CHI HA INVENTATO IL DOCUMENTO INFORMATICO?

Corrado Giustozzi

QUALITÀ, SICUREZZA, INTEGRITÀ ED IMMODIFICABILITÀ DEL DOCUMENTO INFORMATICO

CORRADO GIUSTOZZI

PSG, ENISA

I temi che tratteremo

Corrado Giustozzi

- Il documento digitale
- Requisiti e meccanismi di garanzia
- Firma digitale
- Catene di fiducia: i certificati digitali
- Problemi interessanti e paradossali

Corrado Giustozzi

IL DOCUMENTO DIGITALE

IL "MATTONE FONDAMENTALE" DELLA SOCIETÀ SENZA CARTA

Evoluzione del "documento"

Corrado Giustozzi

- Il documento "classico":
  - è un oggetto *materiale* che coincide col suo *supporto*
  - è unico e originale, si distingue dalle copie
  - richiede una modifica fisica per la validazione
- Il documento "moderno" (digitale):
  - è un oggetto *immateriale* (contenuto informativo) separato ed indipendente dal particolare supporto che lo ospita
  - ogni copia è un originale, anche su altro supporto
  - non ammette modifiche fisiche
- L'informazione digitale può facilmente essere:
  - intercettata, copiata, trasportata, spostata, diffusa
  - modificata, contraffatta, falsificata, alterata
  - distrutta

17 novembre 2011 Collegio Ghislieri 4

Le garanzie necessarie

Corrado Giustozzi

- Da sempre l'uomo ha chiesto ai documenti alcune importanti certezze:
  - **Autenticità**: certezza che il documento sia stato fatto da chi sostiene di esserne l'autore
  - **Integrità**: certezza che il documento non abbia subito modifiche all'insaputa delle parti interessate
  - **Non ripudio**: certezza che il reale autore non possa disconoscere la paternità del documento
  - **Confidenzialità**: certezza che chi non è autorizzato non possa prendere conoscenza del documento
- Per ottenere queste certezze si è sempre fatto ricorso a modifiche fisiche al documento:
  - firme, sigilli, timbri, punzoni, filigrane, ologrammi, …
- …ma il documento moderno è *immateriale*!

17 novembre 2011 Collegio Ghislieri 5

I problemi dei documenti

Corrado Giustozzi

- Per il documento classico (materiale):
  - contraffazione, intercettazione, falsificazione…
  - difficoltà di riconoscere un falso
  - difficoltà di attribuire la paternità
  - difficoltà nella datazione
- Per il documento moderno (immateriale):
  - facilità di copia e trasporto
  - facilità di modifica
  - mancanza di un supporto fisico
  - mancanza del concetto di originale

17 novembre 2011 Collegio Ghislieri 6

Nuove forme di garanzia

Corrado Giustozzi

- Per fornire ai documenti digitali garanzie analoghe a quelle che valgono per i documenti tradizionali si usano tecniche di *validazione* ottenute come effetto collaterale delle moderne tecniche di *protezione* delle informazioni
- La *crittografia a chiave pubblica*, nata per proteggere le comunicazioni di massa, consente anche di attribuire *certezze* ad un documento digitale
- Non si valida il *supporto* del documento bensì il suo *contenuto informativo*
- Nasce così la cosiddetta *firma digitale*

17 novembre 2011 Collegio Ghislieri 7

Validazione dei documenti

Corrado Giustozzi

- La "firma digitale":
  - è verificabile da chiunque
  - non è falsificabile, non è ripudiabile
- Non è una "firma" in quanto:
  - è il risultato di un *calcolo* sul "documento"
  - è separata dal documento e non lo modifica
  - rivela modifiche al testo originale
- "Digitale" significa "numerico":
  - le dita e le impronte digitali non c'entrano (quasi) nulla!
- Prerequisiti:
  - un sistema di *crittografia a chiave pubblica*
  - una *funzione hash* standard

17 novembre 2011 Collegio Ghislieri 8

La crittografia a chiave pubblica - 1

Corrado Giustozzi

- È un sistema di codifica basato su una *coppia* di "chiavi" e su un procedimento di calcolo (cifratura) che fa uso dell'una o dell'altra chiave
- Il sistema è tale che:
  - conoscendo una chiave non si può ricavare l'altra
  - un messaggio cifrato con una chiave si può decifrare solo con l'altra, e viceversa
- In un sistema del genere:
  - una delle due chiavi (K_p) viene resa *pubblica*
  - l'altra (K_s) rimane *segreta* ossia è nota al solo proprietario
- Il meccanismo sottostante è puramente matematico:
  - le chiavi sono numeri primi molto grandi (>200 cifre)
  - la codifica implica calcoli molto lunghi e complessi

17 novembre 2011 Collegio Ghislieri 9

La crittografia a chiave pubblica - 2

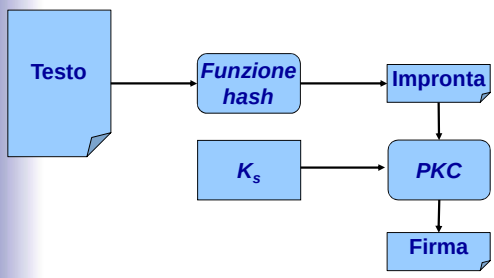
Corrado Giustozzi

- Il sistema è fortemente asimmetrico:
  - chiunque può cifrare un testo con la chiave pubblica A_p di un soggetto A appartenente al sistema
  - tuttavia solo A può decifrare un messaggio cifrato con la sua chiave pubblica A_p, perché egli solo è in possesso della corrispondente chiave inversa A_s (la sua chiave segreta)
- Vale anche il viceversa:
  - chiunque può decifrare un testo cifrato da A con la propria chiave segreta A_s perché la chiave inversa corrispondente è la A_p ovvero la chiave pubblica di A
- Vige il principio fondamentale del *non ripudio*:
  - se nessuno conosce la chiave segreta A_s di A, all'infuori di A stesso, allora ogni testo cifrato con A_s è *necessariamente* stato prodotto da A, e chiunque può verificarlo facilmente

17 novembre 2011 Collegio Ghislieri 10

Firma digitale: creazione

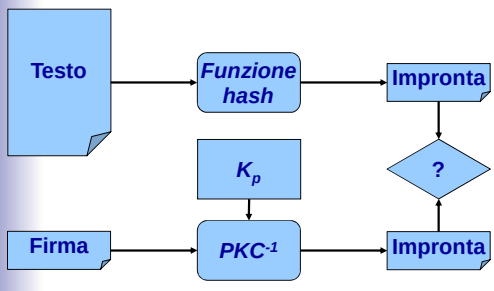
Corrado Giustozzi



17 novembre 2011 Collegio Ghislieri 11

Firma digitale: verifica

Corrado Giustozzi



17 novembre 2011 Collegio Ghislieri 12

Perché la firma digitale funziona?

Corrado Giustozzi

- Il principio su cui si può dare validità "forte" (anche legale) ad un documento sottoscritto con firma digitale si basa su una concatenazione di fatti tecnici inequivocabili e di assunzioni ragionevoli
- Fatti:
  - l'impronta è un riferimento "forte" al documento D originale
  - l'impronta è stata cifrata con la chiave segreta A_s di A
- Assunzioni:
  - è impossibile che l'impronta si riferisca ad un altro documento D'
  - nessuno oltre A conosce e può usare la chiave segreta A_s
- Conclusioni:
  - solo A può aver generato quella firma
  - essa vale solo per il documento D cui si riferisce con certezza

17 novembre 2011 Collegio Ghislieri 13

Proprietà della firma digitale

Corrado Giustozzi

- È separata dal documento cui si riferisce
- Non modifica il documento firmato
- Dipende dal contenuto del documento:
  - non può essere falsificata, imitata, ripetuta
  - non può essere apposta "in bianco"
- È "autentica" anche in copia
- È verificabile da chiunque
- Rivela modifiche al documento successive alla firma
- Non può essere disconosciuta se non per motivi non riferibili alla procedura matematica che l'ha generata

17 novembre 2011 Collegio Ghislieri 14

Corrado Giustozzi

IL CERTIFICATO DIGITALE

IL MECCANISMO DI GARANZIA DELLE FIRME DIGITALI

17 novembre 2011 Collegio Ghislieri 15

L'anello debole della firma digitale

Corrado Giustozzi

- Affinché il sistema funzioni occorre chiarire:
  - chi e come gestisce l'elenco delle chiavi pubbliche
  - chi e come garantisce la validità dell'elenco
  - chi e come garantisce sulla effettiva corrispondenza fra identità dei soggetti e relative chiavi pubbliche
- Queste certezze fondamentali vengono fornite da un sistema cosiddetto di "certificazione"
- La certificazione si attua mediante:
  - entità garanti denominate *autorità di certificazione*
  - strumenti tecnologici denominati *certificati digitali*

17 novembre 2011 Collegio Ghislieri 16

Il processo di certificazione

Corrado Giustozzi

- L'Autorità di Certificazione è un soggetto *super partes*, affidabile per definizione, il quale:
  - attesta la validità di una chiave
  - garantisce l'identità del titolare
  - gestisce l'elenco delle chiavi pubbliche
- Il Certificato Digitale da essa emesso contiene:
  - la chiave pubblica del titolare "firmata" dalla CA
  - ulteriori dati di servizio:
    - scadenza
    - limitazioni
    - …

17 novembre 2011 Collegio Ghislieri 17

Modelli di certificazione

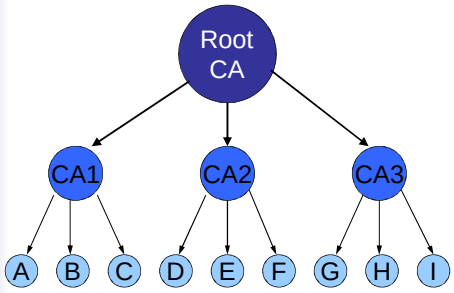
Corrado Giustozzi

- ISO X.509:
  - standard *de iure* basato sulle Certification Authorities
  - struttura gerarchica organizzata formalmente
  - ogni CA certifica quelle al di sotto di lei
  - l'unico valido a norma di legge
- Web of Trust:
  - standard *de facto* affermatosi con PGP
  - modello cooperativo senza struttura formale
  - ogni utente certifica agli altri coloro di cui è certo
  - accettato informalmente ma privo di validità legale

17 novembre 2011 Collegio Ghislieri 18

Certificazione secondo X.509

Corrado Giustozzi



Corrado Giustozzi

CONTENUTO E CONTENITORE

SIAMO PROPRIO SICURI CHE UN DOCUMENTO INFORMATICO SIA DAVVERO COME APPARE?

Dati e metadati

Corrado Giustozzi

- Metadato: un dato che descrive un altro dato
- In passato i file contenevano solo dati, ossia contenuto informativo “puro”; i (pochi) metadati appartenevano al sistema operativo che li usava per gestire i file (posizione, dimensione, date di modifica, ecc.)
- Oggi si tende a organizzare i documenti informatici in strutture complesse, che accanto ai dati conservano molti metadati contenenti informazioni supplementari quali: autore, revisione, statistiche, informazioni tecniche, storico delle modifiche, tag di indicizzazione, informazioni di georeferenziazione, ecc. ecc.
- Molti metadati sono gestibili dall’utente, ma altri sono generati automaticamente e a volte “nascosti”

Dati costanti e campi variabili

Corrado Giustozzi

- Quasi tutti i documenti moderni non sono quindi formati da dati statici ed immutabili ma possono contenere dati *variabili* in funzione del tempo o del contesto
- La variazione del contenuto di un campo può avvenire in modo sia automatico che addirittura programmabile:
  - un campo "data" che viene attualizzato con la data attuale
  - un campo "utente" che viene attualizzato con l'utente attuale
  - un contenuto inserito dinamicamente da un altro documento
  - una stringa di testo che cambia in funzione della data
  - una macro che attualizza un campo col risultato di un calcolo
- In pratica i documenti moderni sono dei *metadocumenti* che descrivono istanze multiple di documenti

17 novembre 2011 Collegio Ghislieri 22

Esempio: i documenti Word

Corrado Giustozzi

- Microsoft Word raccoglie in ogni documento una quantità impressionante di metadati relativi non solo al documento in sé ma anche all'autore, ai revisori, e perfino agli ambienti operativi nei quali il documento è stato elaborato o modificato
  - ad esempio: nomi dei computer, nomi delle directory locali, nomi e posizioni delle stampanti, indirizzo della scheda di rete!
- La funzione "track changes" mantiene inoltre lo stato di tutte le precedenti revisioni del documento
- Questi metadati:
  - sono controllabili e/o eliminabili dall'utente solo con le versioni più recenti di Office o mediante appositi tool
  - molti di essi possono rendere dinamico il contenuto del documento, vanificando il processo di firma digitale

17 novembre 2011 Collegio Ghislieri 23

Dov'è il vero documento?

Corrado Giustozzi

- I campi variabili fa sorgere un interessante problema di natura *semantica* prima ancora che tecnica: cosa significa realmente *modificare* un documento digitale?
- Sono possibili due casi opposti, entrambi paradossali:
  - il documento *non cambia* ma il suo *contenuto informativo* sì
  - il documento *cambia* ma il suo *contenuto informativo* no
- Caso 1: un campo data dinamico, che viene attualizzato al valore corrente, fa apparire il contenuto sempre diverso anche se il file non muta
- Caso 2: stampare un documento non ne muta il contenuto, tuttavia modifica il file in quanto aggiorna il campo interno *"data di ultima stampa"*
- Attenzione quindi a *cosa si sta firmando*!

17 novembre 2011 Collegio Ghislieri 24

Corrado Giustozzi

CONSERVAZIONE

SIAMO PROPRIO SICURI CHE UN DOCUMENTO INFORMATICO SIA ETERNO ED IMMUTABILE?

17 novembre 2011 Collegio Ghislieri 25

La leggibilità dei dati

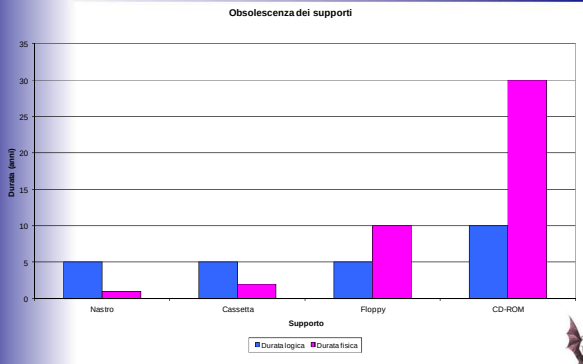
Corrado Giustozzi

- Il documento informatico è immateriale, quindi virtualmente incorruttibile ed eterno... o no?
- Purtroppo, al contrario di quello cartaceo, esso non può essere interpretato direttamente ma va *decodificato*: ossia il suo contenuto informativo va *ricostruito* tramite opportune tecnologie a partire dalle codifiche fisiche apportate al supporto materiale che lo ospita
- Questo doppio passaggio crea nuovi problemi:
  - i supporti fisici utilizzati sono delicati e poco duraturi
  - le tecniche di codifica mutano nel tempo
  - le tecnologie di lettura/scrittura diventano obsolete
- I documenti informatici non sono fatti per durare!

17 novembre 2011 Collegio Ghislieri 26

Quanto durano i nostri archivi?

Corrado Giustozzi



17 novembre 2011 Collegio Ghislieri 27

Trapianto di archivi

Corrado Giustozzi

- Non si può pensare che un supporto informatico risulti leggibile e intelligibile dopo decine di anni:
  - la tecnologia sarà cambiata
  - il supporto non sarà più affidabile
  - non ci saranno più i dispositivi idonei a interpretare i dati
- Provate oggi (2011) a:
  - leggere un floppy 5¼" del 1981
  - utilizzare un documento di WordStar o VisiCalc del 1981
- Se occorre mantenere "vivo" un archivio storico di dati digitali occorre *riversarlo*, ossia periodicamente trasferirne i contenuti su un supporto più moderno
- Per farlo però servono hardware e software appositi, ma soprattutto… tanto tempo!

17 novembre 2011 Collegio Ghislieri 28

Un mondo senza carta?

Corrado Giustozzi

- Il passaggio alla società senza carta è possibile ma:
  - sarà lungo
  - sarà complicato
  - non sarà mai totale
- Il documento digitale è:
  - una grande conquista
  - una fonte di efficienza e snellimento dei processi
  - una semplificazione in molti casi
- ma:
  - non è privo di scomodi effetti collaterali
  - non elimina la necessità di una sua gestione nel tempo
  - sarà probabilmente inutilizzabile dai futuri storici!

17 novembre 2011 Collegio Ghislieri 29

Chi ha inventato il documento informatico?

Corrado Giustozzi

Grazie dell'attenzione

Qualità, sicurezza, integrità ed immodificabilità del documento informatico

17 novembre 2011 Collegio Ghislieri 30