



Rapporto tecnico OR 2

Codice deliverable di progetto: D_2.3

Obiettivo realizzativo di riferimento: OR2 – Progettazione e sviluppo di A@GRIBASE

Documento pubblico: SI

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA
Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management



Regione Puglia - Area Politiche per lo Sviluppo, il Lavoro e l'Innovazione
Servizio Ricerca Industriale e Innovazione - Ufficio Servizi e-Government e ICT
P.O. FESR PUGLIA 2007-2013 - ASSE I - Linea di Intervento 1.4 - Azione 1.4.2 - Investiamo nel vostro futuro
SUPPORTO ALLA CRESCITA E SVILUPPO DI PMI SPECIALIZZATE NELL'OFFERTA DI CONTENUTI E SERVIZI DIGITALI
Apulian ICT Living Labs - Verso Puglia Digitale 2020



INDICE

1	Sintesi attività svolte	3
1.1	WP 2.1: Progettazione e implementazione del Data Ingest	3
1.2	WP 2.2: Progettazione della struttura di A@GRIBASE e sviluppo delle componenti di popolamento della banca dati.....	4
1.3	WP 2.3: Progettazione e implementazione delle componenti costituenti la catena di elaborazione e retrieval del dato.....	4
1.4	WP 2.4: Progettazione e implementazione dei tools di gestione e monitoraggio della catena di acquisizione-elaborazione-providing del dato	5
1.5	WP 2.5: Progettazione e realizzazione dell'infrastruttura IT sulla cui eseguire il deployment del sistema centrale.	5
1.6	WP 2.6: Deployment della banca dati	6
2	Obiettivi raggiunti.....	7

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA
Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management



1 Sintesi attività svolte

L'Obiettivo Realizzativo 2: *Progettazione e sviluppo di A@GRIBASE* è dedicato alla progettazione e realizzazione di A@GRIBASE e delle componenti che costituiscono la catena di acquisizione, validazione e retrieval dei dati.

L'OR1 si è articolato nei seguenti *Work Packages*:

WP 2.1: Progettazione e implementazione del Data Ingest

WP 2.2: Progettazione della struttura di A@GRIBASE e sviluppo delle componenti di popolamento della banca dati.

WP 2.3: Progettazione e implementazione delle componenti costituenti la catena di elaborazione e retrieval del dato

WP 2.4: Progettazione ed implementazione dei tools di gestione e monitoraggio della catena di acquisizione-elaborazione-providing del dato.

WP 2.5: Progettazione e realizzazione dell'infrastruttura IT su cui eseguire il deployment del sistema centrale.

WP 2.6: Deployment della banca dati

1.1 WP 2.1: Progettazione e implementazione del Data Ingest

L'attività svolta per questo work package ha riguardato la progettazione della banca dati e il suo interfacciamento.

L'interfacciamento con dati esterni avviene attraverso interfacce di tipo RESTful applicate sia per il collegamento diretto con database, sia per il data entry manuale attraverso file testuali strutturati. L'idea progettuale principale è stata quella di creare delle API esposte ad applicazioni esterne e utilizzabili dalle varie sorgenti per alimentare la banca dati.

Una delle funzionalità principali di A@GIBASE consiste nell'omogeneizzazione di dati provenienti da sorgenti eterogenee, questo ha comportato la trasformazione dei dati grezzi in input secondo uno schema di rappresentazione unico. Tale operazione, tuttavia, ha reso necessaria la creazione di un'opportuna area di staging in cui il dato grezzo è comunque conservato al fine di mantenere inalterate tutte le informazioni sulla proprietà e sull'origine del dato grezzo.

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA

Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management



1.2 WP 2.2: Progettazione della struttura di A@GRIBASE e sviluppo delle componenti di popolamento della banca dati.

Il disegno della struttura della banca dati è stato ispirato da due idee progettuali di fondo:

- generalità dei dati memorizzabili
- scalabilità

Ciò si è tradotto nella sperimentazione di tecnologie NoSQL in affiancamento a quelle SQL. Il loro utilizzo ha permesso di ottenere una banca dati dalle buone performance e capace, al contempo, di ospitare una molteplicità di dati.

Il progetto per la banca dati messo a punto ha reso necessario lo sviluppo di procedure ad hoc in grado di gestirne l'architettura. Sono state utilizzate largamente interfacce di tipo RESTful, al fine di creare un livello di astrazione che svincolasse dalla struttura della banca dati, e procedure java e php per le operazioni di elaborazione interne al singolo database che hanno riguardato la trasformazione dei dati grezzi a cui si accennava nel precedente paragrafo.

Grande importanza è stata dedicata al Quality Control. Infatti, l'elaborazione di dati grezzi provenienti da una sorgente esterna ha obbligato a dover gestire il QC in due modi diversi:

- a livello di sorgente dati si sono predisposte opportune procedure di ricalcolo automatico avviate dalla modifica del flag di controllo qualità impostata dalla sorgente. Infatti, tali valori possono essere modificati nel tempo;
- internamente alla banca dati si è definita una policy di controllo della qualità dell'elaborazione, che ha tenuto conto della serie storica nel suo complesso

1.3 WP 2.3: Progettazione e implementazione delle componenti costituenti la catena di elaborazione e retrieval del dato

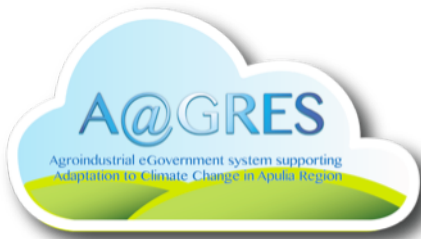
L'obiettivo di A@GRIBASE è la creazione di un catalogo di elaborazioni pre-calcolate e immediatamente disponibili per l'utente. Il catalogo è costituito sia da analisi statistiche sia da indici climatici. Nel dettaglio, le elaborazioni sono:

- Media
- Minimo
- Massimo
- Deviazione standard

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA

Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management



- Cumulata
- Conteggio
- Evapotraspirazione potenziale
- Indice di aridità per la caratterizzazione climatica in rapporto all'aridità
- Indice di termicità di Rivas Martinez per la caratterizzazione climatica in rapporto alla vegetazione.

Per l'elaborazione di queste analisi sono state realizzate specifiche procedure e, inoltre, applicativi dedicate alla gestione delle loro esecuzioni.

1.4 WP 2.4: Progettazione e implementazione dei tools di gestione e monitoraggio della catena di acquisizione-elaborazione-providing del dato

Le attività di questo WP ha riguardato la creazione di procedure ad hoc per l'inizializzazione e la gestione della banca dati e la creazione di interfacce in grado di fornire gli strumenti per "usare" A@GRIBASE. Questo si è tradotto nella creazione di applicativi progettati per rendere l'utente in grado di personalizzare la banca dati attraverso:

- l'aggiunta di aree monitorate definite dall'utente
- l'aggiunta di nuove variabili.

Tali operazioni comportano l'aggiunta di nuove elaborazioni al catalogo e quindi sono fortemente sinergiche con i sistemi di controllo e gestione delle elaborazioni descritte nel paragrafo precedente.

Dal punto di vista del providing del dato, l'interfacciamento con A@GRIFLOW avviene attraverso la tecnologia RESTful: questo per mantenere la modularità delle varie componenti e, dunque, la possibilità di un loro utilizzo in contesti diversi.

1.5 WP 2.5: Progettazione e realizzazione dell'infrastruttura IT sulla cui eseguire il deployment del sistema centrale.

Lo sviluppo del prototipo è stato eseguito presso il CED aziendale. Attraverso l'uso esteso della virtualizzazione è stato possibile creare l'ambiente, hardware e software, in cui sviluppare i

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA

Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management



prototipo al fine di poterlo replicare sull'infrastrutture hardware da installare presso il CED di Assocodipuglia.

1.6 WP 2.6: Deployment della banca dati

Il WP 2.6 ha segnato la conclusione delle attività previste per questo OR. Ha riguardato il deployment di A@GRIBASE, il suo popolamento e i test di funzionamento.

Sono state inoltre approntate opportune policy di sicurezza che riguardano non solo la creazioni di API keys che consentono solo agli utenti registrati l'interfacciamento della banca dati, ma la classificazione degli utenti al fine della gestione di dati sensibili.

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA
Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management



2 Obiettivi raggiunti

L'obiettivo principale del presente OR è stato la realizzazione del prototipo della banca dati, attualmente operativo, la creazione del catalogo delle procedure.

Di seguito alcuni dati indicativi sulla banca dati:

- tecnologie utilizzate: 2 (SQL, NoSQL)
- numero variabili totale 12
- numero sorgenti gestite 2
- numero aree monitorate: 370
- numero cadenze delle analisi: 5 (annuale, mensile, decadale, giornaliero, mese climatico di riferimento)
- Serie storica di riferimento: 1994-2014
- numero elaborazioni a catalogo: 92872
- numero dati grezzi memorizzati: circa 1 500 000
- numero di record afferenti le elaborazioni: circa 16 000 000.

A@GRES Agroindustrial e-Government system supporting adaptation to climate change in Apulia Region

Governo Elettronico per la PA

Identificazione di procedure e protocolli attraverso sistemi di Knowledge Management