

SCABIS

Manifesto della Value Chain

“SCAlable Big data
Infrastructure for innovative
Services”

Manifesto della Value Chain

“SCAlable Big data Infrastructure for innovative Services”

1. NOME/ACRONIMO DELLA VC

SCAlable Big data Infrastructure for innovative Services (**SCABIS**)

2. BREVE DESCRIZIONE DELLA VC

Questa VC focalizza tutte le tematiche relative al progetto, alla produzione, al supporto di servizi che siano intrinsecamente scalabili, efficienti, ed ottimizzati per i requisiti utente delle diverse aree applicative, con particolare accento alle infrastrutture interoperabili presenti nei data center o nelle strutture di supporto per il trattamento di Big Data in modo innovativo ed aperto.

3. OBIETTIVI STRATEGICI DELLA VC

1. Supporto alla progettazione di servizi ad alta scalabilità
2. Integrazione strumenti e servizi per la gestione di Big Data e loro ottimizzazione
3. Business Intelligence e Analisi di Big Data
4. Servizi scalabili nei processi decisionali, produttivi e previsionali
5. Servizi scalabili per Smart Cities e Communities

1. Supporto alla progettazione di servizi ad alta scalabilità: promuovere nuovi modelli e supporti per la gestione elastica e ottimizzata di risorse ICT che tenga conto di tutte le recenti evoluzioni nell'area del cloud computing. Andando oltre, si vogliono supportare scenari di deployment evoluti con la possibilità di avere cloud bursting nei casi in cui si vogliono estendere al volo le capacità del proprio cloud privato sfruttando cloud pubblici, e cloud federation con data center distribuiti geograficamente per ottenere non solo scalabilità, ma completa resilienza dei servizi.

2. Integrazione strumenti e servizi per la gestione di Big Data e loro ottimizzazioneLa prima direzione è la gestione delle risorse fisiche e dei cluster per il processing, includendo tutti i relativi problemi di deployment e distribuzione; la seconda è la memorizzazione e lo storage dei dati nel distribuito (distributed file system, NoSQL databases, scalable pub/sub message queues); la terza è la virtualizzazione della rete e degli strumenti di gestione e di controllo. A livello di applicazione abbiamo la progettazione di piattaforme per il processing offline (batch) e online (stream) di Big Data; la integrazione e il supporto di linguaggi e strumenti di alto livello per il trattamento dei Big Data, spaziando ad esempio da librerie di machine learning, al cognitive computing.

3. Analisi di Big Data: impiegare tecnologie avanzate e infrastrutture integrate di cui agli obiettivi precedenti per estrarre informazioni e conoscenza dai dati raccolti dalle aziende al fine di avere un impatto su efficienza, efficacia e strategia aziendali.

4. Servizi scalabili nei processi decisionali, produttivi e previsionali: si rivolge a un primo e importante scenario applicativo e si propone, attraverso l'adozione di tecnologie per la gestione e l'integrazione semantica e l'analisi di Big Data, di abilitare una visione omogenea e unificata dei dati presenti all'interno e all'esterno dell'azienda/istituzione, proponendo previsioni e fornendo strumenti automatici di supporto e segnalazione.

5. Servizi scalabili per Smart Cities e Communities: intende promuovere l'analisi e l'individuazione di best practice, architetture di riferimento, SLA, ecc., per la pianificazione e l'erogazione di servizi scalabili rivolti alle infrastrutture e alle applicazioni (sia di backoffice che per gli utenti) per il rilevante scenario applicativo delle Smart Cities e Communities.

6. Servizi innovativi per Big Data e Internet of Things: poiché si prevede che entro il 2020, da 20 a 30 miliardi di device avranno accesso Internet, questi possono contribuire in misura significativa a rendere concreta la visione di città più sostenibili e efficienti nonché inclusive; questa prospettiva, tuttavia, richiede la definizione, la standardizzazione, la diffusione e la composizione di metodologie, blueprint e tecnologie per piattaforme di erogazione fortemente scalabili che soddisfino anche ulteriori requisiti essenziali, come disponibilità, integrabilità, facilità di uso, e sicurezza.

4. DESCRIZIONE DELLA VC

STANDARD (soggetti interessati: aziende ICT, aziende di servizi con un forte reparto ICT, start-up) i) Standard di usabilità e accessibilità

ii) Certificate of Cloud Security Knowledge (CCSK) della Cloud Security Alliance (CSA)

iii) Federal Information Security Management Act (FISMA) del National Institute of Standards and Technology (NIST) iv) Cloud computing risk assessment della European Network and Information Security Agency (ENISA) v) Open Virtualization Format (OVF) della Distributed Management Task Force (DMTF)

vi) P2413 - Standard for an Architectural Framework for the Internet of Things (IoT) di IEEE

R&D (soggetti interessati: università e centri di ricerca, laboratori rete alta tecnologia, aziende ICT, aziende di servizi con un forte reparto ICT, start-up, associazioni interessate come Big Data Value European Association) a) Sviluppo di nuove tecniche di analisi dei dati e di gestione dei dati.

L'enorme quantità di dati disponibili ha posto nuove sfide:

b) nella realizzazione di infrastrutture per la loro raccolta e memorizzazione (considerando anche le problematiche di privacy e sicurezza);

c) nella definizione e realizzazione di tecnologie di integrazione, elaborazione e analisi che siano effettivamente scalabili al crescere della dimensione dei dati.

TECNOLOGIE (soggetti interessati: aziende ICT, aziende di settori verticali per la integrazione)

a) Cloud Storage: questo servizio permette di riservare spazio su disco per i dati letti o scritti dalle applicazioni dell'utente ed eventualmente di riservare quantità di memoria di massa su disco o cassetta in modo permanente sui sistemi di storage. b) Tecnologie Cloud Computing: questo servizio permette all'utente di utilizzare le risorse di calcolo configurando il proprio ambiente di calcolo virtuale secondo le proprie esigenze con un approvvigionamento di risorse personalizzate 'on-demand'. c) Tecnologie di Data Science:

i) Tecnologie per la Knowledge Discovery: esplorazione e navigazione di sorgenti Big Data;

ii) Tecnologie per la Big Data Integration: integrazione di sistemi informativi aziendali (ERP, CRM, DW) con sorgenti Big Data esterne attraverso tecniche basate sulla semantica;

iii) Tecnologie per la Big Data Analysis: strumenti e tecniche per l'analisi dati massiva di sorgenti Big Data, data warehouse e open data;

iv) Tecnologie di Business Analytics per generare valore dai Big Data: strumenti e tecniche per una facile comprensione d) Tecnologie applicative di supporto:

i) Machine learning/Data mining: tecniche per l'estrazione di conoscenza dai dati;

ii) Analisi statistica: tecniche per l'analisi dei dati secondo i principi della statistica;

iii) gestione dei dati e applicazione degli algoritmi di analisi nel cloud computing;

iv) IoT: raccolta dei dati attraverso l'Internet of Things.

v) condivisione degli indicatori chiave di business all'interno dell'azienda/istituzione;

e) Servizi di calcolo per applicazioni HPC su cluster Intel/infiniband/GPU:....

PROCESSI PRODUTTIVI (soggetti interessati: aziende ICT)

- a) Sviluppo software;
- b) Integrazione in ambiente locale e/o cloud
- c) Monitoraggio ed ottimizzazione dei processi produttivi industriali
- d) Ottimizzazione dei processi decisionali
- e) Ottimizzazione dei sistemi previsionali
- f) Elaborazione e condivisione degli indicatori chiave di business

APPLICAZIONI (settori interessati: l'economico-finanziario, il manifatturiero, dei trasporti, l'energetico, le Smart Cities, le scienze della vita - ambito medico-sanitario e genomico; in generale, tutti i settori in cui si producono e si gestiscono servizi)

- a) Piattaforme software per il monitoraggio e l'ottimizzazione dei processi produttivi
- b) Sistemi avanzati di supporto alle decisioni
- c) Piattaforme software per l'elaborazione di previsioni
- d) Piattaforme di Business Analytics per l'elaborazione e condivisione dei KPI aziendali

ICT & INTEGRATION (soggetti interessati: aziende del settore ICT e di produzione e di gestione dei servizi)

- a) Raccolta dei dati provenienti da controllo di processo, IoT, sensoristica, automazione e robotica, infrastrutture di rete, datacenter, sistemi ICT a supporto della tracciabilità (RFID).
- b) Servizi di Knowledge Discovery
- c) Servizi per Big Data Integration
- d) Servizi di Business Analytics
- e) integrazione e adattamento dei servizi ai requisiti degli utenti.

BUSINESS MODEL (settori interessati: aziende ICT, economico-finanziario, il manifatturiero, dei trasporti, l'energetico, le Smart Cities, le scienze della vita)

- a) Licenze software: l'azienda user acquista la licenza del software per l'analisi dei dati;
- b) Subscription: l'azienda user paga un abbonamento per l'uso del software, spesso offerto con modalità Software as a service (SaaS);
- c) Pay-per-use: l'azienda user paga in base all'uso effettivo del software spesso offerto con modalità Software as a service (SaaS).
- d) modelli di business e di costo per uso di federazione di Cloud e data center.

EDUCAZIONE E FORMAZIONE (soggetti interessati: Università, ITS, enti di ricerca e di formazione, pubblica amministrazione e policy makers)

- a) Corsi nell'ambito di Corsi di Studio Triennali, Magistrali, Master 1° e 2° livello, Corsi di formazione, Corsi di perfezionamento, riguardanti le tecnologie per la memorizzazione, gestione e analisi di big data. Le figure professionali sono sia del settore informatico e tecnico, ma includono anche molteplici altri settori multidisciplinari vicini (data scientist, esperti di educazione) e anche interessati in modo laterale (digital humanities, design delle smart cities, ...).

ECONOMIA CIRCOLARE (soggetti interessati: tutti gli stakeholders)

- Riduzione dell'impegno di risorse equivalenti per la produzione dei servizi nei diversi settori
- Ottimizzazione del consumo e della CO2 equivalente in data center e altre organizzazioni
- Modelli di riutilizzo dei dati e di tutte le risorse correlate in ambito delle Smart Cities

5. IL POSIZIONAMENTO DELLA REGIONE RISPETTO ALLA VC NEL CONTESTO NAZIONALE ED INTERNAZIONALE

I temi della VC hanno un'importanza strategica in coerenza con le tematiche delle diverse iniziative in atto a livello regionale, nazionale ed internazionale. In particolare, nel seguito si riporta

l'allineamento rispetto alle direzioni più significative di progettazione a livello regionale e europeo.

- A livello regionale:
 - Priorità C - direttrice di cambiamento società dell'informazione e della comunicazione della S3 regionale o Priorità D - servizi alle imprese di informazione e comunicazione della S3 regionale
 - Agenda digitale della regione Emilia Romagna.
 - Il territorio della RER è molto ricco dal punto di vista delle aziende di servizi, e in particolare di servizi ICT, sia come numero di aziende di varie dimensioni, con un rilevante numero di addetti e un significativo assorbimento in termini occupazionali. Inoltre, le RER presenta un numero elevato di nuove startup innovative, tipicamente organizzate intorno a temi e servizi che fanno riferimento alla value chain.
- A livello europeo:
 - Nella nuova architettura delle piattaforme S3 (Smart Specialization Strategies Platforms - S3P), l'ICT compare sia in modo forte ed esplicito nella piattaforma Digital Growth (o Digital), sia trasversalmente come fattore abilitante della Digital Agenda for Europe che a sua volta rappresenta una delle sette flagship a livello europeo.
 - Il settore ICT è molto consistente con la logica di fornitura di servizi scalabili e con un impiego di molti dati, raccolti in modo differenziato ed integrati, per la fornitura di servizi. Tale settore viene riconosciuto come settore verticale e finanziato "direttamente" nelle cosiddette call Leadership in Enabling and Industrial Technologies (LEIT) in cui complessivamente conta un budget di 7,3 B€, quindi maggiore del 60% dedicato da FP7 a ICT in modo diretto. A tale budget si sommano i finanziamenti che arrivano "indirettamente" all'ICT in quanto enabler orizzontale per molte delle call che indirizzano le cosiddette call "Societal Challenges".
 - Il tema Cloud Computing presenta diverse linee di finanziamento all'interno del pillar H2020 Leadership in Enabling and Industrial Technologies (LEIT) nelle recenti call 2016-2017 (ICT-06-2016, EUJ-02-2016, EUB-01-2017). La Commissione Europea ha recentemente firmato una Public-Private Partnership (PPP) con la Big Data Value Association, con l'obiettivo di permettere all'Europa di giocare un ruolo rilevante sui Big Data a livello globale.
 - Il tema BIG DATA presenta diverse linee di finanziamento all'interno del pillar H2020 Leadership in Enabling and Industrial Technologies (LEIT). Per il solo 2017 sono attive 4 call (ICT-14,15,16,17) e un Prize che coinvolgono complessivamente circa 90 milioni di euro.

A livello Europeo, l'interesse per aree e regioni che siano capaci di fornire servizi è particolarmente sviluppato e si prevede una accelerazione con la proposta di regioni nelle diverse nazioni Europee che possano essere singolarmente trainanti nella fornitura di servizi scalabili. La RER si può proporre per le capacità imprenditoriali nella fornitura di servizi tradizionali ad alta scalabilità (per aree anche non correlate all'ICT), e anche nei settori dei nuovi servizi tecnologici ICT ad alto valore aggiunto.

Analisi SWOT della VC

Strengths

- L'Emilia-Romagna ricopre una posizione privilegiata in ambito computazionale. In regione risiede infatti il 70% delle risorse di computing e storage italiane.
- Competenze universitarie su data science e Infrastrutture (CINECA, INFN)
- Consapevolezza dell'importanza dell'analisi dei big data - La futura disponibilità di diversi Data Center Lepida sul territorio regionale
- Una rete significativa di aziende regionali impegnate nella fornitura di servizi ICT per la gestione dei dati in ambito e-Government e e-Health
 - Presenza di gruppi di ricerca (sia nei centri accreditati che in altre istituzioni di ricerca regionali) con elevate competenze trasversali sui temi dei servizi ICT ad alta scalabilità e su quelli applicativi delle Smart Cities, dell'e-Government e e-Health
- Disponibilità di strutture e expertise consolidate per data analysis, harmonization & storage mediante HPC (High-Performance Computing) e HTC (High Throughput Computing)
- Cooperazioni esistenti con attori primari dello scenario globale della fornitura di servizi Cloud/Big Data
- Opportunità per le aziende di accedere, attraverso la collaborazione con gli enti di ricerca presenti sul territorio, a reti di collaborazione tecnologica di dimensione internazionale

Weaknesses

- Dimensioni aziendali e ridotta propensione alla ricerca di molte aziende consolidate
- Scarsa propensione all'innovazione di processo da parte delle aziende
- Limitate capacità di accedere alle opportunità di finanziamento di progetti di ricerca e di trasferimento tecnologico (sempre più limitate) - Difficoltà finanziarie e gestionali degli enti locali - Difficoltà nel rapporto tra enti pubblici e privati
- Difficoltà nel passaggio dalle versioni pilot dei progetti di collaborazione con le imprese a sistemi stabili, solidi in cui viene mantenuto il livello di efficienza nel tempo

Opportunities

- Possibilità di partecipazione a progetti nazionali ed europei dei gruppi di ricerca presenti in regione
- Innovazioni tecnologiche, organizzative e di prodotto - Disponibilità di tecnologie a basso costo
- Il raggiungimento di una massa critica in consorzi di aziende e centri di ricerca per poter affrontare problemi challenging e fornire soluzioni condivise e di validità generale
- Il progetto regionale Big Data che formerà giovani ricercatori in collaborazione con le aziende
- Collaborazioni internazionali pubblico/privato nel campo del cloud computing e big-data analysis consolidate sulla base di progetti europei FP7 e H2020

Threats

- Difficoltà di reperimento di personale formato sui temi della VC e della data science, per carriere non solo tecnologiche
- La difficoltà della politica di individuare e sostenere concretamente scelte strategiche che possano favorire efficacemente tutti gli ambiti di potenziale sviluppo del settore
- La velleità di reinventare tutto ignorando o sminuendo le best practice internazionali, che in molti casi possono essere adattate e applicate alla realtà presente nella regione ER