

Programma di co-creazione del progetto Water Health Open knowledge - WHOW

2 Meeting – 15/11/2021 - Sintesi dei sottogruppi dei Casi d'uso



Introduzione

Il principale obiettivo del progetto [WHOW](https://whowproject.eu) è lo sviluppo del primo grafo della conoscenza europeo in grado di collegare i dati sul consumo e sulla qualità di acque superficiali, sotterranee e marine, con dati sanitari, provenienti da diverse fonti, europee e di singoli stati membri, rendendoli disponibili come dati aperti e collegati per un loro successivo riutilizzo per qualunque scopo.

Nel suo ambito è prevista l'attuazione di un programma di co-creazione che ha lo scopo di creare una comunità di attori istituzionali e non a livello nazionale e internazionale da coinvolgere nello sviluppo di un percorso condiviso nella definizione dei casi d'uso d'interesse e nella identificazione dei relativi *dataset*, così come nelle fasi di realizzazione dell'infrastruttura tecnica. Attualmente al programma di co-creazione risultano iscritti circa 80 partecipanti. Il programma di co-creazione prevede una serie di incontri tra i partner di progetto e gli stakeholder interessati.

Il secondo incontro¹ del programma di co-creazione ha avuto luogo il 15 novembre scorso e ha visto la partecipazione di Luca Lucentini - Dip.to Ambiente e Salute dell'Istituto Superiore Sanità, la cui presentazione si è focalizzata sull'inquadramento normativo e iniziative nazionali settore acque in particolare, sulla direttiva UE 2020/2184 del 16/12/2020 concernente la qualità delle acque destinate al consumo umano. E' stata evidenziata l'importanza delle due questioni della (ri)definizione delle zone di approvvigionamento idro-potabile e del campionamento funzionale alla prevenzione.

Il meeting ha permesso un confronto tra i partner di progetto e i partecipanti, suddivisi in sottogruppi, sui tre casi d'uso già precedentemente definiti:

Caso d'uso 1: Esposizione umana a inquinanti chimici e biologici,

Caso d'uso 2: Acque destinate al consumo umano;

Caso d'uso 3: Eventi estremi;

Un'ulteriore sessione è stata dedicata agli aspetti tecnici della infrastruttura che permetterà lo sviluppo del primo grafo della conoscenza per le finalità e gli obiettivi previsti dal progetto.

Il presente documento integra le sintesi di quanto discusso durante le attività dei gruppi di lavoro e pone l'accento sugli input emersi dai partecipanti.

I risultati del 2° co-creation meeting hanno contribuito alla modellazione dei casi d'uso e del Deliverable 2.1 – *Use Cases Definition* che integra tutti i contributi raccolti. Anche la sessione relativa all'architettura tecnica si è rivelata molto utile e tutti i temi emersi sono stati presi in considerazione nella revisione del Deliverable 4.1 *Design of the technical services for knowledge graph management*.

¹ <https://whowproject.eu/whow-2nd-co-creation-meeting/>

Caso d'uso 1: Esposizione umana a inquinanti chimici e biologici

Moderatrici: Roberta De Angelis, Patrizia Borrello Emanuela Spada ISPRA

Partecipanti:

Susanna Alloisio ETT - S.p.a

Anita De Cesaris - MiTE

Antonella D'Amore - Arpa Sardegna

Alessandra Casali - ISPRA

Stefano de Corso – ISPRA

Sintesi:

I lavori del sottogruppo sono stati aperti con un approfondimento del Caso d'uso 1 - Esposizione umana a inquinanti chimici e biologici - riguardo l'importanza dello studio delle tossine algali, prevalentemente marine e i loro effetti sulla salute umana. Le vie di esposizione principalmente trattate sono state l'ingestione dell'acqua attraverso la balneazione, ad esempio, e il contatto dermico. Sarebbe stato utile approfondire gli aspetti legati al consumo di prodotti della pesca contaminanti anche da eventuali agenti chimici e patogeni ma non erano presenti interlocutori del settore. L'occasione ha permesso un confronto con Susanna Alloisio di ETT Spa, che ha condotto studi di tossicità in vitro, che saranno disponibili a breve, per la valutazione degli effetti delle fioriture algali dannose (HABs) sulla salute umana (irritazione cutanea e oculare tramite test su tessuti umani) e ambientale (test di ecotossicità su naupli di *Artemia franciscana*). Stefano De Corso (ISPRA) ha ricordato di includere tra le banche dati previste per il caso d'uso anche quelle dei pesticidi e dei siti di acquacoltura disponibile in ISPRA.

Riguardo l'utilità del progetto WHOW e nello specifico del Caso d'uso 1, Anita De Cesaris, rappresentante del MiTE, ne ha evidenziato l'uso in termini di *governance*, sottolineando l'importanza di avere, a livello centrale, banche dati sulle acque tra loro interconnesse per una migliore conoscenza del territorio ad esempio nei casi di attuazione di direttive europee di settore o in risposta a procedimenti amministrativi quali le interrogazioni parlamentari e in risposta ai quesiti da parte dei cittadini. Susanna Alloisio (ETT Spa) ritiene la condivisione di banche dati utile ai fini di attività di ricerca perché permette di risparmiare tempo e un'alta qualità del dato. Antonella D'Amore (ARPAS) reputa utile il fatto di accedere velocemente ai dati per finalizzare progetti, monitoraggi, ecc. Su questo aspetto Alessandra Casali (ISPRA) ha ricordato l'utilità di rappresentare il dato anche ai fini della divulgazione scientifica verso un pubblico vasto anche a supporto di attività già in essere in ISPRA (vedi EcoAtl@nte) che utilizzano le stesse banche dati già citate dal collega Stefano De Corso. L'incontro si è chiuso con tutti i partecipanti concordi sulla completezza del Caso d'uso 1 in termini di argomenti e banche dati proposte, tuttavia sarebbe utile perseguire ulteriormente il coinvolgimento di esperti degli Istituti zooprofilattici, al fine di colmare il gap per i contaminanti dei prodotti ittici.

Caso d'uso 2: Acque destinate al consumo umano

Moderatori: Gianluca Carletti, Giovanni Peditto - ARIA

Partecipanti:

Paola Esposito – Celeris

Davide Del Monte– REACT

Girolamo Da Schio –libero professionista

Sintesi:

I lavori del sottogruppo sono stati aperti con la presentazione dei partecipanti e sono proseguiti con un approfondimento del Caso d'uso 2 - Acque destinate al consumo umano. Sono stati esposti i riferimenti normativi europei, nazionali e regionali che regolano sia il tema delle acque che quello della salute. Sono stati inoltre illustrati i *dataset* desiderati e quelli disponibili per l'implementazione del caso d'uso.

Dai partecipanti sono emersi suggerimenti su come migliorare ed estendere la fruibilità dei *dataset*, dal punto di vista dei data consumer di questo specifico caso d'uso. Uno dei partecipanti ha riportato l'esperienza della propria azienda, una start-up innovativa a vocazione sociale, che focalizza la propria attività di ricerca sui crimini economici, ambientali, contro la pubblica amministrazione e sul crimine organizzato. La start-up usa normalmente open data per valutare gli impatti della corruzione della PA sui servizi offerti ai cittadini. L'obiettivo della partecipazione è contribuire alla definizione del caso d'uso mettendo al tavolo l'esigenza di valutare gli impatti della corruzione della PA rispetto alla salute dei cittadini.

Un altro partecipante, *data scientist* iscritto a titolo personale, ha portato la propria esigenza di utilizzatore dei *linked open data*. L'esigenza è stata riportata sia sotto il profilo di cittadino interessato a valutare la qualità dell'acqua per consumo sia in qualità di *data scientist* con esperienza di comunità che fruiscono e mettono a disposizione dati aperti.

L'incontro ha contribuito a raccogliere i seguenti principali contributi:

- verso il cittadino, si suggerisce l'accesso all'informazione e al confronto della qualità delle acque per il consumo umano attraverso una rappresentazione semplificata. Viene riportato l'esempio della bandiera blu, come sintesi della valutazione di parametri complessi facilmente comprensibile e confrontabile dal comune cittadino. Lo stesso approccio può essere esteso ai prodotti alimentari realizzati con acque sottoposte al monitoraggio, dando facile visibilità della qualità delle acque utilizzate.
- Favorire il coinvolgimento, diffondere informazione e formare consapevolezza verso le comunità locali. Viene citato l'esempio del libro *la Terra di Sotto* (<https://www.laterradisotto.it/>) "Dare atto delle condizioni ambientali del territorio in cui viviamo e diffondere questa conoscenza, stimolando la ricerca, l'attivismo e lo studio di soluzioni riguardo un tema estremamente rilevante". Questo può essere considerato un suggerimento per le attività di disseminazione del progetto. Si suggerisce di coinvolgere attivamente il territorio ad esempio attraverso eventi di tactical urbanism.
- Estendere la fruibilità dei *dataset* alle comunità di *data scientist* e alle piattaforme che raccolgono dati liberi in questi contesti e organizzano eventi e competizioni, tra cui a titolo di esempio drivendata.org e kaggle.com

Caso d'uso 3: Eventi estremi

Moderatore: Marco Picone – ISPRA

Partecipanti:

Giulio Settanta – ISPRA
Carmen Ciciriello – Celeris
Daniela Iasillo – Planetek
Massimo Zotti – Planetek
Alessandra Gelmini – Regione Lombardia
Franco Amigoni – Formez OECD
Francesco Misurale – ETT spa
Enzo Lo Piccolo – Regione Siciliana
Giacomo Scala – Regione Siciliana
Ana Pasian – ESG

Sintesi:

I lavori del sottogruppo sono stati aperti con un approfondimento del Caso d'uso 3 – Eventi estremi -. In particolare sono stati esposti riferimenti normativi e la reportistica nazionale e internazionale che hanno ispirato la formulazione del caso d'uso. La presenza di partecipanti esperti in tematiche eterogenee rappresentanti di enti e aziende di diverse parti del territorio nazionale ha permesso di condividere esperienze relative alla individuazione dei rischi derivanti da eventi estremi che vanno dall'utilizzo di dati satellitari, alla compromissione del reticolo idrografico principale e secondario, all'importanza della redazione dei contratti di fiume, agli aspetti economici legati alla perdita di produzione, attività lavorative, aree.

Gran parte della discussione si è spostata sulla tipologia di dataset da utilizzare per la trattazione completa del caso d'uso. È emersa la necessità di allargare la visuale, accostando ai necessari dataset idrologici, mareografici e sanitari, anche dati relativi a componenti non propriamente rientranti nelle aree "Acqua" e "salute" ma che hanno effetti rilevanti su entrambi gli ambiti. Per questo motivo, oltre ai dataset proposti inizialmente (dataset idrologici, meteorologici, stato fisico del mare, stato chimico dei corsi d'acqua, inquinamento di breve durata, uso del suolo, concentrazione di pesticidi nelle acque, acque di balneazione), e alla necessità di recuperare dati sicuramente estremamente utili, ma di difficile raggiungimento (qualità e quantità delle acque negli impianti di gestione idrica, accessi alle cure mediche per determinate patologie o traumi), sono stati proposti altre informazioni utili a valutare gli impatti degli eventi estremi sul comparto idrico e la salute umana. Si tratta in particolare di:

- Grado di antropizzazione del territorio;
- Censimento di aree vulnerabili, quali discariche, aree industriali dismesse o comunque degradate (cartiere, ex industrie tessili, cave), siti di interesse nazionale. A seguito di eventi rilevanti, tali aree possono essere fonte di maggiore carico inquinante nei corsi d'acqua;
- La presenza di aspetti di pianificazione dei territori a tutela della risorsa idrica (contratti di fiume);
- Uso del suolo costiero (in particolare da piattaforma copernicus) e linea di costa aggiornata;
- Reticolo idrografico primario e secondario, spesso impattato da eventi estremi;
- Impatto economico degli eventi estremi in termini di indennizzi, perdita di produzione e di ore lavorate, perdita di aree, infrastrutture e mezzi, inaccessibilità di siti, perdita di forza lavoro, ecc. Tali impatti possono riguardare anche la scarsità di risorsa idrica, creando ripercussioni nella generazione di energia e nell'utilizzo di acqua nei processi industriali;
- Quantità di fondi necessari per il ripristino ambientale e tipologia di interventi;
- Salvaguardia delle sorgenti;
- Attivazione degli sfioratori dei depuratori.

Sono emerse quindi le seguenti esigenze:

- Completare il caso d'uso tenendo conto anche dell'antropizzazione del territorio, degli aspetti economici e della gestione del territorio
- Attuare un maggiore coinvolgimento delle strutture che detengono informazioni rilevanti.

Sessione: Architettura tecnica

Moderatori: Giorgia Lodi, Anna Sofia Lippolis, Andrea Giovanni Nuzzolese - ISTC-CNR

Partecipanti:

Andrea Borruso - Open Data Sicilia, OnData
Matteo Busanelli - Imola Informatica
Stefano Carboni - Freelancer in the health sector
Marco Casaburi - Planetek Italia srl
Paolo Francescangeli – Istat
Elio Giulianelli – ISPRA
Angelo Gulina – Gitlab
Marco Panebianco - ARIA SpA
Azzurra Pantella - Umbria Digitale
Giovanni Perrotta - Open Data Sicilia
Alessandro Russo - BUP srl
Giulio Settanta- ISPRA

Sintesi:

La sessione del co-creation meeting sull'architettura WHOW è stata divisa in due parti: una breve presentazione dell'attuale deliverable per la milestone #7² e una discussione tra i partecipanti basata su una Jamboard di Google³ al fine di visualizzare e condividere i suggerimenti proposti.

La struttura dell'incontro si è rivelata molto utile e tutti i temi emersi sono stati presi in considerazione nella revisione del deliverable, la cui versione finale, disponibile nel [repository Github di WHOW](#), è stata giudicata eccellente. Grazie alla produttività dello scambio durante l'evento, tutti i commenti dei partecipanti sono stati debitamente riportati all'interno del documento, e i loro nomi elencati come collaboratori del programma di co-creazione. Inoltre, a causa dei limiti di tempo per la discussione, sono stati proposti incontri più frequenti e concentrati unicamente sull'approfondimento dell'architettura di WHOW.

Si sono trattati diversi argomenti che possono essere raggruppati in tre punti principali:

Architettura decentralizzata

Durante la riunione è stato notato come l'esplicazione dell'architettura decentralizzata all'interno del deliverable poteva essere ulteriormente chiarita dal punto di vista dei data provider. Nonostante lo stile architetturale di decentralizzazione riguardi i nodi di fornitura dei dati all'interno dell'ecosistema WHOW al fine di far durare l'infrastruttura oltre il termine del progetto, è stato anche utile capire come evidenziare il ruolo di tali nodi nello scenario distribuito delineato, e per ogni nodo come specificare quali componenti fossero da considerare requisiti di base e quali fossero solo da intendere come raccomandazione. Pertanto, nel documento è stata aggiunta una sezione specifica, la Sezione 5.2, che descrive le implicazioni dell'adozione di uno stile architetturale distribuito e component-based.

² The link to the final version of the deliverable can be found at this link: https://github.com/whow-project/deliverables/blob/main/ArchitectureDeliverable_final.pdf. As it is possible to see in the document, a first revision of the document occurred right away after the co-creation meeting.

³ https://jamboard.google.com/d/1PI1G0nGKkhZCg5wbE5Gtx3kw3mjspqNVPINnu_UGWN4/viewer?f=0.

Validazione dei dati

Come gestire il tema di uno schema di dati minimo nel momento in cui essi vengono forniti alla piattaforma? L'architettura WHOW presuppone che i dati ricevuti seguano uno specifico schema concettuale astratto, ma la validazione può essere eseguita anche una volta che i dati sono stati prodotti in Linked Open Data.

Anonimizzazione

Nel deliverable, l'anonimizzazione è stata menzionata più volte per chiarire il suo ruolo all'interno dell'architettura WHOW quando i dati vi affluiscono. Durante la riunione è stato notato che è molto probabile che i dati che arrivano in WHOW siano stati resi anonimi dai data provider dall'inizio del data management lifecycle, prima di qualsiasi tipo di pubblicazione. Pertanto, è stato suggerito di riformulare il requisito di anonimizzazione concentrandosi sul concetto di "aggregazione", che è già di per sé una forma di anonimizzazione, e che potrebbe essere probabilmente garantito anche quando si pubblicano i dati nel grafo della conoscenza di WHOW. L'aggregazione, nel caso sia richiesta, può infatti essere gestita dal componente di Data Transformation nel Linked Open Data Preparation layer descritto all'interno del deliverable.

Durante l'incontro, è stato incoraggiato il riutilizzo di soluzioni software già esistenti. Ne sono emerse due in particolare:

1. Frictionless Data⁴ per gli utenti che dimostrano un profilo tecnico ma non appartengono al mondo semantico:
2. Il Comsode Open Data Node Project⁵ della Regione Umbria, che la Regione desidera utilizzare contribuendo al grafo di conoscenza di WHOW.

Infine, è emerso un ulteriore suggerimento sulla fornitura di metadati da schema.org quando si pubblica il grafo di conoscenza WHOW nei cataloghi di dati dei data provider, al fine di massimizzare la reperibilità dei dataset (il primo principio nel paradigma FAIR).

⁴ <https://frictionlessdata.io/>.

⁵ <https://opendatanode.org/>.