

Come lacrime nella pioggia

“Verba volant, data quoque”
- *Traduttore automatico di Google*

Mario Pascucci
Responsabile Tecnico Archivio Digitale Radiofonico

Due categorie

- 🧱 Quelli che hanno perso i propri dati
- 🧱 Quelli che li perderanno

Tertium non datur

Il peggior nemico

- ❏ Ogni realizzazione umana ha un nemico, il peggiore di tutti i tempi
- ❏ Chimica e fisica ci insegnano come ogni cosa tenda a trasformarsi, inevitabilmente e inarrestabilmente
- ❏ Quello che sopravvivrebbe al tempo, viene distrutto dall'uomo e dalle sue “regole”

L'esigenza di registrare

- ❏ La memoria umana, pur prodigiosa, ha dei difetti insiti nel suo funzionamento
- ❏ Le Scienze Cognitive dimostrano che i nostri ricordi sono frutto di ricostruzioni e di elaborazioni
- ❏ E' comunque una registrazione soggetta a deterioramento

Pietre, pelli e canne di bambù

- ❏ Esistono registrazioni fatte dall'Uomo fin da quando è su questo pianeta
- ❏ Dipingere, scolpire, incidere sono forme di registrazione
- ❏ La scrittura è stato il primo metodo di registrazione formalizzato su larga scala

Poi arriva l'elettronica

- ❏ In un certo senso si ritorna all'età della pietra: incisioni su una superficie per registrare i suoni (tamburi e dischi di cera o di lacca, poi arriva il vinile)
- ❏ La registrazione magnetica inizia con i registratori a filo, poi si passa al nastro con supporto in acetato.

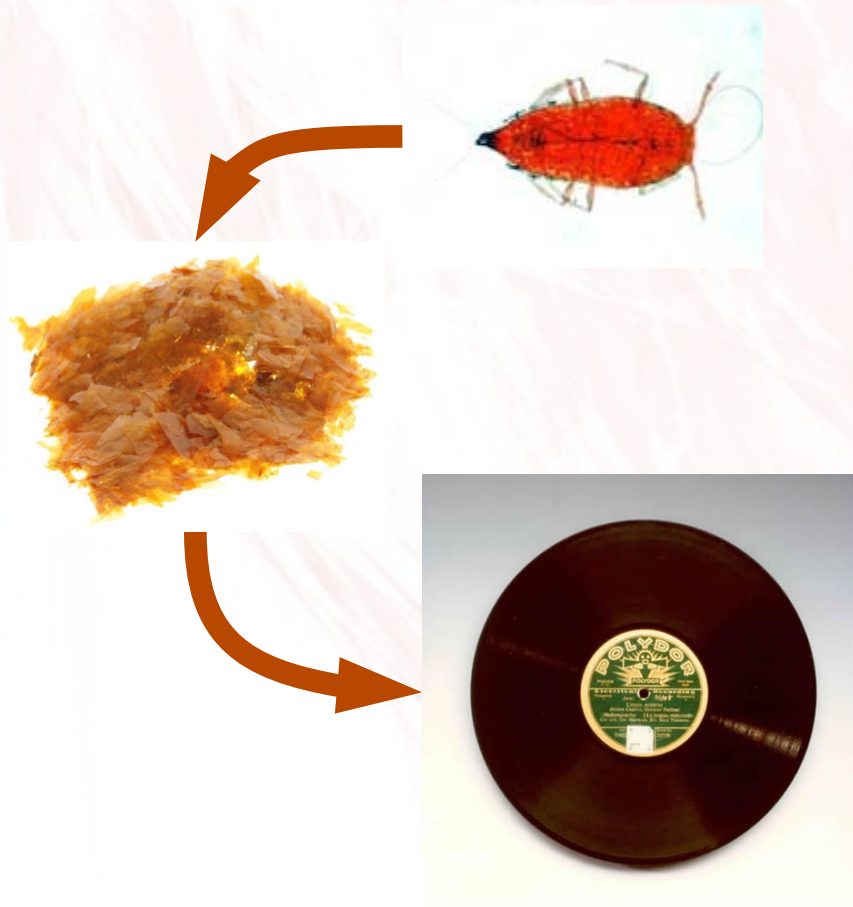
Differenze fondamentali

- ❏ Libri, manoscritti, dipinti, sculture, fotografie possono essere “letti” direttamente, senza intermediari, basta osservarli.
- ❏ La grande differenza con le registrazioni elettromeccaniche, elettroniche o digitali è che serve un intermediario: il **riproduttore**

Prima la meccanica

- ❏ I dischi, e prima i tamburi, erano fatti con i materiali più disparati: cera, lacca, alluminio, *vulcanite*, cartone, bachelite, vinile
- ❏ “Padelloni”, “Lacche”, “Vinili” sono termini comuni fra gli appassionati
- ❏ Il vinile è vivo e vegeto (mercato raddoppiato nel 2013, fonte BBC/NYTimes)

Da dove viene la lacca?



- ❑ E' la secrezione di una varietà di cocciniglia (*Laccifer Lacca*)
- ❑ E' un materiale usato anche in campo alimentare (additivo E904)

Registratori a filo



Immagini prese da <http://www.fracassi.net/>

Il nastro magnetico

- ❏ Utile per registrare praticamente tutto: audio, video, dati
- ❏ Molteplici incarnazioni, numerosi formati meccanici e elettronici
- ❏ Diversi materiali: supporto, substrato, materiale magnetico, leganti, lubrificanti

Dischi magnetici

- ❑ Gli hard disk sono comunemente utilizzati nei nostri computer
- ❑ Alta densità e velocità le caratteristiche migliori
- ❑ Continuo incremento di capacità e velocità

Luce sia

- ❏ Usare la luce come strumento di lettura ha molti vantaggi, in particolare densità e assenza di degrado in lettura, in teoria
- ❏ I supporti sono molto più complessi e delicati
- ❏ I formati sono infiniti: CD, CD-R, CD-RW, DVD, DVD±R, DVD±RW, MO, MiniDisc, BluRay, HD-DVD, OD, HVD, ecc.
- ❏ Quello che la luce scrive, la luce cancella
- ❏ I materiali per i *-R sono organici

Solido, lo stato

- ❑ Gli ultimi anni hanno visto un vero e proprio boom delle memorie a stato solido
- ❑ Niente parti in movimento
- ❑ Velocità, densità e affidabilità crescono in accordo



Ma quanto durano?

Tipo	Supporto	Durata
Pitture rupestri	Pietra, pigmenti naturali	Decine di migliaia di anni
Sculture	Pietra, metallo	Millenni
Manoscritti	Carta, pergamena, papiro	Secoli
Pitture	Tela, pigmenti naturali	Secoli
Stampa	Carta, pigmenti	Secoli
Fotografia (b/n)	Carta, pellicola	100-200 anni
Filmati	Pellicola	Dati insufficienti

Dati insufficienti

Tipo	Supporto	Durata
Foto a colori	Carta, emulsioni chimiche	Dopo 50 anni pellicola e foto sviluppate cambiano colore
Dischi e tamburi	Cera, lacca, bachelite	100 anni o poco più
Dischi in vinile	Cloruro di polivinile	100 anni (forse di più)
Nastri magnetici	Acetato	70-80 anni (anni '50 a rischio)
Nastri magnetici	Poliestere	Dipende dal produttore
Dischi ottici stampati	Policarbonato, alluminio	Dipende
Dischi ottici registrati	Policarbonato, alluminio, ftalocianina	Da sei mesi a 10 anni o più
Memorie Flash	Silicio, alluminio, rame	Casi di perdita a tre mesi
Dischi magnetici	Vario	Estremamente vario

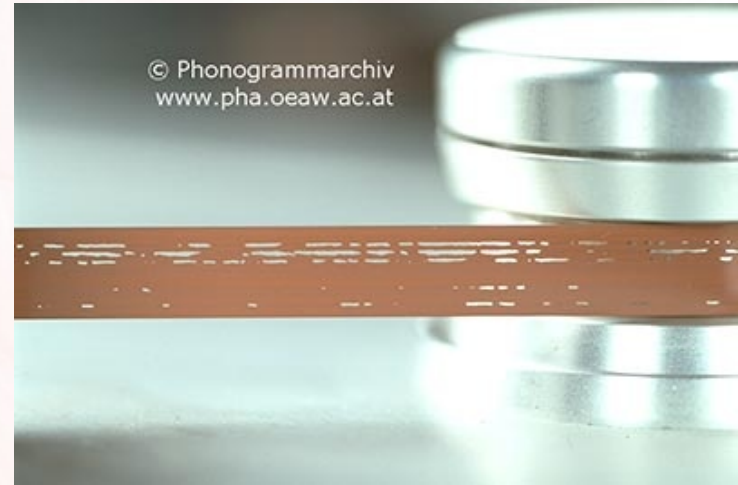
Analogico e digitale

- ❑ Bassa densità
- ❑ Resistenza ai danni meccanici
- ❑ Resistenza a condizioni ambientali sfavorevoli
- ❑ Alta densità
- ❑ Un danno minimo rende inaccessibile l'intero contenuto
- ❑ Altissima sensibilità alle condizioni ambientali

Dischi, nastri e disastri



Dischi, nastri e disastri (2)

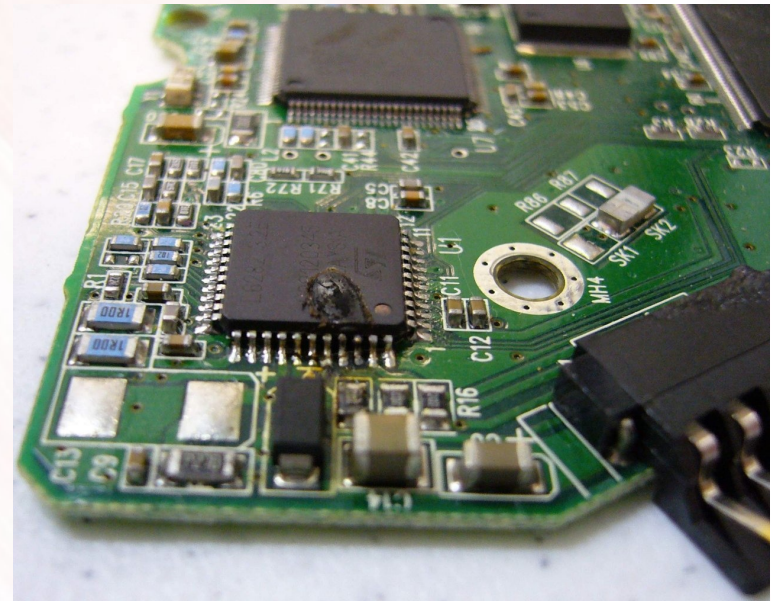


“Hard” disk

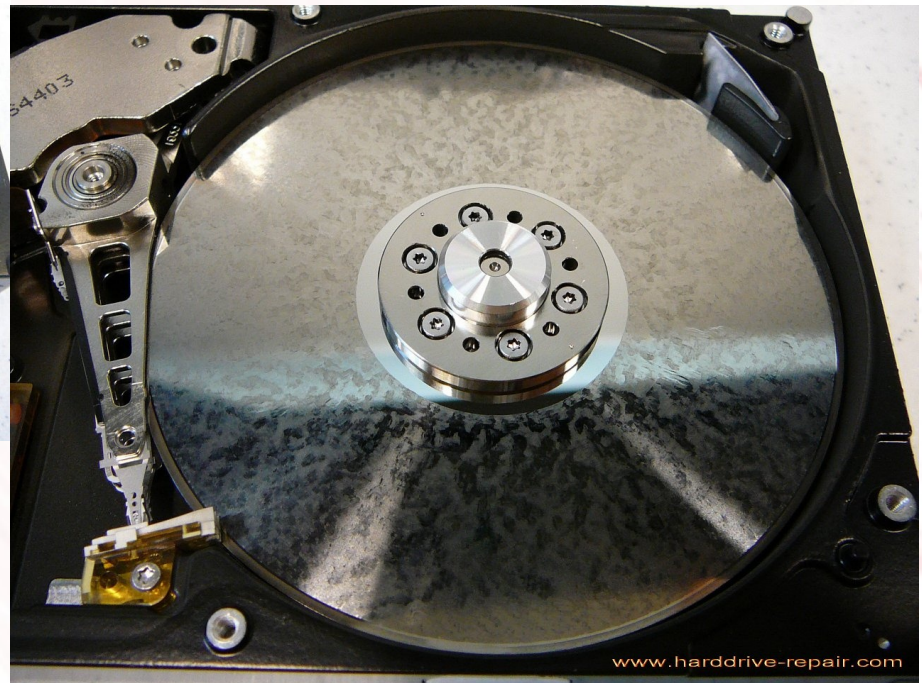


I “piatti” di quelli da 2,5” sono di vetro...

L'elettronica non è perfetta...
e neanche eterna



Troppo “hard”



Immagini prese da:
<http://www.harddrive-repair.com/>

RAID \neq Backup

- ❏ Le probabilità si sommano.
- ❏ In una *stripe* RAID i dischi invecchiano tutti contemporaneamente, anche l'*hot-spare*
- ❏ La ricostruzione di un drive sottopone tutta la *stripe* ad uno stress
- ❏ Alcuni eventi possono guastare più dischi in un solo colpo

Quello che la luce scrive...

...la luce cancella.

I tre CD-R in foto sono piuttosto recenti. Dall'alto:

- 🧱 Nuovo, aperto il 28 aprile 2010
- 🧱 Masterizzato il 26 giugno 2003
- 🧱 Masterizzato il 22 maggio 2007



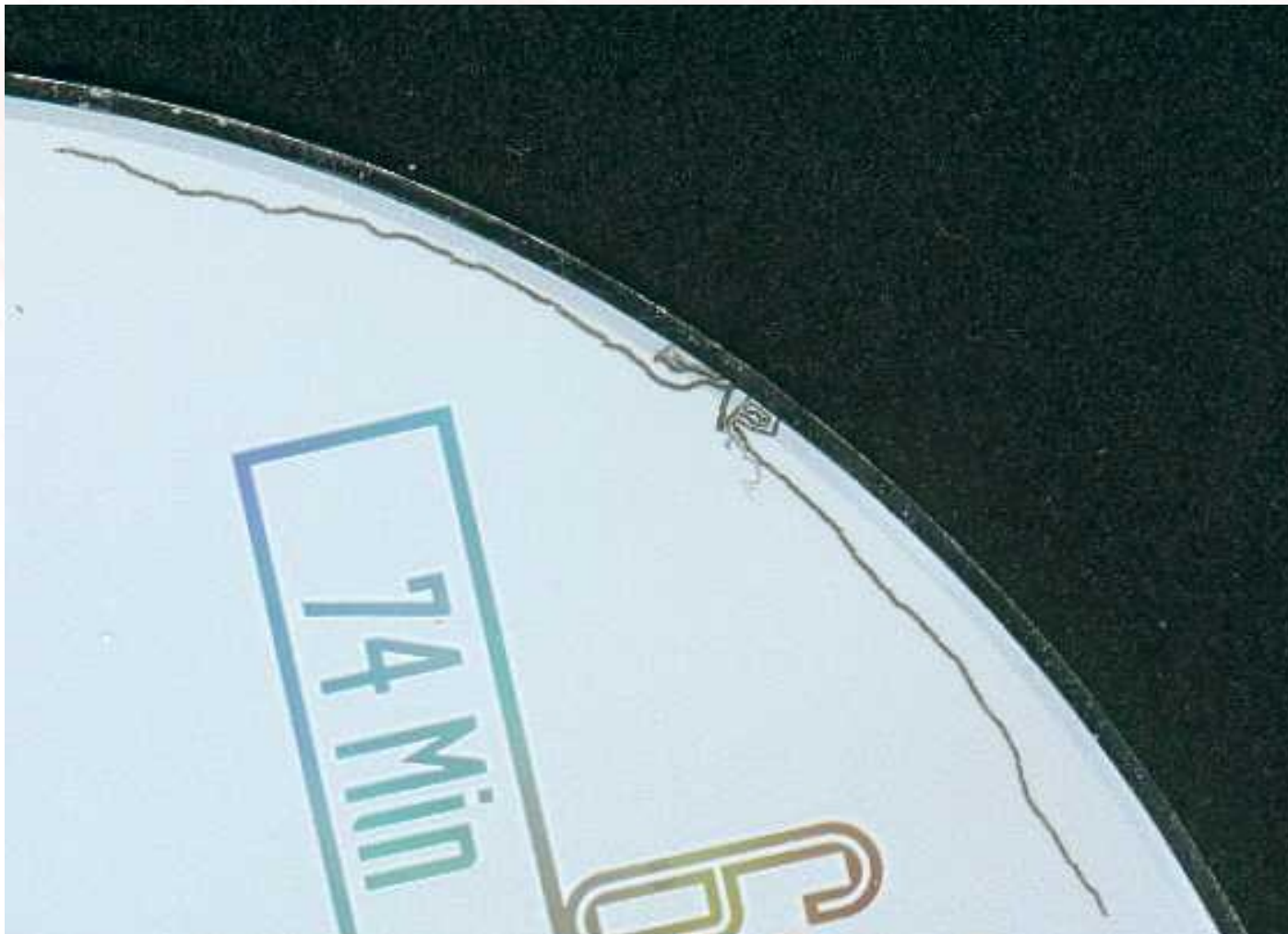
Vi piacciono i formaggi “fioriti”?

- ❏ Nel 2001 si diffuse la notizia di CD illeggibili a causa del deterioramento della superficie riflettente
- ❏ Responsabile: una muffa onnipresente, la stessa sulla crosta commestibile del formaggio Brie e Camembert
- ❏ Il fenomeno si sviluppa in presenza di temperature ambiente sopra i 30° e alti tassi di umidità

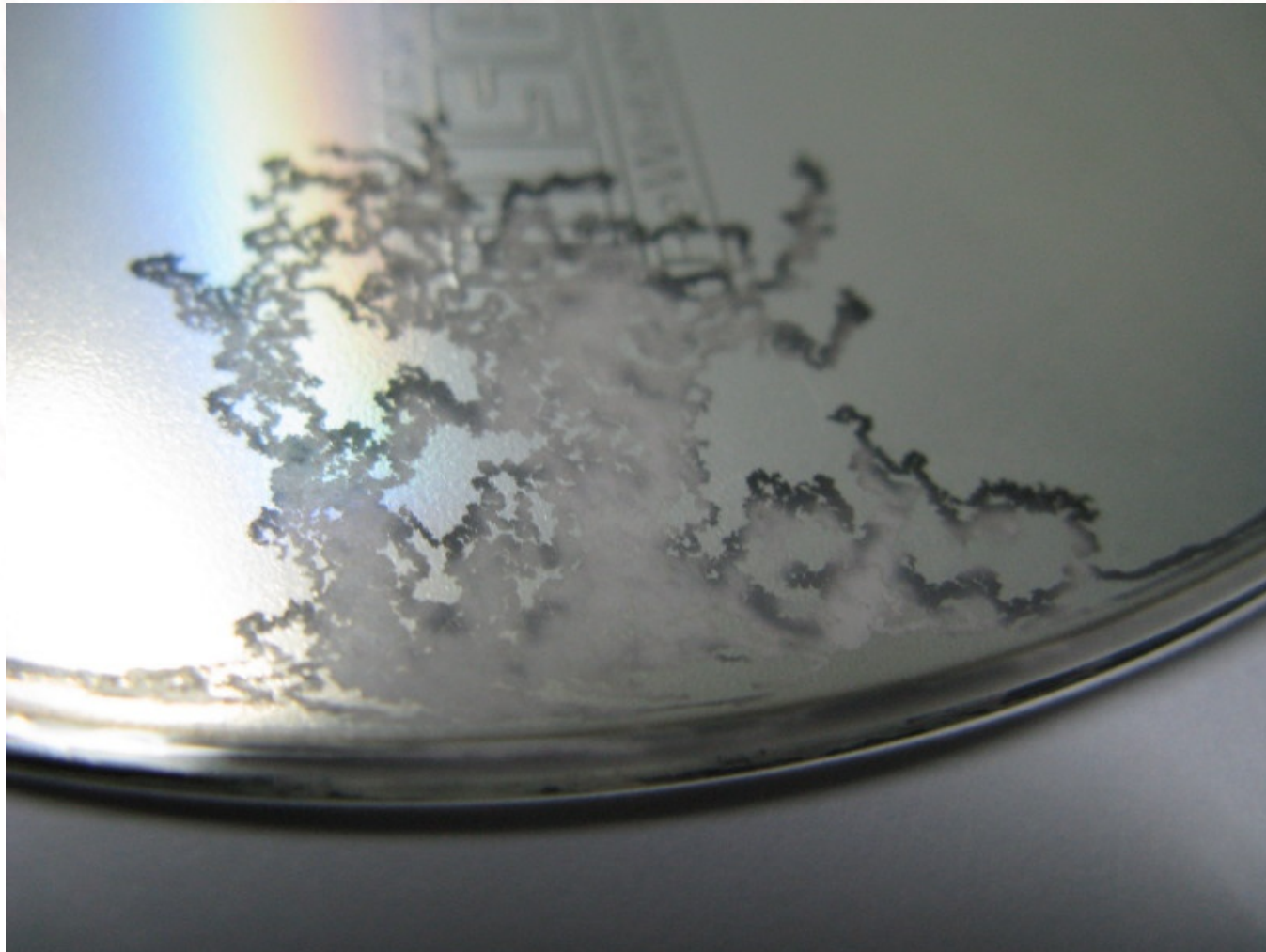
Calore, ...



... umidità, ...



... un po' di tempo...



... e l'infinita forza della vita



Qualità garantita

- 🧱 Nel 2004 dimenticai un CD nel mio lettore
- 🧱 Nel 2011 finalmente ho trovato il tempo e lo spazio per rimettere in funzione il lettore, ed ho ritrovato il CD
- 🧱 Ho pensato subito fosse colpa mia: conservato male, quindi deteriorato
- 🧱 Dopo qualche settimana ne ho trovato un altro, stesso produttore, stessa età, conservato “come si deve”.

Questo è quello “come si deve”



Uno dei tre ha 10 anni di più



I dati si recuperano, a volte...

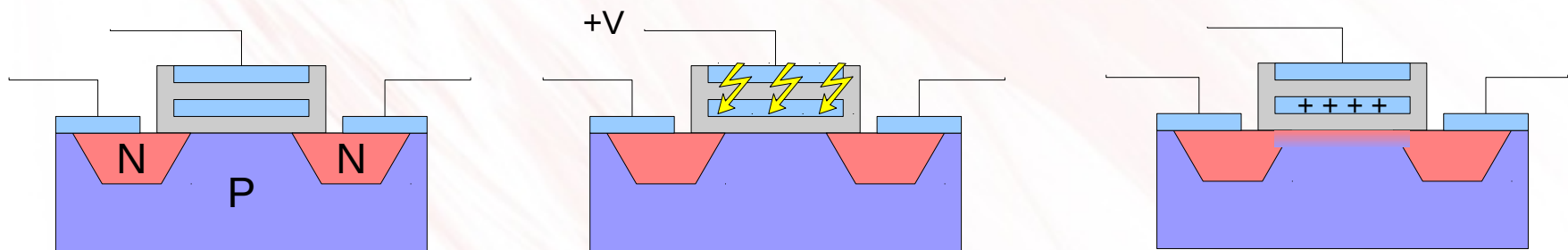
- ❏ Esistono ditte specializzate nel recupero dati, anche da supporti in condizioni disperate
- ❏ I costi sono elevatissimi, tipicamente fra i 5 ed i 10 euro per gigabyte, con un minimo di alcune migliaia di euro per la diagnosi e l'estrazione
- ❏ <http://www.youtube.com/watch?v=IKL6JQXIhaU>

...meglio non perderli

Pos.	Quantità	Unità	Articolo	Prezz.unit.	Sconto	Totale
1	1,0		Recupero dati servizio per partner	6.817,00		6.817,00
2	1,0		Backup disco esterno 1TB	80,00		80,00
3	1,0		Spedizione Italia	20,00		20,00
Totale IVA esclusa						6.917,00
Sconto sul recupero dei dati: 10,00 %						<u>681,70</u>
Subtotale:						6.235,30
+ 22,00% IVA su 6.235,30 EUR						<u>1.371,77</u>
Importo Totale IVA Inclusa EUR						<u><u>7.607,07</u></u>

Flash: dati persi in un lampo

- Le memorie flash usano l'effetto tunnel per scrivere e cancellare dati.
- Il “dato” è un accumulo di cariche elettriche nell'elettrodo fluttuante



La velocità è tutto!

- ❏ Fin dai primi dischi SSD sono state introdotte tecnologie per renderli più veloci
- ❏ Dato che ogni scrittura implica una cancellazione del vecchio dato, sono state introdotte varie tecniche per anticipare la cancellazione, usando tempi morti
- ❏ Due in particolare sono le strategie: una a bordo del disco stesso, una a carico del sistema operativo

Cancellato-cancellato, proprio

- ❏ In molti dischi SSD il firmware provvede in autonomia ad azzerare i blocchi assegnati ad un file marcato come cancellato
- ❏ L'operazione si avvia nei minuti successivi alla marcatura di “cancellazione”


Tre minuti per recuperare un file cancellato per errore


(in sottofondo, la colonna sonora di “Mission Impossible”)


Cancella tu che cancello io

- ❑ Nelle specifiche ATA attualmente rilasciate viene introdotto uno specifico comando, detto “TRIM”, per ordinare ad un disco SSD di cancellare una lista di blocchi
- ❑ La lista è decisa dal sistema operativo
- ❑ Il disco non prende iniziative
- ❑ Il resto è facile da immaginare...

Quanto è grande un secondo?

 La superficie necessaria per registrare un secondo di audio è sempre più piccola

 Il risultato è che per lo stesso danno di piccola entità, si perdono sempre più dati

 Per i supporti digitali la perdita non è mai proporzionale al danno

Supporto	Secondi x cm ²
Nastro analogico	0,02÷1
Vinile (33 giri)	2,5÷6
LTO-3	32
CD	42
DVD-5	285
BluRay	1900

L'elettronica invecchia

❏ Condensatori, resistenze, semiconduttori soffrono il passare del tempo

❏ I condensatori elettrolitici vengono dati per esauriti dopo quindici anni, indipendentemente da quanto siano effettivamente utilizzati

❏ Silicio, rame e alluminio usati nei chip non sono immobili, anzi:

<http://en.wikipedia.org/wiki/Electromigration>

Consumer Vs Professional

- ❑ Basso costo
- ❑ Facile da maneggiare
- ❑ Capacità elevata, ma non troppo
- ❑ Controlli a campione in produzione
- ❑ ...?
- ❑ Costi elevati
- ❑ Procedure specifiche
- ❑ Capacità massima possibile
- ❑ Controlli puntuali sulla linea di produzione
- ❑ Affidabilità e autodiagnostica

Consumer & Professional

- Supporti spesso più longevi dei riproduttori
- Compatibilità solo in presenza di standard reali
- Unico produttore o differenze macroscopiche fra produttori
- DRM, protezione, NDA, soluzioni proprietarie
- Durata? Fino alla prossima generazione di supporti: **obsolescenza programmata**

Paradossi...



- Al mondo non esistono abbastanza testine di lettura per nastri video 2" (Ampex Quadruplex) per rileggere tutti i nastri in archivio
- IOmega Zip&Jaz, floppy da 5", VHS, film 8mm. Li abbiamo ma non possiamo leggerli. Anche floppy da 3,5", nastri Video 8 e MiniDV li seguono.

...scelte di mercato...

- 🧱 Alla fine degli anni '90 apparvero le prime fotocamere digitali
- 🧱 Il formato di registrazione e di codifica delle foto era proprietario
- 🧱 Nessuno sa come “decodificare” le foto, neanche il produttore, che si limita ad un laconico *“this is an unsupported product”*



...ma allora ditelo!

- ❏ Chi ha mai sentito parlare di LS120?
- ❏ Tutti i produttori risarciscono il solo valore commerciale del supporto
- ❏ Qualcuno vende tecnologie già morte, ma non lo dice
- ❏ Aggiungiamo un terzo livello: oltre il riproduttore, serve un “decoder” (DRM/Software proprietario/macchine dedicate)

I dati sulle nuvole

- 🧱 Tutti che corrono ad offrirci gigabyte di spazio *gratis*.

Dareste i vostri dati al vicino di casa, oppure a ME?

Cosa c'è nelle nuvole

- ❏ Una manciata di server ed hard disk
- ❏ Una interfaccia utente via web
- ❏ Qualche volta, un protocollo di storage utilizzabile via rete
- ❏ Una società a scopo di lucro
- ❏ Zero responsabilità ► "Termini del servizio"

Fumo blu, fumo blu...

- ❏ Crittografia per la riservatezza
- ❏ Una password per la sicurezza
- ❏ Server ed infrastruttura ridondati per disponibilità

E' tutto a posto, quindi.

...una nuvola e dentro tu!

- ❏ *Deduplication* (la *buzzword* del semestre) e crittografia non sono compatibili.
- ❏ Le password sono e saranno sempre un punto debole
- ❏ L'unica parte non ridondata è Internet

Onesto == fesso

- ❏ Nel gennaio 2012 il noto (e famigerato) servizio Megaupload viene chiuso, i server sequestrati
- ❏ Il motivo è “l'incoraggiamento alla pirateria” di materiale protetto dal copyright
- ❏ Il servizio era usato da molti per tenere il proprio materiale, legale ed autoprodotta

Ad oggi ***nessuno*** è riuscito a riprendere i propri dati

Cornuti e mazziati

- ❑ Nessuno ha ordinato la cancellazione del materiale (25 Pb = 25.000Tb) e nessuno paga per il mantenimento dei server (1103 server)
- ❑ Nessuno ha ordinato la restituzione del materiale e lo stato ritiene di non avere obblighi di mantenimento dei dati
- ❑ Hai dei dati su Megaupload? Salutali.

Dal sito EFF: <http://goo.gl/7gPHL>

Per tutto il resto c'è... nulla.

- ❏ Le prodezze di Anonymous e di Lulzsecurity, insieme al caso **NSA/Prism** hanno mostrato che nessuno è al sicuro
- ❏ La nuova caccia all'oro è per i dati: sono il principale “prodotto” dei *social network*, e come tale vengono trattati: quotati e commerciati come qualsiasi altro “prodotto”.
- ❏ Niente è gratis

It's a trap!

- ❏ Ottico vs magnetico, e un mare di disinformazione
 - ❏ Velocità: 120Mb/secondo
 - ❏ Rilegge i dati *mentre* li scrive
 - ❏ Si può prevedere il degrado del supporto
 - ❏ Esiste la tecnologia WORM

Ottico? No, nastro.

Soluzioni possibili

- ❏ Per l'archivio della Radiofonia le scelte, obbligate, sono state:
 - ❏ Salvare il contenuto in digitale
 - ❏ Usare formati aperti
 - ❏ Considerare i supporti ***a priori*** inaffidabili
 - ❏ Migrazione ogni 7-10 anni

Trasformare in digitale

- ❏ Ogni supporto (ad esempio nastro da ¼") viene acquisito in digitale, audio e scansione del cartaceo associato
- ❏ I file prodotti sono in formato Wave o FLAC per l'audio e TIFF/JPEG/PNG per le immagini



Il supporto è inaffidabile

- ❏ L'archivio è partito nel 1999 usando supporti DLT da 35Gb
- ❏ Nel 2007 c'è stata la migrazione su LTO-3 da 400Gb (in 5 mesi)
- ❏ Nonostante il breve tempo, alcuni supporti DLT sono risultati illeggibili, per fortuna meno di 20 su oltre 6.500



Migrazione

- ❏ Migrazione in corso su LTO-5
- ❏ La migrazione sarà automatica, e il risultato verificabile
- ❏ Nello stesso spazio si potrà memorizzare il quadruplo dei dati



Bug? *Features!*

- ❑ I drive DLT avevano un piccolo problema non solo meccanico: le cartucce si incastravano nei lettori (1 operazione ogni 50)
- ❑ Il problema era nel sistema di aggancio/sgancio
- ❑ Il produttore del drive non ha mai ammesso nulla, addossando la colpa al produttore dei supporti
- ❑ Il produttore dei supporti forniva anche il costruttore dei drive, che li rimarchiava



Il supporto eterno

- ❏ Cercando in Rete “eternal media” il risultato è esilarante
- ❏ Cercando “long term media” un po' meno: si trovano solo studi sui problemi, nessuna soluzione.



Sitografia

- ❏ Una ditta specializzata in recupero dei nastri magnetici, con foto di nastri disastriati:
<http://www.richardhess.com/tape/index.htm>
- ❏ Conservation Online, dalla Stanford University:
<http://cool.conservation-us.org/>
- ❏ Domande e risposte sul DVD, adattabile a tutti i supporti ottici:
<http://www.dvddemystified.com/dvdfaq.html>
- ❏ Dal "Council on Library and Information Resources" un bel documento sui nastri:
<http://www.clir.org/pubs/reports/pub54/index.html>

Sitografia (2)

-  Foto di hard disk variamente guasti e “sevizati”:
<http://www.harddrive-repair.com/>
-  Phonogrammarchiv Austrian Academy of Sciences:
<http://www.pha.oeaw.ac.at>
-  Training for Audiovisual Preservation in Europe:
<http://www.tape-online.net/>
-  Problemi con i drive DLT:
<http://support.thesmithlife.com/>
-  Il sito di Alliance Storage Technology:
<http://www.plasmon.com/>

Sitografia (3)

- 🧱 Sito amatoriale sul DAT:
<http://www.solorb.com/dat-heads/>
- 🧱 Informazioni amatoriali sul MiniDisc:
<http://www.minidisc.org/>
- 🧱 Sito del FIAT/IFTA: <http://fiatifta.org>
- 🧱 Sito IASA: <http://www.iasa-web.org/>
- 🧱 Il fungo mangia-CD: <http://goo.gl/JwWUU>
- 🧱 La vicenda Megaupload/EFF: <http://goo.gl/7vv54>
- 🧱 Tutto il restante materiale viene da Wikipedia, ove non diversamente specificato

Questo è tutto.

Grazie per l'attenzione!

Per contattarmi: <http://www.ismprofessional.net/pascucci/>

Oppure: mario.pascucci@rai.it

Realizzata con

