque nécessite en moyenne $2^L - 1$ exécutions (L = longueur de clé)
 - La complexité dépend peu de l'algorithme considéré (facteur <10 en pratique). Ordres de grandeur :
 - loi de Moore : Puissance de calcul * 2 tous les ans (= gain de 1 bit/an pour le balayage des clés)
- Attaques par cryptanalyse : spécifiques aux algos, compliquées, compétence mathématique pointue
 - Étude des moyens de casser un algorithme crypto, typiquement déchiffrer un message, sans connaître la clef
 - Utilisation d'effets de bord des algo. pour décoder l'information sans connaître la clef
 - Chiffré connu : on ne connaît que des messages chiffrés.
 - Clair connu : on connaît des messages clairs et leurs chiffrés
 - Clair choisi : on peut choisir des messages clairs et obtenir leurs chiffrés
- Faiblesses liées aux implémentations
 - Principalement dans les générateurs de nombres aléatoires
 - Simplifications ou erreurs dans la programmation

Tailles de clés recommandées

Plus la clé est longue, plus le système est sûr? ...

- « 64 bits » : 18446744073709551616 clés possibles de 64 bits de longueur
- « 256 bits » : 15792089237316195423570985008687907853269984665640564039457584007913129639936 clés

Key Size	Possible combinations
1-bit	2
2-bit	4
4-bit	16
8-bit	256
16-bit	65536
32-bit	4.2×10^9
56-bit (DES)	7.2×10^{16}
64-bit	1.8×10^{19}
128-bit (AES)	3.4×10^{38}
192-bit (AES)	6.2×10^{57}
256-bit (AES)	1.1×10^{77}

Key Size (bits)	Number of Alternative Keys	Time required at 10^6 Decryption/ μ s
32	$2^{32} = 4.3 \times 10^9$	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	10 hours
128	$2^{128} = 3.4 \times 10^{38}$	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	5.9×10^{30} years

- Taille de clés minimale recommandée en crypto symétrique : 128 bits
- DES cracker (EFF deep crack) en 1998 : 5 jours pour trouver une clé
- (90 milliards de clés par seconde, horloge = 2 M Hz



- Imaginons aujourd'hui... ☺

Conseils pour un chiffrement efficace de données

- Choisir un algorithme prouvé solide et adapté au besoin
 - Utiliser les algorithmes qui ont fait leurs preuves et les clés les plus longues possibles.
 - Un chiffrement matériel apporte un plus en accélérant la vitesse de chiffrement.
 - Les données stockées chiffrées restent logiquement isolées des autres.
- Protéger ses clés de chiffrement
 - Sécuriser le cycle de vie des clés de chiffrement: génération, transfert, stockage, renouvellement, effacement
 - Ne pas stocker les clés de chiffrement auprès des données (PIN sur carte bancaire)
- Ne jamais communiquer les clés
 - Chiffrer les données lorsqu'elles sortent de l'entreprise et les déchiffrer lorsqu'elles reviennent.
 - exemple pour mise de données dans le Cloud: VPN + stockage chiffré
- Attention à la manipulation directe des données chiffrées
 - Comment assurer le traitement des données chiffrées et leur sécurité ?
- S'assurer de la sécurité des utilisateurs et de leur environnement de travail
 - employés nomades, utilisation de réseau publics, télétravail
 - coexistence de données perso et pro sur le même équipement,
 - Cloud: qui peut avoir accès aux données , en local au stockage ou à distance ?

Limites des techniques à clé secrète

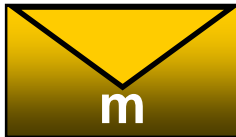
- Partage préalable des clés : Chaque couple doit préalablement s'échanger une clé commune de façon sûre, ou sur un canal autre que le canal à protéger.
- Une clé par couple d'utilisateur : Problème dans un réseau maillé lorsque N utilisateurs veulent dialoguer avec les N-1 autres (messagerie sécurisée) => $N.(N-1)/2$ clés différentes.
- Problème du maintien de la confidentialité des clés secrètes
- Pas de véritable signature vérifiable par un tiers : pas de non répudiation de signatures => pas d'arbitrage possible en cas de «disputes» entre deux interlocuteurs.

Fonction de condensation

- Une fonction de hachage « H » (on trouve aussi « hash function » ou « fonction de condensation ») transforme une entrée de données de taille variable « m » et en une sortie de taille inférieure et fixe : h
 - $h = H(m)$
- Une fonction de condensation doit remplir les conditions suivantes :
 - L'entrée peut être de dimension variable.
 - La sortie doit être de longueur fixe.
 - $H(m)$ doit être relativement facile à calculer.
 - $H(m)$ doit être une fonction à sens unique (non inversible).
 - $H(m)$ doit être sans collision (difficile de trouver deux messages ayant même haché).
- Une fonction de condensation permet de détecter toutes modification de m, fortuite ou malfaisante
 - un code détecteur d'erreur a des propriétés semblable mais ne permet pas de détecter des modification malfaisantes

Fonction de condensation

Ensemble des
messages

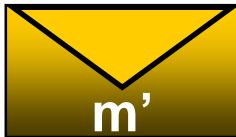
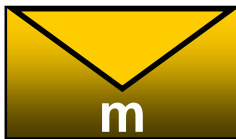


H

Ensemble des mots de p
bits (ex. : $p = 128$)

$H(m)$

Peut on signer $H(m)$ à la place de m ?

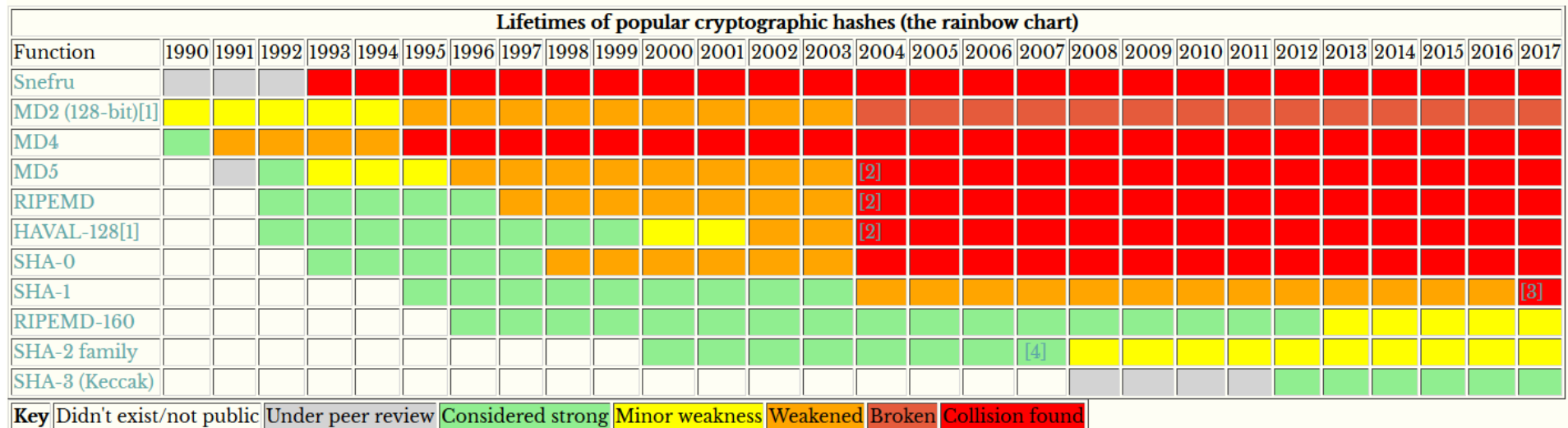


Collision !

$H(m) = H(m')$

**H est une fonction de condensation s'il est
« informatiquement » impossible de construire une collision**

Fonctions de hachage: comparaisons



Source : <http://valerieaurora.org/hash.html>

Source :
<https://www.enisa.europa.eu/publications/algorithms-key-size-and-parameters-report-2014>

Primitive	Output Length	Classification	
		Legacy	Future
SHA-2	256, 384, 512	✓	✓
SHA3	256,384,512	✓	✓
Whirlpool	512	✓	✓
SHA3	224	✓	✗
SHA-2	224	✓	✗
RIPEMD-160	160	✓	✗
SHA-1	160	✓ ^[1]	✗
MD-5	128	✗	✗
RIPEMD-128	128	✗	✗

Table 3.3: Hash Function Summary

Anonymisation par fonction de hachage

- calculer une valeur numérique à partir des données directement ou indirectement nominatives « textuelles » d'un individu
- cette valeur est ensuite substituée aux données à partir desquelles elle a été calculée.
- Le caractère irréversible de l'anonymisation
 - taux très faible des collisions
 - bonnes performances informatiques des algorithmes de hachage

Cryptographie asymétrique (à clé publique)

Algorithmes Asymétriques ou à clé publique

- 1976 : Diffie – Hellman
 - Échange de clés






- 1977 : RSA (Rivest Shamir Adelman)
 - Signature électronique
 - Chiffrement (lent et lourd)



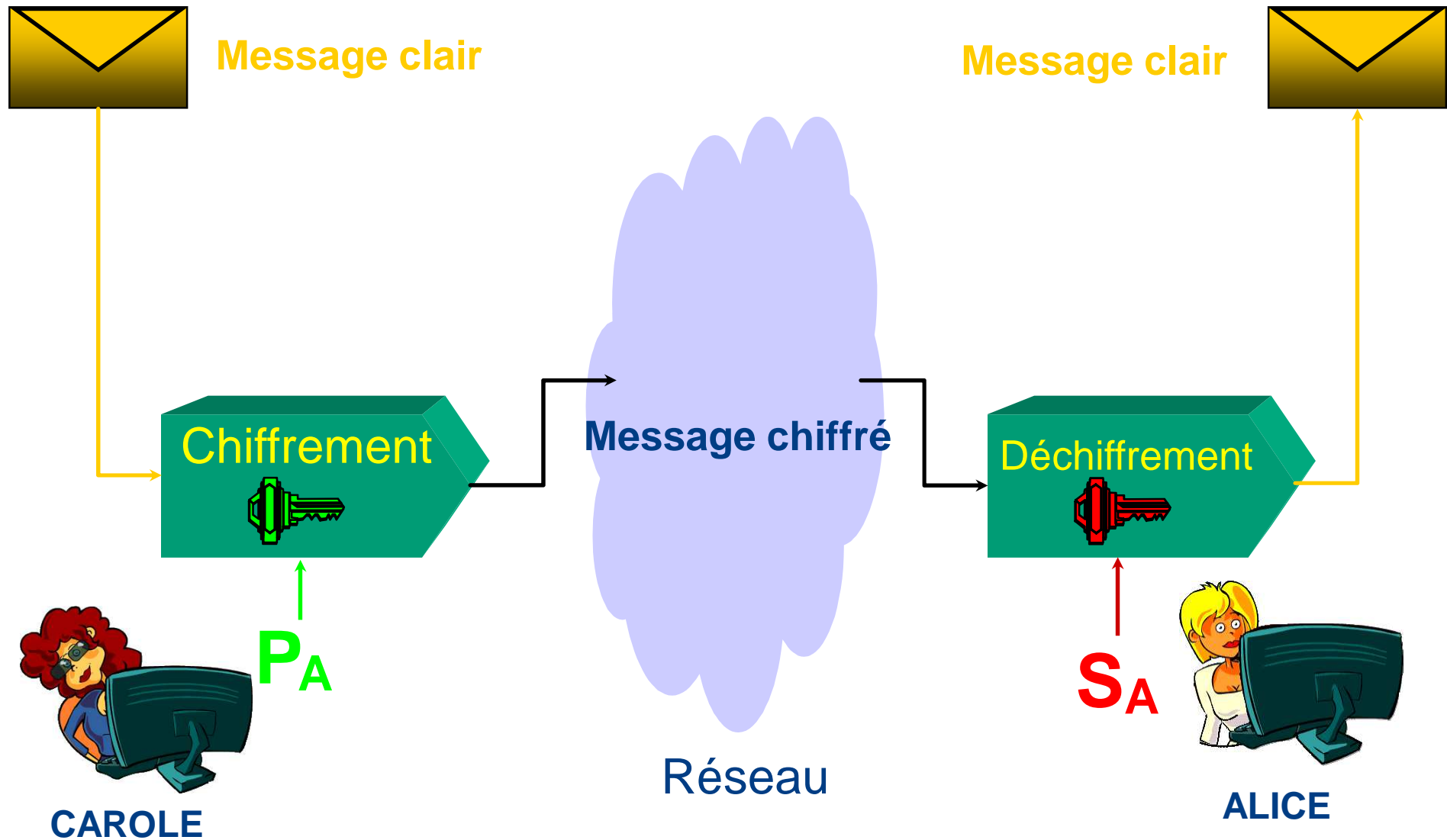
- Protocoles zéro-knowledge
- Certificat numérique et infrastructure à clé publique
- Aujourd'hui :
 - Signature aveugle, signature de groupe...
 - Anonymat révocable
 - Vote, enchère, concours anonyme, jeux en ligne

Cryptographie à clé publique

- Repose sur des techniques mathématiques complexes
- Pas de partage au préalable d'un secret pour chaque couple d'utilisateurs
- Un couple de clés personnelles (S , P) pour chaque utilisateur 
- P est dérivée de S par une fonction à sens unique
- P est rendue publique et définit la fonction de chiffrement utilisée par toute personne désirant établir une communication sécurisée avec celui qui l'a publiée 
- S est gardée secrète/privée par son propriétaire et définit la fonction de déchiffrement des messages reçus 

Dès qu'un utilisateur a choisi S et P, et publié P, toute autre personne peut lui envoyer des messages confidentiels...

Chiffrement (asymétrique)



À retenir

- Chaque individu a :
 - sa propre clé privée qu'il garde secrète
 - sa propre clé publique dérivée qu'il diffuse à ses interlocuteurs
 - ce couple de clés est différent d'un utilisateur à l'autre
- Je chiffre avec la clé publique de mon interlocuteur
 - Afin que mon interlocuteur soit le seul à pouvoir déchiffrer mon message avec sa propre clé privée qu'il est le seul à connaître
- Je signe avec ma propre clé privée, que je suis le seul à connaître
 - Afin que tout les interlocuteurs puissent vérifier ma signature à l'aide de ma clé publique que je leur ai donnée

RSA (1977)

- Pour générer ses clés, Alice
 - Tire au sort p et q premiers [> 512 bits chacun]
 - Choisit e premier avec (p-1)(q-1)
 - Calcule $n = pq$ [> 1024 bits]
 - Calcule d tel que $ed = 1 \mod ppcm((p-1), (q-1))$

Clé publique d'Alice : $P_A = (n, e)$
 Clé privée d'Alice : $S_A = d$
- Pour chiffrer un message pour Alice ou vérifier une signature d'Alice :

Bob utilise la fonction publique P_A d'Alice
 $X \rightarrow Y = P_A(X) = X^e \mod n$
- Pour déchiffrer un message de Bob ou signer un message :

Alice utilise sa fonction privée S_A
 $Y \rightarrow S_A(Y) = Y^d \mod n = (X^e)^d \mod n = X^{ed} \mod n = X^1 \mod n = X$

On a la propriété Fondamentale :

$$\mathcal{P} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{P} = Id$$

Car :

$$(X^e)^d = (X^d)^e = X^{ed} = X$$

Grâce au théorème de Fermat (ou d'Euler) :

$$X^{(p-1)(q-1)} = 1 \mod pq$$

Exemple numérique

$$p = 3, q = 11$$

$$n = pq = 3 \times 11 = 33$$

$$(p-1)(q-1) = 2 \times 10 = 20$$

$$e = 3; d = 7; ed \equiv 1 \pmod{(p-1)(q-1)}$$

$$\text{Car : } 3 \times 7 \equiv 1 \pmod{20}$$

On a alors : $\forall X \in \{0, 1, \dots, 32\}$

Si $Y = X^3 \pmod{33}$ alors $Y^7 = X \pmod{33}$ et

Si $Y = X^7 \pmod{33}$ alors $Y^3 = X \pmod{33}$

Exemple : $X = 29$

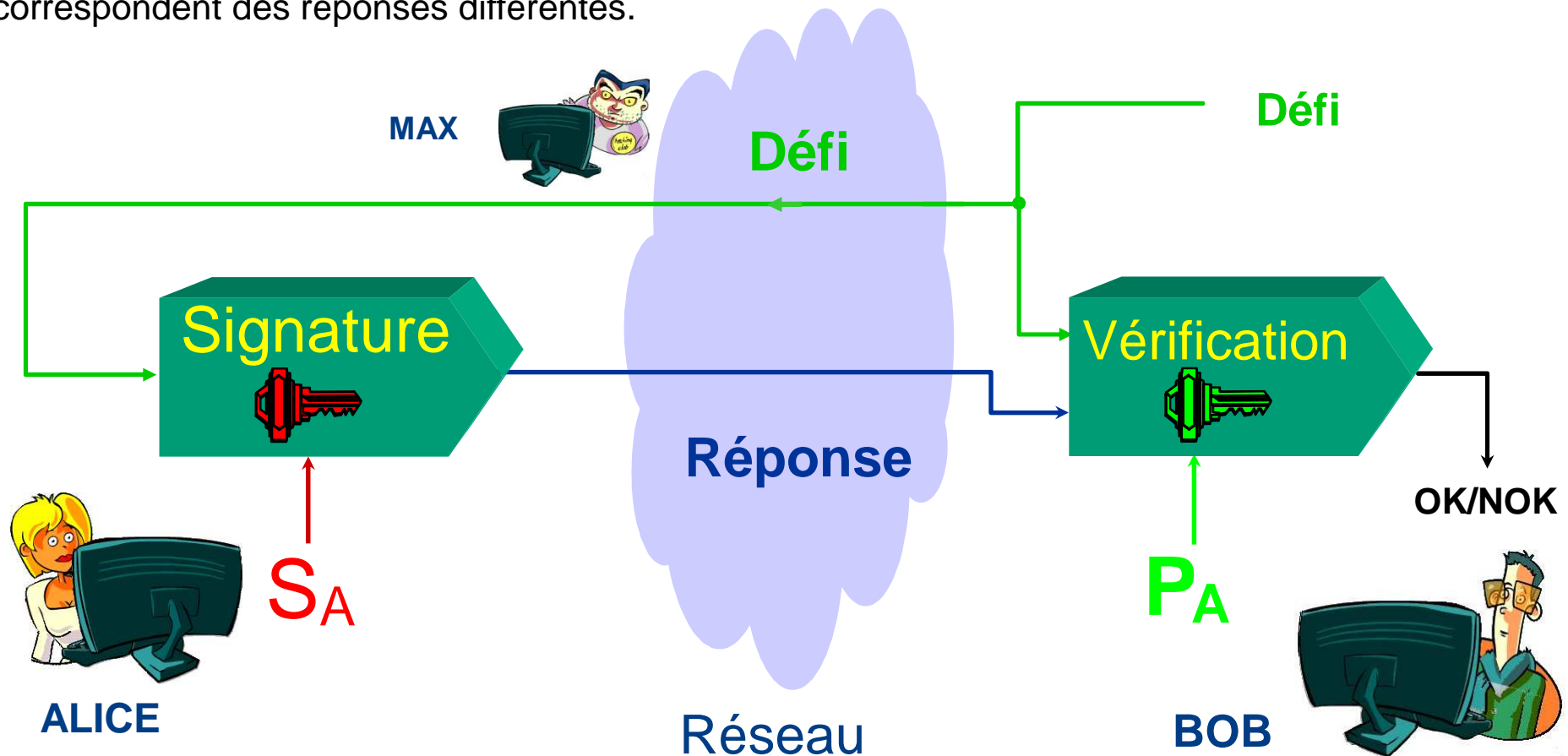
$$Y = X^e = 29^3 = 24389 = 2 + 24387 = 2 + (739 \times 33) \equiv 2 \pmod{33}$$

$$Y^d = 2^7 = 128 = 29 + 99 \equiv 29 \pmod{33} = X$$

L'authentification (forte) à clé publique

Seul le possesseur du secret S_A (Alice) peut construire la réponse au défi de Bob.

En observant les transactions, Max ne peut répondre à aucun défi car à des défis différents correspondent des réponses différentes.



Exemple avec RSA:

Alice génère Réponse = $(\text{Défi})^d \bmod n$

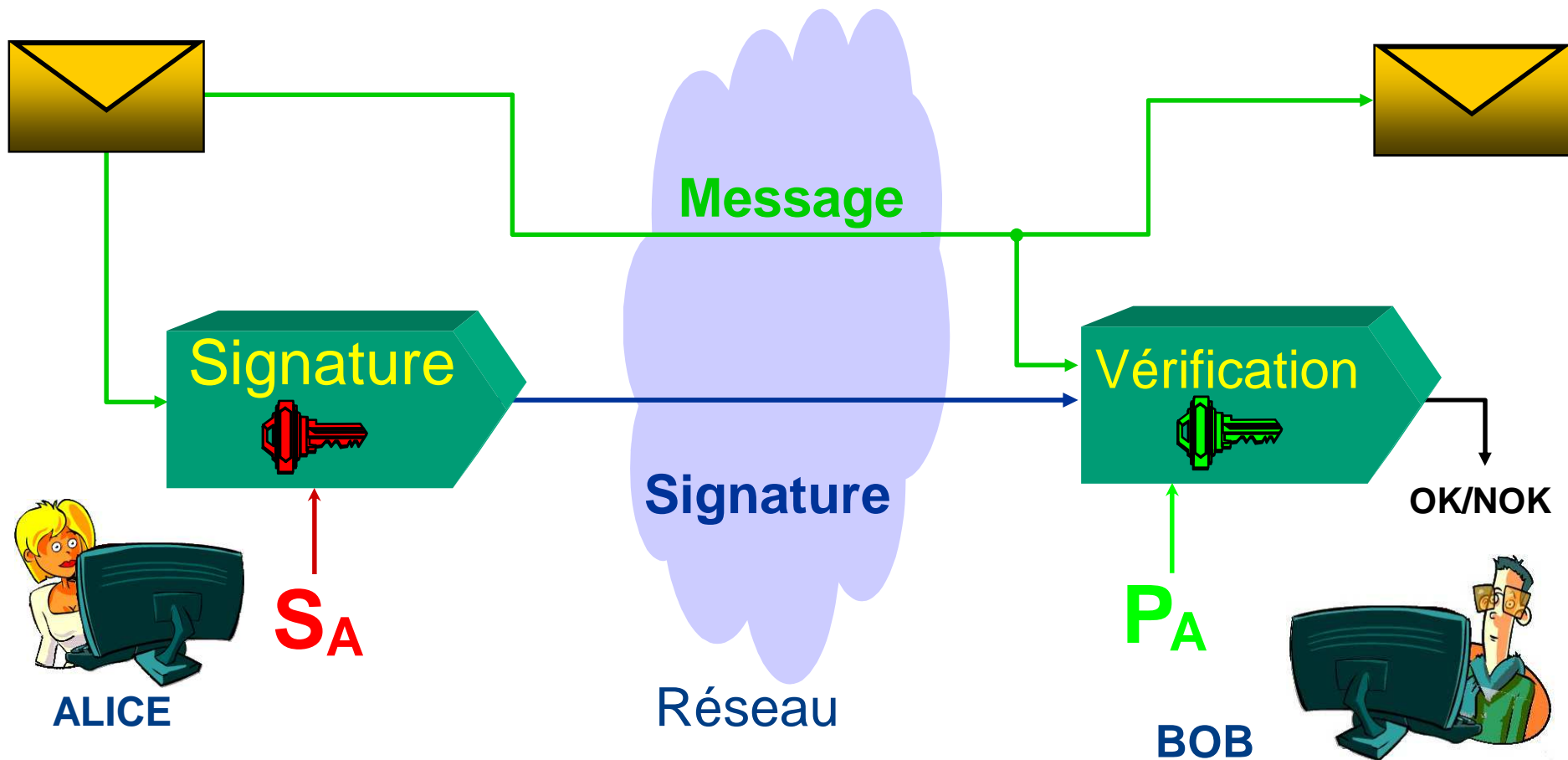
Exemple avec RSA:

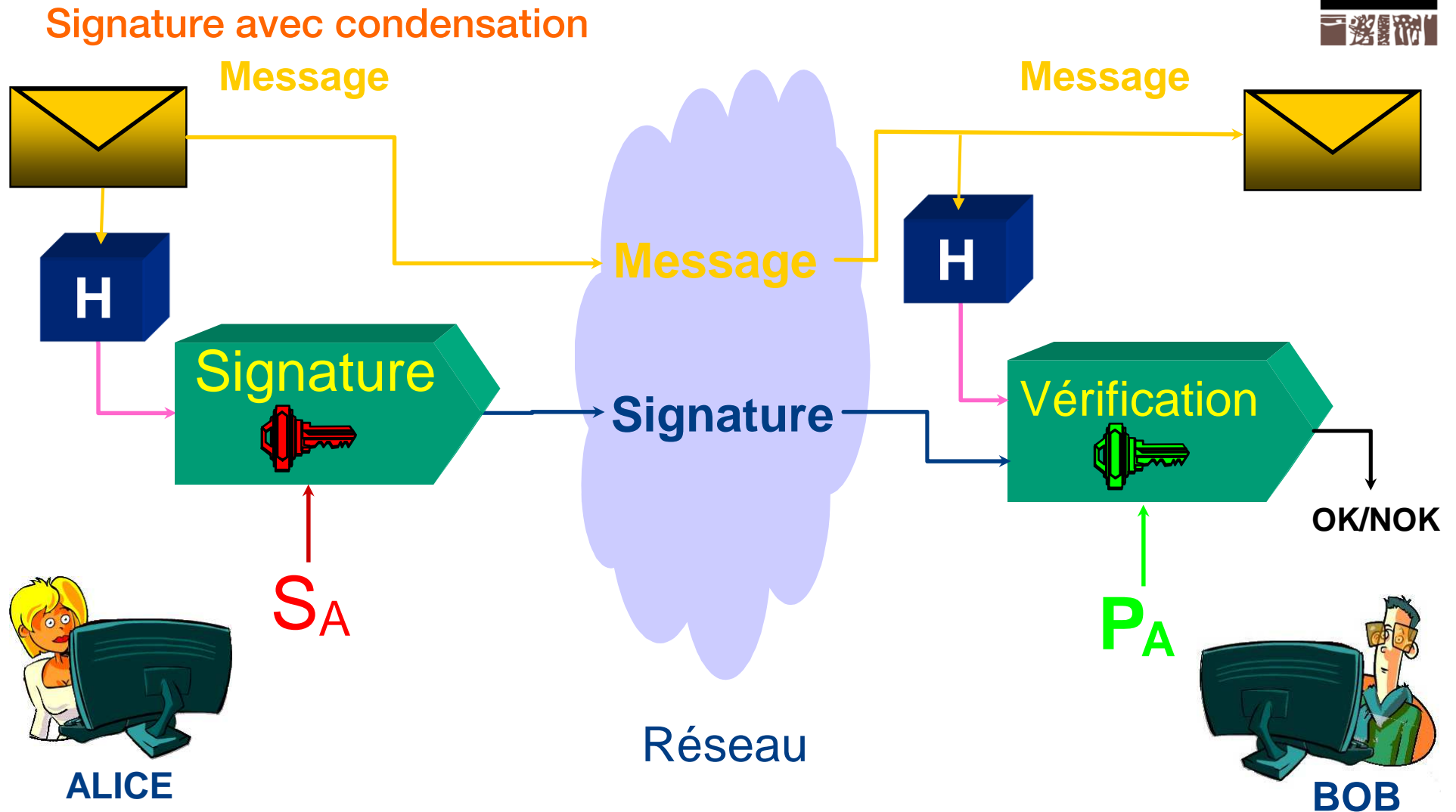
Bob vérifie $(\text{Réponse})^e \bmod n \stackrel{?}{=} \text{Défi}$

Signature

Seul le possesseur du secret S_A (Alice) peut construire la signature du message.

Tout le monde peut vérifier : il suffit de posséder la clé publique P_A





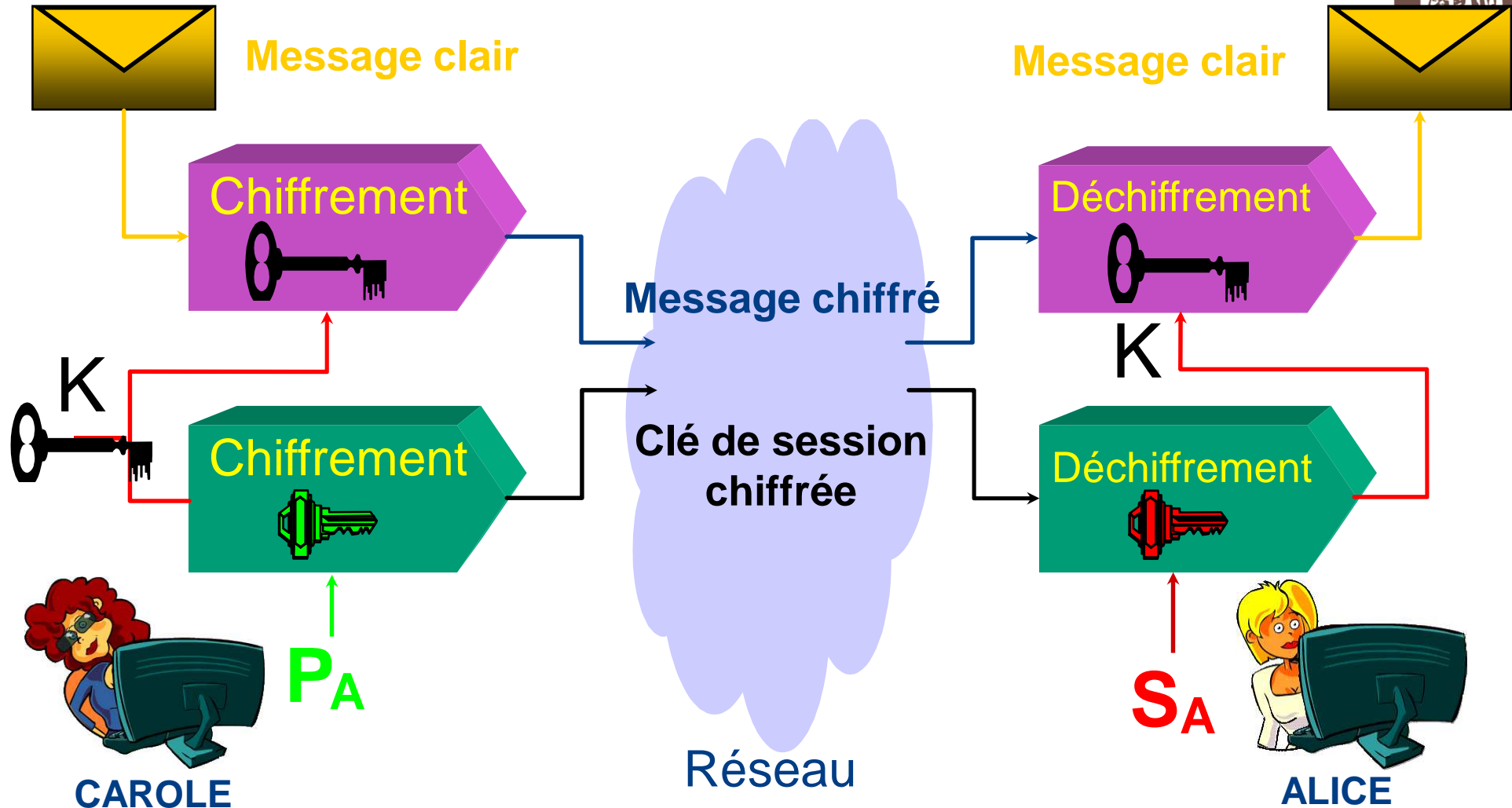
Exemple avec RSA:

Alice génère **Signature** = $(H(\text{Message}))^d \bmod n$

Exemple avec RSA:

Bob vérifie **(Signature)^e mod n** = $? = H(\text{Message})$

Chiffrement avec échange de clé de session



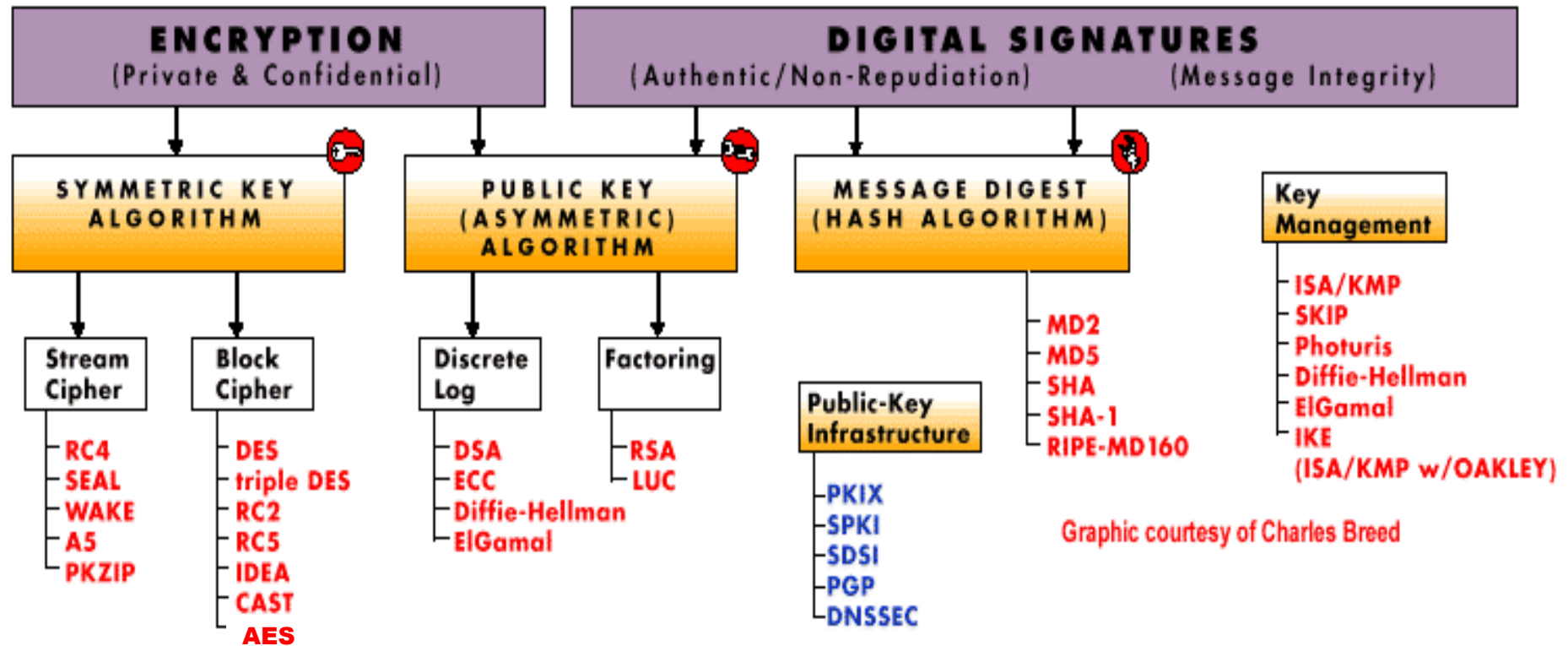
Exemple avec RSA:
Clé de session chiffrée = $K^e \bmod n$

M. chiffré = $C_K(\text{M. clair})$

Exemple avec RSA:
K = $(\text{Clé de session chiffrée})^d \bmod n$

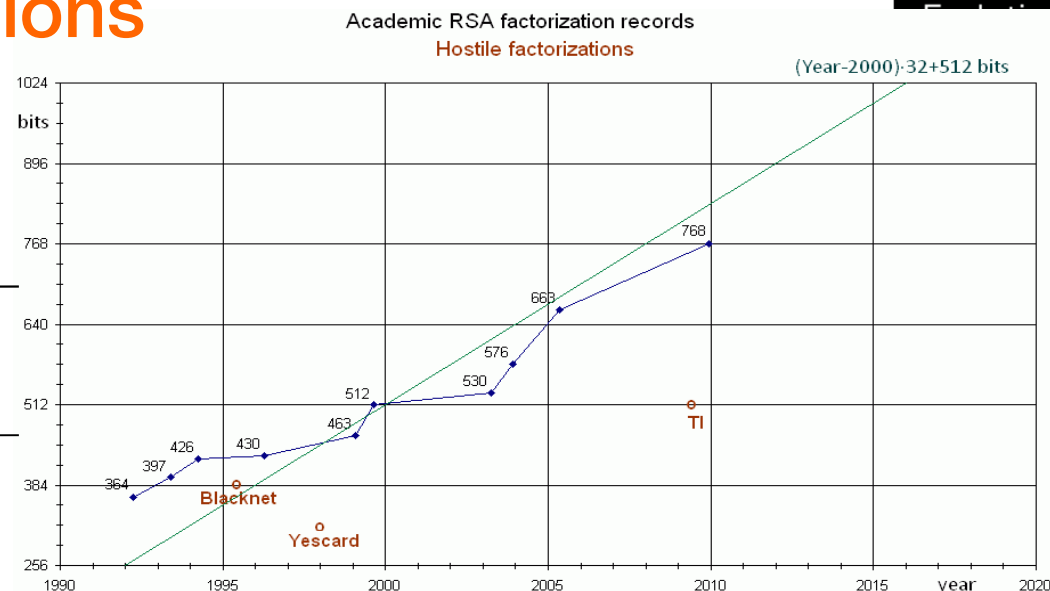
M. clair = $D_K(\text{M. chiffré})$

Algorithm overview



Key size recommendations

Symmetric	DH or RSA
56	512
80	1024
112	2048
128	3072
192	7680
256	15360



Method	Date	Symmetric	Asymmetric	Discrete Logarithm Key	Discrete Logarithm Group	Elliptic Curve	Hash
[1] Lenstra / Verheul ?	2020	86	1881 1472	151	1881	161	171
[2] Lenstra Updated ?	2020	82	1387 1568	163	1387	163	163
[3] ECRYPT II	2016 - 2020	96	1776	192	1776	192	192
[4] NIST	2011 - 2030	112	2048	224	2048	224	224
[5] ANSSI	2010 - 2020	100	2048	200	2048	200	200

Utilisation des schémas à clés publiques



- Chiffrement de données de taille limitée (clés, secret, Identifiant,)
- Gestion (échange) de clés
- Authentification forte
- Signature électronique, non répudiation
- Signature (intégrité) de fichiers
- Idéal dans un environnement ouvert

- Exemples d'applications
 - Internet (Authentification forte, transfert de clés de chiffrement...)
 - Commerce électronique et messagerie
 - Téléprocédures..
 - Tous services nécessitant le recours à une signature non répudiable

Légitimité d'une clé publique

- Les techniques à clé publique ne peuvent fonctionner que si les interlocuteurs d'Alice sont en mesure d'acquérir la certitude que la clé publique P_A appartient effectivement à Alice, i.e. que le détenteur de la clé secrète S_A est bien Alice.
=> problème de l'intégrité d'une clé publique.
- Faute d'une telle certitude, un usurpateur d'identité peut se faire passer pour Alice.
- Problème : Comment fournir la preuve infalsifiable de la légitimité de cette clé ?

Comment lier « identité » et « clé »

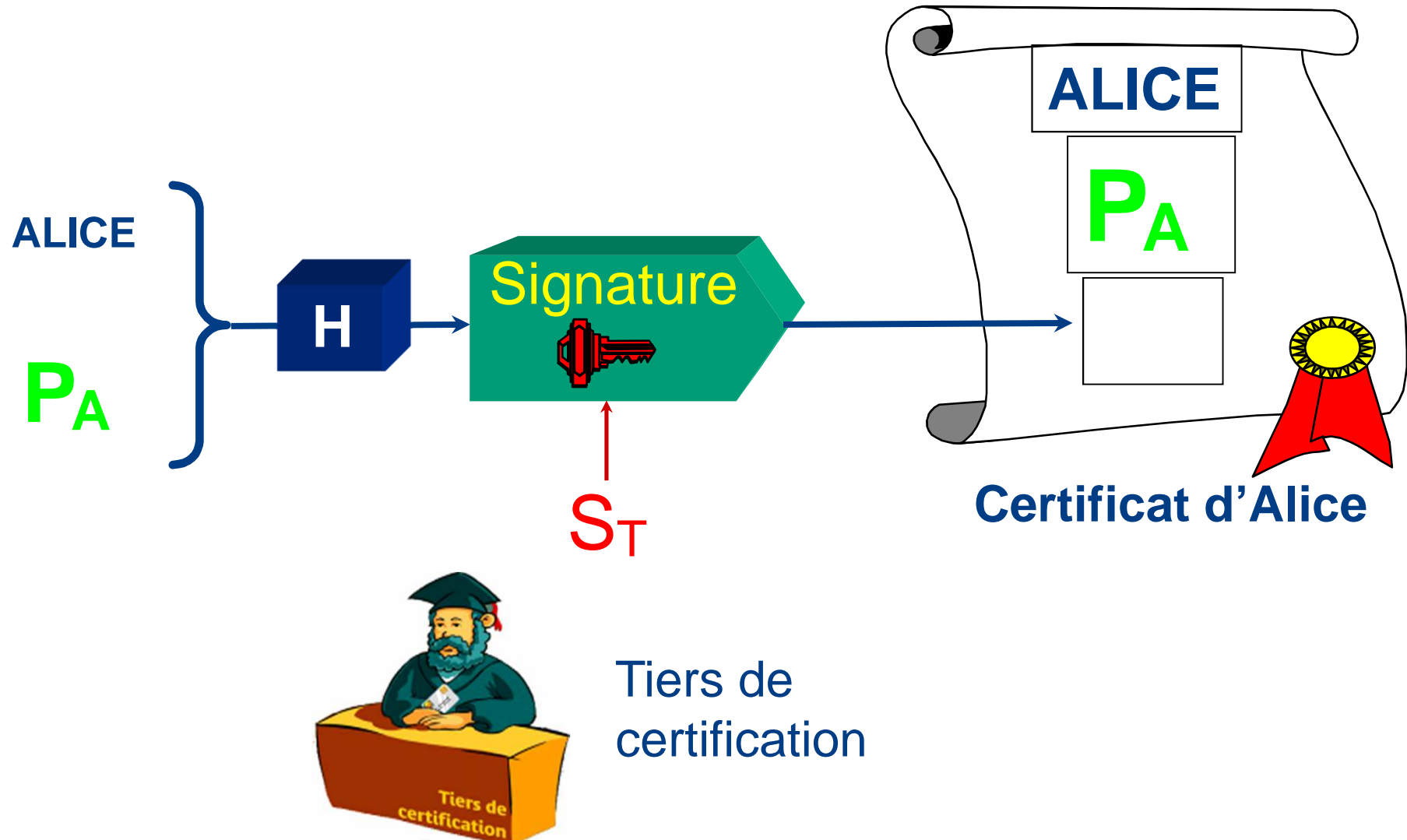


- L'identité est une chaîne de caractère qui représente concrètement l'individu ou l'objet qui réalise la transaction sécurisée.
- Une clé (secrète, privée ou publique) est une chaîne de bits qui n'a à priori aucun lien avec l'individu ou l'objet qui réalise la transaction. Pourtant c'est cette clé qui permet de sécuriser la transaction.
- **Il est donc indispensable de lier « identité » et « clé ».**
- Deux cas à distinguer :
 - Cas symétrique : on utilise la « ***diversification*** »
 - Cas asymétrique : on utilise la « ***certification*** ».

NB : dans le cas asymétrique il existe une technique particulière : les « ***schémas basés sur l'identité*** » où on a :
clé publique \equiv identité
et où la liaison clé - identité devient ainsi inutile.

Comment lier identité et clé : le cas asymétrique

« la certification »

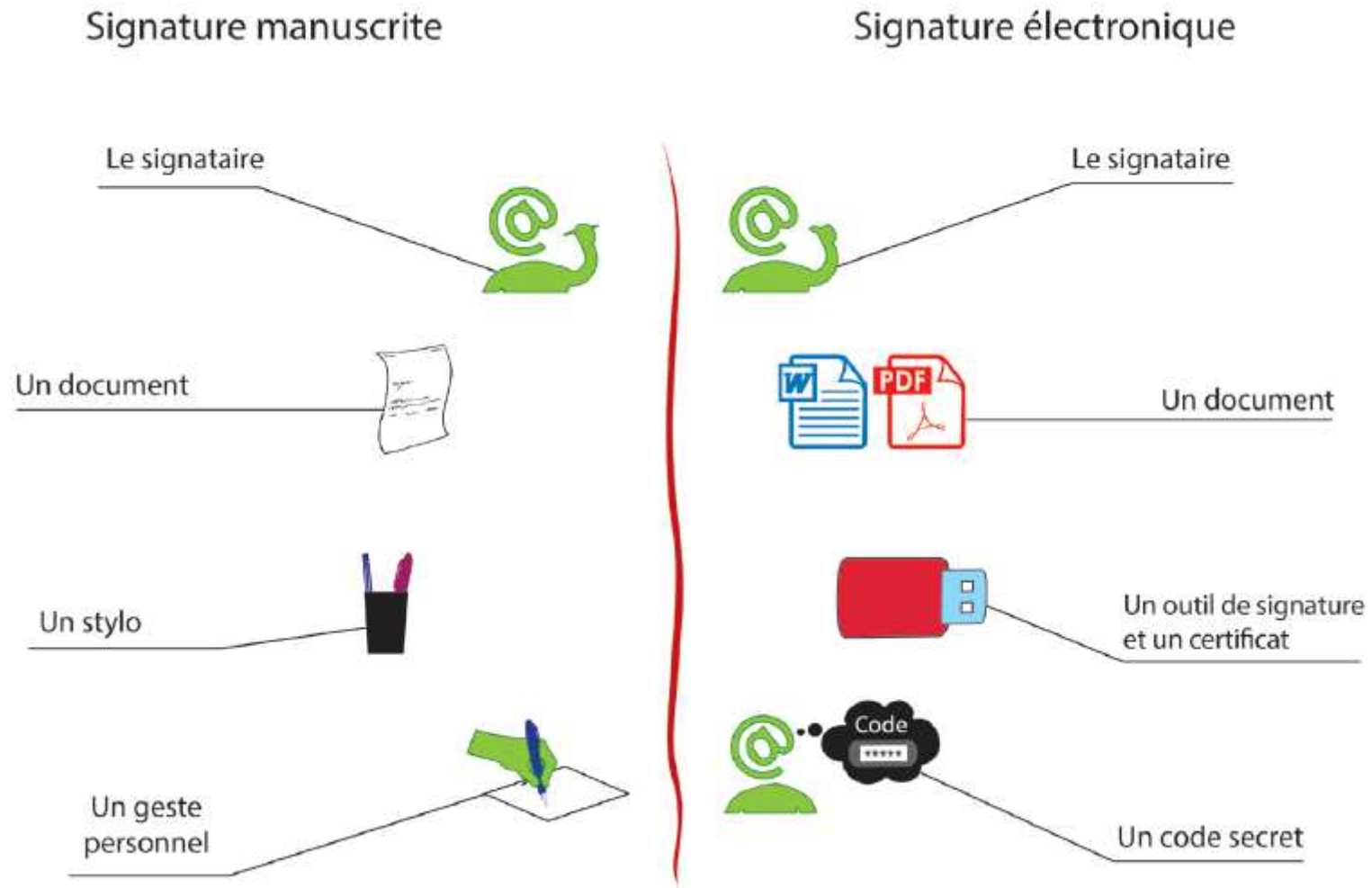




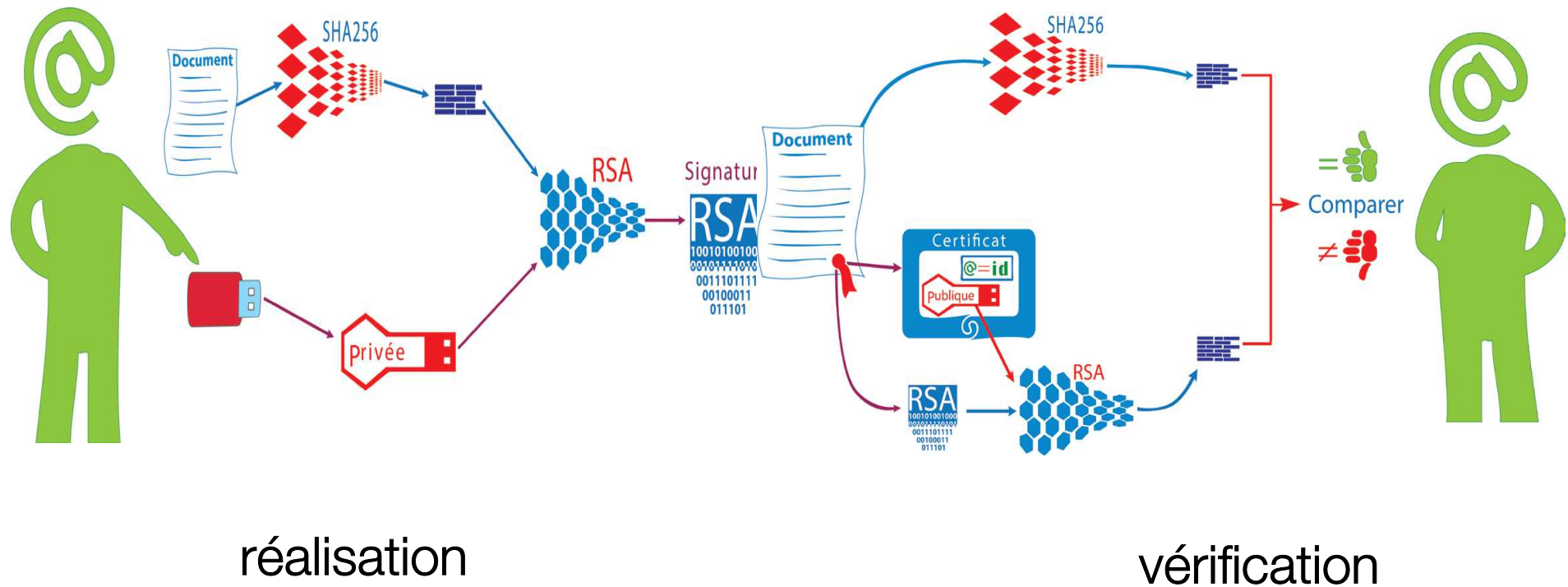
Signature électronique

- La signature électronique remplit deux rôles majeurs, qui tendent à établir les conditions de la confiance dans les échanges numériques et donc à rendre possible la dématérialisation :
 - la signature électronique d'un document (un contrat par exemple) confère à celui-ci une **valeur juridique équivalente à celle d'un document papier signé de manière manuscrite, en marquant l'engagement de la personne qui a apposé la signature** ;
- des fonctions connexes à la signature électronique (cachet, horodatage...) servent à offrir des conditions de **sécurité technique en garantissant sa** provenance, son intégrité, ou encore la date de sa réalisation.
- une signature électronique apporte :
 - la garantie de l'intégrité du document ;
 - un lien certain avec l'identité du signataire.

Comparaison avec le monde réel...

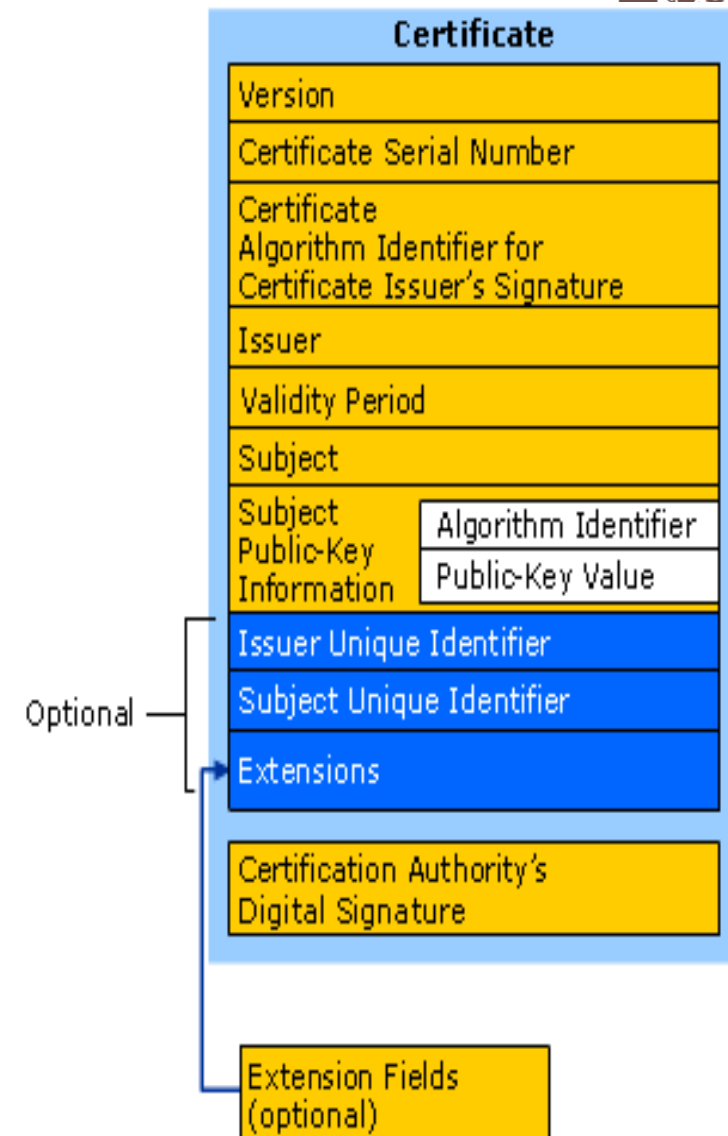


D'un point de vue technique



La norme X509

- X509 : standard ISO pour les infrastructures à clés publiques
 - un format standard de certificats : modèle de référence pour les architectures PKI
- historique :
 - X509 v1 : 1988 (ISO) ; X509 v2 : 1993 (ISO)
 - X509 v3 (extensions possibles) : 1996 (ISO - ANSI - ITU)
 - CRL v2 (Certificate Revocation List)
- Certificat X509 :
 - certifié par une entité de confiance,
 - crée un lien entre l'identité de l'entité et sa clé publique,
 - garantit l'identité du propriétaire à un tiers et l'authenticité de sa clé publique.



Infrastructure à clés publiques (PKI)

- Rappel de la définition d'une infrastructure à clés publiques:
 - L'ensembles des moyens matériels, logiciels et organisationnels permettant la gestion des clés et des certificats
- Le certificat est la pièce d'identité électronique du client, partenaire, fournisseur.
 - Il permet d'authentifier son porteur qui est la condition nécessaire pour accéder à un service.
 - Cette technologie apporte une sécurité fiable et à caractère légal.
 - Elle est disponible sous forme synchrone (SSTL/TLS) et asynchrone (S/MIME).
- Plus concrètement un système permettant de:
 - Délivrer, révoquer, publier renouveler, les certificats de leur porteurs
 - Générer ou séquestrer et remettre les clés numériques à leurs porteurs
 - Générer et publier la liste des certificats révoqués
- Mais:
 - Ce n'est qu'un moyen. Le but est le développement de services sur cette technologie.
 - Ce n'est pas que de la technique. Il est important qu'une concertation voire une coordination au niveau groupe soit mise en place

Les infrastructures à clé publique

- L'algorithmie asymétrique permet :
 - Simplification de la gestion des secrets
 - Procédés efficaces de distribution de clé
 - Mais surtout : une propriété d'ouverture...

- Les PKI : un système ouvert
 - Pour réaliser une transaction en toute confiance entre Alice et Bob, il faut et il suffit :
 - qu'Alice ait confiance dans l'autorité de certification de Bob et
 - que Bob ait confiance dans l'autorité de certification d'Alice.
 - Il n'est pas nécessaire :
 - qu'Alice connaisse Bob ou
 - qu'ils aient la même autorité de certification

- La confiance sur Internet :
 - Les PKI permettent d'établir une relation de confiance entre deux personnes qui ne se connaissent pas.
 - C'est l'outil indispensable au commerce électronique, aux téléprocédures, à toutes les transactions privées, commerciales ou professionnelles.

Fonctions liées à la certification

- *Enregistrement*

- Enregistrement des utilisateurs et arrivée d'un nouvel utilisateur
- Vérification des identités des demandeurs de certificats
- Génération de certificats et remise du certificat à l'utilisateur

- *Gestion des clés*

- Génération et distribution des clés
- Séquestre et recouvrement des clés

- *Gestion des certificats au quotidien*

- Renouvellement du certificat à fin date validité
- Gestion des mutations, mobilités, départs
- Gestion des oubli de code porteur, détérioration/perte/vol de cartes...
- Attribution de carte invité, délégation de signature

- *Révocation*

- Révocation de certificats
- Distribution ou émission périodique de la CRL

- *Publication*

- Publication des certificats de la communauté
- Publication de la dernière CRL à jour

➔ *Nécessité de définir de nouveaux acteurs, rôles et responsabilités*

➔ *Des autorités de confiance (AC, OC, AE) doivent intervenir dans chacune des fonctions de certification*

Architecture technique d'une PKI

- Un serveur de certificats hébergé par l'OC
 - Serveur dans une enceinte sécurisée détenant la clé de l'autorité de certification
 - Assure la production des pièces signées : le certificats et les listes de révocation
 - Maintient la référence de tous les certificats produits, leur état actuel et leur historique

- Un serveur d'enregistrement utilisé par l'AE
 - Serveur offrant l'interface destinée aux opérateurs
 - Assure l'enregistrement des demandes de certificats et leur acquittement

- Un annuaire administré pour le compte de l'AC
 - Référentiel des informations des utilisateurs
 - Assure la publication des certificats et des listes de révocation

- La sous-traitance des AC, AE, OC au sein d'entreprises externes est à considérer avec attention car une entreprise doit :
 - rester maître des informations relatives à ses employés,
 - ne doit en aucune façon déléguer la gestion des clés secrètes de ses employés à une entité extérieure,
 - peut par contre envisager de déléguer le rôle d'opérateur de certification à une entreprise extérieure.

- La délégation des rôles pose des problèmes de *responsabilités juridiques* et de *divulgation* d'information très sensibles, et est à étudier très précisément avant toute délégation de responsabilité ou sous-traitance.

Chiffrement homomorphe

Problématique: la manipulation des données chiffrées

- En environnement Cloud et Big Data, on stocke de gros volumes de données sur des serveurs distants, non maîtrisés.
 - Ces données peuvent être interceptées lors de leur transfert ou stockage
 - Les prestataires de services ont aussi accès à ces informations

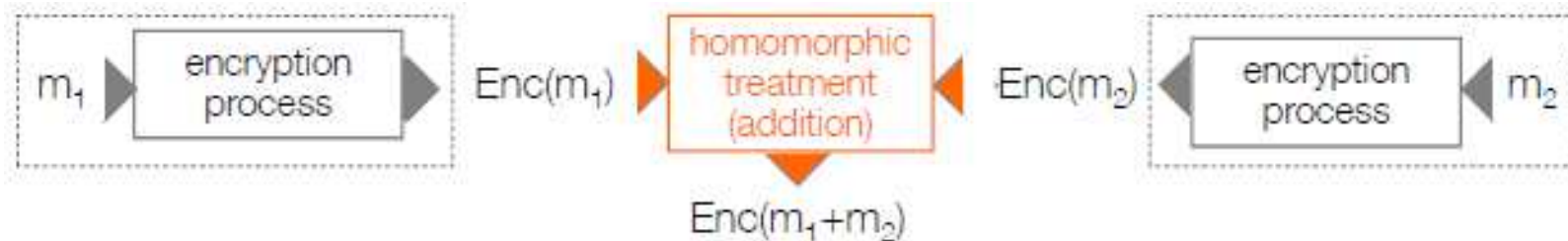
- .Solution** : chiffrer les données avant leur transfert



- Inconvénient**: on ne peut plus manipuler ces données à distance (retoucher ses photos, chercher des mots dans un texte, effectuer des calculs...).
- Une nouvelle piste : la cryptographie homomorphe** : fournit la sécurité des données en permettant que les données chiffrées restent manipulables par les personnes autorisées.

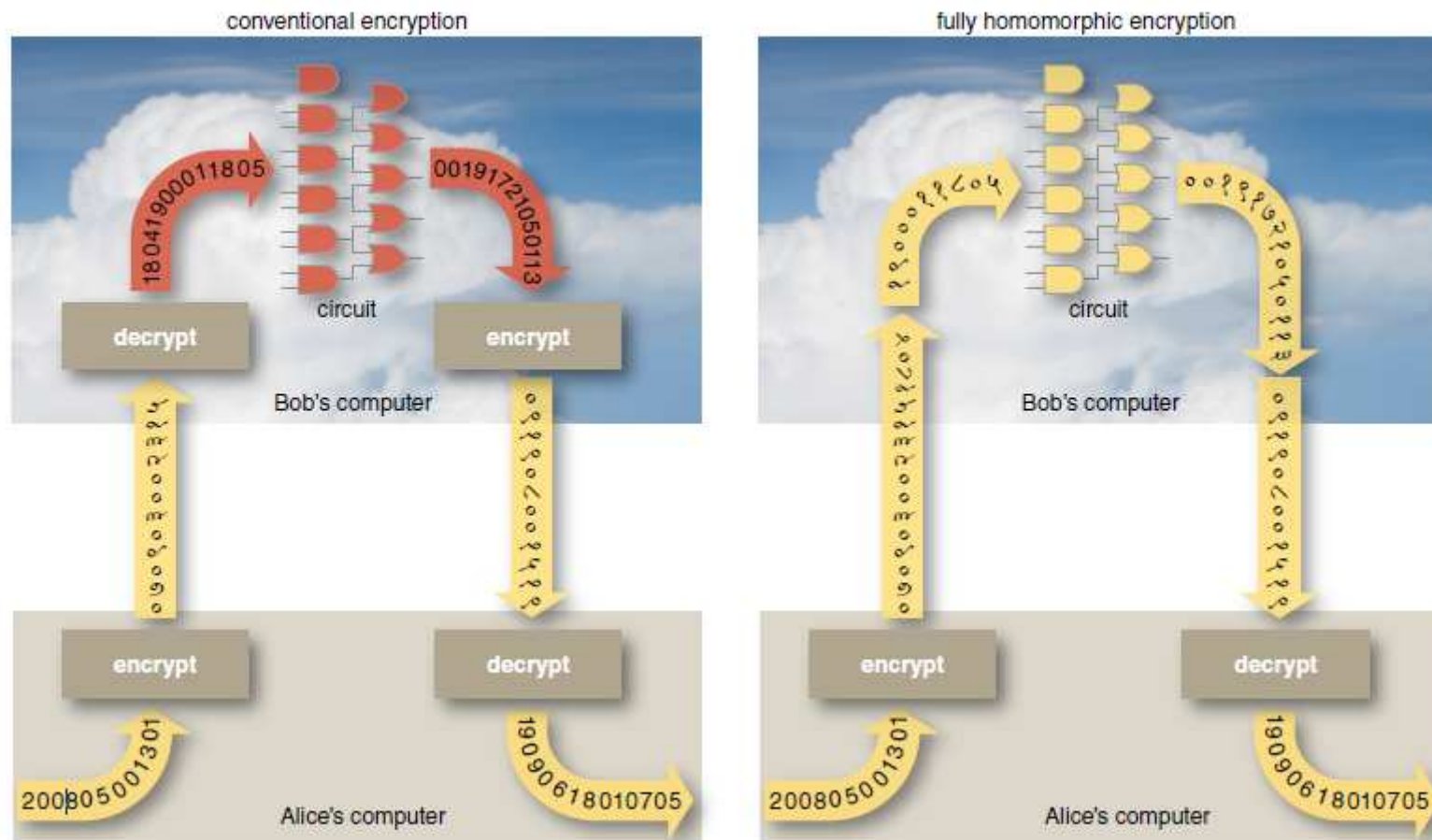
Chiffrement homomorphe : définition

- Un système de chiffrement homomorphe permet d'effectuer divers traitements sur un texte chiffré sans avoir à le déchiffrer.

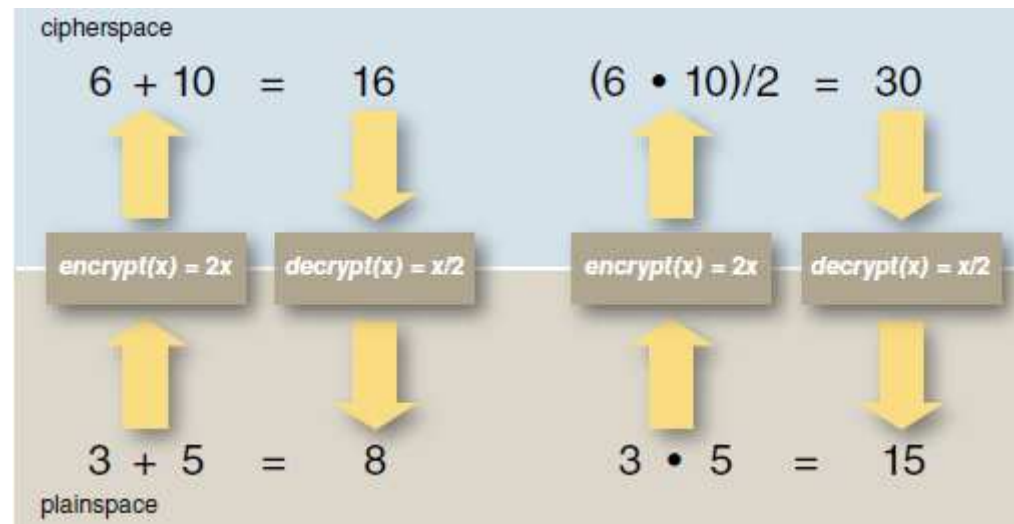


- Fondamental pour le traitement de données en environnement Cloud Computing, Big Data, Internet des Objets.
- Domaines d'applications :
 - Recherche par mot-clé, concaténation
 - Statistiques et datamining sur données chiffrées
 - Vérifier si deux fichiers chiffrés sont identiques
 - Détection de logiciel malveillant au sein de trafic chiffré
 - Fournir des publicités ciblées sans rien connaître du destinataire
 - Faire des recherches sur Internet et recevoir les réponses sans que le moteur de recherche ne sache quel était l'objet de notre requête

Chiffrement homomorphe



Une science encore immature mais prometteuse



exemple

- Les algorithmes connus sont essentiellement partiellement homomorphes
- Les algorithmes génèrent une inflation des données
- Lenteur des calculs (voir exemple)
- Adaptabilité à des cas concrets
- ...mais beaucoup de progrès récents prometteurs...

Bibliographie

- « Merveilleux nombres premiers » JP Delahaye (Belin)
- Article RSA : <http://citeseer.nj.nec.com/rivest78method.html>
- Cours de cryptographie : G. Zemor (Cassini)
- Cryptographie appliquée B. Schneier (Vuibert)
- ANSSI : <http://www.ssi.gouv.fr/>
- CNIL : <http://www.cnil.fr/>
- AFNOR : Livre Blanc – Données massives/ Big Data
- ENISA : Big Data Threat Landscape and Good Practice Guide
- ENISA : State of the Art Analysis of Data Protection in Big Data Architectures (Privacy by Design in Big Data)