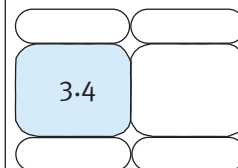




DALL'INTERNET DELLE PERSONE A QUELLO DEGLI OGGETTI IL FUTURO DELLE RETI IP

Questo contributo si inserisce nell'ambito di una serie di articoli pubblicati in precedenza sulla rivista con l'intento di dare una visione degli scenari e delle direzioni di sviluppo delle architetture di rete viste nella prospettiva, da un lato, dell'evoluzione di Internet e, dall'altro, dell'evoluzione delle reti di telecomunicazioni verso la NGN (*Next Generation Network*).

Danilo Ciscato
Vittorio Trecordi



1. IL WEB 2.0 E LA SUA PORTATA INNOVATIVA

1.1. Premessa

Con il termine "Web 2.0" si è voluta indicare una chiara evoluzione del modello di diffusione ed accesso all'informazione tipico del Web tradizionale caratterizzato dal paradigma dell'accesso tramite *browser* (*client Web*) all'informazione multimediale disponibile su numerosi server distribuiti in rete (Siti Web) e, attraverso i *link* ipertestuali, orientato ad una consultazione "non-lineare" dell'informazione [6]. Si deve al Web tradizionale l'innescò dello straordinario successo e diffusione di Internet come strumento fondamentale delle comunicazioni dell'era moderna, proprio per la straordinaria fruibilità ed efficacia espressiva del paradigma. È opportuno inoltre osservare che nel modello Web tradizionale si separano in modo netto i ruoli dei generatori e dei fruitori di infor-

mazione. I primi sono generalmente meno numerosi dei secondi e introducono, rivestendo un ruolo attivo, nuove informazioni in rete inserendo nei siti Web nuovi contenuti corredati del "guscio" ipertestuale che li rende consultabili attraverso i browser. I fruitori di informazione sono invece la gran massa degli utenti della rete Internet che, nella diffusione dell'informazione, rivestono un ruolo "passivo" di consumo dell'informazione inserita in rete dai generatori.

La discontinuità principale introdotta dal nuovo paradigma 2.0 del Web sta proprio nella generalizzazione del ruolo di generatore di informazione nel Web che diventa prerogativa propria di ogni utente della rete Internet. E ciò avviene non semplicemente con la vasta accessibilità di strumenti per la creazione e la messa in linea di contenuti Web (*Web-authoring*), ma proprio con la concezione di nuovi modelli di contribuzione diffusa, di condivisione (socializzazione tra utenti) e di accesso all'informazione, quali per esempio i *Blog* (contrazione di "web-log", che significa tracciato su rete, è lo strumento per la redazione di diari personali on-line consultabili e commentabili da parte

Nota - Gli autori ringraziano sentitamente il Prof. Guido Vannucchi che ha suggerito l'argomento e ha offerto continui stimoli e preziosi suggerimenti per migliorare il lavoro.

dei lettori) o la *Wikipedia* (enciclopedia libera on-line gestita da editori volontari praticamente in tutte le lingue).

La propagazione del ruolo di contribuzione nei contenuti della rete conferisce ad esso un assetto caratterizzato da un minore livello di professionalità (si pensi ai video amatoriali inseriti in *Youtube* o ai commenti - sempre più spesso anche ai contenuti - introdotti nei *Blog*) dando vita ad una conformazione magmatica della base informativa catturata dalla rete, frutto di un concorso di forze accrescitive sempre più polverizzate e diffuse. Si pensi per esempio alle notizie incanalate in alcuni ambiti (quali per esempio quello di *Wikipedia*) che nascono in virtù di un processo di intelligenza collettiva e di una base di conoscenza in continuo accumulo con il contributo correttivo/migliorativo dei generatori più edotti e illuminati.

Il contributo diffuso dal modello Web 2.0 assume pertanto le forme congeniali all'espressione delle personalità individuali più improntate a spirito di protagonismo (è il caso dei *Blog*), ma anche le forme comunitarie che hanno creato le condizioni per la diffusione dei *Social Networking*, intrecci cioè di servizi e di strutture informative in cui si sviluppano e valorizzano a scopi di intrattenimento o di affari le relazioni tra utenti (per esempio, *Facebook* o *MySpace* in ambito *entertainment* e *Linked-in* o *Plaxo* in ambito business).

La figura 1 mostra in modo concettuale l'evoluzione del paradigma da Internet/Web a Web 2.0. Nella rappresentazione del nuovo paradigma vengono riportati alcuni elementi che lo caratterizzano in modo distintivo e hanno fatto entrare nel gergo nuovi termini. Con il termine *User Generated Content* (UGC) si identificano, ad esempio, i contenuti multimediali (generalmente foto e video) che vengono prodotti dagli utenti e iniettati in rete per favorirne la condivisione, attraverso luoghi di scambio dedicati come *Flickr* e *Youtube*. Il termine **prosumer** è stato coniato per contrassegnare la doppia valenza dell'utente della rete: al medesimo tempo consumatore e produttore di informazioni e contenuti. Il termine **ma-shup** è stato coniato per identificare applicazioni basate su piattaforma Web e costruite componendo dati e informazioni provenienti da differenti applicazioni che sono predisposte all'operazione di miscelatura. Ciò può avvenire attraverso il richiamo di interfacce programmatiche accessibili via Web (API o *Application Programming Interface*) o attraverso altri mezzi di esposizione (come per esempio le alimentazioni - *feed* - di informazioni dinamiche diffuse con meccanismi di *publish & subscribe*, ossia di diffusione in pagine Web che riportano l'inclusione del canale informativo dinamico).

È opportuno osservare che la svolta del *Social*

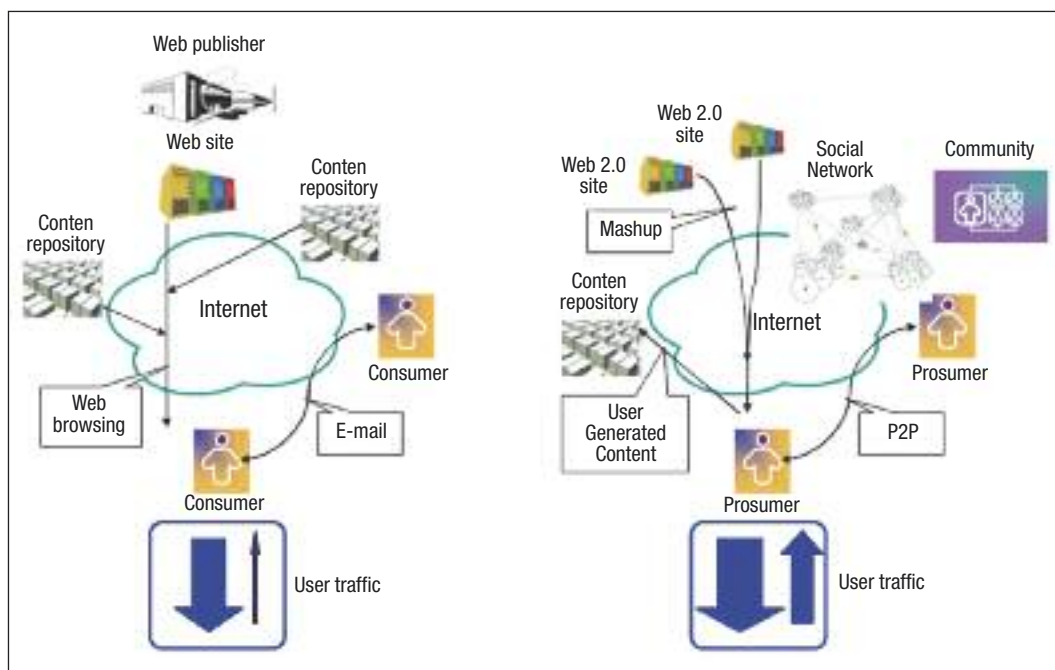


FIGURA 1
Dal paradigma
Internet/Web
al Web 2.0

Networking si concretizza attraverso la “cristallizzazione” dell’informazione e delle relazioni per renderle fruibili in modo permanente. Si superano in tal modo le forme di socializzazione di base proprie di Internet - fin dalla sua origine e ancora oggi assolutamente valide - di comunicazione effimera e privata (quali la posta elettronica e la comunicazione vocale via Internet) nonché le forme meno effimere e pubbliche di *mailing-list*, *newsgroup*, *chat* e del *content-sharing peer-to-peer*.

Anche i professionisti dell’informazione e dei media (inclusi i grandi gruppi editoriali e le *media-company*) sono stati contaminati da queste evoluzioni generali e stanno progressivamente ristrutturando i loro assetti operativi e modelli strategici in relazione al contesto, cercando di preservare la linea di remunerazione del proprio *core business*. In questo contesto fortemente destabilizzante, soggetti provenienti da business tradizionale e soggetti nati con i geni di queste innovazioni sopravviveranno solo se continueranno a rinnovare le forme di remunerazione del valore. Per esempio, i soggetti che hanno la remunerazione della proprietà intellettuale e dei diritti d’autore come fonte principale di valore sono sottoposti a fortissime pressioni dall’avvento di nuovi modelli abilitati dall’evoluzione delle reti e da Internet: basti pensare alla cannibalizzazione dei ricavi dalla vendita delle licenze d’uso del software, consegnato al cliente col classico CD di installazione, da parte dei modelli di vendita rete-centrici di tipo *pay-per-use* (secondo il paradigma del *Software As A Service – SaaS*) o ancor di più dalla diffusione del paradigma dell’*Open Source*.

1.2. Il Web 2.0: la “coda lunga” e il ruolo della pubblicità

Un risvolto particolarmente interessante di questo fenomeno di sviluppo di nuove forme di relazione è la creazione spontanea di comunità di utenti aggregati in base a comunione di interessi e quindi molto ben connotate per affinità. Si apre così un versante particolarmente innovativo e promettente dell’approccio commerciale che punta a colpire una miriade di *target* di mercato, ciascuno di dimensioni modeste, ma molto ben caratterizzato e, per di più, dotato di capacità di interazione e di chiusura della transazione com-

merciale on-line. Il fenomeno definito **long-tail** da Chris Anderson [7] apre la via alla pubblicità mirata per prodotti di nicchia e per colmare il fabbisogno crescente di margini di personalizzazione dell’offerta, per i quali i media e la pubblicità generalisti non offrono una risposta.

È opportuno osservare che lo spostamento del business della pubblicità dai media tradizionali a Internet è in forte accelerazione negli ultimi anni, proprio per l’efficacia dello strumento nello stabilire interazione col cliente e nell’abilitare la transazione commerciale, e, nel caso del Web 2.0 anche nel mirare al target non generalista in modo preciso. Tuttavia coloro che sperano di ricostruire fonti di ricavo sostitutive a quelle tradizionali degli operatori di telecomunicazioni a partire dagli introiti pubblicitari è opportuno che tengano conto del fatto che la dimensione complessiva dei mercati dei media digitali in Italia (*eContent*), è di gran lunga inferiore alla dimensione complessiva del mercato delle telecomunicazioni tradizionali [8]. Solo grazie ad un tasso di crescita straordinario del mercato dell’*eContent* nel suo complesso si potranno riuscire a compensare i cali dei ricavi tradizionali delle Tlc.

Secondo i dati NetConsulting, infatti, nel 2007 il mercato dell’*eContent* in Italia è stato di 5.157 Milioni di Euro, contro i 44.200 Milioni di Euro del mercato Tlc. La crescita anno su anno dell’*eContent* è stata del 21% mentre le Tlc sono cresciute complessivamente dello 0,4%, con un calo del 1,2% nelle Tlc fisse: la perdita di valore anno su anno del 2007 delle Tlc fisse italiane, pari a 245 Milioni di Euro, si confronta con la crescita complessiva di 895 Milioni di Euro del mercato dell’*eContent* (+21% anno su anno), di cui il 50% circa è attribuibile a TV Digitale Terrestre e Satellite, mentre la pubblicità digitale da sola è cresciuta anno su anno di 288 Milioni di Euro (+47,5% anno su anno).

Nella disamina della trasformazione dei modelli di business dei player tradizionali, operatori e *vendor* di tecnologia delle Tlc, va considerato che i mercati dell’*eContent* in forte crescita sono terreno competitivo su cui operano, oltre ai *player* della TV, altri *player* in trasformazione (come i *player* del mercato dei media tradizionali, per esempio gli editori della carta stampata) oltre ad una schiera di nuovi *player* particolarmente brillanti. A questo

proposito, non va dimenticato che giganti del settore ICT mondiale (come Google e Yahoo) si sostengono, per ora, in larghissima parte sugli introiti della pubblicità *on-line*. Tuttavia, il lancio da parte di Google di servizi che pongono un'alternativa all'acquisto tradizionale delle licenze software per messaggistica, collaborazione e gestione documentale a prezzi molto aggressivi (talvolta gratuiti), aprendo la strada a fonti di introito addizionali per Google, impongono serie riflessioni sulla sostenibilità della *value proposition* dei soggetti che operano in modo più tradizionale per offrire soluzioni coperte dall'offerta *in-the-cloud* di Google (intesa come accessibile attraverso la rete Internet senza certezza della dislocazione delle risorse e dei dati).

1.3. Il Web 2.0 e lo sviluppo delle applicazioni e dei servizi

L'evoluzione del paradigma di utilizzo della rete introdotto dal Web 2.0 si completa con l'introduzione di pratiche per lo sviluppo di applicazioni e servizi che considerano la piattaforma Web evoluta come l'ambiente nativo di sviluppo ed esecuzione di applicazioni e servizi (inclusi quelli orientati alla diffusione, condivisione ed accesso all'informazione). Il paradigma del contributo distribuito alla crescita delle applicazioni e dei servizi in questo contesto si concretizza nelle *Service Oriented Architecture* (SOA) e nel *mashup*: due modalità per connotare la componibilità delle applicazioni usando le tecnologie Web, in modo da consentire l'incapsulamento di componenti non-Web-nativi la prima e Web-nativi la seconda¹. La portata di queste, apparentemente semplici, evoluzioni è stravolgente per l'industria ICT e del software in particolare: le applicazioni e i servizi saranno sempre più retentrici (la rete sarà Internet) ed improntati alla logica dell'accrescimento continuo basato sull'apertura a terzi di spazi di composizione di nuove applicazioni e servizi.

1.4. Quasi tutto accade ai bordi della rete

Tutti i fenomeni descritti nella trattazione riguardano evoluzioni che si attuano "ai bordi"

della rete Internet e sono azionate da soggetti che si configurano come utenti della rete, intesa come strumento per il trasporto dei dati, ed interessano pertanto, dal punto di vista tecnico, entità che si configurano come terminali della comunicazione (secondo la nomenclatura Internet: gli *host*).

Un fenomeno particolarmente vistoso è il ricchissimo business sviluppato su Internet dai player *Over The Top* (OTT) – soggetti che operano ai bordi della rete Internet per l'appunto – come i grandi motori di ricerca (tra cui spiccano Google e Yahoo) e i grandi *player* di commercio elettronico (tra cui citiamo Amazon e Ebay). Per completezza si osserva che alcuni di essi adottano accorgimenti particolari per avvicinare i propri servizi agli utenti: ad esempio Google ha investito in una propria rete a larga banda parallela ad Internet e in propri *data center* distribuiti nel mondo e collegati ai principali punti di interconnessione di Internet, realizzando una propria *Content Delivery Network* (CDN) – rete per replicare i servizi in centri distribuiti nel mondo e raggiungibili dagli utenti con un numero limitato di *hop* (router intermedi attraversati).

Il modello di business delle CDN, rappresentato in modo paradigmatico dal *player* di riferimento Akamai, si basa proprio sull'offerta ai soggetti Internet (interessati alla distribuzione di contenuti voluminosi e relativamente stabili) di un servizio di distribuzione per replica geografica dei contenuti, veicolati via Internet con meccanismi di ridirezione delle richieste verso l'archivio più prossimo al richiedente. Gli operatori CDN sono equipaggiati con reti ad altissima velocità, che collegano *data center* distribuiti nel globo ed attrezzati per la replica dei contenuti dei propri clienti e che sono interconnesse con Provider Internet o con punti di interconnessione Internet.

Il paradigma delle reti specializzate, sovrapposte ed interconnesse ad Internet, per la distribuzione dei contenuti si appresta ad essere generalizzato, con il paradigma del *cloud computing* alla distribuzione di risorse informatiche in senso più generale (capacità elaborativa, memoria, servizi applicativi ecc.). Appare evidente che le soluzioni degli operatori CDN si possono classificare come operanti "quasi ai bordi della rete" poiché si basano su meccanismi di controllo applicati-

¹ Recentemente si è coniato il termine *Web Oriented Architecture* (WOA).

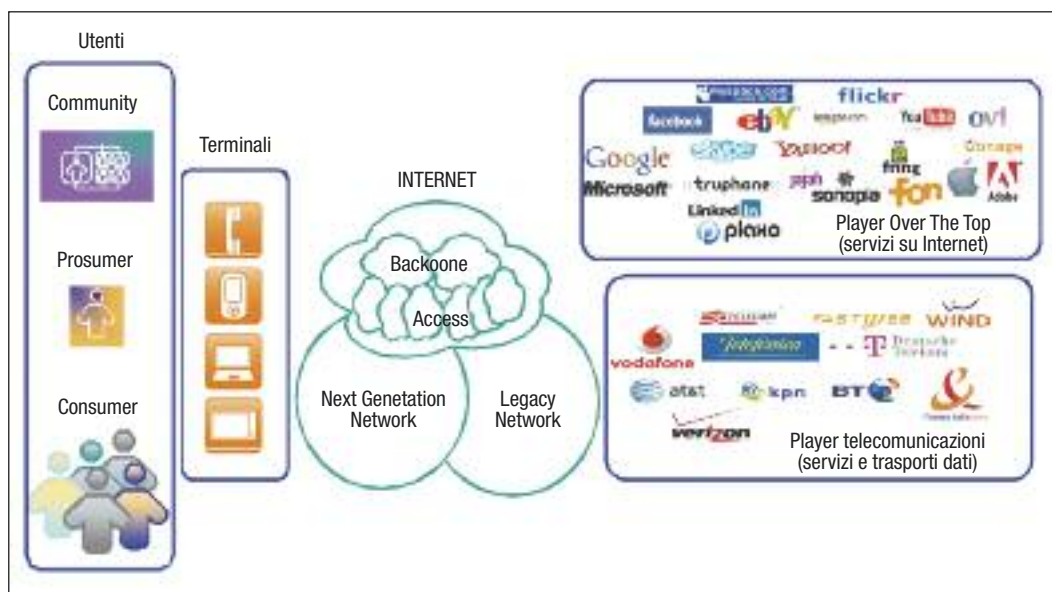


FIGURA 2
Evoluzione
dei servizi e
delle infrastrutture
di Internet:
i protagonisti
del futuro

vi o di supporto (ridirezione DNS, *cache* e *proxy* applicativi, repliche applicative), ma esse richiedono l'interconnessione alla rete Internet in punti non strettamente terminali (nodi di *peering* e nodi di interconnessione Internet).

Nella figura 2 si schematizzano diverse aree di evoluzione del paradigma di Internet verso il Web 2.0. Nella sezione degli utenti, fruitori dei servizi della rete, si mettono in evidenza consumatori (*consumer*), *prosumer* e le comunità (*community*) di utenti – intesi come insieme di soggetti che attraverso la rete coltivano una comunione di interessi e che rappresentano di fatto un soggetto aggregato (caratterizzato da comportamenti di Gruppo) operante in rete.

La sezione dei terminali mette in evidenza una moltitudine di terminali iconizzati in modo da rappresentarne la loro molteplicità e diversità con potenzialità differenziate di ricchezza espressiva e con finalità differenti (esempio, prevalenza di un servizio come l'accesso alla TV rispetto alla navigazione Internet o viceversa e l'orientamento a prediligere l'accesso in mobilità ai servizi telefonici o dati).

La sezione della rete evidenzia la rete Internet, nelle sue varie evoluzioni, come sovrapposta rispetto alle reti degli operatori, di nuova e vecchia generazione (*Next Generation Network* e *Legacy Network* rispettivamente), separando al proprio interno le strutture di accesso (intendendo i *Provider*

di Accesso ad Internet) da quelle di dorsale (intendendo i *Provider* delle dorsali di interconnessione per la lunga distanza di Internet). È necessario notare che la rappresentazione della rete vuole evidenziare la distinzione tra Internet (e sue evoluzioni) e le NGN realizzate dai diversi operatori di rete che offrono - oltre all'accesso ad Internet a larga banda generalmente con tariffazione piatta - una serie di servizi "premium" a qualità garantita come reti aziendali basate su MPLS (*MultiProtocol Label Switching*) ovvero su meccanismi di LAN-Estesa, o servizi IPTV a qualità garantita per gli utenti residenziali. A destra della figura sono stati indicati i soggetti più rappresentativi che operano a livello dei servizi di rete (con infrastrutture proprie di tipo cablato o radi) e operatori che operano ai bordi della rete o *Over-The-Top*. Tra i primi ritroviamo gli operatori ex-monopolisti e i nuovi entranti delle telecomunicazioni, mentre nei secondi troviamo una ben più ampia gamma di tipologie di soggetti, che vanno dai grandi protagonisti del mercato IT (*Microsoft*, *Google*, *Apple* ecc.) e dei servizi in rete (*Amazon*, *Ebay* ecc.), alle società sviluppatesi repentinamente proprio grazie all'evoluzione Web 2.0 (*Facebook*, *Youtube*, *Flickr* ecc.). Ma troviamo anche aziende manifatturiere che hanno tradizionalmente gli operatori delle telecomunicazioni come propri clienti e che si affacciano sul mercato dei servizi *Over the Top* creando situazioni di potenziale

conflitto/competizione (il caso più evidente è quello di OVI² di Nokia-Siemens).

Ai bordi della rete Internet agisce Skype, controllata da Ebay, che si avvale in modo sorprendentemente efficace del modello *Peer-to-Peer* per realizzare, con una rete *overlay* canali di comunicazione audio e video tra i suoi utenti. Coppie di terminali Skype comunicano attraverso flussi di dati crittografati e trasportati in pacchetti IP che attraversano *host* che svolgono le funzioni di nodo di commutazione – “super-nodo” nella nomenclatura Skype – operante ai bordi della rete. Si realizzano in tal modo comunicazioni vocali a costo zero tra utenti Internet. Per abilitare chiamate da utenze Skype Internet verso utenze di rete telefonica fissa e mobile con tariffe locali si adotta il *break-out* (ossia, l'uscita da Internet verso le reti telefoniche tradizionali solo in prossimità dell'utente chiamato) diffuso in tutto il mondo mentre il trasporto su lunga distanza avviene sempre via Internet con qualità accettabile ma senza garanzia delle prestazioni.

1.5. Web 2.0 e la rete Internet

Agli effetti dell'evoluzione Web 2.0, la rete Internet continua ad essere, dal punto di vista del trasporto di dati (da intendersi non in senso letterale e riduttivo rispetto alle altre forme di informazione tra cui video e voce) una rete di reti indipendenti interconnesse, a dimensione e diffusione planetaria, basata sulla famiglia dei protocolli IP (*Internet Protocol*) e capace di consegnare pacchetti IP da una terminazione ad un'altra della rete Internet, con una variante non trascurabile rispetto al fabbisogno di traffico. Il modello dei contributi diffusi (unitamente all'affermazione delle applicazioni *peer-to-peer*) mette in discussione lo sbilanciamento del fabbisogno di traffico a favore del downstream (dalla rete verso il generico utente) considerato finora nell'ADSL decisamente e stabilmente maggiore dell'*upstream* (dal generico utente verso la rete).

Tuttavia, il fenomeno contributivo del Web 2.0 ha già contaminato gli operatori di telecomunicazioni tradizionali dal punto di vista

dell'approccio allo sviluppo dei nuovi servizi: attraverso il meccanismo della *Service Exposure* (esposizione di servizi), gli operatori aderiscono al modello della costruzione dei servizi evoluti con una logica di composizione aperta a molti soggetti e basata sulla piattaforma Internet (già quella attuale). La *Service Exposure* rende accessibili a soggetti terzi la possibilità di invocare funzionalità esposte dall'operatore di telecomunicazioni attraverso interfacce programmatiche aperte, accompagnate da *Software Development Kit* (ambienti di supporto allo sviluppo delle applicazioni distribuite) ed inserite in programmi di certificazione e di sviluppo concorsuale all'innovazione.

2. INTERNET 2.0: VERSO UN NUOVO PARADIGMA DI INTERNET

Le tendenze promosse dal Web 2.0 si uniscono ad altre direttrici di sollecitazione del modello di Internet tradizionale e proiettano verso un modello di Internet evoluta, che indicheremo con il termine Internet 2.0. Si noti che all'interno del più ricco corpo della conoscenza su Internet rappresentato da Internet stessa, non si trova una definizione condivisa e quindi riproponibile come riferimento di Internet 2.0 e pertanto nella trattazione si intendono delineare le linee di trasformazione che, aggiungendosi a quelle “ai bordi” della rete descritte come afferenti al fenomeno Web 2.0 (e successive ondate significative di innovazione), interessano anche il “cuore” della rete, ossia la sezione interna di quella che viene normalmente rappresentata graficamente come una nuvola capace di trasportare pacchetti IP tra terminazioni di rete (riquadro 1 a p. 9).

2.1. Direttrici di evoluzione

Tra le direttrici foriere di discontinuità di maggior rilievo [9], spesso indirizzate con nuovi protocolli che si aggiungono alla famiglia IP e forzano i tradizionali paradigmi della clessidra (tutte le applicazioni su IP su tutti i tipi di rete) e dell'*end-to-end* (i nodi di rete non trattano protocolli di livello superiore al livello 3 del modello di riferimento OSI – *Open System Interconnection* - lasciandone

² Si veda per qualche maggior dettaglio alla fine del paragrafo 2.5.

RIQUADRO 1 - La crescita inarrestabile del traffico IP

In un'epoca di crisi economica, caratterizzata da grafici delle vendite e del PIL che puntano verso il basso, c'è un indicatore che continua imperterrito la sua crescita monotona verso l'alto: il volume di globale del traffico IP. Questo aumento apparentemente incessabile dei volumi di traffico ha rappresentato da sempre la motivazione principale per l'introduzione di nuove tecnologie di rete, sia per la trasmissione che per lo smistamento dei pacchetti IP. Analizzare i fattori che hanno prodotto l'aumento è quindi utile per formulare delle ipotesi sull'andamento futuro del traffico su Internet e quindi dello sviluppo della rete.

Un primo fattore di crescita è dovuto ovviamente al numero di utenti, che ha ormai superato il miliardo e mezzo di individui (<http://www.internetworldstats.com/stats.htm>). È un numero elevato, che rappresenta circa il 24% della popolazione mondiale, ed è interessante osservare che tutte le stime di crescita fatte in passato, nonostante sembrassero esagerate o figlie della bolla speculativa delle .COM, sono state ampiamente superate nella realtà. Il numero di utenti di Internet potrà continuare a crescere? Certamente sì, basta ricordare che è stata recentemente superata la soglia dei 4 miliardi di utenti della telefonia cellulare, anche se occorre tener conto delle maggiori barriere economiche e soprattutto culturali che l'accesso ad Internet comporta rispetto all'uso di un telefono. Tuttavia è evidente che questo fattore di crescita non può continuare a progredire agli stessi ritmi ed destinato a rallentare. Il secondo fattore è costituito dal traffico generato dal singolo utilizzatore, anch'esso cresciuto in misura esponenziale, nel passaggio dai primi rudimentali ipertesti del Web 1.0 ai ricchi contenuti multimediali del Web 2.0 di oggi.

È proprio questo sostanziale cambiamento nei profili di traffico dell'utilizzatore medio che determinerà la crescita futura di Internet. Per fare previsioni sulle dinamiche del traffico Internet Cisco ha sviluppato una metodologia, che si basa sulle proiezioni dei maggiori analisti di mercato (numero di utenti Internet, connessioni broadband, abbonati a servizi video, connessioni mobili, tassi di adozione delle applicazioni su Internet), ma anche sulla validazione attraverso dati raccolti da sonde installate presso clienti operatori di telecomunicazioni. Secondo l'ultimo rapporto, nel periodo dal 2007 al 2012 il traffico IP globale aumenterà di sei volte. Per fare un paragone Internet nel 2012 avrà una dimensione pari a 75 volte quella che aveva nel 2002, raggiungendo un volume mensile di 27 exabytes trasmessi, equivalenti ad oltre 6 miliardi di DVD. Questo traffico è associato ad una moltitudine di applicazioni diverse e ormai quello derivante dalla tradizionale navigazione Web, secondo il paradigma 1.0, è stato superato in termini di volume dal P2P (*Peer-to-Peer*), generato da applicazioni che consentono lo scambio tra utenti di files, tipicamente contenenti musica o filmati. Tuttavia la componente del traffico che sta crescendo più rapidamente è il video, nelle sue varie modalità (on demand o multicast), attraverso diversi terminali (oggi prevalentemente il PC, in prospettiva sempre di più il set-top box) e con livelli qualitativi crescenti (da finestre grandi come un francobollo al formato HD). Si prevede infatti che il video supererà in volume il P2P nel 2010 e arriverà nel 2012 a costituire circa il 90% del traffico IP globale.

l'elaborazione ai soli nodi terminali di rete o *Host*) [10], citiamo:

- la pressione sullo spazio di indirizzamento, in particolare in vista di una diffusione del protocollo IP per le comunicazioni tra oggetti e sui limiti del protocollo IPv4 che spingono verso una migrazione al protocollo IPv6 (riquadro 2 a p. 10);
- l'introduzione di servizi come il VoIP e l'IPTV che richiedono qualità di servizio (ritardi garantiti ed entità di banda continua e consistente);
- il crescente fabbisogno di garanzia dei requisiti di sicurezza degli utenti della rete (rispetto a confidenzialità, integrità delle informazioni e disponibilità dei servizi per la salvaguardia della privacy e la tutela rispetto alle frodi);
- la progressiva compenetrazione di mondo reale e mondo virtuale, ben rappresentata dal successo di *Second Life* come ambito complementare di espansione delle attività reali con elevata efficacia espressiva data dalla rappresentazione antropomorfa degli agenti (*avatar*) che operano per conto e in nome degli utenti della rete e degli ambienti in cui essi si muovono: le applicazioni sono di finalità ludica, ma anche di business che si esplicano con transazioni nel mondo reale (esempio, vendita di be-

ni reali), ma anche nel mondo virtuale stesso (esempio, vendita di artefatti o spazi virtuali);

- la distribuzione della capacità elaborativa, dei dati e delle applicazioni in rete e la crescente capacità di renderli accessibili *on-demand* ai suoi utenti, disaccoppiando in modo completo i fruitori delle risorse dalla loro collocazione e dalla loro dimensione (componibile e scomponibile in modo *soft*, ossia senza interventi di operatori che alterino l'equipaggiamento dei sistemi). La virtualizzazione viene usata per gli elaboratori terminali, ma anche per i nodi di rete per plasmare in modo flessibile le risorse di elaborazione e il loro utilizzo: abbiamo cioè il *Software as a Service* (*SaaS*) usato per spostare il modello di business del software dalla commercializzazione di licenze per l'uso di applicazioni da installare presso i sistemi cliente verso un modello di accesso alle applicazioni installate presso *data center on-demand* via rete, con pagamento basato sull'effettivo utilizzo. Si arriva così fino al *Cloud-computing*, estremizzazione dei concetti di astrazione atti ad abilitare la più spinta virtualizzazione di capacità elaborativa (applicazioni e dati attraverso la rete Internet) a beneficio delle applicazioni e degli utenti della rete.

RIQUADRO 2 - Internet delle cose e IPv6

L'Internet delle cose (termine coniato da ITU - "Internet of Things" - nel 2005) si affaccia sempre più insistentemente come una direttrice di evoluzione della rete imminente attraverso le applicazioni di reti di sensori/attuatori diffusi, il cui rappresentante più noto è il sensore RFID per l'identificazione a breve distanza delle merci/beni). Sulla natura di questa evoluzione si confrontano due scuole di pensiero che prevedono profondità differenti di diffusione del protocollo IP: in un caso si pensa di spingere l'uso del protocollo IP fino a livello di sensori (questo comporta la realizzazione di dispositivi miniaturizzati a basso costo e consumo energetico anche dal punto di vista dei protocolli IP inclusi nel singolo sensore), mentre nell'altro si pensa che il protocollo IP venga usato fino a dispositivi di concentrazione/traduzione (*concentrator/gateway*) che aggregano verso la rete IP le comunicazioni di una moltitudine di sensori che usano protocolli specifici per le reti di sensori (il più noto è *Zigbee* della Zigbee Alliance). La IPSO (*IP for Smart Objects*) Alliance lavora per l'impiego di IP come protocollo pervasivo utilizzato fino ai sensori, puntando su implementazioni compatibili con le scarse risorse elaborative/di memoria e la necessità di contenere i consumi [11].

In entrambi i casi, e con maggior impeto nel primo caso, si prefigura le necessità di disporre di una numerosità di indirizzi largamente superiore a quanto rimane nello spazio di indirizzamento dell'IPv4 (32 bit). All'evoluzione verso l'Internet delle cose viene attribuita la spinta dirompente verso l'introduzione dell'IPv6, auspicabile proprio per il suo ampio spazio di indirizzamento (128 bit) [12]. Già dieci anni addietro si era formulata la previsione di esaurimento ai giorni nostri dello spazio di indirizzamento dell'IPv4, ma l'impiego di tecniche di riduzione dell'uso di indirizzi univoci (indirizzi privati e NAT - *Network Address Translator* - e indirizzi dinamici) e l'uso efficiente dello spazio di indirizzamento con granularità di allocazione fine e aggregazione di indirizzi nelle tabelle di instradamento (uso del *Classless Inter-Domain Routing* - CIDR) hanno fatto sì che la previsione sia stata disattesa e le nuove previsioni fissino i nuovi termini per il giugno 2011. Pressoché la totalità dei *vendor* di apparati terminali e di rete dispone della tecnologia IPv6 e molti operatori hanno già dispiegato la tecnologia IPv6 in porzioni di rete, collegate tra loro a realizzare una rete IPv6 di dimensioni mondiali. Il percorso di migrazione e coesistenza dall'attuale rete IPv4 all'uso progressivo dell'IPv6 si presenta come particolarmente critico proprio per la necessità di gestire un parallelismo di uso e di assicurare, ove opportuno l'interoperabilità tra le due reti, dato che il processo verrà azionato in maniera distribuita da una miriade di soggetti ed avrà una durata difficilmente prevedibile [13].

2.2. Internet 2.0 e NGN

Queste direttrici intersecano il percorso di sviluppo delle reti degli operatori di telecomunicazioni tradizionali sia fisse che mobili, in larga parte protagonisti dei servizi di accesso alla rete Internet, impegnati nella trasformazione delle proprie infrastrutture di rete secondo il paradigma della *Next Generation Network* (NGN).

Il modello NGN condiviso generalmente e codificato dagli organismi di standardizzazione si basa su alcuni principi:

- indipendenza delle funzioni di servizio dalle tecnologie di trasporto;
- utilizzo della tecnica a commutazione di pacchetto (basata su IP);
- capacità di erogare una molteplicità di servizi sulla stessa rete, attraverso una moltitu-

dine di accessi (a larga banda, fissi e mobili);

- supporto alla qualità del servizio e alla mobilità generalizzata per consentire l'accesso ubiquo e senza vincoli ad una moltitudine di servizi e di operatori di servizio.

Il motore primario degli operatori verso l'attuazione di un piano NGN è la necessità di sostituire le piattaforme di rete in campo, dedicate ai diversi servizi (telefonia fissa, telefonia mobile, dati incluso Internet), con una piattaforma unica a maggiore potenzialità (sia in termini di banda trasmissiva, ma soprattutto di servizi) e con minori costi operativi. Questo percorso incontra, tra gli altri, due scogli importanti, peraltro non indipendenti:

□ la necessità di dispiegare *piani di investimento* particolarmente onerosi, principalmente tesi al potenziamento delle risorse di accesso alla rete per la larga e larghissima banda (sia fissa che mobile) a fronte di piani di rientro economico non scontati (oltretutto in periodo di crisi economica generale e di annesse svalutazioni aziendali con difficoltà di accesso al credito);

□ la necessità di definire e calibrare le *funzioni di controllo*, associate al trasporto e ai servizi, da calare all'interno della rete, in modo da garantire le funzioni desiderate (quelle enunciate dal paradigma NGN, con particolare riferimento alla convergenza dei servizi) e di assicurarsi i livelli di remunerazione adeguati. In questo quadro non va trascurata la necessità di stabilire le modalità di interconnessione tra operatori, necessarie a garantire che le prestazioni vengano veicolate in logica *end-to-end* (in quest'ambito i comitati di definizione degli standard stanno lavorando per assicurare autonomia di scelta agli operatori nella propria giurisdizione in tema di tecnologie e fornitori.

2.3. La neutralità della rete

Ambedue i versanti di difficoltà ricorrono nel dibattito emergente, finora primariamente discusso negli USA, della "neutralità della rete" [14] in cui gli operatori infrastrutturali perseguono una linea tesa a promuovere, nelle norme di utilizzo della rete, una separazione e un trattamento differenziato dei traffici in funzione della loro natura: il traffico *best effort* dell'Internet generale veicolato con risorse e prestazioni residuali rispetto al traffico a pagamento (*premium*) con garanzie di qualità del

servizio per applicazioni di audio e video *streaming* (inclusi servizi IPTV di qualità *entertainment*). Questa posizione è chiaramente tesa ad assicurare maggiori garanzie di remunerazione degli investimenti in infrastrutture di rete con i proventi dei servizi “premium” a discapito di traffici “parassiti”, ossia traffici cui, rispetto alla quantità di risorse di rete impiegate, corrispondono remunerazioni trascurabili se confrontati con i traffici “premium” (esempio lampante: il traffico delle applicazioni P2P e dei *player OTT*). Ad opporsi a questa posizione, fino ad ora non fatta propria dalle istituzioni USA, i soggetti che attribuiscono a questa forma di discriminazione del traffico una limitazione sull’accesso alle risorse della rete (informazione e servizi), con ripercussioni sullo sviluppo dell’innovazione, di cui Internet è stato fattore abilitante di enorme portata. D’altra parte è evidente che il mercato delle telecomunicazioni tradizionali (impennate su ricavi per telefonia fissa e mobile) ha raggiunto uno stadio di maturità che porta ad una vistosa contrazione dei ricavi. Si impone pertanto una ricerca di nuovi recuperi di margini grazie alle nuove tecnologie capaci di ridurre i costi (pacchetto e NGN) nonché nuove potenzialità di ricavi (servizi dati a larga banda in mobilità e convergenti) di tipo *flat* e sempre più *semi-flat*, al fine di conciliare ritorni minimi garantiti con riferimento ai costi delle risorse di rete effettivamente utilizzate. Si noti tuttavia che, per gli operatori di telecomunicazioni, il bilancio di compensazione dei vettori di crescita e dei vettori di diminuzione dei margini nel caso di servizi telco-centrici porta in ogni caso ad una risultante negativa che spinge alle concentrazioni tra soggetti economici attivi su mercati sovrapposti o adiacenti per ricercare equilibri economici non più accessibili con mercati fortemente frammentati.

2.4. Lo sviluppo dell’accesso “wireless broadband” nell’ambito della convergenza fisso- mobile

La saturazione del mercato della telefonia mobile e la maturità del mercato telefonico hanno progressivamente spinto gli operatori mobili verso la ricerca di valore nei servizi dati. Con l’evoluzione delle tecnologie cellulari dal 2G (basato sul GSM) verso il 2,5G (GPRS), ma so-

prattutto verso il 3G (UMTS) e le sue varianti evolutive atte ad accrescere la banda accessibile nella tratta radio (dapprima *down-link* e successivamente *up-link*), i diffusissimi servizi di *short messaging (SMS)* sono stati affiancati da servizi di accesso ad Internet a larga banda. *Smartphone* sempre più ergonomici e stilisticamente pregiati (guidati senza discussione dall’iPhone di Apple per il mercato residenziale e affiancato da *Blackberry* di *Research In Motion (RIM)* per il segmento affari) e “chiavette” con interfaccia USB per il collegamento di PC a reti 3G e 3,5G si stanno diffondendo rapidamente nel mercato come veicoli di accesso ad Internet, favoriti, oltre che dalla praticità d’uso, dall’adozione di piani commerciali *semi-flat*³ particolarmente accattivanti.

Va sottolineato che le prestazioni accessibili da queste tipologie di accesso sono fortemente dipendenti dall’uso promiscuo della risorsa radio da parte degli utenti della stessa rete e che le prestazioni di picco messe in evidenza dagli operatori rappresentano un riferimento distante da quelle continuamente accessibili a regime. In ogni caso, tuttavia, il successo del *mobile broadband Internet access* costringe gli operatori mobili a potenziare significativamente le risorse di rete (in particolare la rete di accesso radio ricorrendo anche a tecnologie che consentono l’*off-load* della risorsa radio come le *femto-celle* che instradano su ADSL il traffico mobile che parte dall’interno delle abitazioni) ed a potenziare la rete di *backhauling* (passando dalle tecnologie di rilegamento TDM basate su canali E1 alle tecnologie Ethernet ad alta velocità più adatte al trasporto efficiente del traffico IP). Al rafforzamento delle risorse di rete si affiancano anche interventi sulle tratte di interconnessione ad Internet e l’introduzione di strumenti di ottimizzazione delle prestazioni in grado di intervenire sui protocolli di trasporto e di applicativi specifici (inclusendo funzioni *proxy* e *caching*) per migliorare la prestazione erogata agli utenti.

³ Le proposte commerciali prevedono abbonamenti con compenso mensile indipendente dall’uso, purché non si eccedano soglie di volume di dati o di tempo di utilizzo, oltre le quali l’uso viene tariffato a volume (dette anche offerte *flat* con “*capping*”).

2.5. Funzioni della rete: controllo e sua organizzazione

Anche il tema delle caratteristiche della **porzione di controllo** da sviluppare nella NGN è controverso, pur se in ambito di standardizzazione si è sviluppata una certa convergenza di vedute attorno al **modello di riferimento architetturale IMS** (*IP Multimedia Subsystem*), inizialmente proposto dal 3GPP per l'evoluzione della parte di controllo dei servizi della rete mobile basata su IP e successivamente adottato anche da ETSI-TISPAN come modello per la "core network" della rete NGN convergente (incluso accesso fisso a larga banda). Il controllo dei servizi indirizzato da IMS si pren-

de carico della gestione delle sessioni avvalendosi del **protocollo SIP** (*Session Initiation Protocol*) definito dall'IETF come protocollo standard per la segnalazione indipendente dal servizio. Le funzioni principali assolve da SIP sono la ricerca del partner di una comunicazione di cui è noto l'identificativo (secondo una logica che supera quella del numero di telefono associato ad una linea), verificarne la disponibilità ad instaurare una sessione multimediale, negoziare i parametri e gestire la sessione fino all'abbattimento. L'architettura IMS (per qualche maggior dettaglio si veda anche il riquadro 3) propone di articolare nel nucleo della rete una serie di componenti fun-

RIQUADRO 3 - Il modello di riferimento dell'architettura NGN dell'ITU-T

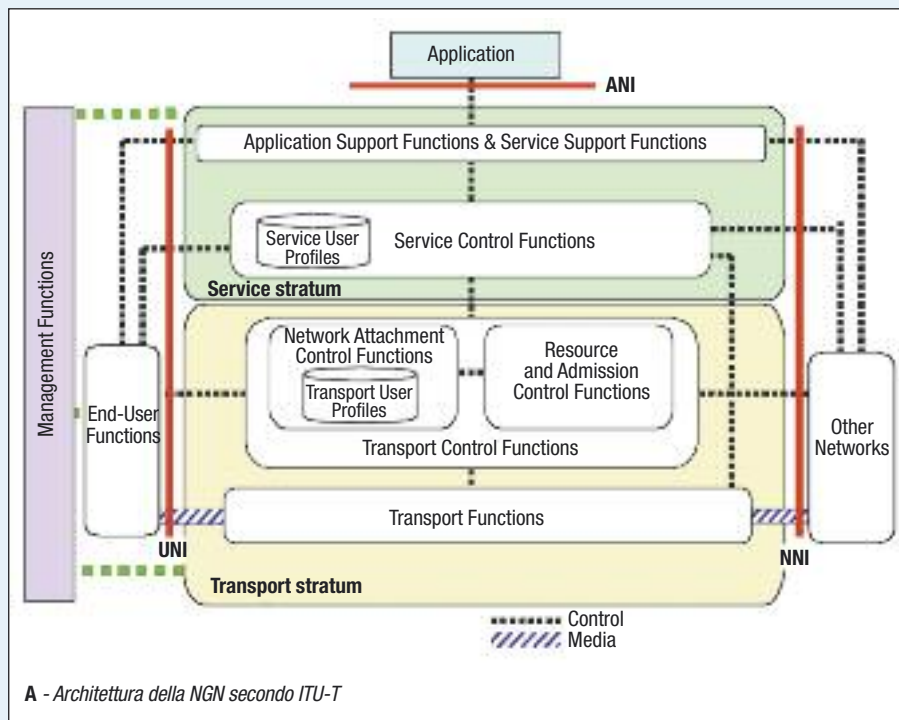
L'ITU-T avviò nel giugno 2004 un Focus Group per l'NGN (FGNGN) col mandato di sviluppare attività di normalizzazione su temi urgenti collegati all'emergente NGN: requisiti di servizio, architettura funzionale, mobilità, sicurezza, qualità del servizio, controllo e segnalazione, evoluzione della rete di trasporto a pacchetto. Il lavoro dell'FGNGN è terminato nel Novembre 2005 ed è confluito nella NGN Global Standardization Initiative (NGN-GSI), istituita per consolidare gli standard dell'NGN secondo un piano articolato su più fasi.

L'ITU-T, compresa la complessità del progetto di sviluppo delle norme per la NGN e la molteplicità dei flussi di lavoro concorrenti, ha messo in linea uno strumento di gestione del progetto NGN [17] che consente di condividere lo stato avanzamento e i piani di lavoro di tutte le iniziative parallele di molti gruppi di lavoro, interni ed esterni all'ITU, che hanno rilevanza per l'iniziativa NGN-GSI. Il lavoro dell'NGN-GSI valorizza gli sforzi degli organismi di standardizzazione regionale attivi sull'argomento (ETSI, sia TISPAN o Telecommunications and Internet converged Services and Protocols for Advanced Networking, che 3GPP, per l'Europa, ATIS per gli US, CJK-NGN per Cina, Giappone e Corea), con fortissime contaminazioni del lavoro di IETF (l'organismo di standardizzazione di Internet) e dei vari Forum (tra cui il Multiservice Forum o MSF e l'IMSForum, particolarmente attivi sull'interoperabilità in ambiente multivendor, ma anche il MetroEthernet Forum o MEF, il riferimento per lo sviluppo delle soluzioni di trasporto evolute, basate su tecnologia Ethernet).

ITU-T ha completato le aree di lavoro riconducibili alla Release 1 della NGN: Requisiti di servizio e architettura funzionali di Fase 1, funzioni di controllo di accesso alla rete e di controllo delle risorse di fase 1, requisiti di convergenza fisso-mobile e impostazione di un framework per la sicurezza. Attualmente l'ITU-T sta lavorando alla Release 2 nelle seguenti aree: Requisiti di servizio e architettura funzionali di Fase 2, funzioni di controllo di accesso alla rete e di controllo delle risorse di fase 2, requisiti funzionali ed architetturali per il traffico multicast, framework per la mobilità, architettura aperta per lo sviluppo dei servizi.

L'architettura di riferimento della NGN consolidata già nei documenti dell'ITU-T per la Release 1, evidenzia, come da figura A, la separazione dello strato di trasporto (a sua volta ripartito tra piano di controllo e di attuazione del trasporto), strato di servizio (con evidenza delle funzioni di controllo del servizio) e strato delle applicazioni.

Il ruolo dell'IMS nel modello NGN dell'ITU è ben chiarito nella Raccomandazione Y.2021 (09/2006) in cui si affida all'IMS il ruolo di controllo per la fornitura di servizi multimediali basati su SIP e l'e-



A - Architettura della NGN secondo ITU-T

segue

Rimandando al testo della raccomandazione per i dettagli, si richiamano nella figura B solo i blocchi funzionali della “core IMS” coinvolti e le loro funzioni: *Call Session Control Function* (CSCF) attiva, monitora e abbatte sessioni multidiali gestendo l’interazione con l’utente, *Media Gateway Control Function* (MGCF) controlla la *Trunking Media Gateway Functional Entity* (TMG-FE) allocando e disallocando risorse per la gestione dei media, *Multimedia Resource Function Controller* (MRFC) alloca risorse per il trasporto dei dati della sessione nella rete, *Breakout Gateway Control Function* (BGCF) seleziona la rete PSTN verso cui inoltrare le chiamate, selezionando l’opportuno MGCF. La direzione strategica dell’NGN-GSI è verso un’architettura che aumenti le proprie funzionalità includendo quelle del mondo Web e abilitando la convergenza di una serie di nuovi servizi che, oltre alla emulazione dei servizi tradiziona-



alla emulazione dei servizi tradizionali e all'offerta dei servizi IPTV, includa per esempio le *Universal Sensor Network* (USN).

L'adozione diffusa di un'architettura di controllo IMS faciliterebbe in linea di principio il raggiungimento di obiettivi di interconnessione *end-to-end* per servizi complessi, anche se tuttora è abbastanza evidente la difficoltà di integrare semplici funzioni tra le comunità degli utenti di reti diverse come la rete mobile e quella fissa o, a maggior ragione quella Internet. Iniziative recenti e anticipatrici del modello IMS completo in forma più modesta - quali *Rich Communication Suite* (RCS) di GSMA, 3GPP/NGN-IMS *Multimedia Telephony Service* (MMTel) ed *Open Mobile Terminal Platform*

Il dibattito sulle funzioni che ci si attende vengano svolte dalla rete e dall'operatore che la realizza (e che gestisce e le funzioni che possono essere svolte ai bordi della rete dagli operatori OTT) è lontano dal trovare una risposta definitiva. Nel frattempo gli operatori infrastrutturati che sviluppano i propri piani NGN applicano meccanismi di garanzia della qualità del servizio per i clienti che accedono ai propri servizi *premium* (telefonia, *streaming video* di qualità *entertainment*), mentre continuano a mantenere l'accesso ad Internet come un servizio ben distinto, erogato in modalità *best-*

effort. Le garanzie da commercializzare per l'accesso ad Internet in modo differenziato riguardano le caratteristiche dell'accesso (in termini di banda massima *down-stream* ed *up-stream*) e di banda Internet minima garantita, dall'attacco utente fino al IXP (*Internet Exchange Point*): ovviamente sulla prestazione ottenibili con la navigazione di tipo generalista non vi sono garanzie di prestazioni *end-to-end*.

L'ipotesi semplicistica di una coincidenza tra NGN e Internet 2.0 appare molto lontana dal realizzarsi, anche se la tecnologia di base comunemente adottata è IP, principalmente per la difficoltà di superare le difficoltà evidenziate e di raggiungere accordi tra una moltitudine di soggetti che operano in questo settore ed i cui interessi non sempre sono coincidenti o comunque facilmente allineabili [15].

Il modello di Internet attuale privilegia, infatti, la capacità di mettere in comunicazione coppie di terminali della rete Internet senza preoccuparsi della qualità dello scambio (modello *best-effort* dell'IP). Il passaggio dell'introduzione di garanzie di qualità delle comunicazioni tra due utenti della rete Internet di nuova generazione richiede invece l'adozione di tecniche di controllo della qualità del trasporto da parte di tutti gli operatori Internet sulla catena che si interpone tra i due utenti, inclusi gli operatori di interconnessione Internet. Questo obiettivo è ancora lonta-

no e per ora i servizi di qualità (esempio, telefonia, IPTV) sono offerti ai clienti di operatori di accesso ad Internet con soluzioni proprietarie e basate sull'attuazione di tecniche di gestione della QoS nella rete sotto il controllo dell'operatore stesso.

La figura 3 offre una **rappresentazione semplificata dell'architettura emergente della rete NGN**, elaborata con un'interpretazione evolutiva di quanto emerge dagli organismi di standardizzazione (con particolare riferimento all'ITU-T).

L'architettura mostra la separazione degli stati di trasporto, servizio ed applicazione (vedi anche figura A del riquadro 3) ad indicare che si persegue l'obiettivo di astrarre le caratteristiche implementative di ciascuno strato, mantenuto come indipendente, rendendo disponibili interfacce condivise e standardizzate per richiamare i servizi offerti da uno strato all'altro.

Lo **strato di trasporto** è a sua volta ripartito in funzioni di pura movimentazione dei dati e funzioni di controllo, distinte in funzioni di controllo del collegamento alla rete e funzioni di controllo delle risorse. Le funzioni svolte da questo strato sono rappresentate sulla destra e comprendono l'indirizzamento, l'instradamento unicast e multicast, le prestazioni di sicurezza (esempio, autenticazione), le funzioni di qualità del servizio differenziata e di affidabilità del

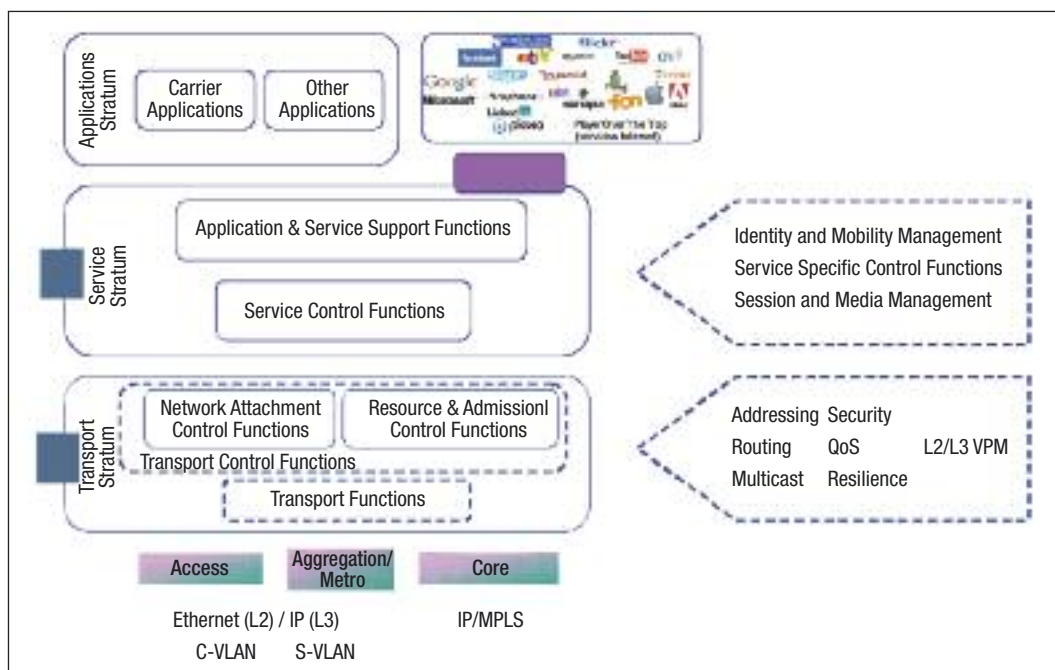


FIGURA 3
Architettura
di riferimento
emergente per
la NGN

servizio di trasporto, nonché le funzioni per realizzare reti private virtuali con tecniche di livello 2 e/o 3 del modello di riferimento OSI (ossia utilizzando servizi di LAN-estesa basati su Ethernet e su funzioni di bridging o servizi di VPN-IP). Sotto lo strato di trasporto si è inteso rappresentare la struttura gerarchica della parte di trasporto distinguendo tra **accesso** propriamente detto, la parte di **aggregazione** con le tecnologie di livello 2 e 3 offerte come interfaccia di servizio, quali Ethernet e/o IP - e si evidenzia la tendenza nel caso dell'uso dell'Ethernet a distinguere tra VLAN per cliente (C-VLAN) e VLAN per servizio - (S-VLAN) nonché infine la **parte di dorsale o core** in cui si tende ad impiegare MPLS, *MultiProtocol Label Switching*.

Lo **strato di servizio** offre funzioni specializzate per i servizi offerti dalla rete, come le funzioni di controllo della sessione (applicabili per servizi orientati alla sessione) o funzioni di controllo dei media impiegati (come nel caso del *media streaming* per i servizi di IPTV) e funzioni generali, come ad esempio la gestione dell'identità dell'utente. Lo strato di servizio offre inoltre funzioni di supporto ai servizi stessi ma anche alle applicazioni: tra questi viene rappresentata la funzione di exposure di funzioni telco-centriche a beneficio dello sviluppo di applicazioni di terze parti che le richiamano e le compongono in logiche di servizio innovative e indipendenti dall'operatore. Nello strato delle applicazioni sono posizionate le applicazioni dell'operatore e di altri soggetti, ivi inclusi i *player OTT*. Per gli strati di trasporto e di servizio sono rappresentate a destra i blocchi funzionali di interlavoro (IWF, *InterWorking Function*) che hanno lo scopo di garantire il collegamento tra strati realizzati da operatori diversi, assicurando l'interlavoro delle funzioni che si desidera vengano garantite anche oltre lo strato di trasporto e di servizio del medesimo operatore (sia in condizioni di roaming, che nel caso che le terminazioni di un servizio di rete si avvalgano di strati di trasporto e servizio di operatori diversi).

Degne di nota per l'impatto che potrebbero avere sullo sviluppo dei modelli di business e del mercato dei servizi ICT sono le iniziative di due grandi *player* del mercato che espandono la propria *mission* indirizzando in modo deciso ruoli non propri, creando nuove op-

portunità di sviluppo del business, in qualche caso con conflitti verso i loro stessi *core business*:

□ Nokia ha avviato l'iniziativa OVI⁴ che si pone come una piattaforma di erogazione di servizi ad utenti per il mercato residenziale degli utenti che accedono da terminale mobile evoluto (*smartphone*): il recente accordo per l'integrazione di Skype apre una via di conflitto con i modelli di business perorati dai maggiori clienti di Nokia (operatori mobili);

□ Google, sull'onda del successo dell'iPhone di Apple, e convinta del ruolo crescente del terminale nello sviluppo dei nuovi servizi per utenza mobile, ha immesso sul mercato il proprio sistema operativo per terminali mobili evoluti, *Android*, puntando su un modello aperto di sviluppo (contrapposto al modello tradizionalmente più chiuso di Apple) su cui costruire la propria strategia per la platea dei clienti dell'accesso mobile ad alta velocità destinata ad espandersi notevolmente nel futuro;

□ l'azienda canadese *Research In Motion* (RIM), nota in tutto il mondo per gli *Smartphone BlackBerry*, si è allineata ai concorrenti aprendo un *marketplace* per agevolare l'incontro tra domanda e offerta di nuove applicazioni sviluppate secondo un modello aperto alle aziende di sviluppo e con meccanismi di supporto allo scaricamento delle applicazioni e al *push* sicuro degli aggiornamenti con pagamenti *on-line* basati su *Paypal*.

3. CONCLUSIONE

La rete Internet e la sua manifestazione più visibile il World Wide Web hanno cambiato il modo di comunicare, di fare le reti e di vivere e di fare business in pochissimi anni. Discontinuità importanti spingono verso la nuova generazione di Internet e del Web (la 2.0), e questa trasformazione si interallaccia in modo complesso con lo sviluppo delle nuove infrastrutture ed architetture di rete (NGN) da parte degli operatori di telecomunicazioni. Gli scenari che emergono da questo concorso di trasformazioni sono complessi ed interessanti soprattutto per gli importanti riflessi

⁴ OVI significa Porta in lingua finlandese

sulla struttura del mercato dell'ICT e di tutti i settori della vita privata e lavorativa che sempre più ne fanno uso.

Bibliografia

- [1] Decina M., Giacomazzi P.: Il futuro del protocollo Internet. *Mondo Digitale*, giugno 2007, p. 17-29.
- [2] Ciscato D.: Rete IP ed Internet: il quadro evolutivo. *Mondo Digitale*, marzo 2008, p. 30-38.
- [3] Spinelli C., Vannucchi G.: New Generation Network: Strategie di Accesso. *Mondo Digitale*, dicembre 2008, p. 17-35.
- [4] Decina M., Giacomazzi P.: WEB 2.0 Tecnologie abilitanti e nuovi servizi. *Mondo Digitale*, dicembre 2008, p. 3-16.
- [5] Giornata di Studio AICT: *Internet 2.0: gli effetti del Web 2.0 sul futuro delle reti e dei servizi*. Roma, 4 dicembre 2008.
- [6] Riley Mark: *What has Web 2.0 got to do with me?* http://www.ncvovol.org.uk/uploadedFiles/NCVO/Events/Events_Archive/2008/What_is_web_2.0.ppt - The National Council for Voluntary Organisations, Novembre 2008.
- [7] Anderson Chris: *Long Tail, The Revised and Updated Edition: Why the Future of Business is Selling Less of More*. Hyperion, luglio 2008.
- [8] Confindustria Servizi Tecnologici e Innovativi in partnership con il Dipartimento per l'Innovazione e le Tecnologie e con la collaborazione scientifica di Net Consulting. [http://www.confindustriasi.it/files/File/Docu-](http://www.confindustriasi.it/files/File/Documenti/Studi%20e%20Documenti/eContent-2008_exesum.pdf)
- menti/*Studi%20e%20Documenti/eContent-2008_exesum.pdf* - Il Terzo Rapporto sul Mercato dei Contenuti Digitali in Italia, Settembre 2008.
- [9] Vinton Cerf: *Internet Collaboration for Change*. PTC'09, gennaio 2009, http://www.ptc.org/ptc09/images/papers/PTC%2709_Cerf_Slides.pdf
- [10] Aguiar Rui L.: *The Bloated IP Protocol*. NeXtworking'07, aprile 2007, <http://www.net.tlabs.tu-berlin.de/ARCADIA/talks/Aguiar.pdf>.
- [11] Internet Protocol for Smart Objects (IPSO) Alliance, Settembre 2008: <http://www.ipso-alliance.org/Documents/IPSO%201%20%20IP%20for%20Smart%20Objects.pdf>.
- [12] Quaderni di Telema: *La domanda di comunicazione chiede di aggiornare Internet*. Coordinatore del Quaderno Francesco Matera, Media 2000, n. 242, XXIV - 10 - dicembre 2006/gennaio 2007, dicembre 2006.
- [13] Gaetano Roberto: *From IPv4 to IPv6: impact and transition*. Seminario Bordini: "Internet del futuro: evoluzione della piattaforma tecnologica", febbraio 2009, <http://www.key4biz.it/files/000110/00011078.pdf>.
- [14] Carter Kenneth R., Scott Marcus J., Wernick Christian: *Network Neutrality: Implications for Europe*. Wik Consult, dicembre 2008, http://www.wik.org/content/diskus/diskus_314.pdf.
- [15] Koichi Asatani: *Trends of NGN and Its Issues*. Ottobre 2008, http://www.ieeeboston.org/Presentations_Society/DLT08BostonAsataniDist.pdf.
- [16] ITU-T: *NGN Management Project*. <http://www.itu.int/ngnproject/>, 2009.

DANILO CISCATO laurea con lode in Ingegneria Elettronica al Politecnico di Milano nel 1985. Ha trascorso nove anni in IBM, ricoprendo vari ruoli tecnici, commerciali e manageriali, sia negli USA che in Italia. È entrato in Cisco nel luglio 1994, nella fase di startup della Società in Italia, in qualità di consulente tecnologico nello staff europeo. Ha operato a stretto contatto con i maggiori clienti italiani nella realizzazione dei più importanti progetti di rete e ricoperto ruoli manageriali di crescente importanza. Attualmente è Direttore "Business Development & Marketing" ed è responsabile di tutte le attività relative allo sviluppo del business, con particolare riferimento all'introduzione sul mercato di tecnologie innovative, che hanno da sempre costituito il motore di crescita della Cisco.

E-mail: dciscato@cisco.com

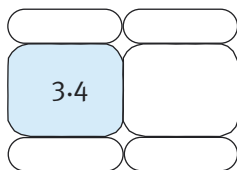
VITTORIO TRECORDI laureatosi con lode in Ingegneria Elettronica al Politecnico di Milano nel 1986, ha conseguito il Master CEFRIEL in tecnologia dell'informazione nel 1988. Dopo brevi periodi in Laben e Pirelli Cavi, è stato impiegato presso il CEFRIEL fino al 1998, occupandosi di ricerca, formazione e consulenza nel settore delle reti innovative. Nel 1998, lasciò l'incarico di Direttore Tecnico del CEFRIEL, per fondare insieme al Prof. Decina e all'Ing. Randi una società di consulenza, la ICT Consulting della quale è attualmente Direttore Generale. Dal 1993 è docente presso la sede distaccata di Cremona del Politecnico di Milano ove attualmente tiene il corso di "Infrastrutture e protocolli per Internet". È stato autore di una cinquantina di pubblicazioni di ricerca per congressi e riviste in ambito internazionale e di numerose pubblicazioni nazionali, tra cui un libro: "Nuovi modelli di business: evoluzione delle architetture di rete e dei servizi", Franco Angeli, 2002.

E-mail: trecordi@ictc.it



LA FEMTOCELLA UNA STAZIONE RADIO MOBILE A LARGA BANDA IN CASA

Rocco Casale
Giuseppe Catalano
Sandro Dionisi
Daniele Franceschini



Le femtocelle¹, trasmettitori lillipuziani (nelle dimensioni e nelle potenze) su banda licenziata, da installare in modo plug&play casa per casa o in ufficio, riducono il cammino del segnale, minimizzando in tal modo la possibilità di degradazioni e massimizzando la qualità percepita. Questo aspetto, unito all'assegnazione dedicata della capacità, concorre a migliorare le prestazioni effettive che il cliente può sperimentare. Grazie alla femtocella il terminale mobile diviene un nuovo elemento dell'home networking; questo dispositivo consentirà agli operatori di sviluppare nuovi servizi basati sull'interazione tra telefono cellulare e gli elementi della casa.

1. INTRODUZIONE: COS'È UNA FEMTOCELLA?

Il servizio radiomobile è nato con l'obiettivo di assicurare assoluta libertà agli utilizzatori, e quindi di essere connessi sempre, ovunque e in piena mobilità, specialmente all'aperto, anche lontano dalle infrastrutture fisse di telecomunicazioni. L'ampia diffusione delle reti e dei terminali GSM e UMTS/HSxPA hanno, poi, reso quotidiano e normale a ogni utente del servizio di telefonare, inviare messaggi e collegarsi a Internet con il telefono in ogni momento e in ogni luogo. Questi gesti non destano oggi alcuna meraviglia perché sono diventati consueti e abituali.

Lo standard 3GPP (*3-rd Generation Partnership Project*), che definisce le specifiche tecniche dei sistemi radiomobili, sulla spinta del

successo mondiale della soluzione HSDPA (*High Speed Packet Data Access*), ha già posto le basi per un'ulteriore evoluzione di tale tecnologia verso il servizio mobile a larga banda. La cosiddetta *HSPA Evolution* sarà basata su diversi elementi: una nuova modulazione radio in grado di fornire un bit rate massimo di 21 Mbit/s, l'uso di più antenne in trasmissione e in ricezione (MIMO, *Multiple Input Multiple Output*), che permetterà di raggiungere un bit rate massimo di 28 Mbit/s e infine l'uso combinato di più portanti, con cui le velocità menzionate aumentano in proporzione alla banda di frequenza (portante) disponibile. L'evoluzione dell'HSxPA proseguirà poi con l'introduzione della tecnologia denominata LTE (*Long Term Evolution*), di cui il 3GPP ha concluso la definizione della prima

¹ Femto è un prefisso che indica un milionesimo di un miliardesimo. La parola deriva dal termine danese *femten* che indica una potenza di 10^{-15} . Il prefisso femto fu utilizzato per i sistemi trattati in questo articolo per indicare che essi coprono un'area di dimensioni estremamente ridotte, sensibilmente inferiori a quella coperta dalle picocelle. Con tale termine si intende quindi un piccolo dispositivo che permette agli utilizzatori del sistema di connettersi da casa o in ufficio, come è chiarito nel testo, alle reti cellulari UMTS, CDMA-2000, WiMAX e alle successive evoluzioni dei sistemi radio.

versione capace di abilitare bit rate teorici sino a 100 Mbit/s.

Le tecnologie mobili divengono pertanto potenzialmente capaci di fornire servizi a larga banda ma, sebbene siano etichettate con la velocità di trasmissione teorica di picco, presentano in pratica prestazioni che dipendono dalla qualità radio e dal numero di clienti che simultaneamente accedono alla risorsa. L'avvicinamento dell'antenna di trasmissione al terminale migliora le caratteristiche del canale radio e quindi le prestazioni effettive che il cliente può sperimentare. La realizzazione di una copertura in qualche modo dedicata a un cliente, riservandogli l'intera capacità di una cella, massimizza, infatti, la qualità del servizio da lui percepita in termini per esempio di velocità di trasmissione.

La disponibilità di un'elevata banda radio all'utente richiede, dal punto di vista dell'architettura di rete, una capacità adeguata del cosiddetto *backhauling*, della connessione cioè delle antenne e delle stazioni radio base ai nodi di rete presenti in centrale e quindi la disponibilità di sistemi per il trasporto dei dati a larga banda (basati su portanti ottici o su soluzioni xDSL) che divengono divengono perciò, un pre-requisito del sistema.

Voler offrire servizi mobili realmente a larga banda impone, dunque, l'introduzione - in aggiunta ai consueti dispiegamenti geografici denominati "macrocellulari" - di dispiegamenti micro e femto cellulari (Figura 1), con antenne installate vicino all'utente e con una disponibilità sulla rete di trasporto di tecnologie a larga banda (fibra nel caso di dispiega-

mento con stazioni radio micro, fibra o xDSL nel caso di dispiegamenti femtocellulari).

In particolare, con il termine femtocella s'intende una cella di piccole dimensioni, quali possono essere quelle di un'unità abitativa o di un ufficio. La femtocella (alias *Femto Access Point* o *Home NodeB*), è perciò realizzata con un apparato wireless normalmente caratterizzato da dimensioni, costi e consumi ridotti, a bassa potenza e autoinstallante, che non richiede manutenzione da parte del cliente ed è gestito a distanza dall'operatore e che, a differenza del WiFi, trasmette su bande di frequenza soggette a licenza (su quelle, per esempio, dell'UMTS). Una femtocella è in grado di garantire la copertura 3G, utilizzando per la connessione in rete la connettività xDSL dell'utente, attraverso il modem a banda larga, oppure mediante un modem integrato (Figura 2).

Una soluzione di questo tipo consente all'operatore di ottimizzare la fornitura del servizio in ambiente indoor, che ha assunto grande importanza in quanto sede di una grossa porzione del traffico mobile. In accordo con tutte le principali analisi di mercato, la maggior parte del traffico cellulare si svolge, infatti, nelle abitazioni o nei posti di lavoro.

A questa esigenza si aggiunge un altro elemento di rilievo: il 2008 appena concluso ha visto una significativa crescita del traffico HSxPA, fruito dai clienti in condizioni di mobilità - più precisamente di quello nomadico - per esempio attraverso l'impiego di chiavette HSxPA inserite in un laptop. Gli analisti prevedono, infatti, nei prossimi anni una cre-

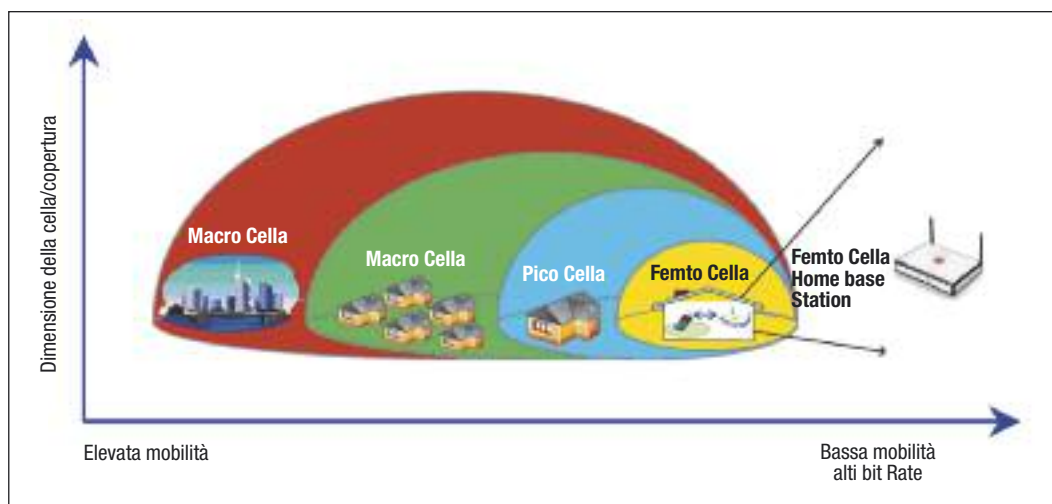
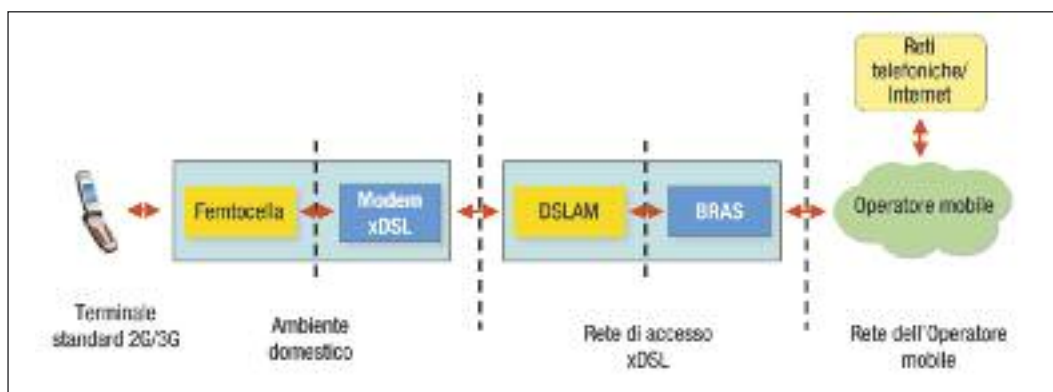


FIGURA 1
Relazione tra estensione della copertura e caratteristiche di servizio

FIGURA 2
Architettura di riferimento del sistema femtocellulare



scita molto significativa del traffico da terminali mobili a larga banda, una cui parte significativa avverrà proprio in condizioni di nomadicità in casa e in ufficio (come già è avvenuto per il traffico fonico).

Da questo punto di vista la femtocella sembra intercettare questa crescita del mercato, fornendo una soluzione tecnicamente efficiente all'operatore per il trasporto del traffico con notevoli benefici per il cliente finale.

Per dare un'idea della trasformazione che si potrebbe osservare nella rete pubblica delle telecomunicazioni con l'introduzione delle femtocelle può essere ricordato che il numero delle stazioni radio base oggi in servizio presso un singolo operatore è dell'ordine di alcune migliaia e che le abitazioni che hanno già un collegamento alla rete xDSL alla fine del 2008 hanno superato la decina di milioni e quelle che si prevede di connettere con la rete a larga banda nel prossimo futuro sono destinate a crescere ulteriormente. Il mercato delle femtocelle non coinciderà naturalmente con il numero di tutte le abitazioni collegate alla rete pubblica a larga banda, ma i dati sopra riportati mostrano che - se questa tecnologia si affermerà, come mostrano già oggi alcuni indicatori di interesse da parte di numerosi operatori alla tecnologia - le dimensioni del mercato di questa nuova categoria di terminali potrebbero essere significative.

In aggiunta, la femtocella, ancor più se pensata nella versione che integra 3G e WiFi, consente al terminale mobile di fruire di tutti i servizi fissi. Il telefonino entra così nelle abitazioni come un ulteriore elemento dell'home networking.

Gli operatori individuano, infatti, come possi-

bile elemento di crescita del mercato dell'accesso fisso il cosiddetto *home networking*, ossia l'interconnessione degli apparecchi di casa alle reti informatiche domestiche e ad Internet.

Il mercato delle reti domestiche sta seguendo, infatti, un processo di evoluzione basato sull'offerta di apparecchi domestici già predisposti per connettersi alla rete Internet, siano essi *entertainment devices* (televisori, riproduttori di musica digitali, terminali hi-fi, videocamere e camere digitali ecc.) oppure elettrodomestici tradizionali (frigoriferi, lavatrici, lavastoviglie, forni elettrici ecc.).

Questi prodotti sono basati su una molteplicità di alternative tecnologiche realizzate da diversi enti di standardizzazione, che operano per realizzare specifiche tecniche in questo ambito, rivolgendosi ad apparati e reti di telecomunicazioni vere e proprie (TISPAN/3GPP), agli *home gateway* (Broadband Forum, HGI) e agli apparati della *consumer electronics* (UpnP Forum). Per essere completamente interoperabili, le soluzioni commerciali devono quindi rispettare le caratteristiche indicate nelle numerose specifiche in vigore.

Lo sviluppo industriale di questi terminali, insieme alla crescente diffusione di PC, è visto con grande interesse dagli operatori poiché può stimolare ulteriormente la diffusione di accessi a banda larga.

Di sicuro, un punto su cui si vanno focalizzando gli sforzi degli enti di standardizzazione e dei progettisti delle industrie di riferimento è quello di individuare soluzioni - dette *no-new-wires* (reti power-line e wireless) - che, per quanto possibile, aggirino la necessità di cablare o ricablare gli appartamenti. La nuova tecnologia delle femtocelle si presenta

dunque come una soluzione idonea a promuovere il mercato delle reti domestiche e a contribuire a costruire il cosiddetto ecosistema della banda larga.

L'articolo si propone di illustrare gli aspetti tecnici più rilevanti della tecnologia emergente delle femtocelle mirando a chiarire gli elementi di maggior rilievo delle soluzioni adottate e i problemi che dovranno essere analizzati e risolti nei prossimi mesi.

2. FEMTOCELLE: UN NUOVO PARADIGMA TECNOLOGICO E DI SERVIZIO

La copertura tipica di una femtocella è fortemente dipendente dalle caratteristiche della propagazione nell'ambiente in cui essa è installata. Con i livelli di potenze in gioco, il raggio di una cella è in genere di alcune decine di metri e varia in funzione del particolare ambiente considerato. In linea di principio si possono attuare sia coperture con femtocelle isolate, di solito appropriate per un impiego in appartamenti, sia coperture multiple, da integrare con una LAN, più adatte agli uffici e alle aree industriali (Figura 3). Finora, tuttavia, i principali costruttori si sono concentrati sul prodotto per un impiego nelle abitazioni; sono disponibili però anche prodotti progettati per applicazioni business/enterprise che, come sarà indicato più avanti, sono oggi ancora in una fase di consolidamento e di messa a punto finale.

Si possono anche ipotizzare, in futuro, scenari di impiego delle femtocelle in ambienti esterni e misti purché in condizioni di limitata mobilità degli utenti, ma in presenza di grandi volumi di traffico aggregato (quali per esempio in centri commerciali, aeroporti, stazioni ferroviarie, percorsi pedonali cittadini ecc.) e anche di richieste di capacità elevate da parte del cliente finale. Le applicazioni delle femtocelle in ambienti esterni sono finora poco investigate (sebbene vi sia interesse da parte di alcuni grandi player, tra i quali Google).

Si può però già supporre un utilizzo analogo, in modo complementare, alle architetture ROF (*Radio Over Fiber*), che sono considerate tra le opzioni d'interesse per le comunicazioni a grande capacità della futura generazione radiomobile grazie alla disponibilità capillare della fibra. Uno scenario, che in futuro potrebbe essere abilitato dalle femtocelle in aree urbane densamente popolate, impiega piccole unità remote su lampioni e a parete con backhaul su una coppia della rete pubblica in rame (xDSL) o su una fibra ottica. Una soluzione di questo tipo potrebbe consentire una più agevole ed economica erogazione di servizi fissi e mobili da parte dell'operatore (Figura 4) o anche di trasmettere da una postazione remota informazione a larga banda. In termini di offerta di servizio, tre sono gli scenari applicativi più promettenti realizzabili con le femtocelle.

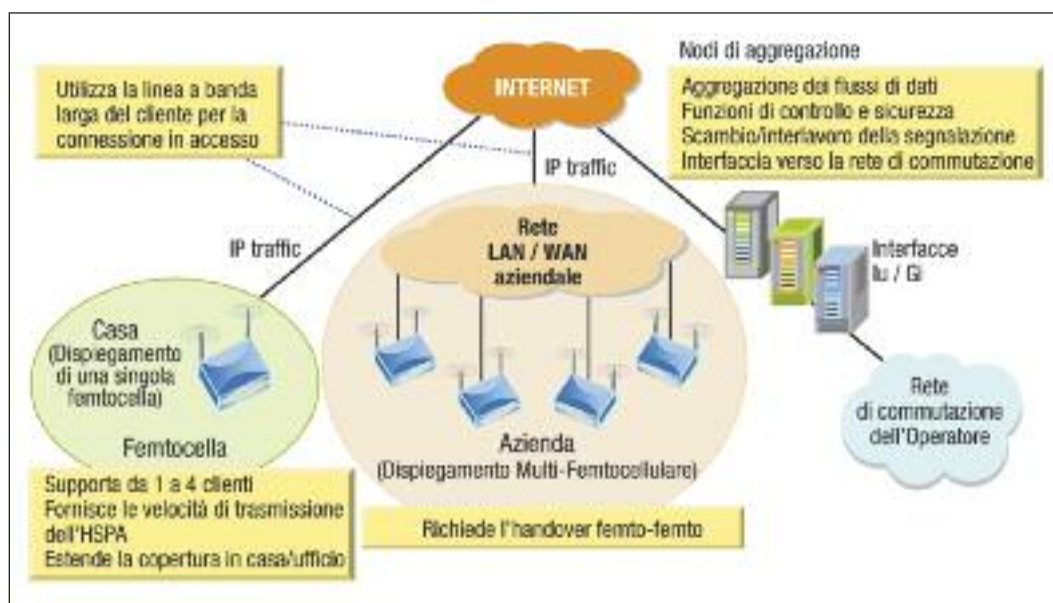


FIGURA 3

Scenari di sviluppo delle femtocelle in casa e in azienda

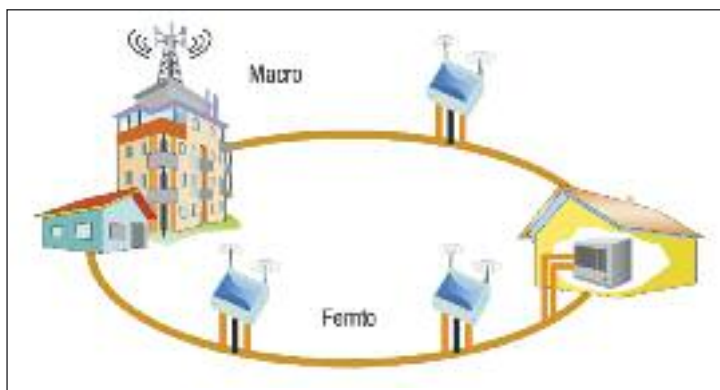


FIGURA 4

Possibile scenario di sviluppo delle femtocelle in ambienti esterni



FIGURA 5 *a. Scenario di utilizzatori privati*

Femtocelle e Home Networking

L'applicazione della soluzione femto 3G in ambiente residenziale, per una clientela consumer, costituisce lo scenario di servizio da offrire da subito, poiché esso è stato l'obiettivo primario perseguito nella definizione e nello sviluppo di questa nuova tecnologia. L'architettura di riferimento coincide con quella descritta nella figura 2: con essa si assume, infatti, che il backhauling della femtocella avvenga mediante una connessione xDSL del cliente. Le caratteristiche del trasporto in termini di banda disponibile, latenza e meccanismi di QoS, devono essere in questo caso compatibili con l'erogazione degli stessi servizi radiomobili che il cliente è abituato a utilizzare sotto una copertura radio di tipo macro. La femtocella 3G è, infatti, un *access point* alla rete 3G che, una volta installato, è visto da un terminale 3G, cellulare o PC card, come una normale cella di rete. In questo scenario, il

cliente può originare o ricevere chiamate (fonia, SMS, MMS, videochiamate o connessioni dati) in maniera trasparente, su rete femto o su rete macro 2G/3G in funzione dell'area di copertura in cui opera con il proprio terminale. In tale scenario, l'unico requisito (radio) per l'utente, che gli permette di utilizzare la copertura femto dedicata è che esso disponga di un terminale mobile 3G standard, cellulare o PC card, configurato in modalità *dual-mode* o solo UMTS, senza alcun *client ad hoc*.

Le femtocelle, grazie alle caratteristiche evidenziate sopra indicate, consentono agli operatori della rete mobile di fornire con un elevato bit rate servizi mobili a larga banda all'utenza consumer, controllando al contempo il costo per bit, con un valore, che rappresenta dunque un requisito che sarà posto alla base di un'ampia diffusione dei differenti tipi di servizi sopra indicati.

b. Scenario business/enterprise

L'applicazione delle femtocelle 3G per la clientela business, nelle due declinazioni SOHO (*Small Office-Home Office*) e large enterprise, costituisce lo scenario di servizio più sfidante allo stato dell'attuale maturità della tecnologia. Lo scenario business, per le sue caratteristiche peculiari, non può prescindere da alcune prestazioni, quali per esempio: l'handover tra due femto, un maggior traffico smaltibile in termini di utenti abilitati e chiamate contemporanee, una maggiore copertura, il controllo dei percorsi del traffico all'interno della rete aziendale, che possono effettivamente decretarne il successo o l'insuccesso. È ragionevole ipotizzare che questa architettura dovrà essere necessariamente più flessibile rispetto a quella relativa allo scenario di tipo consumer, in modo che possa adattarsi e integrarsi in scenari di tipo *large enterprise*, le cui dimensioni e caratteristiche richiedono quasi sempre sviluppi *ad hoc*.

c. Non solo Mobile Broadband...

Come già evidenziato, la femtocella porta nell'ecosistema dell'home networking (costituito da vari elementi presenti nelle abitazioni, quali *access gateway*, *set top box*, *televisioni* ecc.) l'interfaccia radio 3G in aggiunta a quella WiFi (Figura 5). Grazie a questa caratteristica la femtocella può comunicare via WiFi con tutti i terminali della casa e mediante interfaccia radio 3G con gli stessi terminali

mobili che il cliente impiega normalmente. La femtocella diviene così un abilitatore tecnologico di una nuova famiglia di servizi che fanno perno sull'interazione tra il cellulare e i terminali impiegati nelle abitazioni.

Gli oggetti di normale impiego in casa saranno infatti sempre più dotati in futuro di sistemi di calcolo e di memorizzazione integrati e di porte d'accesso wireless, che ne consentiranno la comunicazione con il mondo esterno e potranno colloquiare tra loro mediante le femtocelle. Da questo punto di vista sarà possibile sviluppare nuove tipologie di servizio che integrino aspetti di comunicazione interpersonale (*infotainment, remote control, security and health* e non ultimo *social networking*) rese possibili da terminali che colloquino tramite il cellulare dell'utente che rappresenta quindi il punto di accesso a questo universo di servizi.

In prospettiva, nei prossimi anni, con il progressivo diffondersi delle femtocelle, queste potenzialità tecnologiche potranno essere valorizzate dagli operatori mobili in offerte di nuovi servizi di natura *home zone*, capaci di portare valore aggiunto ai clienti nell'ambiente domestico.

L'offerta dei servizi indicati attraverso le femtocelle consente di fruire di numerosi benefici potenziali sia per i clienti che per gli operatori. Tra i vantaggi conseguibili dai clienti possono essere citati:

a. l'uso del terminale cellulare

Un primo vantaggio nell'impiego delle femtocelle consiste nella possibilità di utilizzare gli stessi terminali mobili 3G, già impiegati normalmente dal cliente, per accedere ai diversi servizi tradizionali offerti dalla rete (voce, SMS, ricerche Web) o a quelli più innovativi (*home networking*), così da potere fare leva sull'ampia diffusione dei terminali e dei servizi mobili già abilitati.

b. Qualità del servizio offerta per il mobile a larga banda

Come si è detto, in un'area chiusa di dimensioni ridotte la femtocella garantisce un'ottima copertura radio e l'intera capacità di una cella è dedicata a un numero ristretto di utenti abilitati. La femtocella assicura così la possibilità di gestire un traffico a larga banda con valori di bit rate sempre comparabili ai valori di picco previsti dallo standard in prossimità

della femtocella. Il supporto della mobilità tra la copertura assicurata dalle femtocelle e quella macro (sia pure in presenza di alcune limitazioni indicate nel seguito) consente inoltre di fruire dei servizi con un grado di mobilità non permesso da altre tecnologie. Una diffusione capillare dell'accesso alla rete con le femtocelle e l'uso di canali radio non condivisi con altri utenti consente, infine, di usufruire dei servizi con tariffe agevolate del tipo *home zone*. È possibile così offrire ai clienti un servizio di qualità elevata in termini di varietà di servizi, prestazioni, condizioni di mobilità e applicare una tariffazione evoluta.

c. Abilitazione di nuovi servizi di home networking

Una femtocella che integri gli accessi radio 3G e WiFi consente di far entrare nell'ambito degli elementi dell'*home networking* (televisioni, gateway dell'utente, modem xDSL, set top box ecc.) lo stesso terminale impiegato normalmente in condizioni di mobilità fuori dall'ambiente domestico. Questa caratteristica potrebbe contribuire a creare, in prospettiva, una famiglia di servizi che sfruttino al meglio le interazioni tra terminale mobile e elementi presenti nelle abitazioni grazie alla femtocella permettendo ai clienti di godere presumibilmente di tariffe più convenienti associate alla fruizione di questi servizi nella propria abitazione.

Tra i vantaggi acquisibili dagli operatori possono essere citati:

a. Incremento della capacità di rete

Il dispiegamento delle femtocelle consente di dedicare la copertura e la capacità radio di una cella a un singolo cliente, che può quindi fruire dei servizi mobili a larga banda grazie alla propria connettività di tipo xDSL e in futuro in fibra ottica. In questo modo la terminazione remota fissa a larga banda contribuisce all'erogazione del servizio mobile grazie al trasporto del traffico di *backhauling* della femtocella verso la *core network*.

In questo modo si attenuano tutti i vincoli che limitano la capacità e la copertura di rete macro (che per la propria natura di copertura geografica estesa permette di servire un numero elevato di utenti potenziali). Essa è infatti sgravata dal traffico instradato dalle femtocelle aumentando così la capacità totale della rete dell'operatore mobile. La cre-

scente penetrazione delle femto e il conseguente alleggerimento progressivo della rete macro comportano anche una minore necessità di installare nuovi siti macro per soddisfare un aumento del traffico in un'area o per migliorare la qualità del servizio percepito dall'utente (per esempio in termini di bit rate), superando in questo modo le frequenti difficoltà logistiche e i costi elevati associati alla predisposizione di nuove stazioni radio base.

b. Incremento della penetrazione dei servizi di mobile a larga banda

L'offerta di servizi mobile a larga banda con un'elevata qualità che utilizzino le femtocelle costituisce un volano in condizioni di mobilità e consente agli operatori della rete mobile di dare maggior valore all'offerta degli stessi servizi nelle singole reti macrocellulari.

c. Fidelizzazione del cliente attraverso offerte di tipo home zone

Sfruttando la copertura radio dedicata della femtocella e limitandone l'accesso ad un uso esclusivo, è possibile costruire facilmente offerte appealing basate su una tariffazione home zone, anche grazie a una migliore accuratezza nell'identificazione dell'area "home" insita nelle caratteristiche di propagazione radio delle femtocelle e nel collegamento alla rete xDSL. Questa caratteristica consente agli operatori di proporre offerte del tipo home zone con un'accuratezza molto maggiore rispetto a quella ottenibile con le coperture macrocellulari.

In sintesi, nello scenario dell'accesso a larga banda, dei servizi home zone e delle reti domestiche la femtocella può dunque rappresentare, specie nella versione integrata, uno dei sottosistemi chiave da associare all'*home gateway* domestico, per offrire una gamma di servizi molto ampia e con qualità elevata in termini di prestazioni, mobilità e tariffazione. Dal punto di vista tecnologico le femtocelle risultano in linea con la tendenza tipica dei moderni sistemi di telecomunicazioni che prevedono che una quota crescente di intelligenza si sposti verso gli elementi posti alla periferia della rete [1].

La maggior parte delle industrie manifatturiere sta producendo femtocelle con interfaccia radio 3G (UMTS/ HSPA). Sono in corso di sviluppo femtocelle anche per altri sistemi, come WiMAX e LTE, che però non presentano ancora soluzioni tecnologiche definite e che non saranno disponibili nel breve termine.

3. PROIEZIONI DI MERCATO

Alcuni analisti di mercato considerano la femtocella la più importante novità nel settore delle comunicazioni mobili, sebbene le proiezioni siano piuttosto diversificate. ABI Research stima circa 52 mila unità consegnate nel 2007 che crescono fino a circa un milione nel 2008, mentre ben 102 milioni di utenti potranno essere serviti alla fine del 2011 mediante 32 milioni circa di punti di accesso (Figura 6) [2]. Se IDATE stima circa 10 milioni di femtocelle consegnate nel 2010 e 18 milioni nel 2011, le previsioni di IDC sono più conservative, giacché riportano un numero di utenti di femtocelle di almeno 5,7 milioni nel 2011. Sempre con riferimento al 2011, OVUM prevede 17 milioni di femtocelle consegnate nella sola Europa occidentale.

Alcuni operatori radiomobili stanno investendo da tempo in femtocelle: per esempio, negli Stati Uniti l'operatore Sprint nel settembre del 2007 ha lanciato a Denver e Indianapolis "Sprint Airave", servizio commerciale basato su femtocelle CDMA della Samsung e ha piani per sviluppare progressivamente una rete nazionale.

Come è mostrato in [8] gli operatori che sono più avanti nella sperimentazione e che hanno cominciato a offrire il servizio sono quelli che

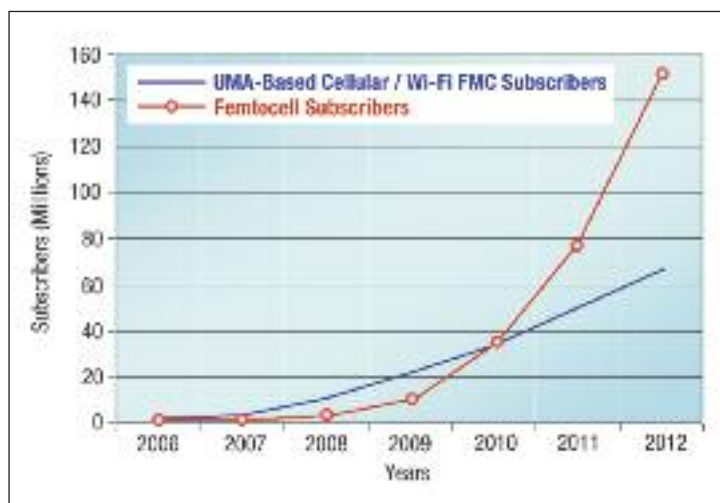


FIGURA 6

Proiezioni di mercato delle femtocelle e del Wi-Fi (Fonte: ABI Research, 2007)

operano nell'Estremo Oriente e negli Stati Uniti: oltre a Sprint Nextel la stessa Verizon ha aperto il servizio al pubblico il 26 gennaio 2009 anche se per ora fornisce solo il servizio fonico e non quello a larga banda. Anche in Europa molti importanti operatori della rete fissa e mobile hanno avviato nel corso del 2008 sperimentazioni con questi sistemi nell'ottica di poter avviare offerte al pubblico, basate su questa tecnologia, nei prossimi anni.

Per poter realizzare le grandi economie di scala necessarie per un'ampia diffusione delle femtocelle è essenziale disporre di uno standard tecnologico unico. A questo obiettivo ha lavorato il 3GPP, che nel corso del 2008 ha definito un'architettura di sistema descritta nel capitolo seguente. Altri fora in ambito internazionale (per esempio, il Femto Forum) hanno contribuito a consolidare una visione omogenea di sistema.

4. ARCHITETTURA DI SISTEMA

Un primo aspetto importante per l'inserimento delle femtocelle nella rete di un operatore mobile riguarda quindi la standardizzazione del sistema costituito dalle femtocelle. Nella tradizione delle reti mobili GSM e UMTS la completa definizione delle interfacce radio, ricordiamo, è stata fondamentale per consentire l'interoperabilità con terminali realizzati da differenti società manifatturiere, in modo da realizzare un'economia di scala nella produzione dei terminali e, soprattutto, di consentire il roaming dei clienti. La

stessa esigenza di standardizzazione non era altrettanto sentita per le interfacce tra la stazione radio base ed il relativo controllore che quindi presentano elementi specifici di ciascuna manifatturiera.

Con l'avvento della femtocella come stazione radio base domestica nasce invece la necessità di avere una standardizzazione completa anche dell'interfaccia tra stazione radio base e rete dell'operatore mobile, poiché questa diviene l'interfaccia di un terminale del cliente (Figura 7).

L'esigenza di standardizzazione è ulteriormente giustificata dal fatto che la femtocella per avere successo, dovrà divenire un prodotto consumer, in modo da assicurare alti volumi di produzione, tipici del mercato dei terminali. Finora il 3GPP ha realizzato una prima fase del processo di standardizzazione sia per il 3G che per la LTE e ha definito, come sarà chiarito in dettaglio più avanti un'architettura di riferimento per la release 8 dei due sistemi.

In anticipo rispetto alle attività dello standard, le società manifatturiere hanno identificato numerose e diverse soluzioni architetturali per inserire le femtocelle nella rete UMTS. La figura 8 mostra le quattro configurazioni principali ad oggi proposte: com'è mostrato nella figura, ogni soluzione prevede l'utilizzo di un nodo di aggregazione delle femtocelle verso la rete di commutazione dell'operatore, che non potrebbe gestire un numero elevato di femtocelle connesse direttamente ad essa. Le diverse soluzioni si differenziano per le fun-

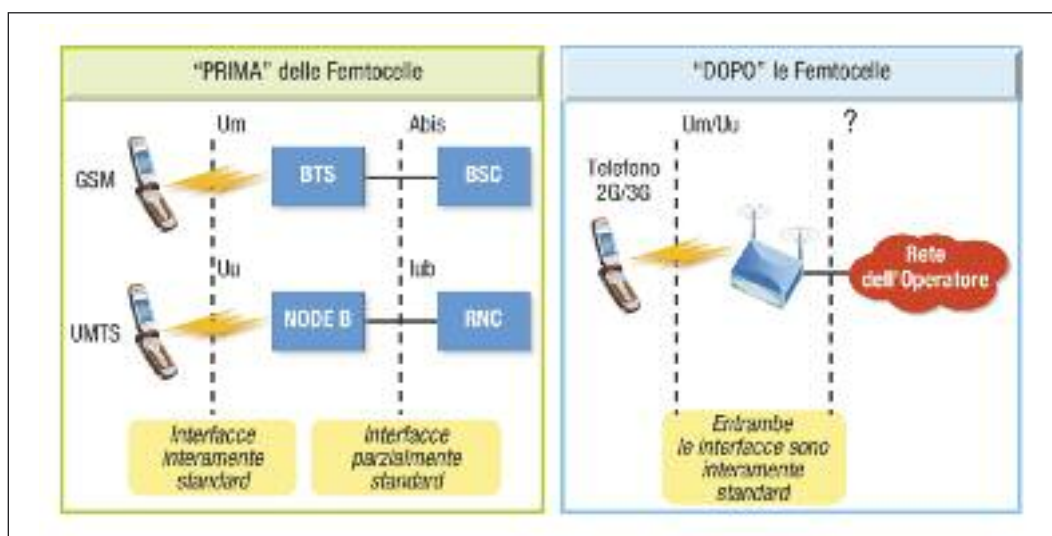
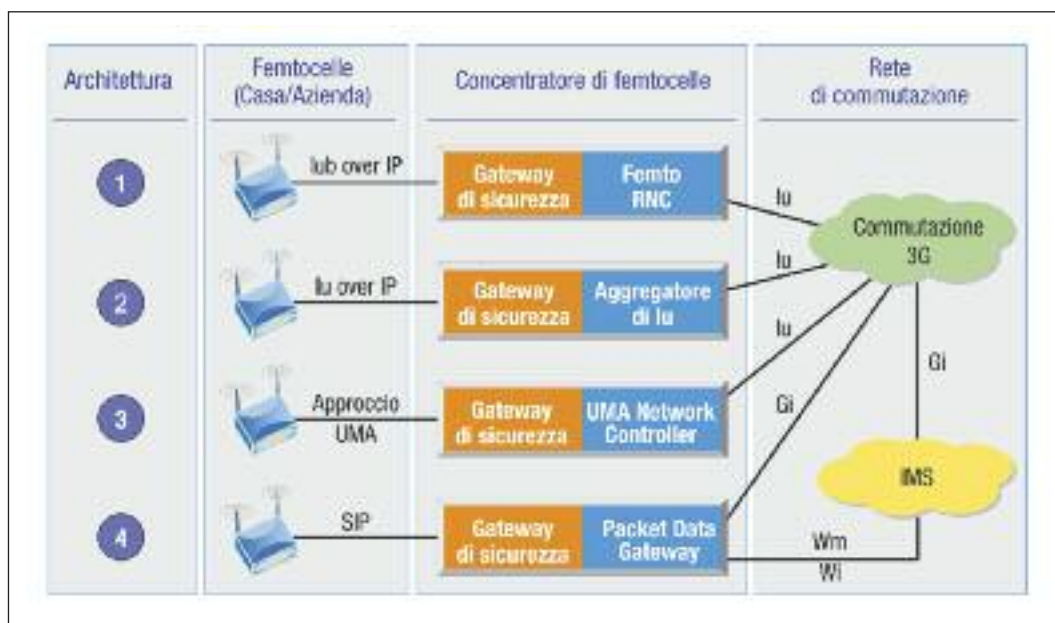


FIGURA 7

Interfacce standard nelle reti di accesso 2G e 3G

FIGURA 8
Confronto
di soluzioni
architeturali



zioni dell'aggregatore e per i protocolli utilizzati all'interfaccia verso le femtocelle.

Un secondo punto, comune alle diverse soluzioni, consiste nell'implementazione di funzioni di sicurezza tra la femtocella e l'aggregatore, dovute sia al fatto che il *backhaul* xDSL può essere gestito da un operatore diverso da quello radiomobile (l'interconnessione può avvenire eventualmente anche attraverso la rete Internet) sia al fatto che, come si è detto, la femtocella è anche un terminale d'utente. Il primo aspetto richiede di garantire al cliente la sicurezza delle comunicazioni; il secondo mira a fornire garanzie all'operatore sulla femtocella che si connette alla sua rete. L'interfaccia, qualunque sia l'architettura di riferimento, deve essere perciò incapsulata entro un tunnel IP e si adottano a questo scopo soluzioni specifiche di autenticazione della femtocella (quali, per esempio, USIM interna alla femtocella, certificati ecc.).

Una prima soluzione (architettura "1" nella figura 8) consiste nel replicare l'architettura della rete esistente, con le femtocelle connesse dunque a un RNC (*Radio Network Controller*) attraverso un'interfaccia lub. La soluzione è stata rapidamente superata sia perché l'interfaccia lub è normalmente un'interfaccia pseudo-standard, sia perché una stazione radio base domestica può comprendere al proprio interno le funzioni di controllo richieste per le risorse radio.

Nell'architettura "2" la femtocella include sia le funzioni di Node-B che quelle di RNC. La femtocella deve utilizzare in questo caso una differente soluzione per interconnettersi alla rete mobile. Una possibile alternativa è la cosiddetta *lu-over-IP*: in questo caso è utilizzata l'interfaccia *lu* dell'UMTS (interfaccia tra l'RNC e la rete di commutazione). L'aggregatore non svolge alcuna funzione di controllo delle risorse radio, ma solo quelle di aggregazione e di interlavoro delle interfacce *lu* tra femtocelle e rete di commutazione.

L'impiego di un'interfaccia standard realmente aperta come la *lu* rappresenta un'alternativa semplice e con elevate potenzialità di interoperabilità tra l'aggregatore e le femtocelle di manifatturieri differenti e ha fornito la base per la definizione della soluzione standard prevista nella release 8. Il limite dell'interfaccia *lu* è costituito dall'assenza di soluzioni di controllo specifiche della femtocella, quali ad esempio la necessità di una registrazione del dispositivo alla rete ad ogni sua accensione.

Un'alternativa nello stesso filone architetturale che integra le funzioni di RNC nella femtocella è mostrata nella figura 8 come architettura "3". Essa si basa sull'UMA (*Unlicensed Mobile Access*)² ed è una soluzione definita dal 3GPP mirata ad assicurare l'interope-

² Detta anche GAN (*Generic Access Network*).

IMS è l'architettura studiata per fornire ai clienti servizi voce, video e multimediali attraverso molteplici reti d'accesso IP, ivi incluse le tradizionali reti fisse e mobili. Con l'IMS l'utente riceve il servizio indipendentemente dalla propria ubicazione fisica, dalla tecnologia di accesso e dal tipo di terminale, purché esso sia dotato di connettività IP. Lo standard è basato sul protocollo SIP (*Session Initia-*

■ **HNB (femtocella)**. La stazione radio, connessa a un servizio di accesso residenziale a larga banda, assicura la copertura in un'area limitata ai terminali rispondenti allo standard 3G. Un HNB include le funzionalità di un No-

49

de-B standard integrate con quelle di gestione delle risorse radio proprie di un RNC.

□ **HNB-GW (aggregatore di femtocelle).** Il gateway, installato nella rete dell'operatore, aggrega e disaggrega sia il traffico sia la segnalazione da e verso un gran numero di HNB e verso la rete di commutazione CN (Core Network), a cui è connesso mediante le due interface standard I_{u-CS} e I_{u-PS} .

Per interconnettere questi due elementi di rete è stata definita l'interfaccia I_{u-h} che effettua il relay del protocollo di controllo delle risorse radio sull'interfaccia I_u , RANAP (Radio Access Network Application Part) tra la rete di commutazione CN e il nodo periferico HNB e introduce il protocollo HNBAP (Home Node-B Application Protocol) per gli aspetti di controllo specifici dell'HNB (HNB e UE Registration).

5. GESTIONE DELLA RETE DI FEMTOCELLE

Il nuovo paradigma delle femtocelle ha comportato di riconsiderare un insieme di caratteristiche di sistema, quali quelle relative alla gestione delle operazioni e della manutenzione (O&M), oltre ad aspetti di RRM (Radio Resource Management) e di MM (Mobility Management).

5.1. Gestione a distanza della femtocella

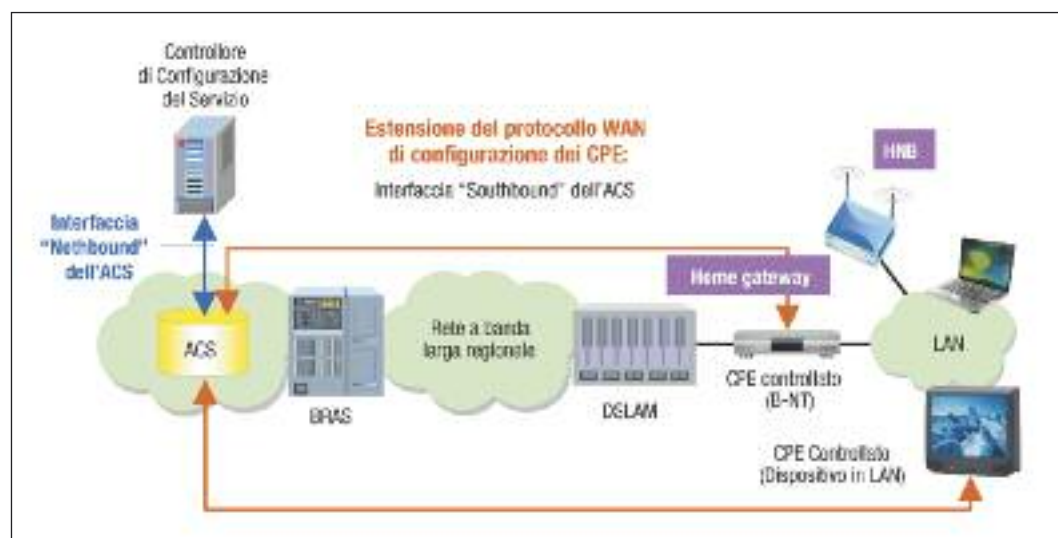
L'approccio alla gestione O&M (Operations and Maintenance) è uno degli elementi che caratterizzano la rete di femtocelle rispetto ai sistemi cellulari tradizionali. La femtocella è,

infatti, anche un dispositivo di utente e, come un modem xDSL, è controllabile da remoto per mezzo di criteri standardizzati, indicati in una specifica tecnica del Broadband Forum (TR-069). La specifica definisce il protocollo Internet CWMP (CPE WAN Management Protocol), che fornisce le funzioni essenziali per gestire a distanza i terminali di utente. In essa sono indicate tutte le caratteristiche necessarie fra cui la gestione della configurazione, l'aggiornamento del software, il monitoraggio delle prestazioni, la diagnostica e la risoluzione dei guasti. La soluzione TR-069 è oggi già diffusa presso tutti i principali operatori europei e del Nord America per la gestione degli *home gateway* connessi tramite le terminazioni xDSL.

La figura 10 mostra l'architettura di gestione: in essa un apposito sistema server di O&M dell'operatore, detto ACS (AutoConfiguration Server) nella terminologia del Broadband Forum, comunica direttamente con l'home gateway e con i dispositivi ad esso connessi, inclusa la femtocella, per svolgere tutte le operazioni di gestione necessarie.

Le funzioni di autoconfigurazione delle femtocelle includono sia la scelta dei parametri di trasmissione, di controllo della mobilità e degli accessi, sia la gestione dell'insieme di celle da monitorare. I parametri di trasmissione di una femtocella non possono essere fissati a priori e configurati manualmente dall'operatore ma devono essere scelti automaticamente durante l'installazione che è generalmente avviata dal cliente e che è effettua-

FIGURA 10
Telegestione
della femtocella
con il CWMP



ta, sempre da parte dell'utilizzatore, ad ogni accensione del terminale.

A tale scopo una femtocella UMTS può includere un ricevitore standard dedicato, detto MMR (*Macro-Mode Receiver*), che riceve le informazioni di cella, inviate in *broadcast* dalla macrocella sovrastante. In analogia a quanto accade per i terminali mobili nello stato dormiente (*idle mode*), la femtocella è quindi in grado di misurare il livello di segnale e di interferenza delle celle circostanti. Grazie a questa caratteristica è possibile introdurre nelle femtocelle opportuni algoritmi di regolazione automatica della potenza che consentono di ottenere il corretto compromesso tra la minimizzazione dell'interferenza isofrequenza, generata dalla femtocella sulla rete macrocellulare sovrastante, e una copertura appropriata dell'appartamento. Questi algoritmi consentono anche di controllare l'interferenza presente tra celle adiacenti nel caso di un'alta densità di femtocelle installate in una determinata area.

5.2. Gestione della risorsa radio

In un sistema cellulare la gestione delle risorse radio persegue l'obiettivo di garantire la qualità del servizio e, allo stesso tempo, di controllare le prestazioni di sistema, in termini ad esempio, di ottimizzazione della capacità e di controllo dell'interferenza. In questa funzione sono comprese le strategie e gli algoritmi atti a controllare parametri quali la potenza trasmessa, l'assegnazione dei canali, i criteri di handover, gli schemi di modulazione e quelli relativi al controllo di errore. Per comprendere la specificità della gestione degli aspetti radio in una rete di femtocelle, da un lato va considerata la libertà di cui gode l'utente nel collocare il proprio terminale HNB nel punto che preferisce in casa, dall'altro quella dell'operatore che può stabilire diverse condizioni operative della femtocella. L'operatore ha, infatti, a disposizione varie opzioni di pianificazione dello spettro radio dello strato di rete di femtocelle. In generale, si impiegano tre soluzioni:

- è possibile attribuire alle femtocelle una frequenza dedicata;
- per le femtocelle possono essere utilizzate le stesse frequenze già assegnate alla rete di macrocelle;

• può essere usata una frequenza in comune fra lo strato di femtocelle e quello di macrocelle e una riservata allo strato di macrocelle, prevedendo che, nel caso di interferenza a un utente della macrocella sia imposto di effettuare l'handover sulla frequenza riservata a questo strato (Figura 11).

Un aspetto collegato è che, a seconda della scelta dell'operatore, l'accesso all'HNB può essere aperto, ossia senza preclusioni per qualsiasi cliente che si trovi nell'area di copertura, oppure può essere ristretto a un CSG (*Closed Subscriber Group*) e cioè a un gruppo chiuso di clienti registrati nella femtocella [3].

I vari scenari d'uso che discendono dalle combinazioni delle scelte operative descritte hanno numerose implicazioni sia per l'operatore - in termini di capacità e di gestione di rete e di servizio - sia per il cliente, in termini di qualità e funzionalità. Qui di seguito si approfondisce il controllo dell'interferenza, ritenuto un aspetto critico per il sistema di femtocelle.

La rete di accesso UTRAN (*UMTS Radio Access Network*) è stata realizzata sulla base del criterio dello sviluppo di una rete ordinata e controllata; lo strato di femtocelle contiene invece un numero molto elevato di HNB tra loro non coordinati, che può generare livelli dannosi di interferenza.

Possono essere distinti tre tipi di interferenza: quella "femto-macro" prodotta dallo strato di femtocelle verso lo strato di macrocelle, quella "macro-femto" in senso opposto e, infine, quella "femto-femto" interna allo strato di femtocelle.

Fra questi tipi di interferenza, almeno nella prima fase di sviluppo delle reti di femtocelle, riveste un interesse primario l'interferenza tra stra-

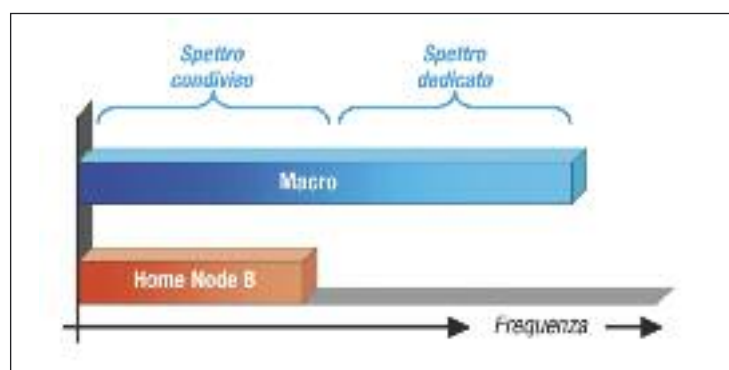
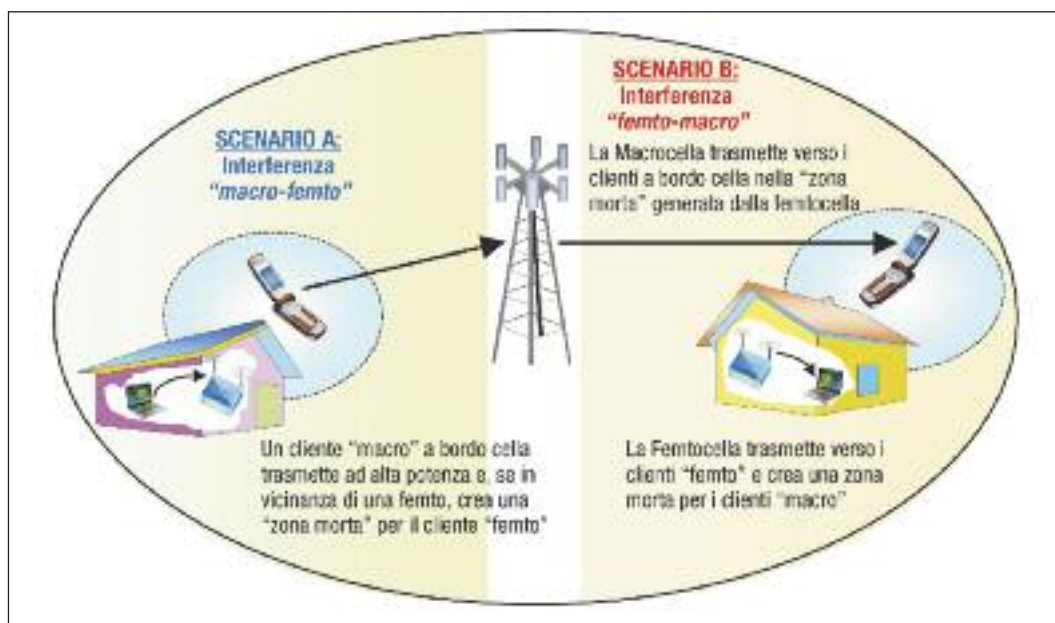


FIGURA 11

Uso dinamico delle frequenze della macrocella

FIGURA 12
Zone morte
per effetto
dell'interferenza
tra strati



ti diversi (ossia femto-macro e macro-femto). Nelle reti cellulari in tecnologia W-CDMA (*Wideband-asynchronous Code Division Multiple Access*), come il sistema UMTS/HSPA, può infatti presentarsi il problema *near-far*, che consiste nell'accecamento dei terminali collegati a stazioni radio base lontane, da parte di altre più vicine, alle quali tuttavia quei terminali non sono collegati. Nell'UTRAN, per compensare l'attenuazione di percorso e il fading lento e veloce (che determinano il fenomeno di *near-far*), si impiegano specifiche soluzioni di controllo della potenza e in particolare il cosiddetto *fast power control*.

Quando alla rete preesistente di macrocelle si aggiunge uno strato di femtocelle e si adotta un profilo di servizio di tipo CSG (il servizio è ristretto, come si è detto, a un gruppo chiuso e il cliente macro non può accedere allo strato femto), il controllo di potenza può generare "zone morte" (Figura 12) sia per i terminali che utilizzano la copertura con le femtocelle sia per quelli collegati alla copertura macrocellulare. Il fenomeno si caratterizza differentemente nei collegamenti in salita UL (*Up Link*) e in discesa DL (*Down Link*) [4]:

a. Interferenza "macro-femto" e zona morta in UL

I terminali collegati a una macrocella e collocati al bordo dell'area da essa servita trasmettono a potenza elevata nel collegamento in UL, causando un'interferenza inaccetta-

bile alle femtocelle vicine. Di conseguenza le femtocelle che si trovano sul bordo dell'area coperta subiscono un'interferenza sensibilmente maggiore rispetto a quella che si ha nelle femtocelle collocate all'interno della macrocella. In altri termini l'aggressore è il singolo terminale della macrocella e la vittima è la tratta Up Link della femtocella.

b. Interferenza "femto-macro" e zona morta in DL

I terminali collegati a una macrocella e collocati al bordo dell'area coperta sono affetti da valori maggiori di attenuazione rispetto a quelli che si trovano più all'interno della macrocella e sono pertanto interferiti nel collegamento in DL dalle trasmissioni delle femtocelle ad essi vicine. In altri termini, in questo caso l'aggressore è la femtocella e la vittima è la tratta Down Link della macrocella.

Per combattere l'interferenza che può insorgere tra i due strati cellulari, si possono ipotizzare alcune soluzioni operativamente semplici ma non sempre praticabili. Come si è prima accennato, possono essere per esempio impiegate frequenze differenti per i due strati, adottando così una soluzione che elimina alla radice le zone morte, ma che, oltre a non essere percorribile per gli operatori che dispongono di una sola portante UMTS, obbliga alla ripianificazione della rete macrocellulare e in più comporta una riduzione di capacità. Per gli operatori che dispongono di almeno due ban-

de UMTS, una diversa soluzione impiegabile è quella dell'uso dinamico delle frequenze, con il trasferimento della comunicazione attiva entro la macrocella sulla seconda portante non condivisa con le femtocelle, in presenza di un'interferenza determinata nella femtocella (Figura 11). Infine, per gli operatori 2G/3G è possibile ricorrere alla limitazione della potenza della femtocella per ridurre la zona morta in DL e alla contestuale adozione del GSM in sostituzione dell'UMTS in tale area.

Si ritiene comunque utile che ciascuna femtocella sia in grado di monitorare l'ambiente a radiofrequenza e di adattare e cambiare autonomamente i parametri di trasmissione (potenza, frequenza, codice, ritmo di trasmissione) per evitare, o quantomeno per contenere, l'interferenza verso le macrocelle e le femtocelle adiacenti [5]. Per ridurre l'interferenza tra due femto, quando un utilizzatore vicino ne installi una nuova, le altre femtocelle presenti nell'edificio possono dover riadattare i propri parametri di trasmissione. Può essere necessario introdurre le varianti anche dinamicamente: in linea di principio, infatti, se due femtocelle interferiscono, esse possono rivelare reciprocamente la presenza dell'altro terminale e modificare i propri parametri di trasmissione in qualsiasi momento.

Recenti analisi effettuate in ambito 3GPP [6] hanno mostrato come intorno alla femtocella si possano avere notevoli problemi d'interferenza laddove la potenza della macrocella sia ricevuta sensibilmente attenuata e quando la semplice limitazione della potenza in uscita dalla stazione radio domestica non possa essere considerata una soluzione soddisfacente, poiché limita l'area di copertura. Questi studi hanno, perciò, confermato l'utilità di tecniche adattative di potenza basate sulla stima dell'interferenza.

5.3. Le strategie di gestione della mobilità in presenza di femtocelle

L'obiettivo principale della gestione della mobilità, detta anche MM (*Mobility Management*), rappresenta l'aggiornamento in una rete radiomobile, con precisione adeguata, dell'informazione di posizione in rete dei clienti, per poter loro assicurare correttamente la fornitura dei servizi in mobilità. La mobilità è gestita mediante procedure di riselectio-

ne di cella e di aggiornamento dell'area di localizzazione, LA (*Location Area*), che include di norma più celle. L'aggiunta delle femtocelle nella rete mobile fa nascere la domanda sulla convenienza di gestire lo strato di femtocelle come un dominio a sé stante o in modo integrato con il preesistente strato di macrocelle. In una rete mobile i principali codici usati per realizzare la localizzazione del terminale sono tre: l'MNC (*Mobile Network Code*), che identifica la rete mobile pubblica (PLMN); la LAC (*Location Area Code*) che individua in modo univoco l'area di localizzazione del terminale entro la PLMN e, infine, il CID (*Cell ID*) che identifica la cella all'interno della LA. I codici identificativi sono diffusi periodicamente dalle stazioni radio base. L'operatore mobile potrebbe decidere di associare alle femtocelle lo stesso identificativo MNC scelto per la rete di macrocelle; in questo caso, per discriminare se un utente si trova nella copertura delle macrocelle o delle femtocelle, è necessario associare a queste ultime un insieme esclusivo di LAC, da gestire con un processo di pianificazione separato. In alternativa l'operatore può impiegare, se disponibile, un differente identificativo MNC per le femtocelle, che consente di attuare diverse strategie di servizio, introducendo così un grado di flessibilità in più e riducendo potenzialmente anche la complessità a livello radio, ma al costo di una maggiore complessità della rete.

Una soluzione che consente la selezione delle femtocelle e la verifica dell'autorizzazione all'accesso, consiste nell'assegnare una LA a ciascuna femtocella. Ogni LAC corrispondente a una femtocella deve essere distinta sia da quella a cui appartiene la macrocella sovrapposta sia dalle LAC delle femtocelle adiacenti. Tuttavia, poiché il numero potenziale di femtocelle in rete potrebbe essere molto grande e i codici LAC hanno una codificazione concepita per associare gli elementi ad aree geografiche di dimensioni molto maggiori dell'area di copertura di una femtocella, è sempre necessario prevedere il riuso a distanza dei codici LAC. In genere il nodo aggregatore delle femtocelle (HNB Gateway nella Figura 13) svolge anche la funzione di nascondere alla rete la complessità legata all'elevato numero di LAC, traducendo ad esempio i LAC associati alle femtocelle in un unico LAC mostrato alla rete

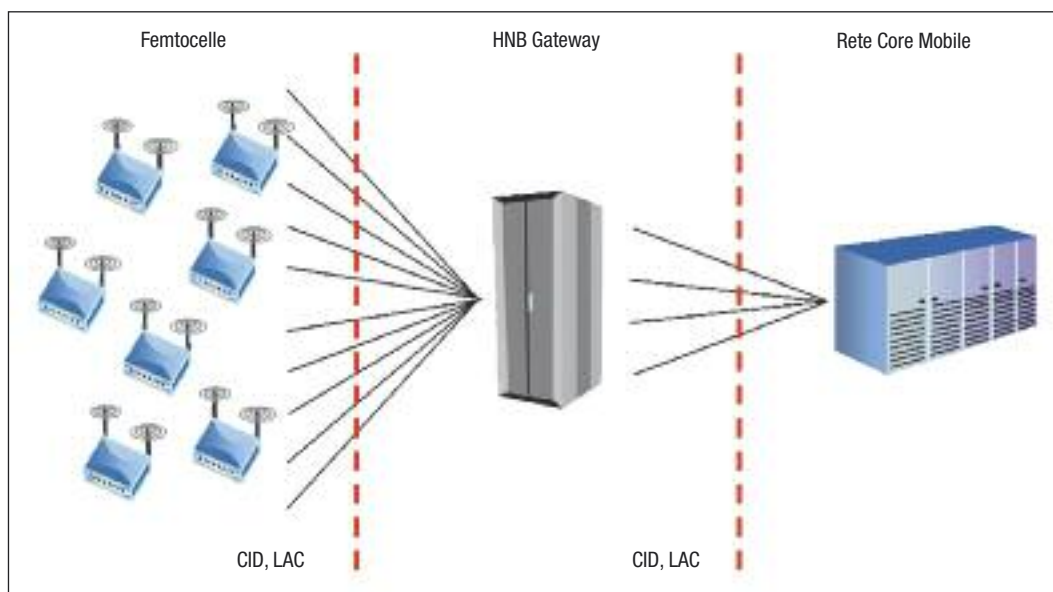


FIGURA 13
Aggregatore
di femtocelle

di commutazione. Si semplificano così sia la configurazione dei nodi di commutazione sia le procedure come il paging.

Nel caso di uno scenario di accesso esclusivo, la femtocella o l'aggregatore, a seconda della particolare soluzione adottata, memorizza una lista di clienti abilitati all'accesso alla femtocella, detta *white list*, sulla base dell'identificativo IMSI (*International Mobile Subscriber Identity*). Quando viene avviato un aggiornamento dell'area di localizzazione, LAU (*Location Area Update*), per poter accedere alla femtocella, la rete chiede contestualmente al terminale l'invio del proprio IMSI. L'identificativo è quindi intercettato dalla femtocella o dall'aggregatore ed è incrociato con l'insieme degli IMSI memorizzati nella *white list*. La funzione di aggiornamento della localizzazione entro la femtocella ha pertanto esito positivo solo se vi è rispondenza tra i dati e, in tal caso, l'utente è accettato.

Per semplificare gli aspetti di configurazione dei LAC e l'attività di aggiornamento della mobilità da parte dei terminali, il 3GPP ha definito nella release 8 soluzioni più efficienti, basate sulla trasmissione in aria da parte della femtocella di un codice di identificazione specifico (*CSG Identifier*). I terminali futuri, realizzati nel rispetto di queste norme avranno al proprio interno la configurazione della lista di *CSG Identifier* cui possono accedere ed effettueranno quindi una procedura di *Cell Reselection*, quando l'identificativo annunciato dalla fem-

tocella coincide con quello configurato sul terminale. Questa soluzione consente di ridurre le procedure di mobilità da parte del terminale, con un risparmio conseguente della carica delle batterie, e rimuove l'obbligo di assegnare un LAC per femtocella, riducendo così in misura significativa l'attività di configurazione e gestione da parte dell'operatore mobile.

6. ALCUNI PROBLEMI TUTTORA ALLO STUDIO

I problemi tecnologici oggi ancora aperti riguardano aspetti del sistema che spaziano dal sottosistema O&M alla gestione della risorsa radio, alla gestione della mobilità. Qui di seguito sono elencati alcuni tra i principali problemi da affrontare e risolvere nel prossimo futuro.

6.1. Sincronizzazione

La sincronizzazione è alla base della corretta gestione delle procedure sia a livello RRM (*Radio Resource Management*) che a livello MM (*Mobility Management*), in particolare di quelle di handover. Secondo le specifiche 3GPP, per rispondere agli standard delle maschere spettrali prescritte per le stazioni radio base, l'accuratezza sulla frequenza che queste generano deve essere di almeno 50 parti per miliardo (0,05 ppm). Per rispettare questi valori, anche su tempi lunghi, occorre utilizzare oscillatori a cristallo stabilizzati o controlla-

ti in temperatura, che richiedono calibrazioni esterne per poter correggere periodicamente (con una frequenza compresa fra un mese e un anno) la frequenza, che è soggetta a deriva a causa dell'invecchiamento del cristallo. Nelle stazioni radio base convenzionali la calibrazione periodica è realizzata tramite GPS o, più spesso, la sincronizzazione è estratta dal livello fisico del trasporto TDM. Nel caso delle femtocelle entrambe le soluzioni non sono utilizzabili, poiché la copertura GPS può essere insufficiente all'interno delle abitazioni e l'impiego di una rete IP per il trasporto implica una struttura asincrona differente rispetto a quella di una rete TDM. Il 3GPP nella release 6 dello standard ha pertanto specificato un valore di 0,1 ppm per l'accuratezza della frequenza per le stazioni radio collocate in zone chiuse e - tenendo presente le ulteriori difficoltà che si incontreranno con le femtocelle - ha introdotto nella release 8 un ulteriore rilassamento dei limiti normalizzati per i terminali HNB (0,25 ppm). Questo rilassamento è ritenuto accettabile in considerazione della ridotta velocità nello spostamento dei terminali da o verso una femtocella.

6.2. Controllo dell'accesso

Come si è chiarito nel paragrafo 5.3, il 3GPP ha definito soluzioni per il controllo della mobilità e dell'accesso alle femtocelle, ottimizzate per ridurre l'attività di segnalazione dal terminale e quella di gestione da parte dell'operatore. Le soluzioni saranno però disponibili solo dalla release 8 e non saranno perciò utilizzabili per un certo tempo dalla maggior parte dei terminali. Inoltre, anche se i terminali fossero disponibili fra breve, gli operatori dovrebbero comunque gestire entrambe le soluzioni per consentire la mobilità e l'accesso sia ai clienti con terminali rispondenti alle nuove specifiche sia a quelli con terminali conformi alle specifiche precedenti.

6.3. Qualità del servizio e confini di responsabilità tra gestore fisso e mobile

Nella tradizione dei sistemi radiomobili - prima per il GSM e poi per l'UMTS - la rete cellulare di proprietà dell'operatore mobile è del tutto autonoma e non richiede necessariamente l'impiego di parti delle reti di telecomu-

nicazioni fisse. Quando questa esigenza si è presentata, il trasporto è stato svolto con flussi trasmissivi dedicati. Se però si dovesse impiegare la connessione xDSL del cliente per il servizio di backhaul, la femtocella potrebbe dover competere nell'uso della banda con altri utilizzatori domestici. E nel caso di condivisione dei flussi trasmissivi si presenterebbe il problema della gestione della qualità del servizio. Il problema sarebbe di maggior rilievo nel caso in cui il gestore di rete fissa fosse differente da quello della rete mobile. Le soluzioni utilizzabili in questo caso consisterebbero nella separazione dei diversi flussi di traffico e in una caratterizzazione di essi secondo classi differenti per la qualità del trasporto, nell'ambito di quanto consentito dalla tecnologia di trasporto utilizzata. Le soluzioni tuttavia non eludono la necessità di predisporre maggiori risorse di trasporto sull'accesso xDSL. Infine, nel caso in cui i provider dell'accesso alle femtocelle fossero differenti da quelli dell'accesso xDSL, occorrerebbe anche stabilire a chi è demandata la gestione dei guasti.

6.4. Servizi di emergenza

La fornitura del servizio di emergenza richiede di assicurare la localizzazione dell'utente e debbono quindi essere implementate soluzioni che consentano di localizzare la femtocella con una precisione, che potrebbe essere anche particolarmente elevata. Questo livello di accuratezza potrebbe richiedere lo sviluppo di soluzioni di identificazione delle femtocelle opposte rispetto all'esigenza dell'operatore, prima chiarita, di mascherare la numerosità delle terminazioni alla rete di commutazione. Potrebbe, in più, essere necessario lo sviluppo di procedure specifiche per la localizzazione del terminale.

6.5. Handover

Finora è disponibile nel mercato l'handover per il passaggio dalla femtocella alla macrocella, mentre non è consentito nella direzione opposta o nel passaggio da una femtocella ad un'altra. Sebbene questi aspetti non pregiudichino l'erogazione del servizio negli scenari di copertura più comuni, sono in via di approntamento le soluzioni non ancora messe a punto in via definitiva.

7. CONCLUSIONI

Le femtocelle rappresentano una tecnologia nuova che permette di realizzare un cambio di paradigma nell'accesso radio. Esse sono state concepite per migliorare prestazioni e qualità del servizio offerto nelle abitazioni. Una rete di femtocelle presenta il vantaggio, rispetto alle reti di macrocelle, di rendere utilizzabile in maniera esclusiva la capacità radio, disponibile in ogni tipo di ambiente e di dedicarla a un numero ristretto di utenti, dando la possibilità di accedere alla capacità di picco diversamente da una rete di macrocelle che non potrebbe offrire più di una capacità media di traffico. Grazie all'introduzione delle femtocelle si può quindi conseguire un incremento rilevante nel bit rate disponibile per utente. Inoltre, *l'assegnazione dedicata delle risorse radio e la conoscenza esatta della posizione dei clienti consentono, infine, di offrire anche i servizi tradizionali (chiamate foniche ed SMS) secondo profili di servizio speciali in termini di tariffazione.*

In più, come si è messo in luce nell'articolo, sono molteplici i benefici che portano oggi a una diffusa aspettativa che la femtocella possa divenire un componente potenzialmente abilitante dello sviluppo dell'ecosistema domestico della banda larga, offrendo una porta d'accesso - eventualmente in soluzioni integrate con altre tecnologie wireless - a numerosi apparecchi senza il vincolo della connessione cablata. La femtocella si presta a essere un elemento idoneo a consentire una completa integrazione degli ambienti domestici e SOHO nelle reti di telecomunicazioni e un importante driver di mercato per la promozione dell'ecosistema della larga banda, tanto da poter essere considerato un elemento essenziale per accelerare l'impiego delle tecnologie previste per l'edificio intelligente.

La consapevolezza delle molteplici potenzialità insite in questa tecnologia ha spinto di recente il 3GPP ad accelerare il processo di standardizzazione, che ha prodotto alla fine del 2008 il primo blocco di specifiche di sistema per le femtocelle.

Come si è visto nell'articolo, con le femtocelle le funzionalità di gestione della rete radio e della mobilità si spostano in prossimità dell'utente finale e risultano assai più distribui-

te che nelle configurazioni di rete tradizionali. Questa trasformazione dell'accesso è in linea con la tendenza tipica dei moderni sistemi di telecomunicazioni, che vedono spostarsi la capacità elaborativa dei sistemi verso la periferia della rete.

Ma, se le femtocelle aprono nuove opportunità nello sviluppo delle tecnologie 3G e LTE e si prestano anche ad integrarsi nello sviluppo della rete di nuova generazione (la NGAN), esse comportano anche nuove sfide in relazione all'intrinseca natura fortemente distribuita delle funzioni allocate nei terminali remoti. Alcuni problemi posti dalle femtocelle richiedono un'attenta considerazione e le scelte non sembrano essere state ancora del tutto individuate. Com'è stato sottolineato in questo articolo, su questi temi occorrerà che si concentri nel prossimo futuro l'attenzione degli Organismi di standardizzazione e gli studi dei laboratori di ricerca e di sviluppo per definire i punti ancora in sospeso e, soprattutto consentano di realizzare soluzioni allo stesso tempo efficienti e compatibili con l'impiego dei terminali maggiormente impiegati dall'utenza potenziale.

Ringraziamenti

Si ringraziano per l'accurata opera di revisione ed i preziosi suggerimenti Andrea Buldorini, Enrico Grosso, Andrea Laganà, Michele Ludovico, Alesandro Vaillant, Franca Venir, Francesco Vatalaro (Università di Roma "Tor Vergata").

Bibliografia

- [1] Isenberg D.: *The Rise of the Stupid Network*. Computer Telephony, August 1997, p. 16 e successive. Web: <http://www.hyperorg.com/misc/stupidnet.html>
- [2] ABI Research: *Femtocell Market Challenges and Opportunities - Cellular-Based Fixed Mobile Convergence for Consumers, SMEs and Enterprises*, 2Q 2007. Web: http://www.abiresearch.com/products/-market_research/Femtocell_Market_Challenges_and_Opportunities
- [3] 3GPP Doc. TR 25.820 V8.0.0 (2008-03): *Technical Report. 3-rd Generation Partnership Project; Technical Specification Group Radio Access Networks; 3G Home NodeB Study Item Technical Report (Release 8)*, 2008.

- [4] Chandrasekhar V., Andrews J.G., Gatherer A.: *Femtocell Networks: A Survey*. June 23, 2008. Web: <http://arxiv.org/ftp/arxiv/papers/0803/0803.0952.pdf>
- [5] Claussen H., Ho L.T.W., Samuel L.G.: *Self-optimization of Coverage for Femtocell Deployments*. 24-26 April 2008, Wireless Telecommunications Symposium (WTS 2008), p. 278 285.
- [6] 3GPP Doc. R4-071660 - Ericsson: *Impact of HNB with fixed output power on macro HSDPA capacity*. 3GPP TSG-RAN Working Group 4 (Radio), Shanghai, October, 2007.
- [7] ABI Research: *Super Femtocells The Need for a Higher Capacity Femtocell*. 4Q 2008. Web: http://www.abiresearch.com/products/research_brief/Femtocell_Research_Briefs/109/Super+Femtocells#toc
- [8] I'm cellular: *Femtocell Trials and Deployments (to date) capacity*. January 11 2009. Web: <http://imcellular.org/2009/01/11/femtocell-trials-and-deployments-to-date/>

ROCCO CASALE, dopo aver conseguito la laurea cum laude e dopo una breve esperienza in una società manifatturiera di sistemi elettronici, ha svolto la sua attività professionale in SIP (divenuta poi Telecom Italia) dove si è occupato con responsabilità crescenti - fino a divenirne responsabile nel 1985 - di ingegneria delle innovazioni tecnologiche e dell'introduzione in rete di sistemi di telecomunicazione innovativi. Dallo stesso anno ha curato la qualità entrante di tutti i prodotti acquistati da Telecom Italia. Dal 1995 ha diretto per nove anni il Notiziario Tecnico Telecom Italia. Ha svolto un'intensa attività internazionale in UIT, CEPT (oggi ETSI), IEC. Ha presieduto per tre anni il Comitato Tecnico dello CELT (oggi TiLab) ed è stato membro del Comitato tecnico della FUB. Dal 2006 collabora con il Consorzio Università Industria Radio Labs. Oggi si occupa prevalentemente di problemi legati all'accesso alle reti di telecomunicazione e in particolare di quelli riguardanti l'introduzione nelle reti di telecomunicazione dei sistemi di nuova generazione NGAN. Si interessa anche di nuovi sistemi di alimentazione a idrogeno, a impatto quasi nullo sull'ambiente, alternativi alle batterie impiegate nelle centrali telefoniche.
E-mail: rocco.casale@alice.it

GIUSEPPE CATALANO ha ricevuto la laurea in Ingegneria delle Telecomunicazioni presso l'Università di Pisa nel 1998, anno in cui è stato assunto in CELT (Centro Studi e Laboratori Telecomunicazioni), ora TILAB (Telecom Italia Lab). La sua attività si è focalizzata sull'evoluzione dei sistemi radiomobili, quali GPRS, UMTS ed LTE, avendo partecipato alle fasi di standardizzazione presso il 3GPP e di testing e deployment per Telecom Italia ed alcune società controllate del Gruppo. Ha svolto un'intensa attività didattica in corsi destinati al personale del Gruppo ed è stato curatore e co-autore del libro "GPRS - Accesso Radio, Architettura di Rete, Protocolli e Servizi" edito da TILAB. È oggi delegato di Telecom Italia presso il 3GPP TSG RAN WG3.
E-mail: giuseppe.catalano@telecomitalia.it

DANIELE FRANCESCHINI è nato a La Spezia il 21/10 1971. Si è laureato in Ingegneria Elettronica indirizzo telecomunicazioni nell'aprile del 1997. Nello stesso anno ha iniziato la sua carriera lavorativa in CELT (Centro Studi e Laboratori Telecomunicazioni) centro di ricerca e sviluppo del gruppo Telecom Italia in Torino, dove si è focalizzato sull'analisi delle prestazioni della rete GSM/GPRS mediante simulazioni di sistema. Dal 1998 ha seguito attivamente il processo di standardizzazione UMTS, rappresentando Telecom Italia nei gruppi ETSI SMG2 L2&3 e successivamente 3GPP TSG RAN WG2 e WG4. Da Marzo 2000 ha continuato la sua carriera lavorativa in Omnitel dove ha seguito vari aspetti relativi alla progettazione radio UMTS. Dal 2001 è ritornato in Telecom Italia dove ha seguito vari aspetti relativi all'evoluzione della rete di accesso radio e focalizza le sue attività sull'evoluzione delle tecnologie radio beyond 3G (e.g. HsxPA ev., OFDMA, LTE).
E-mail: daniele.franceschini@telecomitalia.it

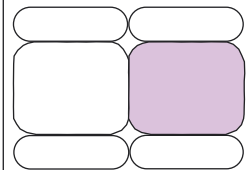
SANDRO DIONISI ha ricevuto la laurea cum laude in Ingegneria Elettronica presso l'Università "La Sapienza" di Roma. Ha oltre 25 anni di esperienza in vari campi delle telecomunicazioni fisse e mobili. Dal 1983 è in Telecom Italia (allora SIP), dove ha lavorato in diverse aree di Rete, ricoprendo vari livelli di responsabilità. Nel 1999 è stato nominato responsabile della Rete di Accesso e di Trasporto Fissa e nel 2002 è divenuto responsabile di un'Area Territoriale di Rete; nel 2003 ha ricevuto l'incarico di coordinare le Infrastrutture e le *Operations* della Rete Internazionale di Telecom Italia per i servizi fissi e mobili per l'Europa e l'America Latina. Ha partecipato attivamente a diversi *forum* e gruppi di standard internazionali in ITU ed ETSI dove, nel periodo dal 1990 al 1996, ha guidato il gruppo responsabile degli standard di radio relay. Da febbraio del 2008 è Direttore di Telecom Italia Lab ed è responsabile delle attività di Ricerca, Innovazione ed Ingegneria per le reti fissa e mobile del gruppo Telecom Italia.
E-mail: sandro.dionisi@telecomitalia.it

L'ICT PER L'EFFICIENZA E IL RISPARMIO ENERGETICO



Paolo Magni

L'articolo concentra l'attenzione su una serie di strumenti ICT utilizzati per controllare e ridurre i consumi energetici e per gestire in modo efficiente ed efficace dal punto di vista delle spese, le risorse energetiche distribuite, che costituiscono un mercato e una tecnologia in forte sviluppo. Viene fornito un quadro sintetico delle proprietà di un sistema per l'*Energy Resource Planning* - denominato "EneRP" - riportando le caratteristiche di alcune soluzioni già disponibili sul mercato. Si fa riferimento, come esempio, a due progetti finanziati nell'ambito dei bandi Meta-Distretti Lombardia ICT.



1. INTRODUZIONE

L'energia sul nostro pianeta non è una risorsa scarsa. Può essere costosa in talune nazioni, può essere causa di conflitti economici e militari, il suo uso o abuso, può determinare mutamenti climatici e ambientali di cui potremmo non essere del tutto consci, ma sicuramente l'energia non è una risorsa scarsa. I problemi tecnologici relativi all'energia sono in limitata parte legati al reperimento e all'estrazione in confronto a stoccaggio, trasporto e utilizzo. Altro discorso è lo spreco che si fa di questa risorsa.

L'unità di misura definita come "barile-di-petrolio" è un cilindro metallico contenente 159l di una sostanza scura, densa e maleodorante che altro non è che quanto resta della decomposizione di organismi vissuti sulla terra qualche milione di anni or sono. Nonostante ciò, si è scoperto che si tratta di una materia prima fantastica, da essa ricaviamo:

- ▣ combustibili che utilizziamo per fornire città e industrie di calore ed energia elettrica;
- ▣ combustibili per auto, navi e aerei facili da trasportare e con un'elevata densità di ener-

gia per unità di volume, e per lo più liquidi a condizioni ambiente;

- ▣ materie plastiche per gli usi più svariati;
- ▣ scarti che utilizziamo come bitume per realizzare le strade.

Anche il costo reale di estrazione è irrisorio: 10-15\$/barile per quello del Mare del Nord.

Il petrolio sarebbe assolutamente perfetto, se non fosse per alcune controindicazioni:

- ▣ i giacimenti, per quanto enormi, non si trovano distribuiti uniformemente, il che porta ad "attriti" tra chi ne ha grandi disponibilità, e non sa bene cosa farsene, e chi non ne ha, e saprebbe bene come utilizzarlo;

- ▣ l'utilizzo come combustibile comporta lo sprigionamento di enormi quantità di gas inquinanti, nocivi e con effetto serra, con conseguenze a medio e lungo termine sul clima globale e a breve e medio termine sulla salute della popolazione, soprattutto di quella (la maggior parte) che vive nelle città;

- ▣ tecnologie ed economie *oil-based* non sono più sostenibili, o meglio, non sono sostenibili nel momento in cui lo stile di vita e di sviluppo utilizzato dai paesi occidentali nel-

l'ultimo secolo venisse ripreso pari pari da quei miliardi di persone che nell'ultimo secolo ne sono state escluse.

Al di là dell'amara ironia, la questione energia è ormai chiara: bisogna consumare meno e meglio, evitare gli sprechi, utilizzare fonti energetiche alternative dal minimo o nullo impatto ambientale, rinnovabili e disponibili su larga scala un po' ovunque.

Una volta preso atto di questa necessità rimane da risolvere il come. È a questo punto che l'ICT, nella sua definizione originale di Information and Communication Technology, può e deve fornire il suo contributo, mettendo a disposizione un insieme di hardware (principalmente sensori e attuatori, oltre ad apparati di trasmissione dati) e software (applicativi ed algoritmi) per la gestione - intesa come controllo, manutenzione ed ottimizzazione - delle fonti di energia, rinnovabili e non, e degli utilizzatori finali, per garantire una gestione energetica globalmente più efficiente ed efficace.

2. IL MERCATO DELLE ENERGIE RINNOVABILI NEL PROSSIMO FUTURO

I dati circa la crescita del mercato internazionale delle industrie del fotovoltaico, solare termico, eolico e geotermico (senza contare gli elettrodomestici di ultimissima generazione, classe A e oltre) sono sotto gli occhi di tutti e impressionanti: crescita a due cifre, con punte di nazioni o segmenti con crescita a tre cifre, negli ultimi tre anni. Questa esplosione, da molti considerata fino a qualche mese fa come precursore di un nuovo *boom* stile "*new economy*", risentirà inevitabilmente degli effetti della crisi globale. Tuttavia, si ritiene che la crescita, anche se rallentata, rimarrà ben superiore al resto dell'industria, per così dire, tradizionale.

I motivi di tale situazione dipendono dal fatto che:

□ i consumatori (ed in molti casi anche i governi) delle nazioni con maggior capacità di spesa hanno sostanzialmente metabolizzato il concetto di energia come risorsa scarsa, preziosa, da utilizzare con giudizio, la cui gestione può a sua volta generare business, per di più etico [1];

□ le industrie dei paesi che si sono attrezzati per primi, si pensi ad esempio alla Germania, alla Spagna e al Giappone per il fotovoltaico, o alla Danimarca per l'eolico, hanno ormai accumulato un *know-how* e una dimensione tale da poter agire anche a livello internazionale sui mercati extradomestici (tra cui l'Italia) proponendo e utilizzando tecnologie affidabili e politiche commerciali aggressive;

□ finanziatori sempre in cerca di valide possibilità di investimento con una rendita ragionevole e rischi prevedibili non potranno restare indifferenti a lungo alle promesse delle energie rinnovabili. Non appena si placherà la crisi torneranno a fluire i capitali e i progetti fotovoltaici avranno probabilmente la precedenza su molti altri investimenti [3].

Gli stessi Stati Uniti nello scorso ottobre hanno definito i dettagli legislativi¹ inerenti gli incentivi fiscali all'installazione diffusa dei pannelli fotovoltaici, col risultato che gli USA si apprestano a diventare nel breve e medio termine il più grande mercato del fotovoltaico, colmando rapidamente un ritardo pluriennale. Uno studio [13] del Dipartimento Federale dell'Energia USA dimostra al contempo, sul fronte dell'energia eolica, come passare dall'attuale 1% al 20% nel 2030 sia tecnicamente ed economicamente fattibile e quindi auspicabile.

Dall'altra parte del mondo la Cina, una tra le nazioni che più inquinano al mondo, si è da tempo preparata per diventare il più grande fabbricante mondiale di celle e moduli fotovoltaici, puntando come di consueto sulla quantità prima per arrivare alla qualità poi. Anche sul fronte dell'eolico la Cina non scherza: secondo il Global Wind Energy Council la Cina diventerà nel 2009 il più grande produttore di turbine eoliche del mondo [13]. Nel contempo, sull'onda delle ultime Olimpiadi, lo stato cinese sta diventando, paradossalmente, teatro di sperimentazioni d'avanguardia per quel che riguarda l'architettura bio/eco-compatibile.

La fonte rinnovabile per eccellenza è il solare fotovoltaico - ovvero la produzione di energia elettrica a partire dalla radiazione solare - per la quale si prevede il raggiungimento della *grid parity* (quando i costi di produzione saranno comparabili con quelli delle fonti energetiche

¹ Investment Tax Credit - ITC.

Da qui al 2050 si potrebbero ottenere (tra parentesi la % rispetto ai 14 TWxs attuali)	Con (tipo di fonte energetica)	Come
8 TWxs (57.1%)	Nucleare	Costruendo 8000 centrali a fissione, una ogni due giorni (in tutto il mondo ce ne sono ora poco più di 400), ma l'uranio si esaurirebbe dopo 10 anni
7 TWxs (50.0%)	Biomasse	Ricoprendo l'intera superficie agricola del pianeta con colture adatte a diventare biocombustibile, ma senza più nulla da mangiare
2.1 TWxs (15.0%)	Eolico	Coprendo tutte le superfici del pianeta su cui soffiano abitualmente venti sostenuti (in gergo, classe 3 o superiore)
0.7 TWxs (5.0%)	Idroelettrico	Costruendo dighe su tutti i fiumi restanti sul pianeta
11.6 TWxs (82.9%)	Geotermia	Utilizzando l'intero potenziale geotermico stimato
60 TWxs (428.6%)	Solare	Coprendo con pannelli FV dal rendimento del 10% lo 0.08% della superficie terrestre, ovvero un quadrato di circa 650 km di lato

TABELLA 1

Disponibilità energetica in funzione della fonte [8]

tradizionali, senza l'ausilio di incentivi) per il 2011-2012, al raggiungimento della soglia dei 0.10 €/kWh, anche se la recente repentina discesa dei prezzi del greggio probabilmente posticiperà tale traguardo (per quanto paradossale, l'industri delle energie rinnovabile ha tratto linfa vitale dalle quotazioni stellari del petrolio dei mesi scorsi).

Per capire il potenziale effettivo delle fonti rinnovabili, fotovoltaico e non solo, è interessante uno studio realizzato da un ricercatore del *California Institute of Technology* [8]: calcolato che oggi sulla Terra si consumano mediamente 14 TWxs (1 TW = 1000 GW = 1'000'000 MW) con una previsione al 2050 28-35 TWxs si è calcolato come si potrebbe coprire tale fabbisogno, in tutto o in parte, attraverso le fonti energetiche disponibili e con l'attuale stato della tecnologia. I risultati sono riportati nella tabella 1.

Dato che non è generalmente possibile avvicinarsi all'autosufficienza² attraverso una sola delle fonti rinnovabili, per lo più per questioni climatiche o di morfologia territoriale, prende sempre maggiormente piede la logica del *mix*: avere a disposizione quante più

fonti energetiche, rinnovabili innanzitutto, in modo da poterne sfruttare le caratteristiche peculiari in tempi diversi e a seconda della prestazione richiesta, tenendo ben presente il bilancio economico complessivo.

Ad una produzione intelligente di energia (sia essa elettrica o termica) deve necessariamente accompagnarsi un uso almeno altrettanto intelligente dell'energia disponibile, eliminando gli sprechi e riducendo in generale i consumi. Troppo spesso i sistemi energetici o di processo funzionano in modo antieconomico, anche se non vi è consapevolezza di ciò perché non sono definite le condizioni di funzionamento "normali" previste e non vi è controllo degli scostamenti del modo di consumo. Nella pratica non è scontata la corrispondenza tra beneficio energetico e beneficio economico e quindi non è da considerare di interesse un intervento che, pur realizzando un vantaggio energetico, non comporti anche un ritorno economico [2].

Nasce quindi l'esigenza di un sistema che gestisca l'intero flusso: dalla produzione ai consumi.

² Quella dell'autosufficienza o niente è in realtà un falso mito: già un risparmio superiore al 50% sarebbe un risultato notevole e di tutto rispetto in qualsiasi altro settore; non si capisce in base a cosa un risparmio del "solo" 60-80% debba essere percepito come deludente quando si parla di fonti energetiche rinnovabili distribuite.

3. ENERGY RESOURCE PLANNING

Già oggi il mercato mette a disposizione di privati particolarmente illuminati o sensibili all'argomento, così come di imprese costruttrici e industrie, tutta una serie di strumenti per il risparmio energetico, inteso in senso lato:

- dispositivi per la produzione di energia da fonti rinnovabili³;
- apparati per il controllo dei consumi elettrici, termici ed idrici.

Sempre più spesso si trovano soluzioni e installazioni multiple, ibride, quali per esempio:

- solare fotovoltaico e termico;
- solare termico e caldaia a condensazione;
- caldaia a condensazione e stufa a pellet;
- solare termico ed elettrodomestici predisposti per l'utilizzo in ingresso di acqua preriscaldata;
- solare fotovoltaico e mini-eolico;
- caldaia a condensazione e pompa di calore;
- geotermico (profondo o superficiale) e solare termico.

Si tratta però di soluzioni estemporanee e comunque non pienamente integrate col resto dell'infrastruttura in cui sono collocate. Quello che ancora sembra mancare, ma di cui si sente urgente necessità, è un vero sistema integrato, e quanto più possibile automatizzato, di sensori, attuatori e applicativi che gestiscano i flussi energetici insieme a quelli economici collegati ai primi. In sostanza, si tratta di prendere in ingresso:

- le richieste energetiche (termiche ed elettriche essenzialmente), anche in termini di prestazioni richieste e livelli di comfort;

- il mix di fonti energetiche disponibili (caratteristiche specifiche, rendimenti, dipendenza dalle condizioni meteo, tempi di attivazione ecc.);

- i costi/ricavi di ciascuna fonte energetica (conto Energia, tariffa bi-oraria, costo unitario, costo del combustibile ecc.);

per fornire in uscita la regolazione più efficiente dal punto di vista energetico ed efficace dal punto di vista economico.

Tali sistemi dovrebbero basarsi su una vasta rete di sensori, con la conseguenza di avere anche un sistema puntuale e distribuito per la rilevazione dei parametri di funzionamento dei vari dispositivi, potendo quindi rilevare istantaneamente, o quasi, lo stato di salute del sistema, per una manutenzione predittiva ed economica.

Diverse imprese, soprattutto grandi, ed enti pubblici hanno al proprio interno da qualche tempo la figura dell'*energy manager*⁴, un professionista che per ruolo deve mantenere sotto controllo e possibilmente ridurre le spese, e quindi i consumi, a parità di prestazione erogata dagli edifici, delle attrezzature e degli impianti del soggetto per cui opera. Quasi mai però dispone di strumenti che anche solo rilevino puntualmente i consumi (per area, stanza, carico, macchinario, sezione di impianto), limitandosi ad agire in funzione dell'aggregato presentato in bolletta.

Anche i pochi privati che dispongono di un sistema domotico più o meno avanzato non sono in grado di stabilire quanto consumi effettivamente il frigorifero o la lavastoviglie o quando un qualsiasi apparecchio inizi a non funzionare più secondo i parametri standard. Perché è possibile avere il dettaglio delle chiamate effettuate dal nostro cellulare o dalla linea di casa oppure il dettaglio delle operazioni avvenute sul nostro conto corrente, mentre per quel che riguarda i consumi elettrici ci si deve accontentare dell'aggregato periodico riportato in bolletta?

Eppure tutte le tecnologie e i componenti necessari sono già presenti sul mercato:

- sensori e misuratori non mancano;
- i sistemi di trasmissione dati sono ormai una *commodity* e potrebbero contare su tecnologie standard a seconda della prestazione richiesta (XML, TCP-IP, ZigBee, powerline, GPRS, UMTS, Wi-Fi, Wi-Max);

- i sistemi per la presentazione delle informazioni ed il loro utilizzo in retroazione per il controllo e/o manutenzione sono facilmente mu-

³ Anche la legna ed il fogliame che un'azienda agricola recupera, come scarto delle proprie coltivazioni o a seguito di pulizia delle aree arboree, specialmente se bruciata in una caldaia ad alta efficienza, può ragionevolmente considerarsi rinnovabile in quanto tanta CO₂ viene liberata in atmosfera quanta ne aveva assorbita la pianta nella sua vita, o quanta ne assorbirà la pianta che andrà a sostituirla.

⁴ La legge 10/1991 prevede l'istituzione di figure professionali che svolgano, sia nel settore pubblico sia in quello privato, attività di razionalizzazione dell'uso dell'energia e di risparmio energetico.



tuabili da altri settori industriali (gli investimenti nel campo dell'automazione industriale, del telecontrollo e della pianificazione dei processi sono ormai stati ammortizzati).

Un esempio banale può essere rappresentato dai sistemi per la contabilizzazione individuale dei consumi termici negli impianti condominiali con riscaldamento centralizzato: con una valvola termostatica ed un sensore di consumo termico su ogni radiatore ciascuno dispone del comfort desiderato e paga solo il calore che ha effettivamente consumato (per la felicità di inquilini e amministratori), con risparmi che variano dal 15 al 24% [6].

Dopo tanti anni dall'introduzione dei sistemi di *Enterprise Resource Planning* (ERP), pare giunto il momento di una soluzione industriale per l'*Energy Resource Planning*, che potremmo contrarre in **EneRP/EnRP/NRP**.

Fatta eccezione per la grande impiantistica, lo scenario attuale pare caratterizzato più che altro da sperimentazioni o comunque soluzioni di nicchia, che utilizzano tecnologie e protocolli spesso proprietari, piuttosto che standard riconosciuti e componenti intercambiabili (un po' come se il nostro lettore di dvd leggesse solo i supporti prodotti da una certa marca o se il nostro pen-drive funzionasse con i computer marca X e non con quelli marca Y).

La tabella 2 riporta un quadro indicativo, anche se assolutamente non esaustivo, di alcuni strumenti o prodotti disponibili attualmente con caratteristiche tali da costituire uno o più elementi di un sistema EneRP. Si tratta di un elenco che non ha alcuna pretesa di essere né completo né critico o categorico: vista l'impossibilità di confrontare ogni singola soluzione presente sul mercato rispetto alle altre, se ne sono considerate solo alcune, giudicate particolarmente significative o ragionevolmente note.

Partendo dai sistemi per il monitoraggio degli impianti fotovoltaici, si è notato come questi in generale comunichino con l'esterno, quando lo fanno, attraverso protocolli proprietari e comunque senza far riferimento ad uno standard, che sia *de facto* piuttosto che ufficialmente riconosciuto. Analogo discorso vale per le caldaie e la produzione di energia termica. Molti produttori anzi si rifiutano esplicitamente di "aprire" i loro sistemi verso altri apparati, per ragioni commerciali magari comprensibili e legittime, ma probabilmente poco lungimiranti.

Sul fronte domotica, nell'accezione più classica del termine di "automazione per la casa", lo standard per il *bus* dati è il già citato KNX, anche se l'effettivo funzionamento di sistemi ibridi (apparati di produttori differenti all'interno dello stesso sistema) non è poi sempre garantito sul campo. In questo caso ciò che pare mancare è l'apertura verso i flussi di energia e informazioni entranti nell'abitazione, nonché l'indisponibilità (con alcune eccezioni, quali ABB per esempio) di apparati wireless plug-in per il controllo distribuito di più apparati o utenze. Questo ne limita fortemente l'impiego in strutture dove il ricablaggio non è, per vari motivi, considerato un'opzione attuabile. Ove presenti inoltre tali oggetti sfruttano protocolli proprietari piuttosto che standard riconosciuti (ZigBee, Wi-Fi) e pongono quasi sempre dei limiti al numero (circa 8 in genere) di apparati che uno stesso "concentratore" è in grado di monitorare (per esempio le soluzioni di ABB e SIPRO) contemporaneamente: questo non ne consente la scalabilità verso ambienti produttivi, commerciali o ad uso ufficio. Esistono sul mercato soluzioni software per la gestione e il controllo dei flussi energetici (Blink di Inspiring Software per esempio), nonché per l'elaborazione dei dati storici e la predisposizione di scenari what-if, tuttavia non fanno parte di un'offerta completa chiavi in mano, comprendente sensori, attuatori e trasmissione dati.

Un caso interessante e per certi versi emblematico per semplicità ed efficacia, oltre che per scalabilità, è dato dal sistema Save Energy di InfoSolutions: si tratta di una soluzione (come altre simili sul mercato) per la contabilizzazione dei consumi termici in caso di riscaldamento centralizzato. In ogni appartamento si dotano i termosifoni di valvola termostatica e di un sensore wireless a batteria che contabilizza il calore; ogni sensore trasmette poi i dati, via ZigBee o altro protocollo wireless, ad un *set top box* collegato alla tv di casa. Sullo schermo l'utente può vedere in tempo reale quanto consumato in termini energetici ed economici da ciascun termosifone. I dati poi passano, ancora in wireless, dai *set top box* degli appartamenti ad un concentratore, unico per l'edificio, che li raccoglie e trasmette ad un data center remoto, dove saranno disponibili per l'amministratore. In questo modo ciascuno paga per i propri consumi effettivi, non ci

TABELLA 2
Alcuni sistemi presenti sul mercato con funzionalità EnerP

Sistema	Produttore	Descrizione	Vantaggi	Limiti
SunGuard	Solarelit	Sistema informatico professionale per il monitoraggio di impianti fotovoltaici	- Controllo remoto via web sfruttando connessioni UMTS, GPRS, Reti LAN e Wi-Fi - Controllo dell'impianto FV e dell'ambiente nel quale è collocato - Permette di controllare inverter di più costruttori	- Limitato al momento ai soli inverter per FV - Non tutti i produttori di inverter rendono disponibile l'accesso ai dati - Il monitoraggio si ferma alla stringa di pannelli solari e non arriva al singolo pannello
IMGSoft	Carlo Gavazzi	Controllo remoto di inverter FV	- Controllo remoto di ogni parametro dell'impianto - Disponibile per il download gratuito	- Limitato ai soli impianti FV Gavazzi - Il monitoraggio si ferma alla stringa di pannelli solari e non arriva al singolo pannello
My Home	Bticino	Sistema domotico	- Permette di controllare quasi ogni dispositivo domestico - Controllo anti black-out: distacco delle utenze secondo un ordine di priorità - Impostazione scenari	- Non è predisposto per gestire fonti energetiche - Non fornisce dati sui consumi delle singole utenze - Trasmissione dati attraverso protocolli proprietari
By-Me	Vimar	Sistema domotico	- Permette di controllare quasi ogni dispositivo domestico - Controllo anti black-out: distacco delle utenze secondo un ordine di priorità - Integrazione con Windows Media Center	- Non è predisposto per gestire fonti energetiche - Non fornisce dati sui consumi delle singole utenze - Non consente analisi storiche e scenari - Trasmissione dati attraverso protocolli proprietari
DomusTech	ABB	Sistema domotico	- Permette di controllare ogni dispositivo - Sono disponibili delle prese elettriche telecomandate (sia da incasso che a frutto) - Trasmissione dati wireless - Controllo anti black-out: distacco delle utenze secondo un ordine di priorità - Impostazione scenari	- Non è predisposto per gestire fonti energetiche - Non fornisce dati sui consumi delle singole utenze - Non consente analisi storiche - Trasmissione dati attraverso protocolli proprietari - Il numero di utenze elettriche controllabili è limitato
Blink	Inspiring Software	Software per analisi e controllo dei consumi energetici	Moduli per: - Caratterizzazione energetica e definizione di un modello previsionale - Carte CuSum per controllo e allarme in Real Time - Budgeting energetico - Scelta e monitoraggio della tariffa energetica - Reporting e analisi - Ottimizzazione della produzione di energia termica ed elettrica	Non include la parte hardware (sensori e attuatori) e la trasmissione dati

Segue

Sistema	Produttore	Descrizione	Vantaggi	Limiti
Kit Load Manager	Sipro	Gestione automatica del consumo di energia elettrica	• Architettura Master-slave con comunicazione powerline (KONNEX/EHS) • La Smart Plug si connette ad una qualunque presa di casa trasformandola in una "presa controllata" • Gestione tariffazione differenziata	• Limitato alle utenze elettriche • Non è predisposto per gestire fonti energetiche • Non consente analisi storiche e scenari • Il numero di utenze elettriche controllabili è limitato • Dati non disponibili in remoto
Save Energy	Infosolutions	Contabilizzazione consumi termici e idrici	• Architettura Slave-Master-Data Center • Zigbee per la trasmissione dai sensori locali • Set top box collegato alla TV di casa per avere in tempo reale i consumi di ogni utenza	• Non integra al momento i consumi elettrici • Non è predisposto per gestire fonti energetiche
Casa Buderus ¹	Buderus	Installazione dimostrativa	• Integra tutte le fonti energetiche rinnovabili o comunque ad alta efficienza • Dimostra caratteristiche e vantaggi di un edificio con valenze bioclimatiche. • Disponibile per visite guidate gratuite • Patrocinio di Legambiente e Ministero Ambiente	I sistemi installati sono stati integrati ad hoc
Agriturismo Vojon ¹	Vari	Agriturismo in attività	• Integra quasi tutte le fonti energetiche rinnovabili o comunque ad alta efficienza • È un'installazione realmente operativa • Ha vinto la bandiera verde di Legambiente • I dati raccolti sono disponibili per divulgazione e ricerca	• Produzione di energia elettrica tramite pannelli solari e/o mini-eolico ancora non implementata • Controllo manuale degli impianti • Non consente ancora analisi storiche e scenari

¹ Casa Buderus e Agriturismo Vojon sono esempi reali in cui è stato implementato un mix di fonti energetiche rinnovabili o ad alta efficienza.

sono sorprese, si risparmia (dal 25 al 50%) e si depenna un argomento di discussione alle assemblee condominiali.

Particolare interesse hanno infine suscitato due installazioni dimostrative: una, Casa Buderus, realizzata dall'omonima azienda tedesca, insieme ad altri partner qualificati, come abitazione allo stato dell'arte, presso la sede di Assago (MI), l'altra, l'agriturismo Vojon a Ponti sul Mincio (MN), condensa, per volontà dei proprietari (che in essa anche risiedono) in una struttura realmente operativa (l'azienda agricola e quella alberghiera) praticamente tutte le fonti energetiche disponibili oggi.

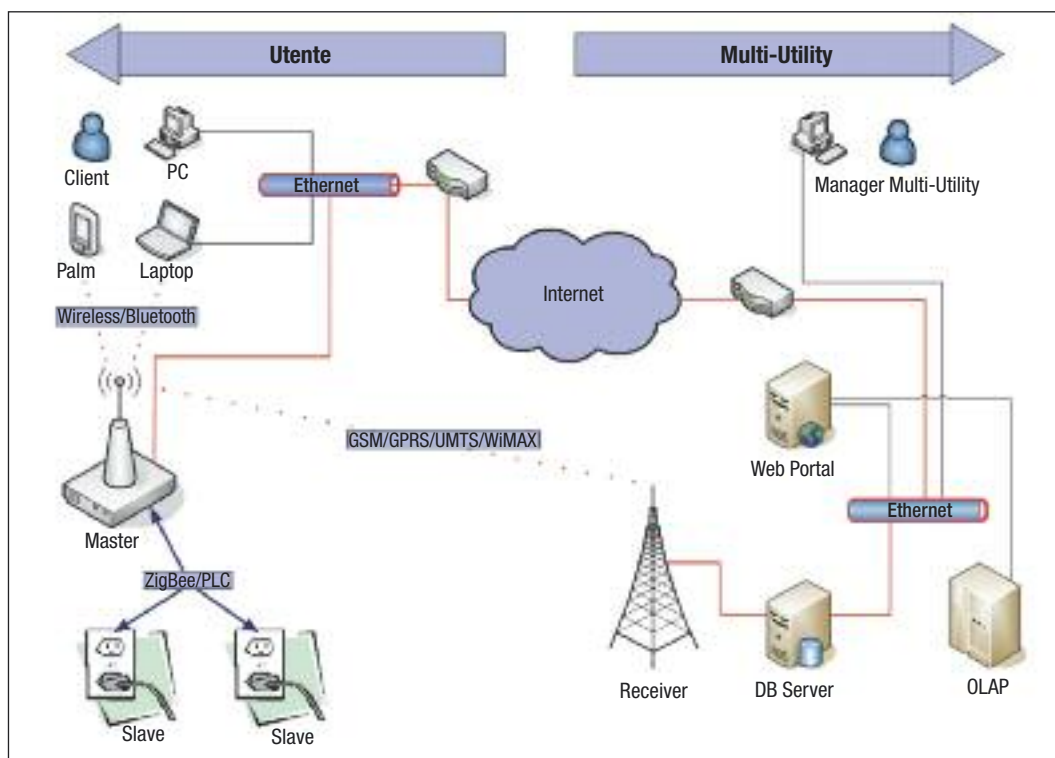
Casa Buderus, visitabile gratuitamente previa prenotazione tramite sito web dedicato, rappresenta l'abitazione ideale per una famiglia di

quattro persone, nel caso di assenza di vincoli, economici innanzitutto. Ogni dettaglio è realizzato allo stato dell'arte in termini di materiali, impianti e tecnica costruttiva ed anche la collocazione dell'edificio rispetto alla direttrice nord-sud, la presenza o meno di porte o finestre su una certa facciata e l'estensione delle pareti murarie e dei tetti è studiata secondo i canoni della bio-architettura ecosostenibile.

La Casa Buderus integra al suo interno:

- pannelli solari termici e fotovoltaici (di questi, una serie in silicio policristallino ed una in silicio amorfo);
- caldaia a condensazione;
- stufa a pellet;
- pompe di calore;
- recupero aria;

FIGURA 1
*Schema UTILTEC
 (nell'ipotesi di
 servizio fornito da
 una Multi-Utility):
 i dati locali vengono
 raccolti dagli slave,
 che li fanno
 confluire al master,
 il quale può
 presentarne una
 parte, sempre
 in locale,
 e interagire con un
 centro servizi*



- distribuzione calore;
- recupero acqua piovana.
- sistemi domotici.

L'agriturismo Vojon, nel mantovano con vista sul lago di Garda, rappresenta invece la declinazione concreta del risparmio energetico applicato ad un casolare che ospita sia un'attività produttiva (l'azienda agricola) sia una struttura turistica, senza dimenticare l'abitazione dei proprietari/conduttori. Al Vojon "tutte le fonti di energia rinnovabili si integrano in cascata utilizzando sempre quella disponibile al minor costo", tanto da potersi aggiudicare nel 2007, unico in Lombardia, la Bandiera Verde Agricoltura. Quanto a impianti, sono presenti: sonde geotermiche (di tre tipi: a terreno, acqua-acqua e terra-acqua); pompe di calore; solare termico; caldaia a biomassa. Gli allacciamenti per solare fotovoltaico e mini-eolico sono già predisposti e l'installazione pianificata.

4. I PROGETTI META-DISTRETTI ICT LOMBARDIA UTILTEC ED ENERTEC

Nell'ambito degli ultimi bandi "Meta-Distretti ICT" della Regione Lombardia, la Fondazione Politecnico di Milano ha partecipato a due

proposte progettuali premiate da finanziamento proprio nell'ambito del risparmio energetico e della produzione distribuita di energia da fonti rinnovabili.

4.1. UTILTEC

Il primo progetto, denominato **UTILTEC** (MD ICT 2005), vedeva come partner tre imprese, concentrate rispettivamente su automazione industriale (sdi automazione industriale), micro controlli e sensoristica (MISARC), sviluppo applicativi software (Beta80), oltre al Dipartimento di Elettronica e Informazione (DEI) del Politecnico di Milano e la Fondazione Politecnico di Milano.

Il progetto UTILTEC era finalizzato alla realizzazione di un prototipo per la gestione ed il controllo intelligente dei consumi elettrici basato su tre elementi (Figura 1):

- una serie di sensori locali interposti tra il carico elettrico e la presa di corrente oppure in un quadro elettrico locale a monte di una serie di carichi. Questi sensori, detti **slave**, hanno il compito di misurare e registrare i consumi dell'apparecchiatura a cui sono collegati, trasmettono quindi ad intervalli regolari (15 min) i dati ad una unità master ed applicano i comandi (distacco carichi e riarmo)

da questa ricevuti; possono essere sia *wired* che *wireless*;

□ una unità (o più d'una a seconda delle necessità) che funge da concentratore, detta **master**, che raccoglie e memorizza i dati dagli slave, provenienti via wireless (ZigBee), via rete o via *powerline* (trasmissione dati attraverso la rete elettrica). Nel master sono programmate anche le politiche di controllo degli slave, a cui comunica se necessario i comandi di distacco. Il master fornisce l'interfaccia locale per la configurazione degli slave e provvede a trasmettere i dati raccolti (ogni ora o ogni giorno a seconda delle impostazioni) ad un centro di controllo remoto (via GSM-GPRS-UMTS-ADSL), dal quale riceve anche la segnalazione di quali politiche di controllo attuare;

□ il terzo componente del sistema è il centro di controllo remoto, o **data center**, nel quale confluiscono tutti i dati raccolti e che vengono utilizzati per gli algoritmi di previsione dei consumi o diagnostici. Nel data center sono presenti ovviamente i dati storici e gli algoritmi di elaborazione, ma anche il cruscotto utente, configurabile su più profili e accessibile via web; è sempre il data center che, in funzione delle elaborazioni effettuate, comanda al master quale delle varie politiche di controllo programmate deve attuare.

Proviamo ad immaginare uno scenario operativo per un sistema come UTILTEC, ipotizzando che sia installato capillarmente nei punti vendita di un grande marchio della *Grande Distribuzione Organizzata* (GDO).

Ogni zona della superficie commerciale è collegata ad un quadro elettrico, con interposto uno slave. In più, su alcune apparecchiature o macchinari particolarmente sensibili (per questioni economiche o funzionali) sono installati degli slave individuali (per esempio, frigoriferi, luci, impianto di condizionamento, impianto di riscaldamento ecc.).

Un responsabile locale del punto vendita accede via web al master e verifica il funzionamento dei vari sistemi monitorati all'interno dei parametri standard; in caso di malfunzionamenti la segnalazione puntuale permette di intervenire a colpo sicuro (esempio: quel frigorifero, quel forno, quella fila di televisori sta avendo un consumo anomalo, per cui occorre provvedere).

L'*energy manager corporate*, dalla sede centrale, può a sua volta:

□ tenere sotto controllo sia l'aggregato dei consumi provenienti dai vari punti vendita, sia indagare in dettaglio ciascuno di essi;

□ impostare centralmente delle politiche corporate da inviare ai master dei punti vendita (esempio, nelle fasi di chiusura le linee luci devono essere parzializzate, oppure le pareti delle tv LCD o plasma in esposizione devono essere sempre spente dalle 22.00 alle 07.00, od anche la temperatura di condizionamento non può mai essere inferiore ai 25°C); in questo modo si avrebbe certezza dell'applicazione delle direttive corporate relative ai consumi elettrici;

□ raccogliere col massimo del dettaglio e dell'affidabilità la curva dei consumi, dai singoli carichi all'aggregato aziendale.

Quest'ultima funzionalità in particolare apre una serie di opzioni finora inesplorate proprio per mancanza dello strumento.

Con la certezza della propria curva dei consumi, sia a livello giornaliero sia a livello orario, l'*energy manager* ha nelle sue mani uno strumento potente con cui fare leva verso il distributore di energia con cui stipula il contratto di fornitura: i trader di energia, infatti, temono più di ogni altra cosa l'incertezza sui volumi e sui tempi degli approvvigionamenti energetici che acquistano sul mercato. Una curva dei consumi dall'andamento fluido e prevedibile è quanto mai raccomandabile, viste le difficoltà nell'immagazzinare o reperire grandi quantità di energia in tempi brevi a fronte di picchi, siano essi verso l'alto o verso il basso. Se il cliente, nella fattispecie il nostro *energy manager* GDO, potesse assicurare una certa curva dei consumi giornalieri (dal giorno 1 al giorno 365), il distributore potrebbe applicare degli sconti non trascurabili, a fronte dell'impegno del cliente stesso a non sfiorare la curva fornita. Qui entrerebbe in gioco la possibilità di controllare i carichi remoti attraverso le strategie inserite nei master. In pratica, il data center, una volta accertatosi che la curva dei consumi rilevati si appresta a eguagliare o superare quella dei consumi previsti, invia al master l'istruzione di applicare una particolare strategia di distacco selettivo dei carichi, tipicamente dal meno

critico in poi. Il master riceve l'istruzione e comunica agli slave interessati di staccare il carico a cui sono collegati, fino a quando non riceverà dal data center l'ordine di riattivarli, una volta scongiurato il rischio di superare la soglia prevista.

La bolletta energetica di una soggetto della GDO con copertura nazionale arriva facilmente ad alcuni milioni di euro l'anno. L'adozione di un sistema come quello descritto genererebbe facilmente un risparmio tra il 10%⁵ ed il 40%.

Non è difficile immaginare come uno strumento del genere possa facilmente essere impiegato anche in altri scenari, dal terziario alla produzione industriale.

Sarebbe possibile costruire delle curve dei consumi in funzione:

- delle condizioni meteo,
- dello storico (lo stesso giorno dell'anno scorso cosa è successo? Nella stessa settimana degli ultimi 5 anni quale è stata la situazione?),
- delle persone presenti nell'edificio (dato reperibile dalle timbrature automatiche e dai registri presenze),
- dei macchinari attivi (numero e tipologia), e utilizzare i dati per algoritmi predittivi che permettessero di ipotizzare con ragionevole certezza i consumi dei mesi successivi. Tale possibilità potrebbe essere utilizzata non solo per prevedere i consumi a 3-6 mesi, 1 anno, ma, molto utile, per realizzare delle previsioni a fronte di scenari:
- quanto si consumerebbe, e quindi si potrebbe risparmiare, cambiando il parco macchine X con una nuova serie Y a risparmio energetico?
- aggiungendo altre N macchine del tipo Y, quanto aumenterebbe il consumo?
- se si ampliasse l'attività produttiva o la superficie di X metri quadrati, quanto inciderebbe nei costi energetici?
- la creazione di un nuovo ufficio con X persone, che consumi extra comporterebbe?

Anche dal punto di vista della manutenzione

i vantaggi sarebbero molteplici: in ambito produttivo si acquistano macchinari, anche molto costosi, con determinati parametri di targa, ma solo in rari casi, e quasi sempre solo all'entrata in esercizio, sono sottoposti a verifica di tali parametri. Non c'è al momento uno strumento diffuso ed a basso costo che, monitorandone i consumi continuativamente, permetta di stabilire se e per quanto il macchinario funziona all'interno dei parametri prestabiliti e garantiti dal produttore. Con un tale strumento a disposizione anche i contenziosi circa malfunzionamenti più o meno presunti o accertabili potrebbero trovare più rapida soluzione.

Per quanto concerne i possibili clienti/utilizzatori del sistema descritto, non è troppo complicato immaginare che una grande impresa, o comunque un soggetto con elevati consumi e diverse sedi produttive o commerciali lungo il territorio, potrebbe essere interessata ad acquisire il sistema completo, portando all'interno della propria infrastruttura IT il data center. Viceversa, una PMI o un artigiano potrebbero essere più interessati ad un servizio "chiavi in mano", fornito in outsourcing da una società appositamente costituita o, ancor meglio, da una di quelle aziende di fornitura servizi energetici che con la liberalizzazione del mercato hanno recentemente affiancato l'*incumbent*, incluse le ex municipalizzate. Da notare che mentre l'*incumbent* ha storicamente tutto l'interesse a far crescere i consumi e quindi i ricavi, i nuovi operatori del mercato libero dell'energia hanno invece interesse a controllare i flussi energetici e ad attrarre nuovi clienti attraverso politiche commerciali basate sul risparmio energetico da una parte e la certificazione di provenienza da fonte rinnovabile dall'altra.

4.2. ENERTEC

Mentre il progetto UTILTEC, conclusosi nel 2007, si è focalizzato sul lato del controllo dei consumi elettrici, il secondo progetto, denominato **ENERTEC** (MD-ICT 2007; presentato dalla medesima cordata di imprese e partner accademici di UTILTEC, con l'aggiunta di Solarday, un produttore di pannelli fotovoltaici) si focalizza sul controllo e gestione di impianti energetici eterogenei, principalmente ma non esclusivamente, rinnovabili, allo scopo di realizzazione un sistema prototipale di misura,

⁵ Tale percentuale di risparmio è rilevata comunemente ogniqualvolta si inserisce un misuratore di qualche tipo: il solo fatto di poter vedere istantaneamente una misura, magari convertita in euro, porta le persone a risparmiare quasi inconsciamente.

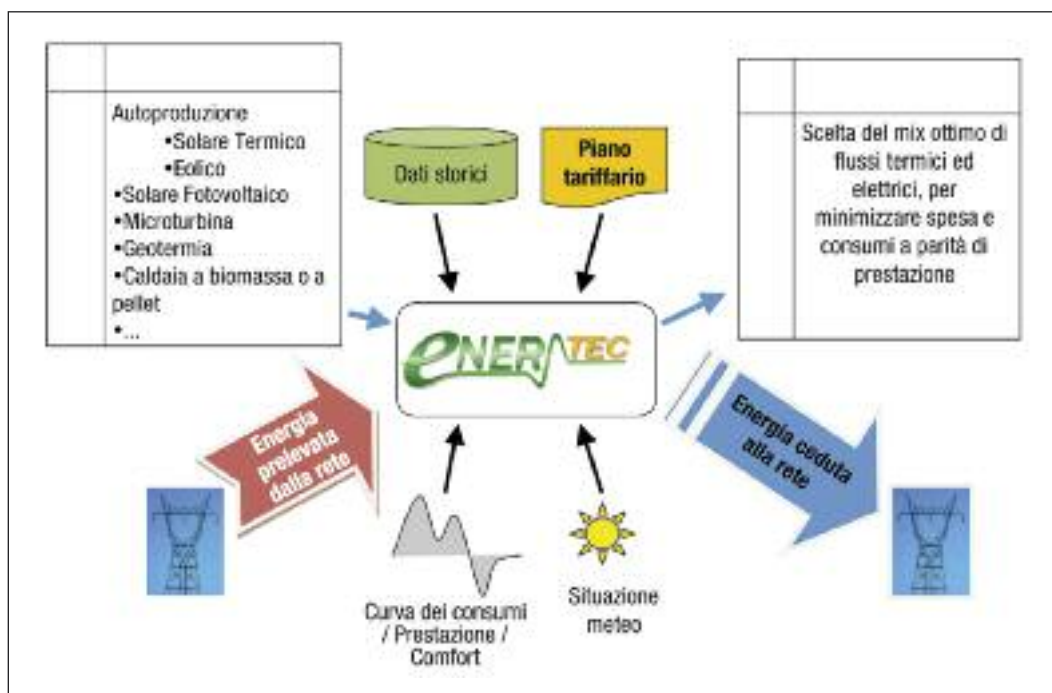


FIGURA 2

ENERTEC: i flussi di informazioni ed energia in ingresso vengono elaborati per avere in uscita le migliori prestazioni energetico-economiche possibili

acquisizione e controllo della produzione e del consumo di energia in applicazioni industriali e di *building automation* (Figura 2).

Lo schema tecnico prevede una parte locale ed una parte centrale.

La parte locale è costituita da dispositivi da installare presso l'utilizzatore, in grado di dialogare tra loro e di intervenire sui dispositivi energetici:

□ **Nodi logici** di acquisizione e controllo

- sensori posizionati strategicamente sui dispositivi energetici, siano essi sorgenti di calore od elettricità o utilizzatori per misurare flusso energetico e trasmissione dati;
- attuatori di vario tipo che manovrano in tempo reale i flussi energetici eseguendo le istruzioni del concentratore;

□ **Concentratore**, per

- acquisizione dati da nodi logici e archiviazione;
- gestione del sistema energetico secondo logica programmabile,
- controllo del sistema energetico tramite nodi logici,
- archiviazione dati locale a breve termine;
- interfaccia utente per configurazione dei nodi logici;

La parte centrale è costituita da strutture (basi di dati, web server, applicativi software) al-

locate presso un centro servizi, in grado di interagire con le varie utenze locali:

□ **Centro elaborazione dati (Centro servizi)**, per

- raccolta dati concentratore e archiviazione (DB Server);
- condivisione dati con Utente (Web Server);
- analisi e reportistica aggregata o di dettaglio.

Concretamente il sistema si prefigge di gestire i **flussi energetici** (corrente elettrica e calore) provenienti da:

- solare fotovoltaico;
- solare termico;
- mini eolico;
- geotermico;
- caldaia a biomassa / pellet;
- pompe di calore;
- micro-generazione;
- co-generazione;
- e in generale qualsiasi dispositivo installabile localmente.

Questi flussi vengono quindi ottimizzati in funzione di flussi di informazioni:

- inerzia termica;
- assorbimenti elettrici;
- rendimenti;
- condizioni climatiche;
- prestazioni o livelli di comfort richiesti;
- condizioni contrattuali in essere;

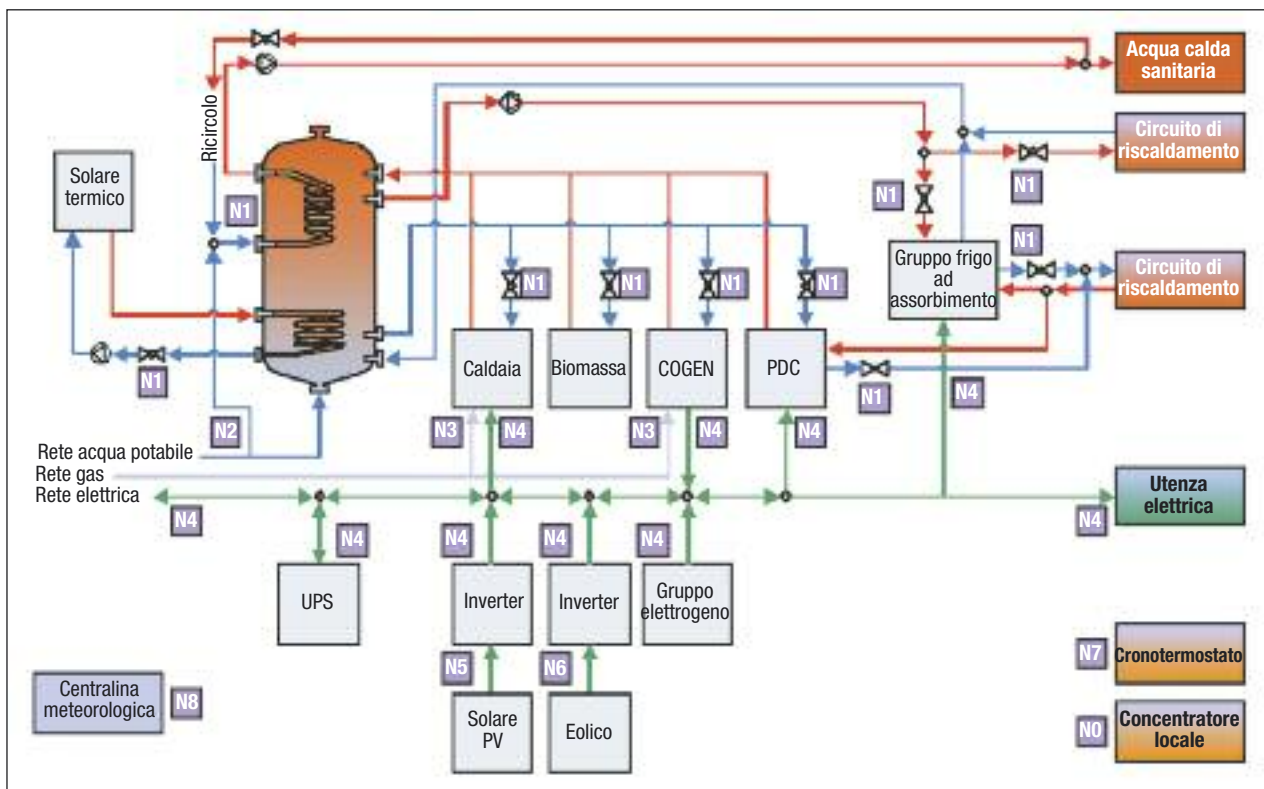


FIGURA 3

ENERTEC: Scenario di riferimento generalizzato

- tariffe e costi.

La figura 3 riporta lo schema a blocchi utilizzato in ENERTEC per individuare uno scenario quanto più generale possibile.

5. CONCLUSIONI

La situazione oggi più frequente è quella di un consumo inconsapevole, dove alla scarsa cognizione dell'assorbimento energetico dei propri impianti si accompagna un'errata gestione degli impianti energetici, con conseguenti elevate perdite, il tutto nell'ambito di una diffusa obsolescenza strutturale.

La necessità di aggiornamento degli impianti porta a considerare le opportunità fornite dalla generazione distribuita: dalla micro-/co-/tri-generazione alle fonti rinnovabili. È in questo contesto che le imprese, pubbliche o private, devono comprendere come gli investimenti in efficienza energetica, specie in nazioni come l'Italia dove il costo dell'energia è particolarmente elevato, ne migliorano la competitività, in quanto beneficiano contemporaneamente della riduzione dei costi di produzio-

ne/gestione e del miglioramento dell'immagine nei confronti dell'esterno.

Prendere coscienza dell'opportunità e stanziare un *budget* non è sufficiente: occorre trovare fornitori preparati che sappiano mettere a disposizione soluzioni integrate e automatizzate:

- un sistema di acquisizione dati distribuito in grado di rilevare e misurare i flussi di energia nei punti di maggior rilevanza;
 - un sistema di controllo, gestione ed integrazione dei dispositivi di generazione e delle utenze per il raggiungimento della massima efficienza energetica.
- La diffusione di sistemi EneRP permetterebbe al privato cittadino, alle aziende del terziario, alle PMI di carattere industriale e alle grandi imprese di avere:
- dettaglio e trasparenza consumi;
 - personalizzazione dei profili di carico;
 - tariffazione agevolata in linea con le reali esigenze;
 - totale controllo delle proprie utenze;
 - interazione a distanza attraverso strumenti ICT tradizionali;

□ un significativo e misurabile risparmio energetico e di costi.

Per chi invece, come nel caso di una *multi-utility*, mettesse a disposizione di terzi un sistema EneRP, il vantaggio competitivo deriverebbe da:

□ servizi innovativi per monitoraggio, analisi, caratterizzazione e gestione dei consumi;

□ personalizzazione e diversificazione dei servizi offerti;

□ efficienza nella gestione dei propri impianti. Rimane ora da vedere chi riuscirà ad arrivare sul mercato con un prodotto adeguato alle esigenze e con che modello di business potrà sostenersi.

Bibliografia

- [1] *Energia – Gli italiani sempre più disposti a investire (anche personalmente) nelle fonti rinnovabili*. Osservatorio Scienza e Società, 15 marzo 2007.
- [2] Periodico D., Lucca M.: Previsione, monitoraggio e controllo dei consumi di uno stabilimento alimentare. *Automazione e Strumentazione*, n. 11, 2008.
- [3] La Rosa M.: Risparmio energetico, tra progetto e controllo. *Technology Review* ed. italiana n. 03, 2008.
- [4] CasaEnergia, n. 6, 2008; www.zeroemission.tv
- [5] FV-Fotovoltaici, n. 6, 2008; www.zeroemission.tv

[6] *Il risparmio energetico negli edifici condominiali*. EnerBuilding.eu – ADICONSUM.

[7] Nova24-*Il Sole 24 Ore*, n. 79.

[8] Nova24-*Il Sole 24 Ore*, n. 99.

[9] Nova24-*Il Sole 24 Ore*, n. 113.

[10] Nova24-*Il Sole 24 Ore*, n. 127.

[11] Nova24-*Il Sole 24 Ore*, n. 141.

[12] *Photon* – il mensile del fotovoltaico, n. 11-2208; www.photon-online.it

[13] *Wind Energy*, n. 6, 2008; www.zeroemission.tv

Siti web:

www.bticino.it

www.buderus.it

www.domustech.it

www.fondazionepolitecnico.it

www.gavazziautomation.com

www.inspiringsoftware.com

www.johnsoncontrols.it/publish/it/it/about/sustainability/good_for_business.html

www.sipro.biz

www.solarelit.it

www.vimar.eu

www.agriturismovojon.it

PAOLO MAGNI, dal 2004 opera come project manager presso la Fondazione Politecnico di Milano, dove si occupa di allestire e coordinare progetti di ricerca e sviluppo, tra imprese, università ed enti pubblici, in ambito automotive, energia, logistica, ICT per l'industria. È professore a contratto di Laboratorio Progettuale di Disegno Assistito dal Calcolatore presso la sede di Piacenza del Politecnico di Milano. Si è laureato in Ingegneria Meccanica presso l'Università degli Studi di Parma e nel 2003 ha conseguito il "Master in the Network Economy – MiNE" presso l'Università Cattolica del Sacro Cuore di Piacenza.

E-mail: magni@fondazionepolitecnico.it

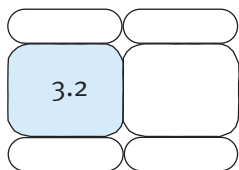


E.W. DIJKSTRA

UNA VITA DA INFORMATICO

Lorenzo Milone

Edsger Wybe Dijkstra è stato un informatico poliedrico e dai numerosi interessi, un pioniere che ha attraversato tutte le fasi dell'informatica moderna, dalla nascita alla maturità. L'articolo ne ripercorre l'intera vita, personale e professionale mettendone in evidenza i periodi fondamentali.



1. INTRODUZIONE

Dijkstra. Avevo ascoltato più volte quel nome strano, pronunciato un po' come veniva¹. Ricordo che stavo studiando gli algoritmi. Ne avevo visti tanti ma uno tra tutti mi aveva particolarmente colpito per eleganza e semplicità: era l'algoritmo dei cammini minimi in un grafo pesato ed era associato a quel nome. Poi è arrivato il tempo dello studio dei sistemi operativi. Problemi complicati e, di conseguenza, soluzioni complesse, eccetto quelle associate a quel nome. Geniale, ad esempio, la soluzione al problema della mutua esclusione: i semafori. A questo punto la curiosità è troppo forte. È giunto il momento di provare a rispondere ad una domanda che spesso, in questi anni, mi sono posto: chi era Dijkstra? Edsger Wybe Dijkstra è stato un informatico poliedrico e dai numerosi interessi. Lo scopo di questo contributo è quello di ripercorrerne, seppure in maniera preliminare, l'intera vita, personale e professionale, e di metterne in evidenza i periodi fondamentali. Ad ogni periodo

sarà dedicato un paragrafo del presente lavoro. Un paragrafo a parte sarà dedicato ai suoi manoscritti, anche noti come la serie EWD. Nel secondo paragrafo analizzeremo gli anni giovanili a Rotterdam. Nel terzo, i primi approcci con l'informatica e il lavoro al Centro Matematico di Amsterdam. Nel quarto, l'insegnamento all'Università di Eindhoven. Nel quinto, la collaborazione con la Burroughs Corporation. Nel sesto, il periodo all'Università di Austin in Texas. Nel settimo, esamineremo i manoscritti, inesauribile e stimolante fonte primaria sulla vita e sull'attività professionale di Dijkstra.

2. I PRIMI ANNI

Edsger Wybe Dijkstra nacque a Rotterdam l'11 maggio 1930, terzo di quattro figli. Sua madre, Brechtje Cornelia Kluyver, era un matematico²; suo padre, Donwe Wybe, un chimico. Dal 1942 al 1948 frequentò il Gymnasium Erasmianum a Rotterdam. Durante l'ultimo an-

¹ "Dijkstra", "Daikstra". La pronuncia corretta, letta in italiano, è "Deikstra": "Etscher Uibe Deikstra".

² Dijkstra ricorda che la madre era dotata di «una grande agilità nella manipolazione delle formule e di un meraviglioso dono per la ricerca di soluzioni molto eleganti». EDSGER WYBE DIJKSTRA, *From my Lyfe*, EWD1166, 1993, p. 0. Il sottotitolo del manoscritto recita: "Scritto perché la gente mi chiede queste informazioni".

no, era il 1947-1948, pensava di andare all'università a studiare diritto nella speranza di rappresentare il suo Paese alle Nazioni Unite: furono gli insegnanti a consigliare gli studi scientifici. Aveva, in effetti, i voti migliori in matematica, fisica e chimica. Dijkstra ricorda che, a sedici anni, aveva acquistato i libri per l'anno successivo e si era particolarmente intimorito per il nuovo libro sulla trigonometria, pieno di seni, coseni e lettere greche. Chiese a sua madre se la trigonometria sarebbe stata difficile. La risposta avrebbe influenzato tutto il suo successivo lavoro: "Oh, no! Impara le formule e ricordati sempre che sei sulla strada sbagliata se hai bisogno di più di cinque righe"³.

3. IL CENTRO MATEMATICO

Finito il liceo, andò a studiare matematica e fisica teorica all'Università di Leiden. Il padre, che aveva letto su una rivista l'annuncio di un corso di tre settimane a Cambridge sulla programmazione per computer (si trattava dell'EDSAC⁴), gli propose di partecipare. Dijkstra pensò che fosse una buona idea nella convinzione che l'uso del computer sarebbe stato utile ad un fisico teorico. Il corso si tenne nel settembre del 1951⁵.

Considerate le nuove conoscenze acquisite, Adriaan van Wijngaarden, allora direttore del Dipartimento di calcolo al Centro Matematico di Amsterdam, gli offrì un posto di lavoro che fu accettato nel marzo del 1952⁶. Al Cen-

tro Matematico, Dijkstra avrebbe lavorato, inizialmente due giorni la settimana⁷, insieme a Bram J. Loopstra e Carel S. Scholten, alla realizzazione di computer: l'ARRA⁸, l'ARRA II, il FERTA⁹ e in particolare l'ARMAC¹⁰. Dijkstra ha scritto:

*"Nessuno di noi era stato formato come matematico. Van Wijngaarden era laureato in ingegneria meccanica, ma con una lunga esperienza nel lavoro numerico; Loopstra e Scholten studiavano fisica sperimentale presso l'Università di Amsterdam e io ancora studio fisica teorica a Leiden. Il nostro comune background era che venivamo da due delle più famose scuole secondarie del paese, van Wijngaarden e io dal Gymnasium Erasmianum, Loopstra e Scholten dal Vossius Gymnasium. Lo dico perché credo fermamente che questo quadro, che comprendeva una solida formazione in cinque lingue straniere, abbia avuto una grande influenza sul modo in cui abbiamo lavorato"*¹¹.

Nel 1952 il Centro Matematico stava realizzando il computer ARRA. Quest'ultimo "era però così inaffidabile da risultare praticamente inutilizzabile¹²" e ben presto si passò a un nuovo progetto: l'ARRA II. La descrizione funzionale dell'ARRA II è stata la prima relazione¹³ che Dijkstra ha scritto come dipendente del Centro Matematico. La relazione è rivolta all'utilizzatore della macchina e descrive ciò che essa fa e non come funziona,

³ EDSGER WYBE DIJKSTRA, *My hopes of computing science*, EWD709, 1979, p. 1.

⁴ Negli anni quaranta del Novecento, l'Università di Cambridge fu molto attiva nel settore dei computer; in particolare, per il progetto di un calcolatore a programma memorizzato detto EDSAC (Electronic Delay Automatic Calculator) basato sulle idee di Von Neumann.

⁵ In un'intervista, Dijkstra ha detto: "Tutto è iniziato nel 1951, quando mio padre mi ha consentito di frequentare un corso di programmazione a Cambridge, Inghilterra". *An Interview with Edsger W. Dijkstra*, Minneapolis, Charles Babbage Institute, 2001, p. 3.

⁶ Quando è entrato al Centro Matematico, Dijkstra aveva ventuno anni e il suo stipendio era circa l'equivalente di 3 \$ al giorno. EDSGER WYBE DIJKSTRA, *A programmer's early memories*, EWD568, 1976, p. 6.

⁷ Il resto del tempo studiava a Leiden. EDSGER WYBE DIJKSTRA, *What led to "Notes on Structured Programming"*, EWD1308, 2001, p. 0.

⁸ Automatische Relais Rekenmachine Amsterdam.

⁹ Fokker Electronische Rekenmachine te Amsterdam. La macchina fu commissionata da un'industria d'aeromobili, la Fokker. Al lavoro partecipò anche Gerard Blaauw che più tardi divenne uno dei progettisti dell'IBM 360.

¹⁰ L'acronimo sta per Automatische Rekenmachine MAthematische Centrum.

¹¹ EDSGER WYBE DIJKSTRA, *A programmer's early memories*, cit., p. 0-1.

¹² EDSGER WYBE DIJKSTRA, *A programmer's early memories*, cit., p. 1-2. Si veda anche Corrado Bonfanti: *Mezzo secolo di futuro: l'informatica italiana compie cinquant'anni*. *Mondo Digitale*, n. 3, 2004, p. 48-68: 56.

¹³ EDSGER WYBE DIJKSTRA, *Functionele beschrijving van de arra*, MR12, 1953.



Edsger Wybe
Dijkstra

quali sono le istruzioni che mette a disposizione e non come sono state realizzate. È evidente come, da subito, Dijkstra era sensibile a quelli che più tardi sarebbero stati chiamati principi di progettazione del software. L'ARMAC è del 1956. Dijkstra scrisse il software¹⁴, Loopstra e Scholten progettarono l'hardware:

*“La prima cosa che abbiamo fatto è stata quella di scrivere un manuale di programmazione per la prossima macchina, compresa la sua **completa** descrizione funzionale. (Per qualche motivo, sono stato quello che ha scritto quel documento). Questo documento quindi è servito da contratto tra me e loro: essi sapevano quello che avrebbero dovuto costruire e io sapevo quello che avrei dovuto sviluppare. E circa un anno più tardi, saremmo stati pronti: essi avevano costruito la loro macchina e io avevo finito tutti i software di base¹⁵”.*

La prima applicazione¹⁶ per l'ARMAC fu scritta da Maria C. Debets, destinata a diventare, un anno più tardi, Mrs. Dijkstra. È in questo periodo che Dijkstra prese una decisione fondamentale.

Racconta:

“Nel 1955 ho preso la decisione di non diventare un fisico teorico ma un programmatore. Ho preso questa decisione perché avevo concluso che tra la fisica teorica e la programmazione, la programmazione rappresentava una sfida intellettuale più grande. Come vedete, in quei giorni non ho sofferto di modestia intellettuale. È stata una decisione difficile, perché ero stato istruito come uno scienziato primo della classe, e diventare un programmatore sembrava un addio dalla scienza. Quando ho spiegato ad A. van Wijngaarden, il mio capo, il mio dilemma, egli mi ha assicurato che i computer erano qui per rimanere e che nel mondo della programmazione avrei benissimo potuto essere quello chiamato a creare la scienza che era ancora carente. Ottenere la laurea in fisica a Leiden divenne una formalità da sbrigare il più rapidamente possibile¹⁷”.

Nel 1956 Dijkstra si laureò in Fisica teorica. Nello stesso anno si svolse l'inaugurazione ufficiale dell'ARMAC. Per dimostrarne la potenza di calcolo, Dijkstra progettò il famoso algoritmo dei cammini minimi, ora noto come l'algoritmo di Dijkstra. Considerati i fini dimostrativi, era necessario realizzare un programma per risolvere un problema che anche i non matematici potessero capire: ad esempio, trovare l'itinerario più breve tra due città dei Paesi Bassi¹⁸. L'algoritmo fu progettato in appena ventuno minuti, senza carta e penna, mentre gustava una tazza di caffè in compagnia della futura moglie in una soleggiata terrazza di un bar di Amsterdam¹⁹.

¹⁴ Fu il primo programmatore del Centro, in un periodo in cui la professione di programmatore non era ancora molto conosciuta. Dijkstra racconta che sull'atto di matrimonio non gli fu consentito di inserire alla voce Professione il termine Programmatore. EDSGER WYBE DIJKSTRA, *The humble programmer*, “Communications of the ACM”, Vol. XV, n. 10, 1972, p. 859-866: 860.

¹⁵ EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 2.

¹⁶ Si trattava del calcolo del movimento trasversale delle carrozze ferroviarie.

¹⁷ EDSGER WYBE DIJKSTRA, *What led to “Notes on Structured Programming”*, cit., p. 1.

¹⁸ Per la dimostrazione, Dijkstra selezionò 64 città e usò una codifica a 6 bit per identificare le città.

¹⁹ EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 2.

Poco tempo dopo, mentre lavorava all'X1²⁰, progettò l'algoritmo per l'albero ricoprente minimo. Il problema da risolvere era quello di collegare una serie di punti sul pannello della macchina utilizzando lo stesso filo di rame in modo che tutti i punti avessero uguale tensione, e riducendo al minimo la quantità di rame usata. Così come per l'ARMAC, anche per l'X1 Dijkstra scrisse il software di base, progettando inoltre un sistema di gestione degli interrupt e fu questo l'argomento della sua tesi di dottorato²¹.

I due algoritmi furono pubblicati, in tre pagine, solo nel 1959²²:

“Non ebbi fretta di pubblicare questi due algoritmi: in quel momento, gli algoritmi non avevano ancora acquistato rispettabilità matematica e non vi erano riviste adatte. Alla fine sono stati offerti nel 1959 alla “Numerische Mathematik”, nel tentativo di aiutare questa nuova rivista ad affermarsi²³”.

Dopo Fortran, ALGOL 60 è stato uno dei primi linguaggi di programmazione ad alto livello. Dijkstra era strettamente coinvolto nello sviluppo, realizzazione e divulgazione dell'ALGOL 60. Tra il 1958 e il 1959 prese parte ad una serie di incontri che portarono alla definizione del linguaggio²⁴. Nell'estate del 1959, mentre era in vacanza con la famiglia, aveva scoperto come implementare *subroutine* ricorsive²⁵,

entro la fine dell'anno aveva definito come usare uno *stack*²⁶ per la valutazione delle espressioni e come tradurle dalla solita notazione infissa a quella polacca inversa²⁷. Tutti lavori preparatori che lo portarono nel 1960 a realizzare, in collaborazione con Jaap Zonneveld, il primo compilatore ALGOL 60²⁸:

“Abbiamo scritto il compilatore in doppia copia, pienamente consapevoli che stavamo facendo la storia: ogni sera ciascuno di noi portava la propria copia del manoscritto a casa in modo tale che, in caso d'incendio presso il nostro Centro Matematico, il compilatore non sarebbe andato perso. (Non c'è stato fuoco)²⁹”.

Il primo caso di test fu:

```
begin
  real a;
  a:=7
end.
```

Ricordando quel periodo, Dijkstra ha scritto:

“La nostra attuazione dell'ALGOL 60 segna la fine del periodo di crescita che intendevo coprire e anche l'inizio di uno nuovo. Finalmente iniziai a considerarmi un programmatore professionista [...]. Essa segnò l'inizio del periodo durante il quale l'attività di programmazione comincia ad evolvere da arte in una disciplina scientifica³⁰”.

²⁰ L'X1 fu progettato dal Centro Matematico e realizzato dall'Electrologica.

²¹ EDSGER WYBE DIJKSTRA, *Communication with an Automatic Computer*, PhD thesis, Amsterdam, 1959. Il supervisore della tesi di dottorato fu A. van Wijngaarden. Qui sono discusse le questioni concernenti la lentezza dei dispositivi di I/O rispetto al processore, è sviluppato il concetto di *interrupt* e sono descritte alcune tecniche di *buffering*.

²² EDSGER WYBE DIJKSTRA, *A note on two problems in connexion with graphs*. Numerische Mathematik, Vol. I, 1959, p. 269-271.

²³ EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 2.

²⁴ Si veda KRZYSZTOF R. APT, *Edsger Wybe Dijkstra (1930-2002): A Portrait of a Genius*. Formal Aspects of Computing, Vol. XIV, 2, 2002, p. 92-98: 93.

²⁵ EDSGER WYBE DIJKSTRA, *Twenty-eight years*, EWD1000, 1987, p. 0.

²⁶ «Al quel tempo, scrivevo in olandese e avevo usato il sostantivo “stapel” e il verbo “stapelen”, che ho poi tradotto in “a stack” e “to stack”. (Evidentemente, la terminologia fu fortunata: nel giro di pochi anni dopo che la ebbi usata nell'articolo “Recursive Programming”, si è affermata nel vocabolario della comunità informatica)». *Twenty-eight years*, cit., p. 1. I termini *stack*, *vector* e *mutual exclusion*, in un contesto informatico, sono attribuiti dall'Oxford English Dictionary a Dijkstra.

²⁷ EDSGER WYBE DIJKSTRA, *Recursive Programming*. Numerische Mathematik, Vol. II, 1960, p. 312-318. Questo lavoro è considerato uno dei primi importanti contributi alla scrittura dei compilatori.

²⁸ Il lavoro durò otto mesi. Dijkstra e Zonneveld concordarono di non radersi fintanto che il progetto non fosse stato completato. Ultimato il compilatore Zonneveld si rasò, Dijkstra tenne la barba fino alla morte. Si veda FRANS E. J. KRUSEMAN ARETZ, *The Dijkstra-Zonneveld ALGOL60 Compiler for the Electrologica X1*, Centrum voor Wiskunde en Informatica, Amsterdam, 2003, p. 6.

²⁹ EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 3.

³⁰ EDSGER WYBE DIJKSTRA, *A programmer's early memories*, cit., p. 11.

L'intera esperienza mi ha reso sensibile a ciò che più tardi sarebbe stata chiamata modularità o divide and rule o astrazione, per la cura con cui le interfacce furono scelte e per il potenziale campo d'applicazione della sfida della programmazione in generale. Essa ha fortemente contribuito alla mia successiva opinione che creare fiducia nella correttezza della progettazione era il più importante ma più difficile aspetto del compito del programmatore³¹.

4. L'UNIVERSITÀ TECNOLOGICA DI EINDHOVEN

Nel 1962, Dijkstra fu nominato professore nel Dipartimento di Matematica dell'Università Tecnologica di Eindhoven.

“Più tardi seppi che ero stato la terza scelta del dipartimento, dopo che due analisti avevano rifiutato l'invito; la decisione di invitare me non deve essere stata facile, da un lato perché non avevo veramente studiato matematica, e, dall'altro, a causa dei miei sandali, della mia barba e della mia “arroganza” (da vedere che cosa può essere). E' anche stato ricordato che dovevo essere molto grato al Dipartimento perché aveva preso un gran rischio a nominare me. Per due anni ho tenuto i corsi di Analisi Numerica perché non avevano nessun altro per farlo, e allo stesso tempo ho iniziato la costruzione di un piccolo gruppo di scienziati del computer³²”.

È qui, infatti, che insieme con alcuni studenti³³ progettò il THE³⁴, il primo sistema operativo con una struttura gerarchica a strati. Il sistema fu progettato per l'X8 dell'Electrologica.

Sono questi gli anni che videro Dijkstra approfondire le problematiche della programmazione concorrente³⁵. Il concetto di semaforo³⁶ e la descrizione delle operazioni P e V³⁷ sono apparsi in un manoscritto del 1962³⁸. L'intuizione del concetto nasce dall'esempio della ferrovia a binario unico: un treno prima di entrare nel binario deve attendere che lo stato del semaforo sia libero, quando entra deve modificare lo stato del semaforo per impedire che altri treni possano entrare, quando esce deve modificare ancora lo stato del semaforo per consentire agli altri treni di usare il binario. L'algoritmo del banchiere è apparso in un manoscritto probabilmente dello stesso periodo³⁹. Nel 1965, in una sola pagina⁴⁰, Dijkstra introdusse il problema della mutua esclusione per n processi e ne propose una soluzione. È stato probabilmente il primo algoritmo concorrente pubblicato. La nozione, ora standard, di sezione critica⁴¹ è stata introdotta in questo documento. Lo scritto *Cooperating sequential processes* fu realizzato nel 1965⁴² e definitivamente pubblicato nel 1968 in un saggio⁴³ di 70 pagine con cui nacque il campo della programmazione concorrente. Poi, in un documento del 1971⁴⁴ illustrò il problema del *deadlock* per mezzo dell'esempio dei “cinque filosofi a cena”. Cinque filosofi, seduti intorno ad un tavolo

³¹ EDSGER WYBE DIJKSTRA, *What led to “Notes on Structured Programming”*, cit., p. 6.

³² EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 4.

³³ Cor Ligtmans, Piet Voorhoeve, Nico Habermann e Frits Hendriks.

³⁴ Acronimo di Technische Hogeschool Eindhoven.

³⁵ Nello stesso periodo furono introdotte le “reti di Petri”, un importante contributo allo sviluppo della teoria dei processi concorrenti. CARL ADAM PETRI, *Kommunikation mit Automaten*, PhD thesis, Darmstadt, 1962.

³⁶ Il semaforo è una variabile intera non negativa condivisa tra più processi concorrenti.

³⁷ *Verhogen* sta per incrementare, *Prolagen* è un neologismo proveniente da *probeer te verlagen* e sta per testare e decrementare.

³⁸ EDSGER WYBE DIJKSTRA, *Multiprogramming en X8*, EWD51, 1962.

³⁹ EDSGER WYBE DIJKSTRA, *Een algoritme ter voorkoming van de dodelijke omarming*, EWD108, s.d.. *Un algoritmo per impedire l'abbraccio mortale*. L'abbraccio mortale (*deadlock*) è la situazione di stallo.

⁴⁰ EDSGER WYBE DIJKSTRA, *Solution of a problem in concurrent programming control*. Communications of the ACM, Vol. VIII, n. 9, 1965, p. 569.

⁴¹ La sezione critica è quella porzione di codice in cui un processo accede a variabili condivise con altri processi.

⁴² EDSGER WYBE DIJKSTRA, *Twenty-eight years*, cit., p. 2. Sono gli appunti delle lezioni tenute all'università di Eindhoven.

⁴³ EDSGER WYBE DIJKSTRA, *Cooperating sequential processes*, in *Programming Languages: NATO Advanced Study Institute*, a cura di F. Genuys, London, Academic Press, 1968, p. 43-112.

⁴⁴ EDSGER WYBE DIJKSTRA, *Hierarchical ordering of sequential processes*, EWD310, 1971.

lo per mangiare gli spaghetti, condividono solo cinque forchette. Gli spaghetti sono scivolo- si e quindi ciascun filosofo si avvale di due forchette per mangiare. Purtroppo se tutti e cinque afferrano simultaneamente una forchetta, nessuno potrà mangiare. Questo problema è poi diventato un classico di riferimento per spiegare nuove primitive di sincronizzazione. Nel 1968, Dijkstra pubblicò uno dei suoi più noti articoli, *Go To statement considered harmful*⁴⁵, nel quale sostenne che la facilità di lettura e di comprensione di un programma erano inversamente proporzionali al numero di go to contenuti e che la dichiarazione go to era una delle principali fonti d'errori e pertanto andava eliminata e sostituita con strutture più evolute.

Verso la fine degli anni sessanta, Dijkstra attraversò probabilmente il momento più difficile della sua vita. Il successo del THE e la crescente autonomia della scienza informatica non erano ben visti all'interno del Dipartimento di Matematica dell'Università di Eindhoven. Il piccolo gruppo di persone che partecipò al THE fu sciolto. Dijkstra soffrì un periodo di profonda depressione. Fu così che nel 1969, "per ragioni terapeutiche"⁴⁶, iniziò a scrivere *Notes on Structured Programming*⁴⁷. Ne inviò solo una ventina di copie ad amici all'estero e ciononostante il documento ebbe una diffusione pressoché istantanea⁴⁸. La programmazione strutturata è una metodologia di scrittura dei programmi basata sulla combinazione e nidificazione di un piccolo numero di strutture di controllo⁴⁹. Dijkstra raccomandava, inoltre, alcuni principi di progettazione, che sono poi divenuti generalmente accettati nella comunità informatica:

□ i grandi sistemi devono essere costruiti partendo da molti piccoli componenti;

□ ogni componente deve essere definito solo dalla sua interfaccia e non dalla sua implementazione;

□ i componenti più piccoli possono essere progettati a seguito di un simile processo di decomposizione, determinando in tal modo un processo di tipo *top-down*;

□ la progettazione dovrebbe iniziare dalla "separazione degli interessi", cioè separando i diversi aspetti del problema e concentrando l'attenzione su ciascuno di loro separatamente;

□ la logica matematica è, e deve essere, la base per la progettazione del software.

Nel 1968 partecipò alla Conferenza sull'Ingegneria del software sponsorizzata dalla NATO Science Committee. Dijkstra fu uno dei protagonisti anche della successiva conferenza NATO che si tenne a Roma nell'ottobre del 1969 e di cui fu presidente il professor Paolo Ercoli.

Nel 1972 vinse il premio Turing dell'ACM, uno dei più importanti premi per l'Informatica. Il suo discorso d'accettazione, intitolato *The Humble Programmer*⁵⁰, affronta l'evoluzione della programmazione come disciplina e riporta una serie di prescrizioni per la sua costante crescita. Dijkstra ricorda:

*"Un tardo pomeriggio nel 1972, il telefono del mio ufficio presso l'Università squillò; ho alzato il ricevitore, era Franz L. Alt dell'ACM, che chiamava dagli Stati Uniti per riferirmi che avevo vinto l'ACM Turing Award. Mi prese di sorpresa, non mi consideravo un potenziale beneficiario del prestigioso premio, e non potevo quasi crederci, ma Alt mi convinse che era proprio vero, quindi abbiammo scollegato*⁵¹".

⁴⁵ EDSGER WYBE DIJKSTRA, *Go To statement considered harmful*. Communications of the ACM, Vol. XI, 3, 1968, p. 147-148. Il titolo originale del documento sottoposto era "A case against the go to statement". Al fine di accelerarne la pubblicazione, il curatore lo aveva trasformato in una *lettera al Curatore* modificandone anche il titolo. Il curatore era Niklaus Wirth.

⁴⁶ EDSGER WYBE DIJKSTRA, *Twenty-eight years*, cit., p. 2.

⁴⁷ EDSGER WYBE DIJKSTRA, *Notes on Structured Programming*, EWD249, 1969.

⁴⁸ «Ho inviato solo circa 20 copie ad amici all'estero, ma in quel momento la fotocopiatrice era diventata onnipresente, e (senza che lo sapessi), il testo ebbe una diffusione istantanea: più tardi ho incontrato persone in luoghi lontani che custodivano una copia di quinta o sesta generazione!». EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 5.

⁴⁹ Il fondamento teorico della programmazione strutturata è il celebre teorema di Jacopini-Böhm: qualsiasi programma scritto usando il go to poteva essere riscritto senza, a patto di avere a disposizione altri tre tipi di strutture di controllo: sequenza, iterazione e selezione. CORRADO BÖHM, GIUSEPPE JACOPINI, *Flow diagrams, turing machines and languages with only two formation rules*. Communications of the ACM, Vol. IX, 5, 1966, p. 366-371.

⁵⁰ EDSGER WYBE DIJKSTRA, *The humble programmer*. Communications of the ACM, Vol. XV, n. 10, 1972, p. 859-866.

⁵¹ EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 5.

La citazione, letta dal presidente del premio, recita:

“Il dono prezioso che questo premio Turing riconosce è lo stile di Dijkstra: il suo approccio alla programmazione come una sfida alta, intellettuale; la sua eloquente insistenza e dimostrazione pratica che i programmi dovrebbero essere composti correttamente, non solo corretti in debug; e la sua illuminante percezione dei problemi alla base della progettazione dei programmi⁵²”.

5. LA BURROUGHS CORPORATION

Il 1° agosto del 1973 Dijkstra entrò, come collaboratore alla ricerca, alla Burroughs Corporation⁵³. Il suo compito consisteva nel visitare i centri di ricerca della compagnia un paio di volte l'anno e nel fare, a sua volta, ricerca, cosa che fece nello studio al secondo piano della sua casa di Nuenen, Paesi Bassi. Viaggiò molto e tenne conferenze in tutto il mondo. Fondò quello che fu chiamato l'ETAC, il **Club del Martedì Pomeriggio ad Eindhoven**: un gruppo di ricercatori⁵⁴ che incontrava tutti i martedì pomeriggio all'università per lavorare su un problema scientifico o per discutere su un documento pubblicato di recente. In questo periodo scrisse un'enorme quantità di articoli di ricerca⁵⁵, la maggior parte relazioni tecniche destinate alla circolazione privata all'interno di un gruppo selezionato. Parte di questi articoli furono pubblicati nel 1982 in una raccolta dal titolo *Selected writings on computing: a personal perspective*:

“A partire dall'estate del 1973, quando diventai Fellow alla Burroughs Research, la vita fu molto diversa da quello che era stata prima. La routine quotidiana cambiò: invece di andare al-

l'Università ogni giorno, dove ero solito trascorrere la maggior parte del mio tempo in compagnia di altri, ora andavo lì solo un giorno la settimana ed ero la maggior parte del tempo - così è, quando non si viaggia! - da solo nel mio studio. Nella mia solitudine, la corrispondenza e la parola scritta, in generale, divennero sempre più importanti. La circostanza che il mio datore di lavoro e io avessimo l'Oceano Atlantico tra noi è stato un ulteriore incentivo a tenere una traccia completa di quello che stavo facendo. La parte pubblica di tale produzione trovò il suo posto in quella che divenne nota come “la serie EWD”, che può essere vista come una forma di corrispondenza scientifica, possibile con l'avvento della fotocopiatrice. (La stessa fotocopiatrice rende difficile stimare la sua effettiva distribuzione: io feci circa due dozzine di copie dei miei testi, ma i loro destinatari accettarono di fungere come successivi nodi dell'albero di distribuzione)⁵⁶”.

Nel 1976 pubblicò la sua *A Discipline of Programming*⁵⁷. In quest'opera, un linguaggio di programmazione è visto come un veicolo per la descrizione formale e astratta di programmi indipendenti dalla macchina su cui questi devono essere eseguiti. Non sempre i programmi di una certa dimensione raggiungono una ragionevole robustezza attraverso il testing e, a volte, questi risultano poco comprensibili anche per i loro progettisti. Non si dovrebbe, quindi, prima scrivere il programma e poi sperare che sia corretto, o, solo successivamente, dimostrarne la correttezza. Dijkstra ritiene che questo metodo di costruzione di programmi sia fondamentalmente errato. Il programma e la prova di correttezza, invece, andrebbero sviluppati in parallelo: in questo modo, il programma risulterebbe corretto per costruzione. Da questo punto di vista, il fatto che un programma possa essere effettivamente ese-

⁵² EDSGER WYBE DIJKSTRA, *The humble programmer*, cit., p. 859.

⁵³ La Burroughs era una delle più importanti aziende di computer di quegli anni.

⁵⁴ Martin Rem, Maarten Boasson, Jan Tijmen Udding, Wim Feijen, Rob Hoogerwoord, Jan van de Snepscheut, Netty van Gasteren, Anne Kaldewaij, Alain Martin, Mohamed Gouda.

⁵⁵ Alcuni lavori di questo periodo sono: *Self-stabilizing systems in spite of distributed control*, EWD391, 1973; *Guarded commands, nondeterminacy, and the formal derivation of programs*, EWD418, 1974; *On-the-fly garbage collection: an exercise in cooperation*, EWD520, 1975; *Smoothsort, an alternative for sorting in situ*, EWD796, 1981.

⁵⁶ EDSGER WYBE DIJKSTRA, *Selected writings on computing: a personal perspective*, New York-Heidelberg-Berlin, Springer-Verlag, 1982, p. V.

⁵⁷ EDSGER WYBE DIJKSTRA, *A Discipline of Programming*, London, Prentice-Hall, 1976.

guito da un computer è del tutto incidentale. Dijkstra lavorò tutta la vita alla professionalizzazione accademica della programmazione e all'autonomia dell'informatica come scienza. Autonomia non solo dalle altre scienze ma anche dal computer inteso come strumento: memorabile è la sua frase «Computer Science is no more about computers than astronomy is about telescopes».

Una disciplina accademica ha però bisogno di un oggetto e di un metodo. Insieme con un collega, Wim Feijen, scrisse *A Method of Programming*⁵⁸. In questo libro la programmazione è presentata come un ramo formale della matematica.

6. L'UNIVERSITÀ DI AUSTIN

Il periodo alla Burroughs fu intenso; tuttavia l'esigenza di un ritorno all'università era sentita. Fu così che quando l'Università del Texas di Austin gli offrì la cattedra di informatica, questa fu accettata. Era l'estate del 1984. Al riguardo, Dijkstra racconta:

L'offerta non fu una sorpresa, né la mia accettazione, conoscevo Austin, conoscevo UT e loro conoscevo me. Attraverso tutti i miei anni alla Burroughs avevo preso l'abitudine di visitare, ovunque andassi, l'università locale e sin dalla fine degli anni '70 la Burroughs aveva avuto un Centro di Ricerca ad Austin, una visita che fu inclusa nella maggior parte dei miei viaggi negli Stati Uniti⁵⁹.

Ad Austin, Dijkstra profuse gran parte delle sue energie nell'insegnamento. Non ha mai seguito un libro di testo, non ha quasi mai utilizzato proiettori, era solito invece scrivere su una lavagna.

Molti dei documenti scritti in questo periodo trattano argomenti già affrontati in passato, ne ripercorrono la storia o aggiungono nuovi contenuti: si riprendono le problematiche

concernenti la metodologia della programmazione⁶⁰ e gli algoritmi sui grafi⁶¹, si descrive l'esperienza che lo portò ad occuparsi dei Sistemi operativi⁶², quella che condusse alla Programmazione Strutturata⁶³.

In una nota del 1988, *A new science, from birth to maturity*, Dijkstra tratteggiò il percorso dell'informatica come scienza, dalla nascita alla maturità, dividendo schematicamente il periodo che va dagli anni cinquanta agli ottanta in quattro decenni. Ad ognuno di questi assegnò un'etichetta.

“Se poniamo la nascita dell'Informatica con l'avvento del computer a programma-controllato, l'argomento è di circa quattro decenni d'età. Questi sono stati decenni entusiasmanti e sono grato per la fortuna di essere stato coinvolto per la maggior parte di tale periodo, grato perché questa partecipazione mi ha permesso di osservare tutto il processo di crescita da vicino [...] per amore di semplicità, ho diviso i quaranta anni in quattro decenni, ciascuno dei quali con una caratterizzazione grezza: potremmo caratterizzare gli anni cinquanta come il decennio dell'hardware, i sessanta come il decennio della sintassi, i settanta come il decennio della semantica, e gli ottanta -- a titolo di prova, essendo ancora giovane -- come il decennio della sintesi⁶⁴”.

Rifondò con un gruppo di studenti laureati, sul modello di quello di Eindhoven, il Club del Martedì Pomeriggio (ATAC)⁶⁵.

In occasione del suo sessantesimo compleanno nel 1990, il Dipartimento di Scienze organizzò un seminario di due giorni in suo onore, con oratori provenienti sia dagli Stati Uniti sia dall'Europa. Un gruppo di colleghi curò, inoltre, la pubblicazione di una raccolta di saggi dal titolo *Beauty is our business: a birthday*

⁵⁸ E. W. DIJKSTRA, W. H. J. FEIJEN, *A Method of Programming*, Boston, Addison Wesley, 1988. Il lavoro fu scritto in olandese nel 1984 col titolo *Een methode van programmeren* e solo più tardi tradotto in inglese.

⁵⁹ EDSGER WYBE DIJKSTRA, *From my Lyfe*, cit., p. 7.

⁶⁰ EDSGER WYBE DIJKSTRA, *Introducing a course on program design and presentation*, EWD1157, 1993.

⁶¹ EDSGER WYBE DIJKSTRA, *On Dijkstra's Lemma and Kruskal's Algorithm*, EWD1273, 1998.

⁶² EDSGER WYBE DIJKSTRA, *My recollections of operating system design*, EWD1303, 2001.

⁶³ EDSGER WYBE DIJKSTRA, *What led to "Notes on Structured Programming"*, EWD1308, 2001.

⁶⁴ EDSGER WYBE DIJKSTRA, *A new science, from birth to maturity*, EWD1024, 1988, p. 0.

⁶⁵ EDSGER WYBE DIJKSTRA, *The ATAC (= Austin Tuesday Afternoon Club)*, EWD927, 1985.

A short proof of one of Fermat's theorems.

Consider the n^p distinct "words" of p "characters" that can be formed from an "alphabet" of n distinct characters. Draw from each word an arrow to the word that is obtained by rotating its characters over one place to the left, i.e. from $n_0, n_1, \dots, n_{p-1}$ an arrow is drawn to $n_1, \dots, n_{p-1}, n_0$. Because p successive rotations transform a word into itself, the arrows form cycles of lengths that are divisors of p .

Hence, if p is prime, 1 and p are the only possible cycle lengths. Because a cycle of length 1 corresponds to a word all characters of which are equal, exactly n distinct words occur in a cycle of length 1. Hence the remaining $n^p - n$ words occur in cycles of length p , i.e. for any n and any prime p , $n^p - n$ is a multiple of p .

Plataanstraat 5
5671 AL NUENEN
The Netherlands

27 May 1980
prof. dr. Edsger W. Dijkstra
Burroughs Research Fellow.

FIGURA 1

Un manoscritto di
E.W. Dijkstra (Fonte:
[www.cs.utexas.edu/
users/EWD/](http://www.cs.utexas.edu/users/EWD/))

salute to Edsger W. Dijkstra⁶⁶. Il titolo, che a prima vista può sembrare un po' curioso, in realtà è la citazione di un articolo di Dijkstra:

"Inoltre, quando bene applicata, l'induzione matematica può portare ad argomentazioni molto compatte ed efficaci - in breve: belle! - e quando riconosciamo la battaglia contro il caos, il disordine, l'indomita complessità come una delle principali vocazioni dell'informatica, dobbiamo ammettere che *Beauty is our Business*⁶⁷".

Il seminario organizzato per il settantesimo compleanno ha, poi, un titolo particolarmente significativo: *In pursuit of simplicity*. Semplicità, concisione e originalità di approccio sono le caratteristiche distintive del metodo di lavoro di Dijkstra. In *A short proof of one of Fermat's theorems*⁶⁸, nel dimostrare il piccolo teorema di Fermat – quindi un problema squisitamente matematico – egli segue un metodo squisitamente "informatico" (Figura 1).

⁶⁶ *Beauty is our business: a birthday salute to Edsger W. Dijkstra*, a cura di W. H. J. FEIJEN, A. J. M. VAN GASTEREN, D. GRIES, J. MISRA, New York-Heidelberg-Berlin, Springer-Verlag, 1990 (Texts and Monographs in Computer Science).

⁶⁷ EDSGER WYBE DIJKSTRA, *Some beautiful arguments using mathematical induction*, EWD697, 1978, p. 1.

⁶⁸ EDSGER WYBE DIJKSTRA, *A short proof of one of Fermat's theorems*, EWD740, 1980. L'articolo apre anche il fascicolo degli Atti di un convegno A.I.C.A. (*Formal methods for software development*, Milano, marzo 1992) in cui Dijkstra fu l'ospite di spicco.

Dijkstra insegnò fino al 1999, quando andò in pensione. Morì a Nuenen, Paesi Bassi, il 6 agosto 2002.

7. I MANOSCRITTI

Dijkstra non ha mai scritto i suoi articoli utilizzando un computer. Ha sempre preferito fare affidamento sulla macchina per scrivere⁶⁹ e sulla sua penna stilografica Montblanc. Questi articoli, in gran parte mai pubblicati, sono stati poi distribuiti in maniera singolare: egli ne inviava delle copie ad alcuni amici e collaboratori che poi fungevano da successivi punti di distribuzione. Questi brevi articoli coprono un periodo di quarant'anni, sono numerati consecutivamente e sono in gran parte note tecniche, relazioni di viaggio, penetranti osservazioni, commenti pungenti, saggi. Essi raramente sono più lunghi di quindici pagine e le pagine sono numerate partendo da zero. L'ultimo, il n. 1318, è del 14 aprile 2002. Il primo dovrebbe essere del 1959. Nel mondo dell'informatica sono noti come le relazioni EWD, o semplicemente le EWDs. Così Dijkstra ricorda com'è iniziata:

“Ventotto anni fa ho scritto EWD0. Fu durante una sorta d'interregno scientifico: la mia tesi e il software di base per l'X1 erano stati completati e non ho potuto iniziare il compilatore ALGOL in quanto ALGOL 60 non era stato ancora definito. Ho fatto tutti i tipi di cose non collegate. È stato, per esempio, l'anno che ho pubblicato - finalmente! - i miei algoritmi per il cammino minimo e l'albero ricoprente minimo. È stato anche un tempo d'esplorazioni. In un determinato momento stavo lavorando a quattro o cinque diversi manoscritti. Naturalmente, le pagine d'ogni manoscritto erano state numerate; dopo che ebbi mischiato le pagine tra i manoscritti per la seconda volta, decisi che sarebbe stato meglio mettere fine alla confusione, e numerai i manoscritti: EWD0, EWD1, EWD3, EWD4. E questo è il modo in cui è iniziata⁷⁰”.

La sua scrittura a mano era così perfetta e distinta che alla fine del 1980 Luca Cardelli, allora della Digital Equipment Corporation, progettò un font “Dijkstra” per computer Macintosh. Poco dopo, Dijkstra ricevette una lettera in questo formato e pensò che fosse stata scritta a mano, ma poi ebbe notizia della creazione di questo font⁷¹.

I manoscritti erano un informale ma efficace veicolo di diffusione delle idee: «La serie EWD sembra un po' inusuale. Se è sufficientemente insolita per rappresentare un nuovo stile di fare scienza, lo sviluppo di questo stile può essere uno dei miei principali contributi»⁷².

8. CONCLUSIONI

In questo lavoro mi sono proposto di evidenziare come Edsger Wybe Dijkstra possa essere considerato uno dei protagonisti dell'informatica moderna. Il suo immenso coraggio e audacia intellettuale, e le profonde, nonché semplici ed eleganti, idee hanno cambiato il corso dell'informatica. Le sue opinioni in materia di scienza in generale e sulla ricerca, in particolare, sono state di notevole profondità e originalità. «Egli è stato un uomo con la luce nelle tenebre. Ha illuminato in pratica ogni problema che ha discusso». In breve, è stato un genio. In informatica siamo tutti i figli di Dijkstra»⁷³.

Dijkstra ha attraversato, davvero, tutte le fasi dell'informatica moderna, dalla nascita alla maturità: il suo contributo alla professionalizzazione della programmazione e all'autonomia dell'informatica come scienza è stato decisivo. Nessun aspetto della disciplina è stato trascurato: dalla progettazione degli algoritmi ai sistemi operativi, dai linguaggi di programmazione ai compilatori ecc.. A lui dobbiamo l'intuizione che la logica e la metodologia matematica dovessero essere la base per la ragionevole progettazione dei programmi. Al suo pionieristico contributo dobbiamo la costruzione di una

⁶⁹ La macchina era una Hermes.

⁷⁰ EDSGER WYBE DIJKSTRA, *Twenty-eight years*, cit., p. 0.

⁷¹ KRZYSZTOF R. APT, *op. cit.*, p. 94.

⁷² EDSGER WYBE DIJKSTRA, *Twenty-eight years*, cit., p. 7.

⁷³ KRZYSZTOF R. APT, *op. cit.*, p. 98.

solida base scientifica e teorica per la scienza informatica.

A lui dobbiamo non solo l'arricchimento del vocabolario dell'informatico (stack, semafori, programmazione strutturata ecc.), ma anche la concezione della programmazione come una sfida intellettuale all'interno di un modello chiaro, semplice e matematicamente rigoroso. A lui dobbiamo l'enfasi posta sulla qualità e correttezza dei programmi (o come preferiva chiamarla: "eleganza").

«Semplicità, bellezza ed eloquenza sono state le sue caratteristiche distintive, e la sua insistenza senza compromessi sull'eleganza nella programmazione e sulla matematica è stata una fonte d'ispirazione per migliaia»⁷⁴.

⁷⁴ R. S. BOYER, W. FEIJEN, D. GRIES, C. A. R. HOARE, J. MISRA, J. MOORE, H. RICHARDS, In Memorium Edsger W. Dijkstra 1930-2002, «Communications of the ACM», vol. VL, 10, 2002, pp. 21-22: 22.

Bibliografia

- [1] DIJKSTRA E.W.: A note on two problems in connexion with graphs. *Numerische Mathematik*, Vol. I, 1959, p. 269-271.
- [2] DIJKSTRA E.W.: Go To statement considered harmful. *Communications of the ACM*, Vol. XI, n. 3, 1968, p. 147-148.
- [3] DIJKSTRA E.W., DAHLO J., HOARE C. A.R.: *Structured Programming*. London-New York, Academic Press, 1972.
- [4] DIJKSTRA E.W.: The humble programmer. *Communications of the ACM*, Vol. XV, 10, 1972, p. 859-866.
- [5] APT KRZYSZTOF R., Edsger Wybe Dijkstra (1930-2002): A Portrait of a Genius. *Formal Aspects of Computing*, Vol. XIV, n. 2, 2002, p. 92-98.

I manoscritti della serie EWD sono disponibili sul sito web www.cs.utexas.edu/users/EWD/ curato dal Dipartimento di Informatica dell'Università del Texas.

Nota. Poiché non risultano esistere traduzioni italiane delle opere di Dijkstra, i passi riportati nell'articolo sono stati tradotti per l'occasione.

LORENZO MILONE è laureato in economia e commercio, lavora professionalmente nel settore del software ed è un cultore di storia dell'informatica.

E-mail: lorenzomilone@libero.it



ICT E INNOVAZIONE D'IMPRESA

Casi di successo

Rubrica a cura di

Roberto Bellini, Chiara Francalanci

La rubrica *ICT e Innovazione d'Impresa* vuole promuovere la diffusione di una maggiore sensibilità sul contributo che le tecnologie ICT possono fornire a livello di innovazione di prodotto, di innovazione di processo e di innovazione di management. La rubrica è dedicata all'analisi e all'approfondimento sistematico di singoli casi in cui l'innovazione ICT ha avuto un ruolo critico rispetto al successo nel business, se si tratta di un'impresa, o al miglioramento radicale del livello di servizio e di diffusione di servizi, se si tratta di una organizzazione pubblica.

Product Lifecycle Management

Il Caso Faital – Fabbrica Italiana Altoparlanti

Francesco Merlo

1. INTRODUZIONE

Il caso riportato descrive le soluzioni IT per la gestione del ciclo di vita del prodotto (PLM, *Product Lifecycle Management*) adottate da un'azienda operante nel settore della produzione di altoparlanti. La duplice necessità di ottimizzazione del processo di progettazione e di integrazione della catena del valore verso clienti e fornitori ha portato Faital – Fabbrica Italiana Altoparlanti – ad una continua ricerca di innovazione nei sistemi informativi adottati, in particolare per quanto riguarda i sistemi verticalizzati specifici del settore a supporto delle fasi di progettazione e industrializzazione del prodotto. L'articolo descrive in dettaglio l'esperienza di Faital analizzandone il core business, il contesto competitivo, l'evoluzione dei sistemi informativi nell'azienda e le valutazioni e l'impatto che ha il nuovo sistema PLM in corso di realizzazione.

2. FABBRICA ITALIANA ALTOPARLANTI

La storia di Faital testimonia la costante evoluzione e la continua crescita del gruppo nel corso di oltre 50 anni di attività. A partire dalla fondazione, avvenuta nel 1958, si sono susse-

guite numerose tappe che hanno portato l'azienda a diventare un punto di riferimento di primo piano nel mercato mondiale.

Faital SpA nasce a Milano nel 1958 come piccola impresa a conduzione familiare. Alla fine degli anni Sessanta, si colloca tra i principali produttori italiani di altoparlanti, soddisfacendo ben il 40% del fabbisogno nazionale. Gli altoparlanti Faital, inizialmente destinati solamente ad impianti radio e a diffusori audio, vengono introdotti anche nel mercato dei sistemi di amplificazione per televisori. L'impresa, in continuo sviluppo sia tecnico che commerciale, si è trasformata in una società di ampio respiro in progressiva espansione.

Negli anni settanta Faital apre al mercato automobilistico e supera i confini nazionali, conquistando i primi clienti europei: al fine di soddisfare le sempre crescenti richieste del mercato e i nuovi clienti acquisiti dalla società, viene inaugurato nel 1974 lo stabilimento di Chieve, nei dintorni di Crema.

Gli anni ottanta rappresentano per Faital un momento di crescita continua: l'azienda diventa il fornitore di altoparlanti di primo impianto per conto di diverse case automobilistiche. La società avvia un processo di sviluppo tecnologico con investimenti mirati alla modernizzazione de-

gli impianti e al potenziamento del settore Ricerca e Sviluppo e delle Risorse Umane. L'azienda sceglie la strada della "qualità totale" e fissa come proprio obiettivo primario la formula "zero difetti". Nel 1985 nasce in Spagna la Fabrica Iberica de Altavoces e si inaugura il nuovo stabilimento di Barcellona.

Agli inizi degli anni novanta nasce in Francia Sofaital S.A. e viene inaugurato lo stabilimento di Argenteuil, vicino a Parigi. Qualche anno dopo viene aperta in Spagna la nuova unità produttiva di Vacarisses, presso Barcellona. Grazie alla costituzione delle due nuove consociate Sofaital in Francia e Fabrica Iberica in Spagna, il Gruppo Faital assume il ruolo di leader europeo nella produzione di altoparlanti per l'industria automobilistica.

Nel nuovo millennio nasce in Messico la Produtora Mexicana de Bocinas, per sostenere le necessità del mercato locale. Il gruppo consolida i risultati conseguiti in mezzo secolo di attività e, al fine di ben introdursi nel mercato statunitense dell'automotive, costituisce nei pressi di New York un nuovo magazzino per dare un supporto logistico ai nuovi clienti del Nord America. Analogamente, sempre con il fine di migliorare il servizio alla clientela, vede la luce in Ungheria un una nuova sede, che ad oggi rientra per importanza tra le prime tre del gruppo. Nel 2008, infine, viene aperta la nuova sede logistica di Hong Kong per rispondere alle nuove esigenze del mercato asiatico. La multinazionale, impegnata in tre continenti, conta circa 500 dipendenti per un fatturato che tocca i 45 milioni di Euro.

3. IL PRODUCT LIFECYCLE MANAGEMENT (PLM)

Un sistema di *Product Lifecycle Management* (PLM) consente la costituzione e la gestione centralizzata in un unico *repository* dell'insieme delle informazioni ed il *know-how* relativi ad un prodotto e delle tecnologie necessarie per avervi accesso. In tal modo, tutte le fasi del ciclo di vita del prodotto vengono digitalizzate: i dati legati alla concezione, progettazione, produzione e servizio vengono gestiti in modo integrato e possono essere fruiti più agevolmente dalle persone e dai processi, aumentandone quindi efficacia ed efficienza in modo sensibile. L'implementazione di un sistema PLM non si risolve quindi nella semplice introduzione di un nuovo

sistema software, ma interessa il lato strategico del business. Se l'implementazione delle tradizionali tecnologie IT (quali sistemi di office automation) ha lo scopo principale di ridurre i costi operativi dei processi, un sistema PLM consente anche di accrescerne la produttività, al pari dei sistemi ERP, CRM o SCM.

I sistemi PLM hanno visto le loro origini nell'ambito del settore automobilistico a metà degli anni Ottanta. Storicamente, il primo prototipo di sistema PLM è considerato essere quello implementato dall'*American Motors Corporation* (AMC) per la progettazione e lo sviluppo del modello *Jeep Grand Cherokee*. Le novità introdotte da tale sistema consistettero principalmente in due punti:

- l'introduzione estensiva del *Computer-Aided Design* (CAD) per la realizzazione dei progetti del veicolo, che consentì di aumentare la produttività dei gruppi di lavoro;

- la realizzazione di un unico *database* per la gestione di tutti i disegni, i progetti e le informazioni relative al nuovo modello, che permise una migliore gestione delle comunicazioni tra i team di sviluppo ed una più efficiente risoluzione delle incompatibilità tra le differenti parti del progetto. Il sistema di gestione dei dati così realizzato venne considerato talmente innovativo e vantaggioso che, quando AMC venne acquisita da Chrysler nel 1985, esso venne adottato a livello di gruppo: l'introduzione di tale tecnologia consentì a Chrysler di divenire, a metà degli anni Novanta, la Casa produttrice di automobili con i più bassi costi di sviluppo e progettazione, pari a circa la metà della media di settore.

Attualmente, i sistemi PLM si rivolgono al supporto di ogni attività facente parte del ciclo di vita di un prodotto; le soluzioni offerte dai principali vendor consentono infatti di gestire le informazioni ed il *know-how* relativi alle seguenti fasi (la cui enumerazione è puramente indicativa di un generico processo di sviluppo di prodotto):

- Concezione
 - specifica dei requisiti;
 - progettazione concettuale.
- Progettazione e industrializzazione
 - progettazione di dettaglio e ingegnerizzazione;
 - analisi e verifica delle specifiche.
- Realizzazione
 - pianificazione e gestione operativa della produzione;

- pianificazione e gestione dei processi di assemblaggio;
- controllo di qualità.

□ Servizio

- vendita e consegna al cliente;
- utilizzo;
- manutenzione e supporto;
- dismissione del prodotto.

Le soluzioni tecnologiche maggiormente adottate si basano sull'utilizzo di strumenti software avanzati per la progettazione. In particolare, le principali tipologie di strumenti sul mercato sono rivolti all'*Electronic Design Automation* (EDA), al *Mechanical CAD* (MCAD), al *Product Data Management* (PDM) ed al *Computer-Aided Engineering* (CAE). In base all'integrazione su scala aziendale, alla numerosità dei sistemi utilizzati ed alla loro interoperabilità è possibile distinguere differenti livelli di implementazione di un sistema PLM [1]:

- a.** il livello base, in cui il sistema consente la semplice transazione di progetti, schemi tecnici e documentazione tra le differenti funzioni aziendali, ed è implementato tramite l'uso di meccanismi basati su scambio di e-mail o, al più, tramite l'implementazione di un semplice *database/data warehouse* con eventuale controllo degli accessi;
- b.** il livello intermedio, in cui è possibile una gestione più sofisticata dei dati e dei metadati di prodotto (per esempio, i dati di distinta ba-

se) tra le diverse funzioni aziendali; in questo caso è tipicamente prevista la presenza di un repository dei dati mediamente complesso, in grado di gestire accessi controllati e riservati;

c. il livello avanzato, per cui il sistema prevede l'utilizzo intensivo e pervasivo di workflow automatizzati tra le diverse funzioni aziendali, per la gestione di dati di prodotto e di altre informazioni strutturate e complesse.

La tendenza evolutiva dei sistemi PLM sta convergendo verso la generazione dei cosiddetti sistemi PLM 2.0 che, mutuando alcune delle tecnologie e dei paradigmi tipici del Web 2.0, propone un approccio basato sugli strumenti di lavoro cooperativo e di condivisione della conoscenza tipici delle reti sociali. Questa nuova tipologia di sistemi consente di soddisfare necessità imprescindibili soprattutto nelle realtà che operano sfruttando il paradigma dell'impresa-rete, per cui i vari processi e la gestione delle varie fasi del ciclo di vita del prodotto sono spesso delocalizzate o esternalizzate a *business partners*.

L'implementazione di sistemi PLM, soprattutto nel caso di sistemi di livello avanzato, abilita inoltre la possibilità di pervenire a soluzioni di *business intelligence* legate ai processi di produzione, costituendo di fatto una delle pietre angolari per la realizzazione dei cosiddetti Engineering Management Information Systems [2] (Figura 1). Mediante l'utilizzo di tecnologie quali *data wa-*

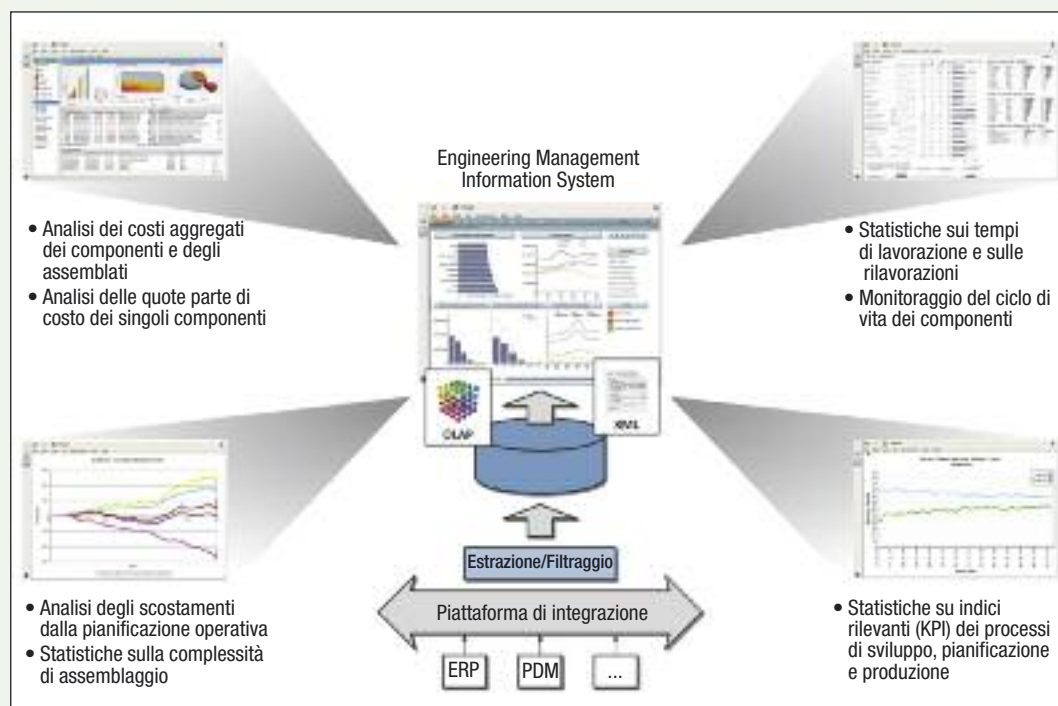


FIGURA 1
Soluzioni di business intelligence abilitate dalla realizzazione di un sistema PLM

rehouse per la memorizzazione dei dati, OLAP per la loro analisi e XML per la definizione di formati comuni per lo scambio dei dati, si ha la possibilità di effettuare analisi di tipo avanzato sul repository comune delle informazioni legate al ciclo di vita del prodotto e dei relativi processi.

Il mercato di *vendor* che a livello mondiale offrono soluzioni per sistemi PLM (Figura 2 A) vede un contesto attualmente dominato da pochi player: la francese Dassault Systèmes occupa la posizione di leadership con una quota di mercato prossima al 37%, seguita dalla tedesca Siemens al 30% circa e dall'americana PTC con il 20%. Il resto del mercato è composto da altri fornitori, come MSC Software, Altair Engineering o Agile Software Corporation (parte di Oracle Corporation). Per quanto riguarda invece la suddivisione dell'offerta (Figura 2 B), come precedentemente accennato il mercato è composto principalmente da soluzioni orientate ai settori *Electronic Design Automation* (EDA), *Mechanical CAD*, *Architecture, Engineering & Construction* (AEC), *Product Data Management* (PDM) e *Computer-Aided Engineering* (CAE).

Per quanto riguarda invece la scena italiana, le soluzioni PLM costituiscono ad oggi un settore in crescita: da un'indagine campionaria svolta da Netconsulting [3], al 2008 il 45% delle imprese analizzate possedevano già o stavano implementando soluzioni PLM (solo l'implementazione di sistemi ERP e di *Business Intelligence* trovavano percentuali maggiori di penetrazione), e comunque la necessità di dotarsi di soluzioni PLM appariva al quinto posto sulle prime otto posizioni dal punto di vista dell'importanza strategica assegnata alle soluzioni prese in considerazione.

4. L'ICT IN FAITAL

L'utilizzo della tecnologia dell'informazione in Faital ha sempre rivestito, anche storicamente, un ruolo fondamentale come fattore abilitante per il successo dell'azienda. L'introduzione dell'IT in Faital avviene a partire dalla metà degli anni Settanta, quando viene informatizzata la sezione amministrativa con sistemi proprietari basati su tecnologie IBM. Tali sistemi sono poi rimasti in uso fino al nuovo millennio: nel 2000 è stato introdotto un nuovo sistema ERP che consentisse una maggiore integrazione delle funzioni aziendali, ma anche, e soprattutto,

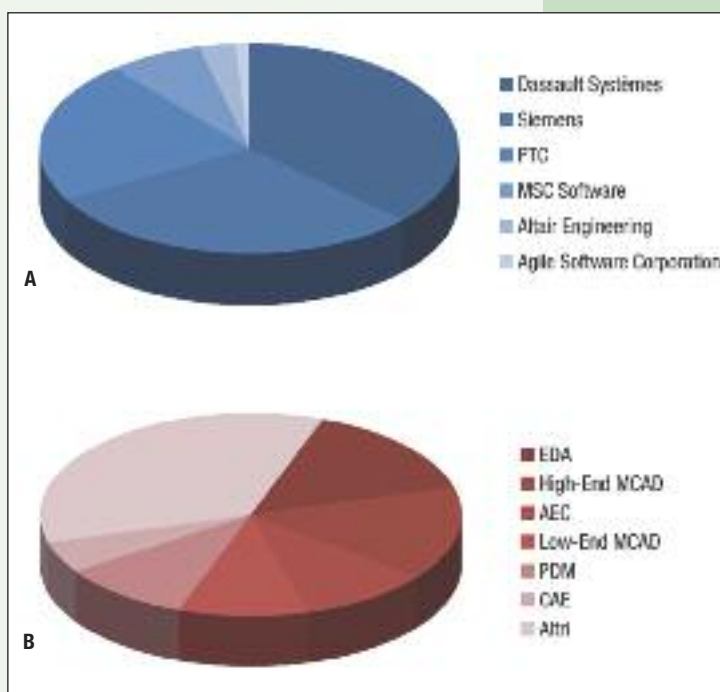


FIGURA 2

A - Principali fornitori a livello mondiale di soluzioni PLM; B - Principali aree di suddivisione dell'offerta dei software costituenti soluzioni PLM

to, con le catene del valore di clienti e fornitori. Date le peculiarità dei prodotti di Faital il sistema ERP è completamente *customizzato* e proprietario: tale scelta è stata adottata per far fronte alle particolari esigenze di "ritagliare" il sistema sui requisiti imposti dalle caratteristiche del processo produttivo. Per esempio, l'ERP comprende un modulo per lo scambio di dati in formato elettronico con i clienti (EDI, *Electronic Data Interchange*). Tale modulo è stato realizzato con caratteristiche specifiche per ciascun cliente con cui Faital interagisce (in particolare nel settore automotive), poiché lo scambio di informazioni viene definito caso per caso ed i dati richiesti sono specifici e raramente standardizzati nel formato.

Una seconda area di utilizzo della tecnologia dell'informazione in Faital riguarda l'automazione del processo produttivo: tutte le linee di produzione sono ad oggi automatizzate e allo stato dell'arte, secondo il paradigma CIM (*Computer Integrated Manufacturing*). Date le peculiarità del prodotto, ma soprattutto la particolare attenzione dell'azienda alla qualità finale degli altoparlanti, al termine delle linee di produzione sono impiegate macchine automatizzate per il controllo di qualità di ogni singolo pezzo

finito, sia dal punto di vista strutturale (per esempio, correttezza dell'assemblaggio) che funzionale (ad esempio, corretta risposta in frequenza a sollecitazioni note).

Infine, l'IT viene estensivamente utilizzato in Faital anche nel processo di ricerca e sviluppo, che comprende la fase fondamentale di progettazione degli altoparlanti. L'ufficio tecnico dell'azienda ha introdotto i primi sistemi CAD in 2D dalla metà degli anni Ottanta, consentendo la migrazione di parte del processo di progettazione dal cartaceo al digitale. Successivamente, in concomitanza con l'acquisizione dei primi clienti nel settore automotive (primi anni novanta), sono stati introdotti i primi sistemi di CAD in 3D. La successiva evoluzione dei sistemi informativi a supporto della progettazione è stata dettata dai requisiti e dalle specifiche esigenze dei clienti, seguendo sempre un'ottica di ricerca della massima integrazione per l'ottimizzazione dei processi. A fianco dei sistemi CAD per la progettazione vera e propria sono quindi stati introdotti nel tempo anche sistemi CAE (*Computer-Aided Engineering*) che consentissero la caratterizzazione specifica dei prodotti per lo svolgimento di analisi predittive di comportamento senza ricorrere alla fase di prototipazione dei prodotti (per esempio, per verificare la risposta degli altoparlanti a sollecitazioni fisiche esterne equivalenti a quelle riscontrate durante la marcia di un autoveicolo). L'introduzione dei sistemi CAE e CAD in 3D ha costituito una tappa fondamentale nell'evoluzione dei sistemi informativi adottati per aggiungere valore al rapporto con i clienti, soprattutto nel settore automotive: tramite questi sistemi è infatti possibile adottare il paradigma della progettazione cooperativa, che, tramite lo scambio di modelli 3D e matematiche descrittive, consente un'analisi più efficace delle migliori soluzioni adottabili e la risoluzione preventiva delle problematiche di compatibilità meccanica.

5. IMPLEMENTAZIONE DEL SISTEMA PLM

Muovendosi nello scenario precedentemente descritto, Faital si è venuta a trovare in una situazione in cui la tecnologia dell'informazione ha raggiunto un elevato livello di penetrazione nei processi, ma sostanzialmente resta ancora

suddivisa nelle tre aree precedentemente evidenziate:

- la progettazione degli altoparlanti;
- le linee di produzione;
- le funzioni supportate dal sistema ERP.

Sebbene le funzionalità offerte dai sistemi informativi in uso consentissero già un buon livello di integrazione delle funzioni aziendali soprattutto verso le catene del valore di fornitori e clienti, ciò di cui Faital avvertiva la mancanza era l'integrazione dei sistemi dal punto di vista interno. Tropo spesso, infatti, grosse inefficienze dei *work-flow* venivano evidenziate dall'esperienza quotidiana. Per esempio, la gestione delle revisioni ai progetti costituiva un processo dispendioso in termini di tempo e materiali: una volta approvata dall'ufficio tecnico, le nuove versioni dei progetti dei componenti o dei prodotti dovevano essere rese note a livello di gruppo, e tale operazione avveniva inviando copie cartacee firmate a tutte le sedi, che le dovevano quindi conservare nei loro archivi.

Al fine di ovviare a tali problematiche e di ottimizzare l'esecuzione dei processi di progettazione dei prodotti, la decisione presa nel corso del 2008 è stata quella di procedere all'implementazione di un sistema PLM integrato. La *vision* del progetto consiste nell'avviare un processo che dovrebbe consentire, nell'arco di qualche anno, un'integrazione completa a livello aziendale dei sistemi informativi utilizzati nelle tre aree precedentemente identificate. La prima fase del processo, attualmente in corso, riguarda l'implementazione del sistema PLM vero e proprio a beneficio dei processi di progettazione e produzione, mentre una seconda fase prevederà l'estensione dell'interoperabilità dei sistemi informativi mediante l'integrazione con il portafoglio applicativo istituzionale.

Il fondamentale vantaggio, derivante dall'implementazione del sistema PLM, percepito da Faital è la possibilità di "allargare" l'accessibilità ai dati tecnici, finora di pertinenza esclusiva dell'ufficio tecnico, anche al resto dell'azienda. In questo modo è possibile quindi ristrutturare la fase di progettazione in un'ottica di collaborazione tra più attori:

- i progettisti dell'ufficio tecnico, *process owner* del processo di progettazione e punto di riferimento interno di Faital per la definizione dei progetti operativi dei nuovi prodotti;
- gli altri utenti aziendali, come per esempio i

tecniche responsabili del processo di produzione, che possono interagire proponendo *feedback* sulle soluzioni progettuali adottate in prospettiva del processo di produzione;

□ i fornitori e i clienti, che, se necessario, possono avere accesso ai dati tecnici di prodotto. La collaborazione di tali attori al processo di progettazione è resa possibile dalle funzionalità offerte dal sistema scelto da Faital. In particolare, il sistema consente l'unificazione e il collegamento di tutti i documenti attinenti alla progettazione di un prodotto: in tal modo, si adotta una prospettiva centrata sul progetto, che dispone di uno spazio logico all'interno del sistema in cui vengono raccolti progetti tecnici, documentazione descrittiva, note, revisioni, e anche documenti accessori come link a pagine web, filmati, o altro. Il sistema consente quindi ai vari utenti, mediante un meccanismo di accessi controllati, di visionare i contenuti informativi relativi al progetto e interagire con essi.

Più specificatamente, l'implementazione del sistema PLM permette a Faital di concretizzare tre obiettivi principali:

□ **Consentire una collaborazione più stretta tra le varie funzioni aziendali durante la fase di progettazione.** La possibilità di estendere l'accessibilità agli addetti delle altre funzioni aziendali è avvertita da Faital come necessità inderogabile. L'implementazione del sistema PLM consente di adottare tecniche di industrializzazione dei prodotti tipiche del *concurrent engineering*, pervenendo quindi ad una situazione in cui le differenti funzioni aziendali non intervengono più sequenzialmente nelle fasi di progettazione e industrializzazione, ma cooperano nella definizione collaborativa dei progetti e nella risoluzione delle incompatibilità progettuali eventualmente riscontrate.

□ **Semplificare i flussi documentali e pervenire ad una completa digitalizzazione dei documenti di progetto.** L'introduzione del sistema PLM implica per Faital un radicale cambiamento di prospettiva nella gestione dei flussi documentali. Per esempio, il già accennato meccanismo di gestione delle revisioni di progetto comporta numerosi scambi di informazioni tra differenti funzioni aziendali e l'invio di grandi quantità di documenti cartacei tra le varie sedi dell'azienda (situate peraltro in tre diversi continenti). La possibilità di disporre all'interno di un unico repository tutte le informazioni relati-

ve al progetto ed alle sue revisioni costituisce un primo ed indubbio vantaggio legato all'implementazione del sistema PLM. A questo si aggiunge la completa dematerializzazione della documentazione di progetto: mediante la definizione di opportune *policy* di accesso ai dati, non è più necessario provvedere alla distribuzione della documentazione definitiva in seguito ad ogni revisione di progetto, poiché le singole sedi potranno avere accesso diretto ai documenti originale contenuti nel repository una volta che questi siano stati approvati dall'ufficio tecnico centrale. Ne consegue una drastica riduzione (che può anche culminare nel completo annullamento) dei volumi di documenti cartacei utilizzati e scambiati tra le varie sedi, evitando l'insorgere delle difficoltà legate al mantenimento degli archivi con la documentazione aggiornata ed i costi legati all'uso ed alla movimentazione della carta stessa.

□ **Ridurre il time to market.** Come già discusso, l'introduzione del sistema PLM consente di ottenere una maggiore integrazione delle operazioni svolte dalle varie funzioni aziendali di Faital durante tutto il ciclo di vita del prodotto. A partire dalla fase di progettazione, la possibilità di disporre di un unico punto di accesso ai dati tecnici permette agli addetti di tutte le altre funzioni aziendali di poter collaborare attivamente alla realizzazione del progetto fin dalle prime fasi. Grazie al *concurrent engineering*, quindi, è possibile introdurre maggiore efficienza nello svolgimento dei processi e, conseguentemente, pervenire ad una riduzione dei tempi complessivi di progettazione, ingegnerizzazione, prototipazione e produzione.

La riduzione del *time to market* rappresenta per Faital un obiettivo particolarmente significativo, anche dal punto di vista strategico. Dai primi anni 2000, infatti, è stato introdotto un nuovo *brand* di altoparlanti, denominato FaitalPRO, con lo scopo di diversificare l'offerta dei prodotti. Il nuovo *brand* si rivolge a clienti professionali con elevate esigenze in termini prestazionali, e prevede differenti linee di altoparlanti: tra le altre, i prodotti a catalogo, sviluppati internamente da Faital e commercializzati in risposta alla richiesta di mercato, ed i prodotti personalizzati, per cui è prevista una stretta interazione con il cliente fin dalle fasi di progettazione. Ad oggi, la linea FaitalPRO costituisce una quota estremamente rilevante del volume di vendite dell'azien-

da, e su di essa si concentra particolarmente l'attenzione del management. Il *time to market* tipico di un prodotto della linea FaitalPRO si attesta intorno ai 6/7 mesi, in linea con la media di settore. L'introduzione del sistema PLM consentirà di ridurre drasticamente tali tempistiche, con un ordine di miglioramento prossimo al 25% del *time to market* complessivo (ottenendo quindi un risparmio di tempo quantificabile in 6-8 settimane). La stessa riduzione del *time to market* è prevista anche per quanto riguarda i prodotti "classici" dell'azienda, cioè gli altoparlanti destinati tipicamente al mercato automotive, il cui *time to market* medio è di poco superiore all'anno: in questo caso, il risparmio in termini di tempo può arrivare fino a tre mesi.

6. CONCLUSIONI

Risulta evidente come la costante innovazione introdotta da Faital nei processi di ricerca e sviluppo e di produzione grazie alle tecnologie dell'informazione sia stata elemento fondamentale per il successo dell'azienda. In particolare, se si considera il settore automotive, l'innovazione di processo attuabile mediante l'adozione di sistemi CAD/CAE prima e PLM poi risulta fondamentale per garantire la permanenza nel mercato. Dato l'elevato potere con-

trattuale dei clienti, ovvero delle case automobilistiche, la possibilità di rispondere alle esigenze dettate in termini di capacità di integrazione dei processi e di adattamento alle specifiche progettuali degli autoveicoli assume importanza strategica per un player come Faital. L'introduzione del sistema PLM risulta quindi scelta obbligata nella prospettiva di riuscire a migliorare la flessibilità dei processi interni, di introdurre maggiore efficienza nel loro svolgimento e maggiore efficacia nelle fasi progettuali tramite l'adozione del *concurrent engineering*. Inoltre, la possibilità di ridurre sensibilmente il *time to market* dei nuovi prodotti costituisce un vantaggio competitivo nei confronti dei concorrenti di assoluta rilevanza, soprattutto nell'ottica di competizione con i nuovi player del mercato asiatico.

Bibliografia

- [1] Terzi S., Garetti M., Balocco R.: Il PLM in Italia. *Sistemi & Impresa*, n. 7, 2007, p. 70-75.
- [2] Abramovici M., Sieg O. C.: *Status and development trends of product lifecycle management systems*. Conference on Integrated Product and Process Development, 2002.
- [3] Assinform: *Rapporto sull'informatica, le telecomunicazioni ed i contenuti multimediali*. 2008.

FRANCESCO MERLO conseguirà il Dottorato di Ricerca in Ingegneria dell'Informazione presso il Politecnico di Milano nel 2010; presso lo stesso ateneo ha conseguito la Laurea Specialistica in Ingegneria Informatica nel 2006 e la Laurea in Ingegneria Informatica nel 2003. Durante gli studi di dottorato è stato visiting researcher presso il College of Management del Georgia Institute of Technology (Atlanta, USA). Collabora con la Fondazione Politecnico di Milano in qualità di consulente in ambito delle tecnologie dell'informazione. Si occupa di ricerca su tematiche legate al Green ICT, alla qualità ed ai costi di sviluppo e manutenzione del software; su questi temi è autore di pubblicazioni nazionali e internazionali. Dal 2006 tiene presso il Politecnico di Milano parte dei corsi di Sistemi Informativi per l'Impresa e di Gestione della Tecnologia dell'Informazione. E-mail: merlo@elet.polimi.it

ICT E DIRITTO

Rubrica a cura di

Antonio Piva, David D'Agostini

Scopo di questa rubrica è di illustrare al lettore, in brevi articoli, le tematiche giuridiche più significative del settore ICT: dalla tutela del *domain name* al *copyright* nella rete, dalle licenze software alla *privacy* nell'era digitale. Ogni numero tratterà un argomento, inquadrandolo nel contesto normativo e focalizzandone gli aspetti di informatica giuridica.



Aspetti giuridici dei social network tra opportunità e rischi

David D'Agostini, Antonio Piva

1. INTRODUZIONE

Con la locuzione “*social network*” si può intendere qualsiasi gruppo di persone connesse tra loro da diversi legami che vanno dalla conoscenza casuale, ai rapporti di lavoro, ai vincoli familiari.

Come noto, fin dai tempi di Platone l'uomo è considerato un “animale sociale”, pertanto la sua storia è caratterizzata dalla continua ricerca di relazioni interpersonali e dalla creazione di reti di individui: si pensi a una comunità di lavoratori che costituisce un'associazione sindacale, a un gruppo di atleti che si incontra per praticare lo stesso sport, agli appartenenti alla medesima confessione religiosa, agli iscritti a un unico club esclusivo e, in generale, a tutti i casi in cui... l'unione fa la forza.

Che questo rappresenti un valore fondamentale e giuridicamente tutelabile è dimostrato dal fatto che le Costituzioni dei Paesi democratici moderni contemplano e difendono la libertà di associazione e il diritto di ogni uomo a realizzarsi non solo come individuo, ma anche nella sua dimensione collettiva.

Ovviamente, fino a ora, ogni rete associativa ha dovuto fare i conti con i propri limiti fisici: il luogo in cui incontrarsi, le distanze tra i consociati, il numero dei membri; tali aspetti possono essere superati grazie alle nuove tecnologie che permettono a un numero maggiore di persone di comunicare tra loro, anche a grandi distanze, in tempo reale e a costi accessibili.

Fin dagli anni '70 - '80 i primi utenti della rete internet potevano fruire di sistemi come Usenet, i *Bulletin Board Services* (BBS); poi si svilupparono siti web con applicazioni che consentivano di creare vere e proprie comunità virtuali in cui i navigatori discutevano dei loro argomenti preferiti.

Così è nato il fenomeno dei “*social network virtuali*” che negli ultimi anni ha conosciuto una vera esplosione grazie alla popolarità di siti come Facebook, Myspace, LinkedIn: il *community networking* sta generando approcci innovativi al modo di conoscersi, lavorare, relazionarsi di molte persone, con nuove tipologie di applicazioni un tempo impensate (riquadro). Ma con alcuni rischi.

I social network maggiormente diffusi

Facebook, MySpace, Hi5, Flickr, Skyrock, Friendster, Tagged, LiveJournal, Orkut, Fotolog, Bebo.com, LinkedIn, Badoo.Com, Multiply, Imeem, Ning, Last.fm, Twitter, MyYearbook, V Kontakte, aSmallWorld, Windows Live, Xiaonei.

2. FUNZIONAMENTO

Per far parte di un *social network online* occorre in primo luogo registrarsi compilando il proprio profilo personale, ossia immettendo informazioni come l'indirizzo e-mail, i propri dati anagrafici, fino ad arrivare agli interessi (per la vita privata) ovvero alle esperienze di

lavoro e relative referenze (per l'ambito professionale).

Una volta creato un account, è possibile contattare colleghi, amici, conoscenti già presenti nel network o invitarli a iscriversi, cosicché il numero dei partecipanti via via si allarga.

Naturalmente è possibile far parte o costituire *communities* tematiche in base alle proprie passioni o aree di lavoro, aggregando altri utenti e stringendo contatti di amicizia o di affari.

Tutto ciò favorisce la diffusione di dati personali relativi agli iscritti, ma anche a soggetti terzi, in una misura che non ha precedenti con pubblicazione di informazioni che divengono disponibili in modo globale, anche attraverso enormi quantità di fotografie e filmati digitali a cui è possibile attribuire un'etichetta (tag da cui il neologismo "taggare" che generalmente definisce la funzione di indicare il nome di una persona all'interno di una foto o di un video).

Rispetto al rischio che quanto contenuto nel profilo-utente venga riprodotto, allo stato attuale non vi sono che scarse tutele: tali dati possono essere copiati da altri membri del network, o perfino da terzi non autorizzati, e quindi venire utilizzati per costruire profili personali oppure essere ripubblicati altrove.

Talora risulta assai difficile, o addirittura impossibile, ottenere la totale eliminazione dei propri dati dalla rete una volta che siano stati pubblicati; possono essere conservati nel server del servizio di *social network* anche dopo la cancellazione dell'utente e, comunque, chiunque potrebbe averli copiati e messi on line anche su altri siti.

Inoltre i dati personali contenuti nei profili possono essere indicizzati da un motore di ricerca, mentre alcuni fornitori di questi servizi consentono a terzi di accedere ai dati relativi agli utenti attraverso API (interfacce di programmazione applicazioni), in maniera da disporne liberamente.

Fra gli esempi di utilizzo ulteriore dei dati, si può citare la prassi invalsa presso molti uffici del personale di varie aziende di ricercare i profili-utente relativi a candidati all'assunzione o singoli dipendenti: molti responsabili delle risorse umane utilizzano informazioni tratte da servizi di *social network* per verificare o completare le informazioni fornite dai candidati all'assunzione.

Le informazioni contenute nei profili-utente e i dati di traffico sono utilizzati anche per l'invio di messaggi mirati di marketing.

3. TUTELA DELLA PRIVACY

Gli esempi portati fanno intendere quanto sia concreta la possibilità che la privacy, non solo degli utenti del *social network*, ma soprattutto di terzi, venga violata.

Il menzionato fenomeno del tagging, utile al fine di riordinare i documenti per argomento o per soggetti, presenta come rovescio della medaglia il pericolo di un'indicizzazione globale inconsapevole.

Sotto il profilo tecnico, non esiste attualmente un valido rimedio al fatto che chiunque possa inserire foto altrui e taggarle; paradossalmente i siti di *social network* permettono una determinata limitazione (modificando le impostazioni di privacy) in modo che l'interessato possa scegliere chi può vedere la foto con il suo tag, ma ciò lascia completamente sfornito di tutela le persone non iscritte che sono completamente ignare della pubblicazione di proprie immagini (o dati in genere).

Vi è un'ulteriore prospettiva alquanto preoccupante, forse non sempre debitamente valutata dagli utenti, vale a dire il rischio che quanto messo in rete ora resti in linea a vita: giova ricordare che in tema di privacy (originariamente intesa come *diritto a essere lasciato solo*¹) sussiste un autentico diritto all'oblio, collocato dalla giurisprudenza tra i diritti inviolabili riconosciuti dall'art. 2 della Costituzione.

In buona sostanza ogni individuo ha il diritto a essere dimenticato, ovvero a non essere ricordato per fatti passati, ormai appartenenti a una sfera privata, che non vuole più rendere pubblici: si pensi al caso di una persona che si converta a un'altra religione e voglia cancellare dalla rete ogni traccia della precedente fede (peraltro si tratterebbe di dati sensibili secondo il dettato normativo).

Ma cosa succederebbe se l'utente non potesse cancellare tali dati? Oppure qualora, anche rimuovendoli dal server del *social network*, fossero stati copiati (anche per funzione di "cache") o riprodotti in altri siti o reti? Ogni utente (e non solo, visto che possono essere inserite anche informazioni di terzi) potrebbe ritrovarsi in un futuro apparentemente lontano a vivere con l'ingombrante ombra di un passato che non passa.

¹ Sull'argomento si vedano gli articoli pubblicati in questa rubrica su Mondo Digitale n. 9 e n. 22.

Considerati tutti questi aspetti, la Conferenza internazionale delle Autorità di protezione dei dati tenutasi a Strasburgo nell'ottobre 2008 ha approvato una "Risoluzione sulla tutela della privacy nei servizi di social network"² in cui, riprendendo i concetti già espressi nel "Memorandum di Roma"³ del marzo 2008, si evidenziano i principali pericoli per la privacy di milioni di persone. Le Autorità Garanti, conseguentemente, per fronteggiare il fenomeno hanno impartito alcune precise raccomandazioni da un lato agli utenti, che devono valutare con attenzione se e in quale misura pubblicare dati personali in un profilo creato sui servizi di *social network* e sono comunque tenuti a rispettare la privacy altrui; dall'altro ai fornitori di tali servizi che devono ottemperare la normativa in materia di protezione dei dati, in particolare informando gli utenti in merito al trattamento dei dati personali che li riguardano e alle conseguenze derivanti dalla loro pubblicazione sul profilo.

È necessario che i fornitori mettano gli utenti in condizione di decidere l'utilizzo dei dati contenuti nei rispettivi profili per quanto riguarda i membri della comunità, consentendo di limitare la visibilità dell'intero profilo, nonché di singoli dati contenuti nel medesimo o ottenuti attraverso funzioni di ricerca messe a disposizione della comunità.

Devono permettere agli utenti di decidere sugli utilizzi dei dati contenuti nei rispettivi profili offrendo la possibilità di negare il consenso (opt-out) rispetto all'utilizzo dei dati non sensibili contenuti nel profilo e prevedendo un consenso preventivo (opt-in) rispetto all'utilizzo di dati di natura sensibile contenuti nel profilo (per esempio i dati relativi a opinioni politiche o all'orientamento sessuale).

I fornitori devono riconoscere alle persone (siano esse membri del servizio o meno) il diritto di accedere e, se necessario, apportare modifiche a tutti i dati personali detenuti dai fornitori stessi, permettendo l'integrale cancellazione del rispettivo profilo e di ogni contenuto o informazione pubblicato attraverso il servizio di *social network*.

Hanno pure l'onere di adottare misure idonee a

impedire che soggetti terzi possano raccogliere attraverso dispositivi di *spidering* e/o scaricare (o raccogliere) in massa i dati contenuti nei profili-utente e che i dati relativi agli utenti siano navigabili da parte dei motori di ricerca soltanto con il previo consenso espresso ed informato da parte del singolo utente.

Pur essendo raccomandazioni non vincolanti (lo sono, invece, le norme contenute nel cosiddetto Codice della privacy approvato con il celebre d.lgs. 30 giugno 2003 n.196), rappresentano in modo significativo la corretta modalità di affrontare la tematica della riservatezza nei *social network*.

Dal canto suo il Garante per la privacy non ha mancato di far sentire la propria voce sul tema: si richiamano alcune recenti pronunce nelle quali è stato stigmatizzato il comportamento di giornali e telegiornali che (al fine di dare un volto a una delle vittime del terremoto in Abruzzo) hanno diffuso fotografie tratte da un *social network*, senza verificare la corrispondenza di identità tra la persona ivi rappresentata, ancora in vita, e l'omonima deceduta nel sisma.

Rilevata l'illiceità di tale trattamento di dati personali e disposto il divieto di diffondere ulteriormente tali fotografie, resta ovviamente impreviudicato il diritto al risarcimento del danno da far valere in sede civile, posto che il Garante non ha competenza in tale ambito.

4. REATI NEI SOCIAL NETWORK

Fra gli altri rischi specifici per la sicurezza già individuati dal menzionato "Memorandum di Roma", si ricorda l'incremento dei furti di identità favorito dalla diffusa disponibilità dei dati personali contenuti nei profili-utente e dalla cattura dei medesimi a opera di terzi non autorizzati. Tale ipotesi può configurare il reato previsto dall'art. 167 del d.lgs. 196/03 che, sempre in tema di trattamento di dati personali, punisce chi al fine di trarne profitto o di recare ad altri un danno procede a un trattamento illecito di dati personali, se dal fatto deriva nocumento, con la pena della reclusione da sei a diciotto mesi; se vi è comunicazione o diffusione di dati, la reclusione va da sei a ventiquattro mesi. Si tratta di un reato procedibile d'ufficio e, naturalmente, la persona offesa può costituirsi parte civile chiedendo la condanna al risarcimento del danno (anche morale) subito.

Potrebbe essere, altresì, contestata la violazio-

² Si veda il sito www.garanteprivacy.it/garante/doc.jsp?ID=1560428 [doc. web n. 1560428]

³ Si veda il sito www.garanteprivacy.it/garante/doc.jsp?ID=1567124 [doc. web n. 1567124]

ne dell'art. 494 del codice penale che sanziona chiunque, al fine di procurare a sé o ad altri un vantaggio o di recare un danno, induce qualcuno in errore, sostituendo illegittimamente la propria all'altrui persona, o attribuendo a sé o ad altri un falso nome.

Molto spesso il furto d'identità è poi funzionale e propedeutico alla commissione di ulteriori illeciti; su questo tema i problemi tuttora irrisolti per quanto concerne la sicurezza dei servizi Internet costituiscono un rischio ulteriore connesso all'utilizzo dei servizi di *social network*.

L'ENISA (*European Network and Information Security Agency*) fa menzione di casi di *scripting* fra siti diversi, virus e "vermi", phishing mirato e forme di phishing specifiche dei servizi di *social network*, l'infiltrazione della rete, l'utilizzo abusivo di profili-utente e attacchi reputazionali basati sul furto di identità, nonché forme di persecuzione personale⁴.

Si tratta di reati di natura eterogenea, dei quali possono rimanere vittime i titolari di un account: tali illeciti, pur con le peculiarità del caso, possono trovare idonea risposta sanzionatoria nelle norme del codice penale che puniscono i reati informatici quali l'accesso abusivo (art. 615 ter c.p.), la detenzione abusiva di codici d'accesso (art. 615 quater c.p.), il danneggiamento informatico (art. 635 bis c.p.), le frodi telematiche (art. 640 ter c.p.)⁵.

Un cenno a parte merita la condotta della cd. persecuzione in rete (definita *cyber-stalking*) in quanto rientra a pieno titolo nel reato recentemente introdotto nel codice penale all'art. 612 bis rubricato "*atti persecutori*": viene punito con la reclusione da sei mesi a quattro anni chiunque, con condotte reiterate, minaccia o molesta taluno in modo da cagionare un perdurante e grave stato di ansia o di paura ovvero da ingenerare un fondato timore per l'incolumità propria o di un prossimo congiunto o di persona al medesimo legata da relazione affettiva ovvero da costringere lo stesso ad alterare le proprie abitudini di vita.

Non è difficile immaginare che tale norma possa trovare applicazione significativa su Internet e, in particolare, nelle comunità virtuali.

Vi sono, poi, alcune ipotesi delittuose che pos-

sono essere commesse dall'utente di un *social network*: si pensi alla pubblicazione di immagini relative alla vita privata come già alcuni episodi di cronaca hanno dimostrato.

Nel caso in esame un'infermiera, volendo condividere la sua vita lavorativa, ha pubblicato su un *social network* alcune fotografie in cui ha immortalato i pazienti intubati durante le operazioni chirurgiche.

La Procura della Repubblica presso il Tribunale di Udine, territorialmente competente, ha ipotizzato la violazione dell'art. 615 bis c.p. secondo il quale viene punito con la reclusione da sei mesi a quattro anni chiunque, mediante l'uso di strumenti di ripresa visiva o sonora, si procura indebitamente immagini attinenti alla vita privata e chi le diffonde, mediante qualsiasi mezzo di informazione al pubblico.

Da ultimo, ma non per importanza, si fa menzione al caso ancora più rilevante per frequenza, vale a dire ai comportamenti e alle affermazioni offensive della dignità e reputazione che configurano il delitto di diffamazione previsto e punito dall'art. 595 del codice penale, considerando che l'uso dello strumento telematico costituisce aggravante specifica di tale reato⁶.

5. SOCIAL NETWORK SUL LUOGO DI LAVORO E NELLE SCUOLE

Tra le conseguenze non volute della diffusione dei *social network*, per la rilevanza del fenomeno si registra anche l'uso che ne viene fatto sul luogo di lavoro.

Che ciò sia visto come un problema di dimensioni rilevanti è dimostrato dall'intenzione annunciata dal Ministro per la Pubblica Amministrazione e l'Innovazione Brunetta di installare un filtro che impedisca l'accedere a Facebook dai computer dei dipendenti pubblici.

Tale iniziativa non rappresenta un'idea nuova, posto che già alcuni enti pubblici (come per esempio il Comune di Torino che pure aveva limitato la durata massima di navigazione durante la pausa pranzo) e società private erano intervenuti con appositi filtri.

Il tema non è di poco conto se si considera la

⁴ Cfr. ENISA Position Paper No.1: "*Security Issues and Recommendations for Online Social Networks*", ottobre 2007.

⁵ Per la trattazione del fenomeno del phishing si rinvia all'articolo pubblicato su questa rubrica in Mondo Digitale n. 20.

⁶ Per l'approfondimento di tale argomento si rinvia all'articolo pubblicato su questa rubrica in Mondo Digitale n. 24.

quantità di ore/uomo che potrebbero essere sottratte alle mansioni lavorative per dedicarsi ad attività meno produttive.

Oltre a tale valutazione, il datore di lavoro risulta tenuto ad adottare misure di sicurezza finalizzate a evitare che attraverso la rete vengano commessi illeciti dei quali l'azienda risponderebbe quanto meno patrimonialmente in solido con il dipendente: si pensi al caso di violazione del diritto d'autore commesse da quest'ultimo diffondendo in un *social network* (o su YouTube) materiale tutelato da *copyright*.

In tale ipotesi il datore di lavoro sarà responsabile dei danni provocati e dovrà farsi carico del risarcimento, salvo poi rivalersi sul dipendente; qualora ne ricorrano i presupposti, potranno essere applicate anche le sanzioni disciplinari previste dall'art. 2106 del codice civile e dall'art. 7 dello Statuto dei lavoratori (legge 300/70).

Secondo alcuni studi⁷, il ricorso al *social network* (ma in generale la navigazione su internet o l'uso di sistemi di messaggistica istantanea) durante l'orario di lavoro migliorerebbe sensibilmente la produttività dei lavoratori. A prescindere dalla fondatezza di tali conclusioni, in realtà - sotto il profilo normativo - il dipendente che utilizzi il computer aziendale per scopi diversi dalle sue mansioni risulta violare il dovere di diligenza e rischia le seguenti sanzioni, che andranno irrogate secondo la gravità e il ripetersi degli illeciti:

- richiamo verbale;
- ammonizione scritta;

□ multa, che non può comunque essere di importo superiore a 4 h di retribuzione;

□ sospensione dal lavoro e dalla retribuzione, che non può comunque superare i 10 giorni;

□ licenziamento disciplinare.

Il datore di lavoro potrà, quindi, legittimamente applicare alla rete aziendale delle limitazioni all'uso di internet e dei *social network* ed eventualmente sanzionare le violazioni (sarebbe opportuno che in proposito venisse redatta e divulgata un'ideale *policy aziendale*); non potrà invece spingersi sino a tracciare la navigazione del dipendente in quanto ciò sarebbe apertamente in contrasto con il menzionato Statuto dei lavoratori, che vieta l'uso di sistemi informatici aventi lo scopo di spiare i comportamenti dei medesimi.

Oltre che nei luoghi di lavoro, vi è la forte preoccupazione della diffusione dei *social network* anche negli istituti scolastici: si ricorda, in proposito, l'opportunità che l'utilizzo degli strumenti informatici, soprattutto da parte degli studenti minorenni nei laboratori informatici, venga disciplinato da un apposito regolamento e che si adottino programmi di *web filtering* al fine di implementare una modalità di controllo preventiva⁸. Ciò risulta oltremodo utile anche per prevenire ulteriori rischi legati all'adesamento di minori e alla pedofilia.

Concludendo, mai come in relazione ai rischi e alle opportunità sopra esposti, risulta auspicabile che gli utenti siano correttamente informati e formati⁹ circa l'uso consapevole dei nuovi servizi resi disponibili da internet.

⁷ Si cita, tra i tanti, lo studio recentemente pubblicato dal dottor Coker del Dipartimento management e marketing dell'Università di Melbourne reperibile sul sito internet www.unimelb.edu.au

⁸ Su tali concetti e, in generale, sulla sicurezza informatica nelle scuole si richiama l'articolo pubblicato su questa rubrica in Mondo Digitale n. 26.

⁹ Il Codice della Privacy prevede l'obbligo di formazione; il programma di certificazione ECDL (syllabus 5) è stato recentemente aggiornato con gli aspetti giuridici sull'uso del computer per fornire competenze adeguate agli utenti

DAVID D'AGOSTINI avvocato, master in informatica giuridica e diritto delle nuove tecnologie, collabora all'attività di ricerca scientifica dell'Università degli studi di Udine e ha fondato l'associazione "Centro Innovazione & Diritto". È componente della Commissione Informatica dei Consigli dell'Ordine del Triveneto, responsabile dell'area "Diritto & informatica" della rivista "Il foro friulano", membro dell'organo di Audit Interno di Autovie Venete SpA. E-mail: studio@avvocatodagostini.it

ANTONIO PIVA, laureato in Scienze dell'Informazione, Vice Presidente dell'ALSI (Associazione Nazionale Laureati in Scienze dell'Informazione ed Informatica) e Presidente della commissione di informatica giuridica. Docente a contratto di *diritto dell'ICT e qualità* all'Università di Udine. Consulente sistemi informatici e Governo Elettronico nella PA locale, valutatore di sistemi di qualità ISO9000 ed ispettore AICA. E-mail: antonio@piva.mobi



PROFESSIONE ICT

Competenze e professionalità per l'innovazione digitale

Rubrica a cura di

Roberto Bellini, Federico Butera, Alfonso Fuggetta

Il tema dell'innovazione e della competitività del sistema Italia è all'ordine del giorno della discussione economica e di quella sulle politiche industriali; sono promosse iniziative istituzionali a supporto dell'innovazione e si auspica un maggiore contributo della ricerca a livello universitario e privato. Anche l'Unione Europea spinge sul tema dell'innovazione, in particolare sul ruolo che le tecnologie ICT possono svolgere sia nei sistemi industriali che nei sistemi di governo e sull'importanza che può avere la definizione di un *framework* comune delle competenze ICT, compatibile con quanto previsto dall'EQF - *European Qualification Framework* - recentemente approvato dall'Unione Europea (2006).

Mondo Digitale vuole sostenere la diffusione di una maggiore sensibilità sul contributo che le competenze e le professionalità relative alle tecnologie digitali possono fornire in termini di innovazione dei servizi e del business dell'Impresa e di servizi per la cittadinanza erogati dagli enti della Pubblica Amministrazione. Questa nuova rubrica è dedicata appunto all'approfondimento sistematico di tutti gli aspetti che riguardano i progetti di analisi e di miglioramento delle competenze per l'innovazione digitale, il monitoraggio dei bisogni di competenza richiesti dal mercato e la valutazione delle offerte di qualificazione e aggiornamento delle competenze proposte dalle istituzioni educative di base e dagli operatori della formazione professionale e permanente.

La rubrica analizzerà l'andamento del mercato del lavoro delle professionalità ICT, i casi di successo nella crescita di competenze del personale dei fornitori di tecnologie e servizi e degli specialisti ICT, sia delle imprese manifatturiere e di servizio che degli enti della Pubblica Amministrazione, nonché l'andamento delle retribuzioni a livello nazionale e internazionale, usando come riferimento i profili e le competenze dello Standard EUCIP che AICA promuove in Italia.

La nuova frontiera dei servizi digitali: quali competenze tecnologiche

Fulvia Sala, Roberto Bellini

1. INTRODUZIONE

La crescita economica e sociale dipende sempre di più dai **servizi**; l'importanza dei servizi sia in termini economici che di personale impiegato, richiede che questi siano accuratamente **progettati, erogati e innovati**: emerge l'importanza delle competenze per tutte le fasi del loro ciclo di vita, in particolare per i servizi accessibili attraverso una qualunque "protesi" digitale: Cellulare, Personal Computer, Televisione Digitale.

Nel breve periodo, per quanto riguarda i Servizi

Digitali, ci possiamo aspettare che si consolideranno due linee di tendenza:

- il crescente volume di acquisto di prodotti e servizi attraverso un **terminale digitale**;
- oltre all'acquisto, la trasformazione dei servizi tradizionali in servizi digitali e l'emergere di nuovi servizi anche **consumabili** per via digitale, dotati cioè di un alto livello di interazione, oggi assicurata ancora da operatori di servizio. In questo contesto assumerà rilevanza crescente la necessità di aggiornare e accrescere in modo sistematico tutta la filiera delle competenze, d'uso e di erogazione, e quindi il circuito

virtuoso della formazione e dell'aggiornamento basati su standard di competenze riconosciute e certificate a livello internazionale.

2. I RISULTATI DELLA RICERCA SUL COSTO DELL'IGNORANZA

Facciamo riferimento al programma pluriennale di ricerca avviato da AICA in collaborazione con SDA Bocconi sul tema del "Costo dell'ignoranza informatica", mirato a approfondire gli obiettivi e la profondità degli interventi di formazione delle competenze adeguate al miglioramento delle prestazioni dei Sistemi Informativi che operano in organizzazioni pubbliche e private. In vari cicli di ricerca su questo tema è stato analizzato prima quanto è il "tempo perso" dagli **utilizzatori** di Personal Computer nelle imprese in generale, e in seguito è aumentato il livello di focalizzazione con approfondimenti presso Banche, enti della Sanità ed enti della Pubblica Amministrazione Centrale. Rimandiamo per maggiori dettagli a quanto pubblicato nel Libro sul "Costo dell'Ignoranza Informatica", Egea, 2008.

In estrema sintesi, il modello di valutazione del costo dell'**incompetenza** informatica è stato mutuato nei primi anni 2000 da un'indagine condotta in Norvegia, Paese ad alta intensità d'uso dell'ICT e quindi particolarmente sensibile al tema della formazione e della produttività delle risorse umane. AICA e SDA Bocconi hanno preso come riferimento la metodologia norvegese e l'hanno ulteriormente arricchita e focalizzata per ottenere risultati spendibili per la realtà italiana.

La determinazione del "Costo dell'Ignoranza", così come è stato definito nella metodologia di ricerca sviluppata fino ad oggi, riguarda l'uso dei normali strumenti informatici, cioè del PC, del software di uso individuale (MS/Office), dei servizi di rete (e-mail, Internet) e delle applicazioni *legacy* (sistemi informativi interni istituzionali).

Secondo le rilevazioni e le analisi disponibili, facendo riferimento a diversi contesti di lavoro, mediamente il tempo lavorativo assorbito dagli utenti di sistemi informatici è, ogni settimana, pari a circa il 60% delle 40 h teoriche di lavoro; delle 24 h di lavoro con il PC, usato come terminale del sistema informativo, gli utenti restano improduttivi per 171 min alla settimana a causa

dell'insorgere di difficoltà che bloccano il normale svolgimento dell'attività in relazione al loro ruolo. Quindi ogni lavoratore che utilizza il PC "perde" settimanalmente il 7,13% del proprio tempo lavorativo (171 min su 5 giorni di 8 h).

Considerando il costo medio annuo del singolo utente in Italia pari a 36.000 € e dividendo tale costo per 220 giorni lavorativi, il costo medio per giornata lavorativa risulta di 163,63 €; AICA e SDA Bocconi hanno quindi stimato un costo di improduttività annuale, a livello di singolo utente, pari a 2.565 €, ottenuto moltiplicando la percentuale di tempo perso in modo improduttivo per il costo medio per utente: questo valore è stato qualificato come "costo dell'ignoranza informatica". Moltiplicando questo valore con i dovuti accorgimenti di tipo statistico per il numero di utenti rappresentati dal campione su cui viene fatta la rilevazione, si può arrivare ad una stima del costo totale dell'ignoranza degli utenti per azienda, per settore, per Paese.

Nelle ricerche approfondite relative a diversi settori economici, al campione di utenti viene anche richiesto di quantificare le cause del tempo perso; cause che si riconducono a **tre** componenti fondamentali: **scarsa conoscenza** degli strumenti dichiarata dall'utente, **aiuto prestato a colleghi** in difficoltà, problemi percepiti come tecnici, quali il **malfunzionamento** dell'hardware e/o del software e della rete. I risultati più recenti della ricerca AICA e SDA Bocconi vedono generalmente in un rapporto uno a tre il tempo perso per **incompetenza dichiarata dell'utente** (35%) e quello perso perché attribuito a **motivazioni tecniche** (65%).

Il tempo perso per incompetenza dell'utente si può ridurre attraverso diversi tipi di intervento, per esempio effettuando una **formazione qualificata**, ossia certificata nei suoi risultati, dei dipendenti che usano sistemi informativi.

Il metodo adottato per la misura del miglioramento della produttività di chi lavora in ufficio con il Personal Computer è basato sul test ECDL: questo test permette di misurare gli impatti della formazione e di stimare nel 10% (ordine di grandezza) il tasso di miglioramento medio della produttività del lavoro di ufficio ottenuto a valle dell'intervento formativo stesso.

Non viene invece stimato il costo dell'incompetenza degli specialisti ICT che presidiano la **prestazione del PC** a livello tecnico e che risulta pa-

ri al 65% del tempo complessivo di fermo macchina rilevato presso gli utenti dei PC collegati. L'obiettivo del presente articolo è appunto quello di arrivare a stimare il **Costo Totale dell'Ignoranza** nella intera filiera delle competenze digitali, tenendo conto di due aspetti:

- analizzare a fondo, secondo un nuovo modello interpretativo, tutti gli elementi che hanno qualche influenza sul miglioramento complessivo della prestazione di un utente qualificato che usa un terminale digitale, misurata con opportuni indicatori; il costo dell'ignoranza diventa con questo approccio uno degli indicatori della qualità del servizio, da misurare anche in termini di soddisfazione dell'utente finale;
- identificare tutte le varie tipologie di operatori coinvolti nell'erogazione di un qualunque servizio che sia utilizzato o dai dipendenti di un'organizzazione pubblica o privata o dai clienti/utenti di servizi all'esterno della stessa.

3. IL NUOVO MODELLO DI FILIERA DELLE COMPETENZE DIGITALI

Per analizzare il contributo alla generazione di valore che può dare il miglioramento delle competenze digitali, adatteremo il modello della Catena del Valore; M. Porter (1988) ha introdotto questo concetto nella teoria microeconomica circa 30 anni fa con l'obiettivo di spiegare il contributo che i vari processi di un'organizzazione pubblica o di un'impresa privata possono dare alla generazione del Margine Lordo, inteso, semplificando al massimo, come il differenziale fra quanto l'organizzazione ricava dal suo mercato di riferimento e quanto spende in termini di acquisti presso i suoi fornitori. Già allora Porter sottolineava come le Tecnologie dei Sistemi Informativi e di Comunicazione fossero da considerare parte integrante dell'operatività di qualunque organizzazione finalizzata all'erogazione di prodotti e servizi al mercato o al sistema sociale.

Sappiamo che ormai le Tecnologie Digitali (Tecnologie dei Sistemi Informativi + Tecnologie dei Sistemi di Comunicazione) sono diventate assolutamente pervasive sia nelle imprese manifatturiere che, ancora di più, nelle imprese finanziarie e commerciali e in quelle che erogano servizi al consumatore e alle imprese, come d'altra parte per gli enti pubblici centrali e locali che erogano servizi al cittadino; in questi settori le

Tecnologie Digitali hanno assunto/stanno assumendo un triplo ruolo: come tecnologia di supporto alla gestione, come tecnologie di produzione dei servizi vendibili e più recentemente come componente fondamentale per l'innovazione dei prodotti/servizi.

Il modello di Porter riconosce per ogni organizzazione nove aree principali che contribuiscono alla generazione di valore per l'impresa: fra i principali processi primari consideriamo quelli della produzione, del marketing&vendita e dell'assistenza tecnica mentre fra quelli secondari consideriamo specificamente quelli relativi alla gestione delle risorse umane e l'amministrazione; i processi di innovazione e gestione delle tecnologie digitali vengono considerati nel modello di Porter come distribuite e legate ai processi di sviluppo del valore già segnalati. I processi primari in particolare entrano a far parte della filiera del valore a cui partecipa l'impresa: per filiera si intende l'insieme dei processi primari di imprese/enti che collaborano in una specifica rete di imprese, attraverso le quali il prodotto/servizio **prende forma e si dirige** verso l'utilizzatore finale, consumatore o cittadino. Il valore di scambio generato nell'atto di vendita al consumatore/cittadino viene assorbito dall'operatore presso cui si è concluso l'acquisto e viene retrocesso per quota parte a tutti gli operatori della filiera che hanno partecipato alla sua costruzione, distribuzione e assistenza (Figura 1).

La crescita della competitività e il miglioramento della efficienza per catturare e servire clienti finali soddisfatti, insieme alla disponibilità di

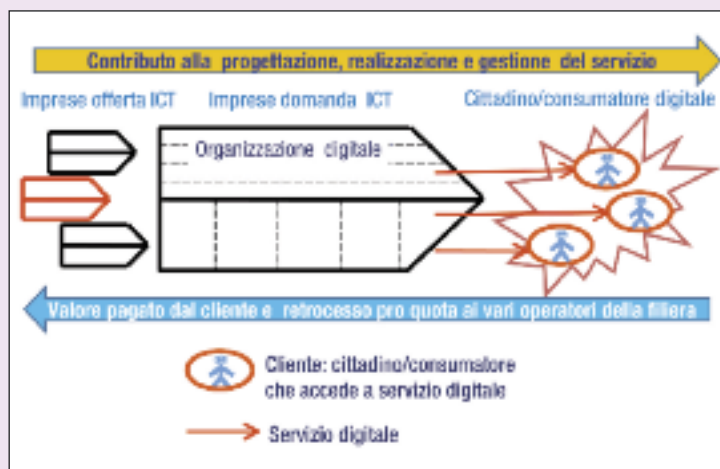


FIGURA 1
Filiera del valore dei Servizi a base digitale

nuovi materiali e di nuove tecnologie di produzione e gestione, sono i fattori che spingono le imprese a specializzarsi e ad associarsi in filiere attraverso le quali sia possibile mantenere i clienti acquisiti e catturarne di nuovi nei mercati in cui si è già presenti o in nuovi mercati.

Nel mondo finanziario le filiere di servizio sono già molto presenti: la **banca** eroga direttamente o indirettamente al cliente finale una serie di servizi finanziari (per esempio la gestione e lo sviluppo del portafoglio azionario o obbligazionario per ogni singolo cliente) piuttosto che il prelievo e il deposito di contante con i servizi Bancomat, in cui arriva ad accollare al cliente stesso anche i costi di transazione, mentre la **compagnia di assicurazione** commercializza e gestisce polizze in formato elettronico, direttamente o tramite agenzie specializzate; in tutti questi esempi il cliente finale è solo interessato a poter selezionare in una gamma di offerta concorrenziale i prodotti/servizi di qualità e prezzo adeguato alle proprie esigenze, indipendentemente dal punto vendita, virtuale o fisico, attraverso cui accede. Sono inoltre presenti e attive reti specializzate in filiere come quelle degli operatori telefonici, delle imprese specializzate in componenti e prodotti finiti elettronici e informatici, dell'abbigliamento, del lusso, dell'automotive, dell'alimentare, dei prodotti per la cura del corpo, del settore chimico farmaceutico. Sono nate/stanno nascendo reti specializzate nelle filiere commerciali (Centri Integrati di Via), in quelle turistiche, in quelle della salute, per i servizi al cittadino e molte altre. Secondo Rullani, l'Economia della Filiera *"distribuisce il lavoro, le conoscenze e gli investimenti fra molti operatori autonomi specializzati per fasi di lavoro e di servizio"*.

3.1. Le specificità dei Servizi

Approfondiamo con qualche indicazione ulteriore la natura e le caratteristiche dei servizi (Normann, ETAS, 1999), completamente diverse dai **prodotti materiali**.

Mentre il prodotto esiste prima di essere venduto e usato, il servizio si materializza solo nel momento della sua erogazione e assume configurazioni specifiche che si definiscono nel corso della relazione fra cliente/utente e fornitore. E ancora, mentre un progettista può descrivere in modo anticipato le funzioni d'uso e l'esatta configurazione degli attributi del prodotto che

verrà realizzato e reso disponibile al cliente finale, non può invece prescrivere con la stessa accuratezza il risultato dell'interazione fra cliente e fornitore del servizio, né tantomeno disegnare la forma e le caratteristiche del valore emotivo prodotto: basta pensare in questo senso alle emozioni suscitate dalla vista di un quadro in un museo piuttosto che alla rappresentazione teatrale o a quella di un concerto, a tutti gli effetti classificabili come servizi culturali ad alto valore aggiunto rispetto al costo di erogazione.

La capacità di suscitare emozioni e la qualità percepita dal cliente/utente percettore del servizio è fortemente influenzata, oltre che dalla capacità relazionale dell'operatore a cui è affidata l'erogazione, dal contesto logistico e organizzativo nel cui ambito il servizio viene rilasciato; per queste sue caratteristiche la progettazione del servizio (Service Design) richiede accorgimenti, competenze e strumenti specifici e diversi da quelli necessari alla progettazione di un prodotto.

Di nuovo Rullani fornisce indicazioni che qualificano con maggiore precisione l'economia dei servizi, come riportato di seguito:

- l'economia dei servizi è basata sulla conoscenza, il cui massimo valore si ottiene dal suo ri-uso con fini di volta in volta diversi e i cui costi sono bassi o trascurabili;
- la conoscenza è moltiplicabile, nel senso che il suo uso può essere propagato in bacini di impiego molto ampi e durevoli;
- la conoscenza è condivisibile e non divisibile, nel senso che la sua produzione e propagazione non può essere affidata a meccanismi anonimi del mercato ma richiede la costruzione di un sistema di regole socialmente condivise;
- la conoscenza infine non è soltanto un mezzo, ma è una risorsa auto-generativa, nel senso che la sua produzione e il suo uso cambiano i fini, sviluppano identità, creano legami modificando il mondo in cui si opera.

Tenendo conto di queste indicazioni andiamo a specificare le modalità di progettazione ed erogazione dei servizi digitali: questi possono essere definiti attraverso **la specificazione e realizzazione di media tecnologici a supporto di relazioni strutturate fra persone, che permettono di erogare capacità di azione il cui valore è riconosciuto da parte di un definito cliente**.

La capacità di azione, per quanto riguarda i Servizi Digitali, assume una forma base di tipo assertivo rispetto alla quale il cliente/utente reagisce; l'operatore accoglie le reazioni, cerca di interpretarle e ripropone (al cliente/utente) funzioni, attributi e benefici che ritiene possano essere di interesse; solo quando il cliente ha capito i benefici del servizio, questo viene acquistato e utilizzato. Nell'ambito sanitario (in cui gli operatori sono costituiti da medico e infermiere) i servizi assumono la forma di una valutazione diagnostica e di una prescrizione; nei servizi educativi (in cui l'operatore è costituito dal docente) questi assumono la forma di una promessa di produrre per il discente una nuova conoscenza e capacità di saper fare che migliora la sua possibilità di lavoro.

Sotto molti aspetti i servizi, specificati come risultata di un'interazione **in presenza** fra un fornitore e un cliente/utente e privi di materialità (intangibili), hanno assunto via via una caratterizzazione in cui si sono recuperati anche componenti di tangibilità. Aziende come eBay, o comunità come quelle che si sono sviluppate intorno a Wikipedia o ad Apple (ci riferiamo a prodotti/servizi come iPod e iTunes) hanno costruito una ricca e sofisticata combinazione di *"deliverables"* che estendono la capacità del cliente/utente ad agire e produrre valore per se stessi e per altri. La progettazione dei servizi sviluppa in modo crescente componenti di servizio sia intangibili che tangibili, nel senso che può includere artefatti che riguardano la comunicazione (incluso il packaging), l'ambiente in cui il servizio viene erogato e le guide ai comportamenti (una per tutte, il distributore di priorità di accesso per la regolazione delle code agli sportelli).

Come conseguenza la progettazione dei servizi è una attività che si esprime attraverso la predisposizione di menu di comportamenti e di *"scripts"* per gli attori che interagiscono nella fase di vendita e in quella di erogazione e uso del servizio, che lasciano comunque un discreto livello di libertà al fornitore in funzione dei comportamenti del cliente.

Più recentemente, con la nascita e la diffusione di strumenti che facilitano l'interazione fra pari, sono emerse, aggiungendosi a quelli di tipo assertivo, nuove tipologie di servizi di tipo sociale (social networks, communities ecc.) in cui ogni utente collegato al web diventa anche genera-

tore di contenuti e sviluppatore di relazioni e interazioni completamente autonome; tutti gli attori collegati via web, compresi quelli tradizionalmente considerati come esterni rispetto ai confini dell'impresa o dell'ente che produce i servizi (clienti e utenti), diventano protagonisti nella interazione e sono spesso chiamati a contribuire per mettere a fuoco appropriate funzioni d'uso e caratteristiche nella progettazione di nuovi servizi in cui vengono analizzate le interazione fra gli attori, pianificate le sequenze degli eventi, i flussi di informazioni e i materiali di supporto; il coinvolgimento di attori non tecnici implica inoltre che il servizio non possa essere progettato solo dal punto di vista dei requisiti funzionali finalizzati alla gestione ottimale delle risorse (ottimizzazione dei flussi e riduzione dei tempi operativi) ma debba tenere in conto anche la componente emotiva del cliente (creazione di eventi che abbiano un significato, motivazione del cliente, comunicazione del servizio). Vengono usate inoltre tecniche come il *"just in time"* e il *"total quality management"* per produrre modelli funzionali del sistema di servizio e per controllare i relativi processi di erogazione.

Fra le molteplici tipologie di servizi, due in particolare assumeranno in futuro maggiore importanza:

□ i servizi della Pubblica Amministrazione per la cittadinanza e le imprese, la cui crescita si può ricondurre al rapido ed esteso cambiamento sociale che richiede una riorganizzazione del welfare: stanno in particolare emergendo, come molto importanti, i servizi relativi alla sanità, alla scuola, al supporto delle istituzioni culturali e infine ai servizi per la sicurezza. In queste nuove aree vengono continuamente sviluppate e adottate nuove e specifiche metodologie di progettazione ed erogazione;

□ i servizi sviluppati come catalizzatori della innovazione: si sta mettendo a fuoco una nuova nozione di Service Design che indica come alcuni prodotti digitali non siano più da considerare come elementi isolati rispetto ai bisogni dei consumatori e dei cittadini; ci riferiamo a prodotti digitali come iPhone che è diventato una sorta di negozio di servizi on line (oggi permette l'accesso a 32.000 servizi diversi!!) e iTunes, che svolge attraverso il PC e lo Smart Phone il ruolo di negozio di musica on line. In questo caso il concetto di offerta si basa sulla com-

binazione di oggetti tangibili e intangibili che lasciano ai consumatori la massima flessibilità perché possano decidere quando e come vogliono usare il servizio.

L'importanza assunta dai servizi nell'economia (si va verso un contributo dei servizi al PIL (Prodotto Interno Lordo) di oltre il 60% a livello di sistema Paese) giustifica l'attenzione posta alle loro caratteristiche e al ruolo determinante che le Tecnologie Digitali hanno e sempre di più avranno per la loro realizzazione ed erogazione attraverso terminali digitali a basso costo. Al rilevante contributo economico dei servizi digitali corrisponde, evidentemente, un altrettanto rilevante impegno nella definizione e nello sviluppo di conoscenze e capacità di azione (competenze) di coloro che sono chiamati a progettare, erogare e utilizzare detti servizi.

3.2. I livelli di specializzazione delle competenze

Le analisi condotte a livello europeo mostrano che le risorse con competenze specialistiche per la progettazione, realizzazione, gestione e innovazione delle Tecnologie Digitali costituiscono l'85-90% del totale delle risorse per le **imprese dell'offerta ICT**, che comprendono un numero limitato di imprese multinazionali che producono hardware, software di base e *middleware* a livello mondiale, ma con cui collaborano altre centinaia di migliaia di *software house*, *system integrator*, *value added reseller* e *service integrator* che operano a livello locale; queste imprese operano come terze parti qualificate della filiera delle *Tecnologie Digitali* (TD), con il compito di personalizzare

le applicazioni, installare le tecnologie e i servizi e assistere il cliente/utente finale; fra le imprese della domanda ICT, che costituiscono il vertice della filiera business delle TD, includiamo le imprese manifatturiere non elettroniche, le imprese del commercio, dei servizi e del settore finanziario e gli enti della Pubblica Amministrazione.

Con questa classificazione, le risorse specialistiche per lo sviluppo e la gestione delle Tecnologie Digitali installate variano dall'1-2% delle manifatturiere al 6-9% di quelle finanziarie: con esclusione di quelle manifatturiere, le Tecnologie Digitali per le organizzazioni di tutti gli altri settori richiedono competenze anche quantitativamente più significative a causa del ruolo che assumono come tecnologie di produzione dei servizi (digitali) vendibili. In questo tipo di organizzazioni la quasi totalità delle risorse non specialistiche hanno comunque bisogno di competenze d'uso delle tecnologie e dei servizi a base digitale (*power user*) relativi sia ai servizi interni all'organizzazione che ai servizi venduti ai consumatori/cittadini finali. Nelle imprese manifatturiere in particolare il numero dei *power user* è relativamente inferiore (non più del 50% dei dipendenti totali) essendo l'altra parte delle risorse costituita da operai che lavorano ai processi produttivi di prodotti materiali (Tabella 1).

Se applichiamo quanto esposto sul modello di Porter con una focalizzazione sulla specifica leva di valore costituita dalle **competenze digitali**, possiamo costruire la **Filiera Cognitiva** per l'erogazione dei servizi a base digitale (Figura 2).

Tipologia di organizzazione / settore merceologico	Risorse umane con competenze specialistiche per sviluppo e gestione delle TDO (% di dipendenti specialisti)	Risorse umane con competenze d'uso delle TDO (% di dipendenti power user)
Imprese manifatturiere	1-2%	45-50%
Imprese settore terziario (servizi e commercio)	4-5%	96-95%
Imprese settore finanziario (banche e assicurazioni)	6-9%	94-91%
Enti Pubblici (centrali e locali)	4-5%	96-95%
Imprese dell'offerta ICT	85-90%	15-10%

TABELLA 1

Il peso delle risorse umane dotate di competenze distintive delle TDO

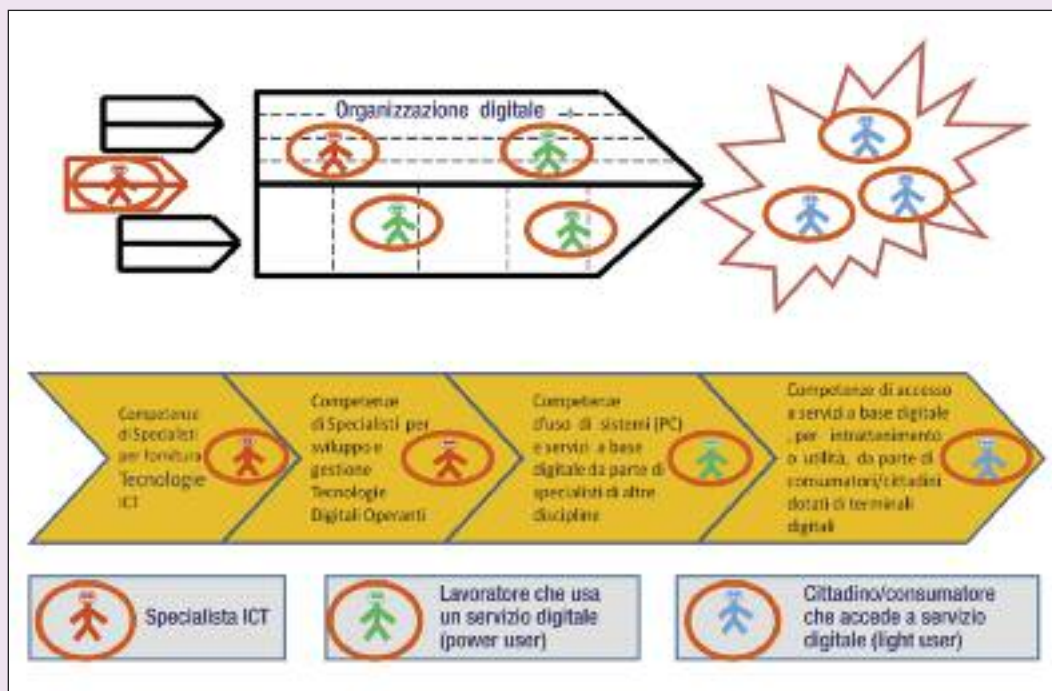


FIGURA 2
*Filiera cognitiva
per l'erogazione
di servizi a base
digitale*

Partiamo dal vertice della filiera costituito dagli utenti interessati (**light users**): costituiscono una immensa platea di cittadini interessati ai servizi pubblici digitalizzati e uno straordinario mercato del consumo digitale. Cittadini e consumatori hanno bisogno di una competenza essenzialmente di tipo abilitante all'accesso e all'uso dei terminali digitali dotati di interfaccia molto semplificata, come sono ormai gli *smart phone*, le televisioni digitali, gli strumenti di navigazione su web, gli e-book ecc..

Il secondo tipo di operatori della filiera è costituito dalle organizzazioni pubbliche e private, che fanno parte della domanda business di ICT, a loro volta interessate a sviluppare ed erogare i propri servizi su base digitale al consumatore/cittadino. Tali operatori hanno bisogno di due tipologie di competenze fra i loro dipendenti:

- quelle del personale che con specializzazioni diverse da quelle ICT ha comunque bisogno di usare sistemi e servizi digitali a supporto delle proprie attività a favore di colleghi e clienti/utenti finali (**power user**);
- quelle degli **specialisti ICT** interni dedicati alla pianificazione, allo sviluppo, alla gestione e alla manutenzione di sistemi digitali operanti e servizi a base digitale.

Il terzo tipo di operatori della filiera è costitui-

to infine dai fornitori di piattaforme tecnologiche e di servizio che hanno come propri clienti le organizzazioni della domanda precedentemente indicati; i propri **specialisti ICT** hanno bisogno di essere sistematicamente aggiornati rispetto alle innovazioni tecnologiche delle piattaforme offerte e di sviluppare competenze di analisi dei business legati ai servizi, di metodologie per lo sviluppo delle personalizzazioni, di ottimizzazione della gestione dei servizi e di manutenzione delle installazioni esistenti.

4. LE COMPETENZE, I PROFILI PROFESSIONALI E LE CERTIFICAZIONI DELLA FILIERA COGNITIVA A SUPPORTO DEI SERVIZI DIGITALI

Siamo arrivati nei paragrafi precedenti a qualificare il tipo di competenze necessarie a supportare, nell'ambito di una filiera cognitiva, la pianificazione e progettazione, lo sviluppo e l'installazione, la gestione e la manutenzione di servizi a base digitale di qualità elevata. Vediamo ora come gli standard proposti dal CEPIS in Europa e a livello internazionale e le rispettive certificazioni trovino una importante modalità applicativa rispetto al modello proposto.

Per quanto riguarda le competenze distintive dell'ICT il CEPIS rende disponibili 3 standard:

□ lo standard eCitizen, relativo alle competenze di chi vuole navigare su Internet per accedere a servizi a base digitale (*light users*) o nell'area dell'intrattenimento, o in quella del consumo di beni e servizi a supporto della vita di lavoro e domestica o in quella della cittadinanza digitale;

□ lo standard ECDL, relativo alle competenze dei dipendenti di qualunque organizzazione, pubblica o privata, che usano il Personal Computer sostanzialmente per lavorare per una quota rilevante del loro tempo (oltre il 50%) su e per servizi interni (a favore di colleghi e clienti in house) e per servizi esterni a fronte di benefici rilevanti per il consumatore e il cittadino (*power users*);

□ lo standard EUCIP, relativo alle competenze degli specialisti ICT, articolato in 22 profili di specializzazione che coprono una buona parte delle competenze necessarie a intervenire sull'intero ciclo di vita dei sistemi e delle piattaforme di erogazione di servizi digitali (specialisti ICT).

Ciascuno di questi standard è costruito per permettere sia l'apprendimento delle competenze richieste che per verificarne la effettiva applicabilità nel contesto dato per ogni dipendente aziendale attraverso la **certificazione**. Analizziamo ora quali sono le specializzazioni coinvolte nella filiera cognitiva e che contribuiscono, in assenza di competenze adeguate, ad incrementare il costo dell'ignoranza informatica sulla intera filiera.

A partire dal cliente finale di un servizio digitale, le competenze richieste per un utilizzo ed una erogazione ottimale del servizio sono:

□ competenze di accesso al servizio, previste dallo standard eCitizen: sono quelle di consumatori/cittadini che accedono a servizi come quelli del Servizio Sanitario Nazionale, della propria Banca e Compagnia di Assicurazione, di una Agenzia Turistica o di un Supermercato ecc.;

□ competenze d'uso di chi lavora nella organizzazione pubblica o privata che confeziona, mantiene e innova il servizio digitale consumato dal cliente/utente; si possono ipotizzare due tipi di servizi:

- quelli erogati direttamente al cliente finale attraverso un terminale digitale mobile (cellulare o PC portatile); in questo caso la configurazione organizzativa adeguata al supporto del cliente/utente prevede la disponibilità di uno specialista

(*Help Desk Supervisor*) in grado di assistere direttamente il cliente stesso in caso di bisogno;

- quelli erogati in house attraverso uno sportello in cui è disponibile un operatore formato e disponibile per il cliente; le competenze d'uso dell'operatore sono quelle previste dallo standard ECDL, che oggi riguarda essenzialmente le funzionalità di macchina, ma che gradualmente si sta arricchendo, per specifico settore applicativo, di ulteriori competenze legate sia ai problemi specifici che l'operatore può incontrare sia al contesto specifico di tipo applicativo;

□ competenze specialistiche di supporto agli operatori: come abbiamo già considerato la configurazione di competenze necessarie al supporto degli operatori di sportello o del cliente finale prevede due figure specialistiche di interfaccia (definite nello *standard EUCIP-European Certification of Informatics Professionals*) che sono rispettivamente l'*IT Administrator* e l'*Help Desk Supervisor*; entrambe queste figure sono quelle previste dalle best practices indicate nello standard dei processi denominato *ITIL-IT Infrastructure Library*), che qualifica il processo su cui questi specialisti intervengono come processo di gestione degli incidenti (*Incident Process Management*) il cui scopo è, a seguito di un malfunzionamento, di ristabilire il normale servizio il più rapidamente possibile e minimizzare gli effetti sfavorevoli sulle attività dell'Organizzazione; ma a queste due figure se ne aggiungono almeno altre due, sempre parte dei profili EUCIP, che sono:

- il *Data Base Manager*, che aggiorna sistematicamente il Data base da cui l'operatore di sportello o l'*help desk supervisor* possono estrarre il profilo dei Clienti/Utenti;
- il *Data Center Configuration Manager* che, nella "fabbrica" del software e del servizio gestisce e presidia il sistema tecnologico su cui si appoggiano sia il servizio utilizzato dal cliente/utente finale sia i servizi interni di supporto ai vari operatori considerati.

Restano da valutare infine i contributi di competenza che dovrebbero avere tutti gli altri specialisti che contribuiscono alla progetta-

zione e realizzazione di nuovi servizi digitali: secondo lo standard EUCIP, potrebbero essere coinvolti gli *IS Business Analyst*, gli *IS Project Manager*, i *Software Developer* e gli *Web & Multimedia Master* ecc..

Gli operatori di sportello, rispettivamente della banca, del servizio sanitario, della agenzia turistica o di quella assicurativa ecc. che forniscono assistenza al consumatore/cittadino in presenza o via telefonica o via web, sono quindi comunque supportati da un sistema con cui sia possibile ottenere la profilatura del consumatore/cittadino chiamante al momento dell'accesso per poterne conoscere le caratteristiche del comportamento di acquisto ed eventualmente i comportamenti di consumo su precedenti servizi percepiti/acquistati oppure per poterne controllare i diritti di accesso rispetto al servizio richiesto.

Si vuole mettere in evidenza quanto sia necessario il supporto degli operatori di sportello, sia fisico che virtuale (in presenza, via telefono o web), per i quali è necessario configurare una struttura tecnico-operativa complessa che permette di mantenere un alto livello di servizio comunque su due situazioni possibili;

□ le risposte adeguate in termini di contenuto e tempestività alle domande del consumatore/cittadino, incluse le possibili registrazioni di

tipo operativo e/o amministrativo per l'acquisto o la conferma di impegno del servizio richiesto (prenotazione di una visita, contabile di una transazione bancaria, registrazione di una polizza, spedizione dei prodotti acquistati via web presso il supermercato all'indirizzo, alla data e all'ora concordata, spedizione del biglietto di aereo via web ecc.);

□ il supporto agli operatori di sportello in caso di guasto del sistema tecnico operativo (*incident*); in questa situazione il supporto agli operatori che si troveranno in difficoltà rispetto all'accumularsi di domande a cui non possono rispondere, starà nella fornitura di indicazioni rispetto alla durata del fermo tecnico e alle indicazioni rispetto a percorsi alternativi se il consumatore/cittadino ne ha bisogno.

La configurazione di supporto alla erogazione del servizio si presenta come indicato nella figura 3.

La struttura tecnico-operativa così delineata permette al responsabile del Sistema Informativo Aziendale di garantire ai clienti esterni dei servizi digitali e ai dipendenti interni (utenti del SI) di poter erogare servizi digitali di alta qualità e su cui sia praticabile un sistematico controllo attraverso periodiche rilevazioni della *customer* e della *user satisfaction* e una misurazione delle prestazioni del sistema, confron-

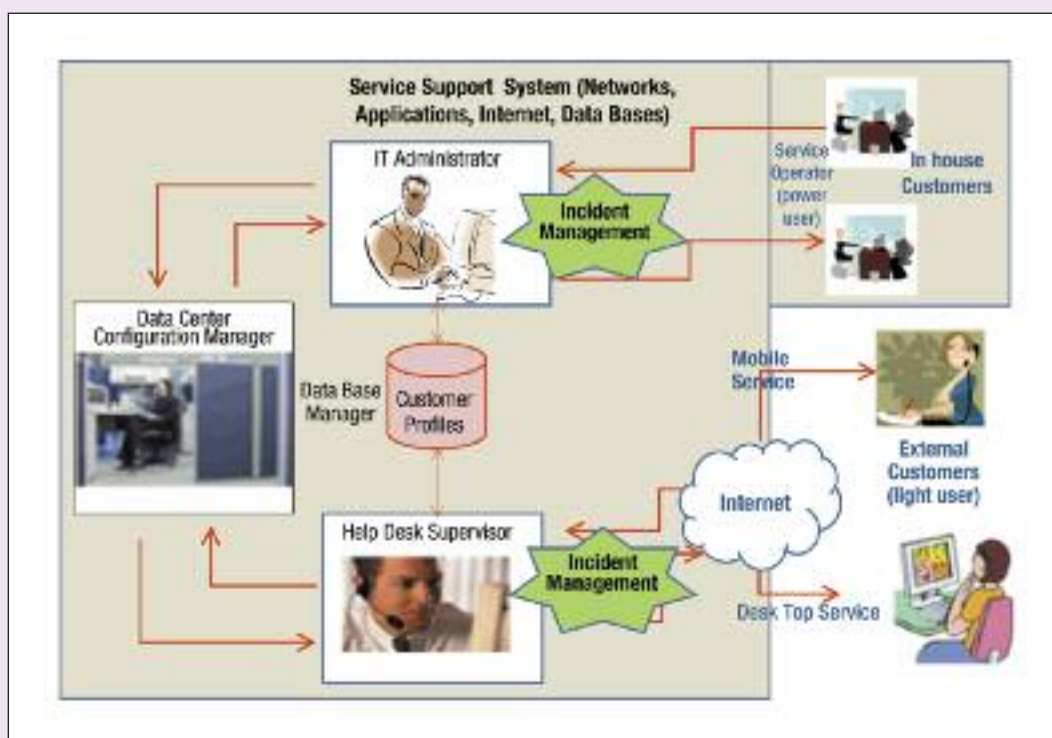


FIGURA 3

Una possibile configurazione del sistema di erogazione dei servizi a base digitale

tabili con quanto indicato in fase di pianificazione dei Servizi; gli indicatori di prestazione saranno la base dei miglioramenti e delle innovazioni che successivamente emergeranno come necessarie mano a mano che il portafoglio dei servizi digitali si andrà ampliando e articolando per rispondere ai nuovi bisogni dei consumatori/clienti.

5. IL COSTO TOTALE DELL'IGNORANZA COME PARADIGMA PER LA REVISIONE E L'OTTIMIZZAZIONE DEI SERVIZI

È stato proposto un modello completo della filiera delle competenze da monitorare e sviluppare per mantenere accettabile il costo totale dell'ignoranza per la erogazione di servizi digitali, lato utente e lato specialista.

Sostanzialmente e molto schematicamente, quello che succede è che Imprese ed Enti della Pubblica Amministrazione assumano/stiano assumendo nuove forme organizzative digitali, più adeguate a sostenere la propria competitività e quella di sistema, per l'influenza dei seguenti fattori:

□ la competizione si sviluppa su una scala globale, basata su leadership che non riguardano solo i fattori tecnologici, ma anche e soprattutto le modalità con cui l'impresa si posiziona sul mercato e può sviluppare nuovi clienti in territori fino a ieri irraggiungibili;

□ i clienti sono sempre più esigenti, interessati a pacchetti di prodotti-servizi per bisogni molteplici, tutti dotati di componenti digitali che, almeno potenzialmente, introducono per ogni prodotto la capacità di essere localizzato e monitorato rispetto alle proprie funzioni d'uso, e per ogni cliente la possibilità di monitorarne le modalità di acquisto e i comportamenti d'uso;

□ si configurano filiere del valore in cui più aziende cooperano fra di loro soprattutto nei business dei servizi digitali per ottenere una maggiore capacità di incidere sul mercato e sostenere la contrazione dei margini;

□ si sviluppano le conoscenze e le competenze digitali, che aiutano l'impresa a mantenere la propria "traiettoria nei processi di innovazione" dei servizi; diventa inoltre opportuno migliorare le capacità di anticipare e gestire aspetti normativi sempre più sofisticati a causa della internazionalizzazione, della salvaguardia dei diritti di proprietà intellettuale e d'uso dei contenuti

digitali e della crescente responsabilità sociale dell'impresa;

□ in questo contesto i servizi a base digitale vengono abilitati dalle tecnologie digitali e si riconoscono comunque tre tipologie di competenze per il successo del servizio;

- Le competenze per progettare, realizzare e gestire la rete tecnologica;
- Le competenze per progettare e supportare la erogazione del servizio in termini di contenuti soddisfacenti per il consumatore/cittadino;
- Le competenze degli altri lavoratori dell'impresa in termini di uso dei sistemi e servizi digitali che richiedono non solo abilità sulla macchina ma anche capacità di *problem solving*.

6. CONCLUSIONI

Avere definito la filiera delle competenze digitali aiuta tutte le figure professionali coinvolte nella erogazione dei servizi digitali a mettere a fuoco l'importanza della preparazione necessaria per ottenere alta qualità nei servizi e prestazioni di valore percepibile da parte del consumatore/cittadino finale; in particolare:

□ Per gli specialisti che progettano, realizzano e installano infrastrutture e sistemi tecnologici digitali, tenere conto del livello di competenza degli operatori che saranno addetti alla erogazione e manutenzione del servizio e alle competenze d'uso dei servizi stessi, fornisce la possibilità di definire anche la qualità e le prestazioni del servizio da erogare e quindi, successivamente, poter misurare la distanza fra quanto progettato e quanto realmente ottenuto, con evidenti positivi input rispetto ai miglioramenti da introdurre;

□ Per gli specialisti della erogazione e della manutenzione del servizio, tenere conto della qualità e della prestazione progettata significa poter predisporre supporti agli utilizzatori finali tarati rispetto ai reali livelli di competenza riscontrati, perseguendo un più alto livello di qualità percepito rispetto a quello progettato e indicazioni mirate sulle prestazioni ottenute;

□ Infine, per i cittadini e i consumatori destinatari dei servizi, l'aver tenuto conto del rispettivo livello di competenza d'uso comporterà un più alto livello di soddisfazione, in quanto si troveranno ad essere guidati comunque ad

una soluzione minimale anche nei casi di mal-funzionamento del sistema tecnico.

Il paradigma del **Costo Totale dell'Ignoranza** potrà così essere utilmente applicato per innescare processi di revisione e miglioramento della **qualità** e della **produttività** sia dei servizi esistenti che dei nuovi servizi.

Bibliografia

- [1] Porter M.: *Il vantaggio competitivo*. Edizioni di Comunità, 1988.
- [2] Normann R.: *La gestione Strategica dei Servizi*. ETAS, 1999.
- [3] Rullani E.: *La fabbrica dell'immateriale. Produrre valore con la conoscenza*. 2007.

Il presente articolo sarà seguito da altri 3 approfondimenti che contribuiranno a mettere a fuoco il sistema delle competenze da presidiare da parte di un IT Trainer e precisamente:

1. il sistema delle competenze dei profili coinvolti nella progettazione, realizzazione e installazione dei nuovi sistemi di servizio;
2. il sistema delle competenze dei profili coinvolti nella erogazione dei servizi e nella loro manutenzione;
3. il sistema delle competenze di cittadini e consumatori coinvolti come utilizzatori finali dei servizi digitali erogati da imprese private o da enti della Pubblica Amministrazione.

ROBERTO BELLINI è socio AICA e ne presiede la Sezione di Milano; per AICA è inoltre responsabile del Cantiere dei Mestieri ICT, per il miglioramento delle competenze e delle professionalità del settore.

Svolge da anni una intensa attività di ricerca e docenza (Politecnico di Milano) sui temi della Innovazione e delle relative Competenze; con la Fondazione Politecnico di Milano sviluppa in particolare analisi e modelli relativi al marketing di servizi erogati attraverso Reti Governate di imprese e organizzazioni pubbliche che operano nell'Economia di Filiera.

È uno degli esperti CNEL per il Tavolo dei Saperi e il Tavolo Reti nell'ambito del progetto Trasformazione di Impresa.

E-mail: roberto.bellini@polimi.it

FULVIA SALA dopo una laurea in Matematica, conseguita a Milano, ottiene un Master in Statistica presso l'Università di Berkeley, California. Lo sviluppo professionale avviene prevalentemente all'interno di una grande azienda informatica. Dirigente a 32 anni, ricopre ruoli direttivi nell'ambito del marketing e della pianificazione, quali: responsabile della Direzione Marketing Operativo e responsabile della Direzione Marketing e Pianificazione Strategica. In qualità di libera professionista ha svolto attività di docenza nell'area marketing e collabora con AICA dove è responsabile dell'area progetti e ricerche. In tale ruolo, congiuntamente a SDA Bocconi, ha avviato, dal 2003, il progetto "Il costo dell'ignoranza informatica" che ha visto, ad oggi, la pubblicazione di più studi relativi a diversi settori di mercato.

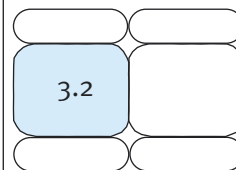
E-mail: fulvia.sala@aicanet.it

MIGRAZIONE DI SISTEMI SOFTWARE LEGACY



Diversi produttori di software devono la loro posizione sul mercato a sistemi software sviluppati parecchie decadi fa, con tecnologie che sono diventate obsolete. Se dal punto di vista economico tali sistemi rappresentano una risorsa inestimabile, in quanto codificano una quantità enorme di regole di business e di conoscenza, dal punto di vista tecnico pongono notevoli problemi. In questo articolo riportiamo una sintesi dell'esperienza maturata nel corso della migrazione verso Java di un grosso applicativo bancario, considerando, in particolare, la strutturazione del flusso di controllo e del modello dati.

Paolo Tonella
Mariano Ceccato
Davide Marchignoli
Cristina Matteotti
Thomas Roy Dean



1. INTRODUZIONE

Molti sistemi software legacy (tipicamente in ambito bancario) sono difficili e costosi da mantenere ed evolvere a causa delle tecnologie e dei linguaggi di programmazione usati. Spesso tali sistemi derivano da software scritti diverse decadi fa, quando le tecnologie disponibili non erano ancora fortemente orientate alla strutturazione del flusso di controllo e del modello dei dati. La riscrittura di tali sistemi da zero non è di solito praticabile, in quanto si tratta di sistemi di grosse dimensioni (anche svariati milioni di linee di codice) che contengono in forma implicita tutte le regole di business relative al dominio su cui operano. Inoltre non è di solito possibile fermare la manutenzione di questi sistemi per il tempo necessario a riscrivere il software. L'alternativa che minimizza il rischio e consente un passaggio maggiormente controllato verso tecnologie più moderne ricade nella categoria del *re-engineering* o della migrazione [9].

I problemi principali che si incontrano in un progetto di migrazione da un linguaggio di programmazione e una tecnologia obsoleti verso

una piattaforma moderna hanno a che vedere con il grado di strutturazione delle istruzioni e dei dati. La disponibilità di istruzioni di salto incondizionato (GOTO) consente di scrivere codice estremamente non-strutturato (codice a “spaghetti”), non consentito dai linguaggi di programmazione più recenti. La possibilità di posizionare le variabili con un *overlay* arbitrario in memoria consente d'altro lato di definire un modello dati non-strutturato, in cui la definizione (scrittura) di una variabile influenza il contenuto di altre variabili che si sovrappongono ad essa parzialmente o completamente.

In questo articolo descriviamo i problemi relativi alla strutturazione del flusso di controllo e del modello dati, e gli algoritmi e le strategie che si possono adottare per affrontarli. Le tecniche descritte nell'articolo sono piuttosto generali e possono essere applicate a sistemi software scritti in linguaggi diversi e basati su tecnologie diverse, ma accomunati da un flusso di controllo e/o da un'organizzazione delle strutture dati di tipo non-strutturato. Per quanto riguarda il flusso di controllo, diversi linguaggi di programmazione ampiamente usati

in sistemi legacy fanno ricorso a istruzioni di salto non condizionato. Per esempio il Cobol e l'RPG. Per quanto riguarda le strutture dati, il problema si riscontra principalmente nei linguaggi di più basso livello (per esempio, *main-frame assembly*); in parte anche in linguaggi di più alto livello come il Cobol (istruzione REDEFINE). Contrariamente a quanto si potrebbe pensare, linguaggi di basso livello come l'*assembly* dei mainframe, sono ampiamente usati nei sistemi legacy, talvolta scritti interamente in tali linguaggi, ma più spesso contenente moduli specifici scritti in tali linguaggi. Dunque, entrambi i problemi affliggono un numero elevato di sistemi legacy e necessitano delle tecniche loro dedicate nei progetti di *re-engineering* e migrazione verso linguaggi di programmazione moderni.

In questo articolo affronteremo i due problemi menzionati qui sopra con riferimento a un caso di studio particolare. Si tratta di un'applicazione bancaria legacy di grosse dimensioni (8 milioni di linee di codice). Il linguaggio di partenza è il BAL (*Business Application Language*), un linguaggio di programmazione proprietario che richiede un ambiente di sviluppo e una piattaforma di esecuzione proprietari (B2U, *Business under Unix*). Il linguaggio target è Java su *application server* e data base relazionale. Per la migrazione di questo applicativo verso Java abbiamo sviluppato uno strumento di trasformazione automatica del codice, *bal2java*. Questo strumento, scritto nel linguaggio di manipolazione del codice TXL [4], implementa gli algoritmi di eliminazione dei GOTO, descritti nel paragrafo 2 di questo articolo, e l'algoritmo di strutturazione del modello dati, delineato nel paragrafo 3. Nonostante gli algoritmi siano descritti con riferimento al linguaggio BAL, la loro applicabilità va ben oltre tale linguaggio, estendendosi sostanzialmente ad ogni linguaggio di programmazione che ammetta flusso di controllo o modello dati non strutturati. Nel paragrafo 4 sono riportati alcuni dati sperimentali raccolti nel corso della migrazione. L'articolo si conclude con una discussione dell'esperienza maturata e degli insegnamenti che se ne possono trarre. Insegnamenti che riteniamo possano essere utili a qualunque azienda si trovi a dover affrontare un progetto di *re-engineering* simile quello descritto come caso di studio.

2. STRUTTURAZIONE DEL FLUSSO DI CONTROLLO

I sistemi software legacy sono spesso scritti in linguaggi che non supportano costrutti iterativi quali: WHILE, DO-WHILE, FOR, BREAK, CONTINUE. In tali sistemi, il flusso di controllo viene definito mediante l'uso del costrutto di salto incondizionato: GOTO. I programmatori sono dunque costretti a ricorrere a codice non strutturato per ottenere la semantica dei costrutti mancanti.

Il linguaggio BAL, usato per lo sviluppo del nostro caso di studio, il sistema software Gesbank, originariamente non supportava alcun costrutto iterativo tranne il FOR. In seguito sono stati aggiunti al linguaggio costrutti iterativi quali il WHILE, DO-WHILE, BREAK e CONTINUE. Tuttavia diverse porzioni del sistema Gesbank erano già state sviluppate e i programmatori erano oramai abituati a ricorrere al costrutto GOTO in diversi contesti, avendo elaborato una serie di idiomi di programmazione basati sul GOTO. L'introduzione dei nuovi costrutti non ha cambiato radicalmente le abitudini dei programmatori, i quali spesso realizzano nuove funzionalità tramite "copia-ed-incolla" di frammenti di codice esistente. La versione attuale di Gesbank, circa 8 milioni di linee di codice, contiene 0.5 milioni di istruzioni GOTO.

In letteratura esistono diversi approcci per l'eliminazione automatica delle istruzioni GOTO dal codice (per esempio, [1, 6, 8, 10]). Tuttavia le soluzioni disponibili hanno spesso un impatto molto negativo sulla qualità del codice risultante, soprattutto in presenza di GOTO cosiddetti "irriducibili" (coppie di GOTO aventi direzioni opposte e che si intersecano). Purtroppo tali GOTO sono presenti in numero non trascurabile nel codice Gesbank. Poiché il nostro obiettivo è anche quello di evitare il deterioramento della qualità del codice, non è stato possibile applicare uno degli approcci esistenti così come disponibile. È stato necessario valutare attentamente l'impatto di ciascun approccio sulla qualità del codice risultante, in modo da scegliere il migliore o la combinazione dei migliori sulle caratteristiche specifiche del codice Gesbank. In generale, ogni specifico approccio all'eliminazione dei GOTO avrà una serie di casi in cui produce codice leggibile e simile a quello che verrebbe prodotto manualmente da un intervento di rimozione dei GOTO,

ma potrà anche avere una serie di casi “degeneri” in cui i GOTO vengono eliminati al prezzo di un deterioramento della leggibilità del codice. Per questo motivo, riteniamo che una strategia efficace consista nel valutare l’impatto sulla qualità del codice risultante di un insieme di algoritmi alternativi, per poi scegliere la combinazione che garantisce il risultato migliore. Nel nostro caso di studio, abbiamo preso in considerazione quattro approcci alternativi, descritti in maggiore dettaglio nell’articolo presentato a CSMR 2008 [3]:

□ **Pattern-based:** in questo approccio i GOTO vengono eliminati identificando idiomi (pattern) di codifica ricorrenti basati su GOTO. Per ciascun idioma viene specificata la traduzione più naturale e diretta in codice privo di GOTO. L’elenco dei pattern trattati da questo approccio è stato prodotto manualmente, sulla base di ispezioni del codice Gesbank e di interviste fatte ai programmatori BAL. Un esempio di pattern di eliminazione dei GOTO è fornito nella figura 1 (*While_goto pattern*). In questo pattern, i programmatori ottengono l’equivalente di un costrutto iterativo di tipo WHILE tramite un primo GOTO che consente di uscire dall’iterazione quando la relativa condizione (*cond*) è vera. Quando tale condizione è falsa, viene eseguito il corpo dell’iterazione (*Stmt_seq_0*), seguito da un salto non condizionato (*GOTO L0*) che di fatto realizza l’iterazione (figura 1).

□ **Bohm-Jacopini top level:** uno dei primi approcci all’eliminazione dei GOTO è dovuto a Bohm-Jacopini [1]. Questo algoritmo di trasformazione è in principio in grado di eliminare tutte le occorrenze di GOTO (compresi i GOTO irriducibili). Tuttavia, in presenza di salti a istruzioni innestate all’interno di altre strutture di controllo, il risultato che si ottiene da questa trasformazione è piuttosto intricato e di difficile comprensione, in quanto la struttura innestata viene completamente distrutta dall’algoritmo. Quando invece i salti si riferiscono a istruzioni non innestate (ovvero, di *top-level*), la qualità del codice risultante è accettabile. Pertanto abbiamo deciso di applicare questa trasformazione solo quando sia l’istruzione GOTO che la label corrispondente hanno livello di *nesting 0* (*top-level*).

□ **Erosa:** anche l’algoritmo proposto da Erosa [6] consente in principio di eliminare tutti i

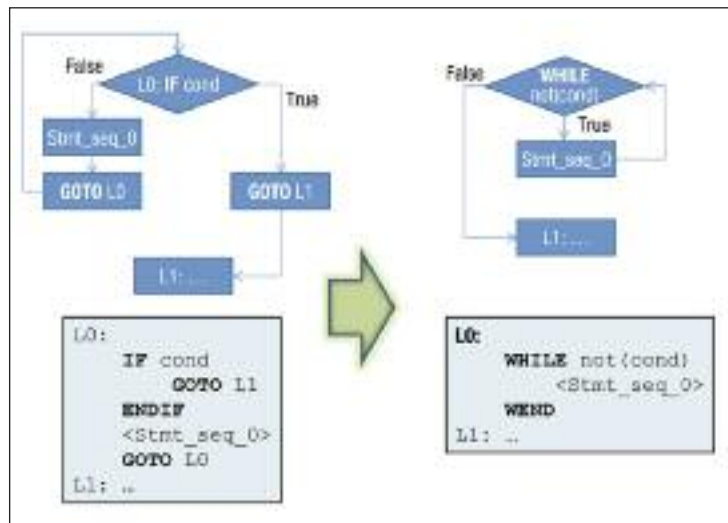


FIGURE 1

While_goto pattern

GOTO, compresi quelli irriducibili. L’algoritmo è stato sviluppato nel contesto di un compilatore ottimizzante e parallelizzante. Di conseguenza, la comprensibilità e la familiarità del codice risultante non erano un obiettivo primario. Il problema principale che talora deriva dall’applicazione di questo approccio consiste nel numero elevato di istruzioni condizionali (IF) necessarie a produrre il flusso di controllo desiderato. Talora queste istruzioni hanno un livello di *nesting* così elevato che il codice risultante è completamente illeggibile. Altre volte la trasformazione risulta assolutamente naturale e simile al codice che un programmatore scriverebbe manualmente.

□ **JGoto:** l’ultima strategia si basa sulla trasformazione del byte code Java dopo la sua compilazione. Anziché trasformare il programma sottoposto a migrazione, viene arricchito il linguaggio target. Si tratta di una strategia a cui ricorrere quando tutte le altre possibilità falliscono. Consiste nel generare invocazioni fittizie (*jgoto(label), jlabel(label)*) dove originariamente erano presenti GOTO e label. Dopo la compilazione, il byte-code risultante viene trasformato: ogni invocazione a *jlabel* viene sostituita da una *label*, inserita nel byte-code Java, e ogni invocazione a *jgoto* diventa un’istruzione GOTO, non disponibile in Java ma ammessa nel byte-code Java.

Le quattro strategie descritte qui sopra sono state implementate all’interno del tool di trasformazione del codice bal2java. Il risultato

della loro applicazione al sistema Gesbank è riportato in sintesi nel paragrafo 4.

3. STRUTTURAZIONE DEL MODELLO DATI

Il problema di strutturazione dei dati che abbiamo affrontato nel caso di studio ha origine dalla possibilità di posizionare le variabili in maniera arbitraria in memoria. Tale possibilità è offerta da svariati linguaggi di programmazione che espongono un modello dati di basso livello, molto vicino al posizionamento dei dati nella memoria fisica, vista come sequenza di byte. Per esempio, ciò accade per i linguaggi assembly, quali l'assembly dei mainframe. In tali linguaggi, identificare un insieme strutturato di dati, consistente di record contenitori (strutture o unioni di strutture) e dei campi contenuti, è piuttosto difficile. In letteratura, questo problema ha ottenuto scarsa attenzione, mentre invece gli autori si sono concentrati sul problema del riconoscimento di tipi complessi basati sull'uso di tipi di base (esempio, scalari) [5] e del riconoscimento di discriminatori per le unioni di strutture [7]. Il problema di estrarre un modello dati strutturato in presenza di variabili dislocate in memoria in maniera arbitraria non appare (per quanto a nostra conoscenza) nella letteratura specialistica del settore, nonostante abbia un'importanza

enorme nei progetti di migrazione e re-engineering di sistemi legacy scritti (completamente o in parte) in assembly o in linguaggi che espongono un modello piatto della memoria, vista come sequenza di byte ove posizionare le variabili. Il linguaggio di programmazione BAL, usato nel nostro caso di studio, ricade in questa categoria.

In BAL le variabili sono disposte sequenzialmente in memoria. Le variabili globali vengono disposte in uno spazio di memoria globale, mentre variabili locali e parametri di funzione vengono posizionati sulla pila dati. Il raggruppamento di variabili in record viene ottenuto in BAL collocando esplicitamente in memoria le variabili che costituiscono la stessa struttura dati, dando loro posizioni che si sovrappongono ai relativi contenitori. Il costrutto BAL principale con cui si ottiene questo risultato è il *FIELD=M*.

Nella figura 2 appare un esempio di dichiarazione di un record per una ipotetica struttura dati. Il codice (compreso tra *#ifdef* e *#endif*) comincia col dichiarare la variabile *conto*, di tipo stringa e di lunghezza 9 byte. L'istruzione *FIELD* che segue la dichiarazione di *conto* serve a riposizionare il puntatore alla prossima posizione disponibile in memoria per l'allocatione di variabili. Nel caso specifico, andrà a puntare all'inizio (primo byte) della variabile *conto*. Di conseguenza, le due dichiarazioni

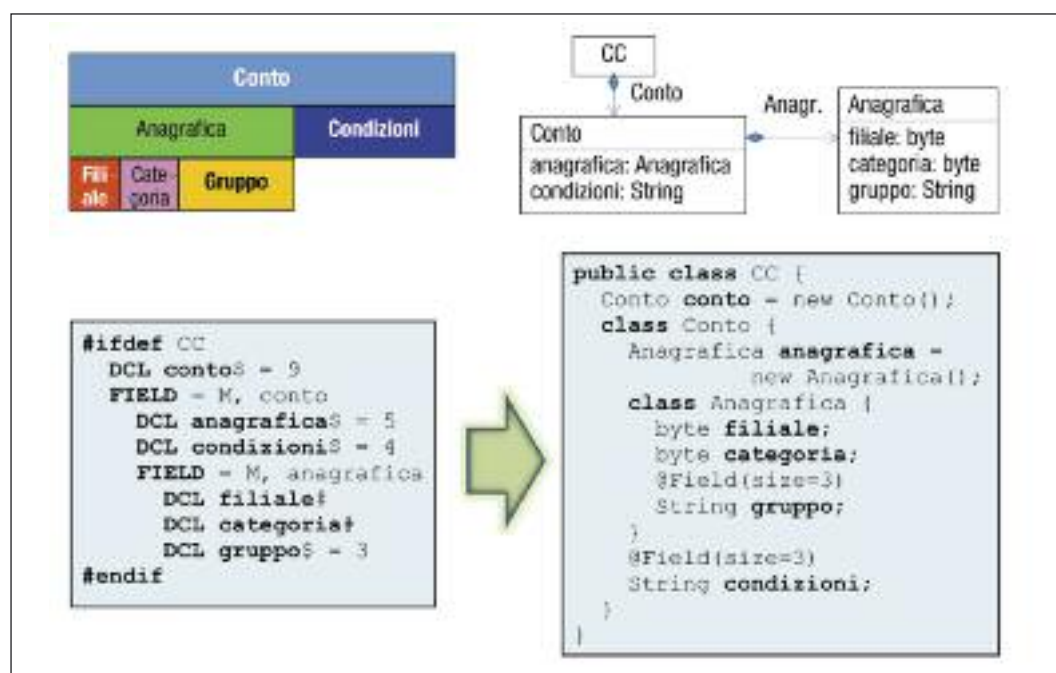


FIGURE 2
Inclusione
semplice con
dimensionamento
esatto

successive (*anagrafica* e *condizioni*) verranno a sovrapporsi completamente a *conto*. Ciò significa che la stessa sequenza di 9 byte può essere acceduta da programma come un tutt'uno (tramite la variabile *conto*) o come due segmenti separati (*anagrafica* e *condizioni*). In modo analogo, i 5 byte che costituiscono la variabile *anagrafica* vengono ridefiniti (tramite *FIELD=M*) come una prima variabile *filiale* di tipo byte, seguita da *categoria* (byte) e infine da *gruppo* (stringa di lunghezza 3). L'effetto risultante dalle dichiarazioni nella figura 2 è che, ad esempio, un assegnamento alla variabile *filiale* produce il cambiamento anche del primo byte di *anagrafica* e del primo byte di *conto*. Queste tre variabili sono infatti nomi diversi dati a regioni di memoria che partono dalla stessa posizione di memoria e dunque condividono una medesima sequenza di byte. L'uso di strutture dati posizionate direttamente sulla memoria fisica (sequenza di byte) ha diverse conseguenze negative:

1. i record non introducono alcuno scope lessicale aggiuntivo: lo spazio dei nomi è piatto e non c'è l'equivalente della notazione puntata (esempio, *conto.anagrafica.filiale*);
2. è responsabilità dei programmatori garantire il corretto dimensionamento delle variabili;
3. ci sono molti modi diversi di esprimere esattamente la stessa struttura dati.

Tali caratteristiche del modello dati di partenza rendono difficile la ricostruzione di un modello dati strutturato, quale quello imposto da Java. Per esempio, la dimensione della variabile *conto*, che contiene l'intero record, potrebbe essere erroneamente dichiarata pari a 8 o 10 byte. Le istruzioni *FIELD=M* che definiscono i campi dei sotto-record potrebbero apparire in ordine diverso e con variazioni, pur risultando nella stessa dislocazione in memoria delle variabili.

Nella parte destra della figura 2 mostriamo la traduzione in Java della struttura dati BAL di sinistra. Le dichiarazioni di variabili innestate (tramite *FIELD=M*) diventano classi contenute, come illustrato dalla relazione di composizione (*part-of*), rappresentata nel diagramma delle classi che nella figura 2 compare sopra alla traduzione in Java. Le stringhe BAL usate come contenitori di record diventano oggetti Java. Per esempio, le stringhe *conto* e *anagrafica* diventano due oggetti (*conto* e *anagrafi-*

ca) dichiarati come attributi delle classi *CC* e *Conto* rispettivamente. La traduzione nella figura 2 si basa sull'ipotesi che i record siano acceduti attraverso i loro campi (le foglie dell'albero di composizione delle classi generate) o come un tutt'uno attraverso un contenitore. Quando un intero record viene letto o scritto, il codice Java generato ricorre a primitive di serializzazione (*readFrom*, *writeTo*) degli oggetti generati.

Non sempre la disposizione delle variabili in memoria rispetta la struttura illustrata nella figura 2, in cui ogni ridefinizione si limita a spezzare un contenitore più grande in sotto-campi più piccoli. Può accadere che su uno stesso contenitore siano definite più viste indipendenti. Quando ciò accade, il contenitore definisce una *union* di più viste (*varianti*) alternative. In Java non è disponibile alcun supporto nativo per le union, pertanto nella traduzione è necessario ricorrere a un insieme di interfacce (per le varianti) e a una classe che realizza la union implementando tutte le interfacce. Ciò consente di accedere ai dati contenuti nella classe attraverso una qualsiasi delle viste disponibili. La consistenza tra queste viste è garantita da codice generato automaticamente, in grado di cambiare la variante attiva, quando viene acceduta una variante diversa da quella correntemente attiva, e in grado di copiare i dati dalla variante corrente su quella che deve essere attivata, se necessario.

Oltre alle union, le strutture dati di partenza possono presentare altri problemi, relativi a:

1. **inversioni:** contenitore e campi contenuti (esempio, *conto*, {*anagrafica*, *condizioni*}) nella figura 2) possono essere dichiarati in ordine inverso (prima le parti e poi il contenitore);
2. **contenitore mancante:** il contenitore potrebbe essere assente;
3. **dimensionamento scorretto:** le dimensioni dei sotto-campi potrebbero non corrispondere a quelle dei rispettivi contenitori.

La soluzione al problema 1 consiste nel riconoscere l'inversione e nel ripristinare automaticamente l'ordine corretto delle dichiarazioni. Per il problema 2 si ricorre all'introduzione automatica di un contenitore fittizio. Il problema 3 può essere affrontato riconoscendo una serie di pattern e applicando un'opportuna euristica di strutturazione dei dati BAL per ciascun caso [3]. Per esempio, l'ultimo campo conte-

nuto in un record sfiora la dimensione del record, ma è seguito immediatamente dalla dichiarazione di un altro record. In tale caso è ragionevole ipotizzare che il campo appartenga interamente al record precedente.

4. CASO DI STUDIO

Il sistema Gesbank sottoposto a migrazione da BAL a Java è un applicativo bancario che supporta tutte le funzionalità necessarie all'operatività di una banca: gestione dei conti correnti, gestione dei prodotti finanziari, operazioni allo sportello, comunicazioni alla banca centrale e alle altre authority, comunicazioni interbancarie, generazione di statistiche e di reportistica. L'interfaccia utente è a caratteri e l'architettura complessiva è client-server. L'ambiente di esecuzione è una piattaforma proprietaria detta B2U.

L'applicativo è di dimensioni notevoli (circa 8 milioni di linee di codice). Il linguaggio BAL consente direttive di *pre-processing*, che comportano un fattore di espansione del codice pari a circa 1,63. I dati persistenti risiedono all'interno di tabelle ISAM, per un totale di 1.339 file ISAM e 5.893 tabelle ISAM. I dati persistenti sono descritti in una tabella ISAM particolare, det-

ta *dizionario*. Poiché il modello dati usato nel dizionario è lo stesso supportato dal linguaggio BAL (basato su FIELD=M), la generazione delle classi Java che modellano i dati persistenti utilizza gli stessi algoritmi definiti per le strutture dati BAL (vedi paragrafo precedente).

Tutte le strategie di eliminazione dei GOTO discusse nel paragrafo 2 sono state applicate a tutto il codice Gesbank. Il codice risultante è stato analizzato in modo da identificare la strategia o la combinazione di strategie maggiormente appropriate. I risultati sono discussi in dettaglio nell'articolo presentato a CSMR 2008 [3]. Nella tabella 1 mostriamo una sintesi di tali risultati. Per le quattro strategie considerate (per Erosa distinguiamo l'applicazione *forward* dall'applicazione *backward* dell'algoritmo) possiamo confrontare la percentuale di GOTO rimossi, il massimo livello di *nesting* presente nel codice trasformato ed il tempo di computazione richiesto dalla trasformazione automatica del codice. La strategia JGoto non elimina alcun GOTO dal codice sorgente, in quanto trasforma il byte-code risultante dalla compilazione del codice Java. Il relativo *overhead* di compilazione è trascurabile (per tanto non compare alcun tempo di trasformazione per questa strategia).

Dai dati della tabella 1 possiamo concludere che le strategie Pattern-based e Bohm-Jacopini top-level sono estremamente efficaci, in quanto consentono di rimuovere una percentuale significativa di GOTO (ovviamente, non tutti), con un impatto minimo sulla qualità del codice risultante (il livello di *nesting* aumenta al più di 1 o 2 per le due strategie). Al contrario, Erosa è sì in grado di rimuovere tutti i GOTO, ma produce un incremento inaccettabile del livello di *nesting*. Per questo motivo, si è deciso di scartare questo approccio e di ricorrere a JGoto per i casi residui in cui Pattern-based e Bohm-Jacopini top-level non sono in grado di eliminare i GOTO.

La tabella 2 consente di valutare l'efficacia delle euristiche di strutturazione dei dati BAL presentate nella sezione precedente (ulteriori dettagli sono forniti in [3]). Il caso 1, in cui si ha un dimensionamento esatto (come nella figura 2) dei contenitori, copre la maggior parte delle istanze di strutture dati presenti nel codice utente e nel dizionario. Le euristiche relative ai casi 2 e 3 consentono di aumentare significativa-

Strategy	GOTO rimossi	Max nesting	Tempo
Pattern-based	21,4%	29 (+1)	1 h 2 m
Bohm-Jacopini top-level	79,3%	30 (+2)	22 m
Erosa (backward)	100,0%	162 (+134)	117 h 38 m
Erosa (forward)	100,0%	207 (+179)	144 h 22 m
JGoto	0%	28 (+0)	-

TABELLA 1

Risultati prodotti dalle quattro strategie di eliminazione dei GOTO

Casi	Codice utente	Dizionario
Caso 1: dimensionamento esatto	425.988	8.279
Caso 2: sottodimensionamento	27.100	65
Caso 3: sovradimensionamento	13.850	4
Caso 4: errore	5.286	45

TABELLA 2

Numero di occorrenze dei vari casi di strutture dati BAL

	Codice utente	Dizionario
Classi	510.108	12.402
Interfacce	148.621	753
Union	29.394	263

TABELLA 3

Codice Java generato per le strutture dati BAL

mente la quota di strutture BAL tradotte automaticamente in classi Java. I casi di errore residui sono in numero sufficientemente basso da consentire un intervento manuale. Ai programmatori BAL è stato chiesto di fissare i problemi all'origine delle incompatibilità di dimensione che hanno prodotto questi errori. Spesso l'intervento necessario è stato banalmente un ridimensionamento del contenitore. In alcuni casi è stato necessario introdurre un contenitore mancante o un campo FILLER, avente lo scopo di completare la ridefinizione di una variante di una union o di un sotto-record.

Nella tabella 3 viene mostrato il risultato della traduzione in Java delle strutture dati BAL. La maggior parte delle strutture dati di partenza può essere tradotta tramite classi Java "standard". In un numero limitato di casi è stato necessario simulare le union in Java, tramite l'implementazione di interfacce, associate alle diverse varianti della union. Nella tabella 3 riportiamo il numero di union generate, nonché il numero di interfacce generate per le varianti delle union.

5. CONCLUSIONI

I progetti di migrazione, specialmente per sistemi software di grosse dimensioni, sono intrinsecamente ad alto rischio e diverse sono le cause che possono portare al loro fallimento. Dall'esperienza di migrazione riportata in questo articolo si possono ottenere le seguenti regole empiriche:

1. La valutazione dei risultati prodotti dalla migrazione del codice coinvolge più dimensioni, spesso in conflitto l'una con l'altra. Per esempio, nell'eliminazione dei GOTO abbiamo riscontrato una dicotomia tra la capacità di trattare in modo automatico tutti i possibili casi di GOTO e la manutenibilità del codice trasformato.

2. Dato un sistema software complesso, come Gesbank, non è possibile risolvere tutte le istanze di un problema con una singola soluzione. La varietà dei casi presenti impone la valutazione e la combinazione di una molteplicità di alternative.

3. Soluzioni con proprietà teoriche interessanti (esempio, capacità di trattare GOTO irriducibili) possono rivelarsi in pratica inusabili, a causa delle caratteristiche specifiche del codice sottoposto a migrazione.

4. È importante identificare sempre soluzioni di *backup*, (esempio, JGoto) a cui ricorrere quando si incontrano ostacoli insormontabili nel corso del progetto di migrazione, nonostante queste possano introdurre del codice non del tutto soddisfacente. Rappresentano infatti una tappa importante verso la soluzione "ideale".

5. I progetti di migrazione hanno obiettivi di breve e obiettivi di lungo periodo. Nel breve termine non è possibile fissare obiettivi troppo elevati in quanto irrealistici e spesso bisogna accettare dei compromessi. Tuttavia, rimane importante avere chiara la strategia di lungo termine, in modo da innescare un processo di miglioramento continuo, che continua anche quando la migrazione è terminata.

Bibliografia

- [1] Bohm C., Jacopini G: Flow diagrams, turing machines and languages with only two formation rules. *Communications of the ACM*, Vol. 9, n. 5, 1966, p. 366-371.
- [2] Ceccato Mariano, Roy Dean Thomas, Tonella Paolo, Marchignoli Davide: *Data model reverse engineering in migrating a legacy system to Java*. In: Proc. of the 15-th Working Conference on Reverse Engineering. IEEE Computer Society, Antwerp, Belgium, October 2008.
- [3] Ceccato Mariano, Tonella Paolo, Matteotti Cristina: *Goto elimination strategies in the migration of legacy code to Java*. In: Kontogiannis K., Tjortjis C., Winter A., editors, In: Proc. of the 12-th European Conference on Software Maintenance and Reengineering, p. 53-62. IEEE Computer Society, Athens, Greece, April 2008.
- [4] Cordy Jim: The TXL source transformation language. *Science of Computer Programming*, Vol. 61, n. 3, 2006, p. 190-210.
- [5] van Deursen Arie, Moonen Leon: *Exploring legacy systems using types*. In: Proc. of the 7-th Working Conference on Reverse Engineering, 2000, p. 32-41.

- [6] Erosa A.M., Hendren L.J.: *Taming control flow: a structured approach to eliminating goto statements*. In: Proc. of the Int. Conf. on Computer Languages, 1994, p. 229-240.
- [7] Komondoor R., Ramalingam G.: *Recovering data models via guarded dependences*. In: Proc. of the 14-th Working Conference on Reverse Engineering, 2007, p. 110-119.
- [8] Ramshaw L.: Eliminating goto's while preserving program structure. *Journal of the ACM*, Vol. 35, n. 4, 1988, p. 893-920.
- [9] Ricca Filippo, Tonella Paolo: Reverse Engineering di sistemi software. Limiti e potenzialità. *Mondo Digitale*, n. 3, 2006, p. 52-62.
- [10] Zhang F., D'Hollander E.H.: Using hammock graphs to structure programs. *IEEE Trans. on Software Engineering*, Vol. 30, n. 4, 2004, p. 231-245.

Migrare, riscrivere o incapsulare? Strategie aziendali di gestione del rischio

Le aziende che operano nel campo delle tecnologie dell'informazione in generale, e dello sviluppo del software in particolare, si trovano costantemente sottoposte ad una forte pressione che spinge verso l'innovazione e l'adozione di nuove tecnologie. Aziende di successo possono vedere notevolmente ridotta la loro quota di mercato o possono addirittura essere estromesse dal mercato dai concorrenti a causa di un mancato adeguamento dei loro prodotti alle novità tecnologiche.

Nel settore bancario, a causa della criticità dei dati manipolati, l'innovazione si accompagna necessariamente ad altri requisiti importanti quali l'affidabilità, la robustezza, la scalabilità e la maturità delle tecnologie usate. Per questo motivo la spinta all'innovazione è spesso mitigata da una tendenza a mantenere soluzioni tecnologicamente superate ma affidabili (per esempio, Cobol su mainframe). D'altra parte gli utenti finali dei sistemi bancari stanno diventando sempre più competenti circa le tecnologie informatiche che usano e di conseguenza chiedono funzionalità e prestazioni in linea con lo stato dell'arte (interfacce grafiche, accesso via Web services, interoperabilità sui dati ecc.), senza ovviamente rinunciare all'affidabilità. Le aziende che operano nel settore del software bancario (come IBT) si trovano pertanto costantemente di fronte al dilemma se innovare, assumendosi il rischio legato all'adozione di nuove tecnologie, o mantenere il sistema nella sua forma attuale, dando però così un margine di sorpasso tecnologico ai concorrenti. Talvolta il dilemma si risolve in favore dell'innovazione, soprattutto se ci sono motivi per ritenere che la perdita di quota di mercato dovuta a prodotti concorrenti maggiormente innovativi sia una possibilità concreta. La valutazione strategica di IBT circa le tecnologie in uso (piattaforma proprietaria B2U, linguaggio proprietario BAL ed ambiente di compilazione ed esecuzione proprietario) è stata esattamente di questo tipo: innovare per rimanere competitivi sul mercato. L'innovazione si accompagna però a enormi rischi. Sono documentati diversi progetti di passaggio a nuove tecnologie che hanno avuto esiti fallimentari e talora disastrosi [9]. La scelta di una strategia di innovazione adeguata è pertanto fondamentale e di solito si riduce a tre alternative principali:

□ **Incapsulamento (wrapping):** è l'alternativa a rischio minore, ma è anche quella che dà meno vantaggi nel processo di sviluppo. Il software esistente viene mantenuto e i programmatori continuano a lavorare con ambiente, strumenti e linguaggi usuali. Per rendere però disponibili all'utente finale funzionalità più avanzate, il software viene incapsulato in modo da esportare procedure e dati necessari ai moduli esterni (magari sviluppati con tecnologie allo stato dell'arte), in grado di fornire le funzionalità richieste dall'utente finale.

□ **Riscrittura:** è l'alternativa a maggiore rischio. I problemi principali legati a questa alternativa sono:

1. a documentazione delle cosiddette "business rules", spesso carente o del tutto assente;
2. la disponibilità di risorse aziendali adeguate a un progetto di riscrittura.

Nei sistemi software legacy, spesso il codice sorgente è anche la documentazione più fedele delle regole di business implementate nel sistema. La riscrittura deve pertanto partire da un lungo e complesso processo di acquisizione di conoscenza a partire dal codice esistente prima che si possa progettare il nuovo sistema. I rischi maggiori sono relativi alla difficoltà di ottenere un insieme di specifiche completo e accurato. Le risorse necessarie a tale attività vengono inoltre sottratte al normale processo di sviluppo del software, che continua sul sistema legacy mentre viene prodotto il sistema nuovo. Di conseguenza tali risorse sono spesso inadeguate e sovraccaricate di altri compiti. È dunque difficile pianificare i tempi e i costi della riscrittura, che deve comprendere anche una fase di test tale da garantire la stessa affidabilità che possedeva il sistema in via di sostituzione. Ciò può richiedere un tempo almeno tanto lungo quanto la riscrittura stessa.

□ **Migrazione:** è un'alternativa a rischio intermedio, anche se all'interno di questa alternativa troviamo uno spettro di opzioni che si avvicinano da un lato all'incapsulamento e dall'altro alla riscrittura. Il vantaggio prin-

segue

La principale della migrazione è che in principio può avvalersi di strumenti automatici di analisi e trasformazione del codice. Ciò consente di limitare il coinvolgimento di risorse umane sia nella fase di transizione che in quella di testing, anche se un processo totalmente automatico di solito non è praticabile. A un estremo della migrazione troviamo la generazione di codice in un linguaggio moderno (per esempio, Java da Cobol) senza alcuna pretesa di leggibilità e manutenibilità del codice. L'assunzione è che (come nell'incapsulamento) si continuerà a lavorare con il vecchio codice, ma il codice generato consentirà di agganciare in esecuzione ambienti e tecnologie moderni. All'altro estremo, si punta sull'analisi automatica (reverse engineering) per facilitare l'estrazione delle business rules, ma si lascia ai programmatori la maggior parte dell'attività di ri-codifica. La scelta strategica di IBT è stata quella di migrare, cercando di minimizzare l'attività manuale. Si è fatto pesantemente ricorso a strumenti di trasformazione automatica del codice e di analisi del codice in supporto agli inevitabili interventi manuali. Un'altra scelta critica riguarda il rilascio all'utente finale. La strategia IBT è basata su un ambiente di esecuzione misto, che consente il rilascio incrementale dei programmi migrati.

PAOLO TONELLA è a capo dell'unità di ricerca SE (Software Engineering) presso la Fondazione Bruno Kessler-IRST di Trento. Si è laureato e ha ottenuto il dottorato di ricerca presso l'Università degli Studi di Padova. È autore del libro "Reverse Engineering of Object Oriented Code", Springer, 2005, oltre che di svariati articoli pubblicati su riviste e presentati a conferenza, su temi che spaziano dal reverse engineering e re-engineering, al testing del software, alla programmazione ad aspetti. Nel 2007 è stato elencato tra i primi 50 ricercatori nell'area del Software Engineering in un articolo pubblicato su *Communications of the ACM* (Vol. 50, n. 6, June 2007, p. 81-85).
E-mail: tonella@fbk.eu

MARIANO CECCATO è ricercatore presso la Fondazione Bruno Kessler-IRST di Trento e insegna "Laboratorio di Analisi e Testing del Software" presso l'Università degli Studi di Trento. Ha ottenuto la laurea in Ingegneria Informatica dall'Università di Padova nel 2003 e il dottorato di ricerca in Informatica dall'Università di Trento nel 2006. È autore di numerosi articoli pubblicati su riviste e presentati a conferenze internazionali. I suoi interessi di ricerca comprendono l'analisi automatica e la trasformazione del codice sorgente, la sicurezza, il testing, e l'ingegneria del software empirica.
E-mail: ceccato@fbk.eu

DAVIDE MARCHIGNOLI è responsabile del progetto di migrazione Gesbank presso IBT. Ha conseguito nel 2002 un dottorato di ricerca in informatica presso l'Università di Pisa occupandosi principalmente di teoria dei tipi. Da allora si occupa principalmente di progettazione di sistemi software.
E-mail: davide.marchignoli@ibttn.it

CRISTINA MATTEOTTI ha conseguito nel 2007 la laurea triennale in Informatica presso l'Università degli Studi di Trento. Da allora lavora come software engineer presso la Fondazione Bruno Kessler-IRST di Trento, ove realizza tool per la trasformazione automatica del codice.
E-mail: matteotti@fbk.eu

THOMAS ROY DEAN è professore associato presso la Queen's University di Kingston, Canada. In passato ha condotto attività di ricerca su sistemi di controllo del traffico aereo e sulla formalizzazione dei linguaggi. Per oltre cinque anni ha lavorato presso la Legasys Corp. come ricercatore senior, occupandosi di trasformazioni avanzate del codice e di tecniche di evoluzione in un contesto industriale. I suoi interessi attuali includono la trasformazione del software, l'evoluzione dei siti Web e la sicurezza delle applicazioni di rete.
E-mail: tom.dean@queensu.ca