

Modello Dati e API - Osservazioni COVID 19

Proposta progetto pilota a cura del Gruppo di lavoro 2 - Data collection and Infrastructure

13 luglio 2020

Indice

Introduzione	3
Obiettivi	3
Dati, Governance e Standardizzazione	4
DATASET COVID-19	4
Disponibilità del DATO e indicazioni nazionali di riferimento	4
Flusso informatico a livello nazionale - NOTIFICA DELLE MALATTIE INFETTIVE	4
Tracciato dati aggregati COVID-19 del Dipartimento della Protezione Civile	6
Criteri relativi alle attività di monitoraggio del rischio sanitario	9
Considerazioni su un possibile modello comune dei dati per gli indicatori	9
Definizione dell'ontologia degli indicatori	11
Elementi di calcolo degli indicatori	13
Elementi amministrativi e/o territoriali	14
Elementi raw (osservazioni e proprietà)	14
Open Data COVID-19 - Iniziative regionali	16
Buone pratiche dal territorio: il caso della Regione Umbria	17
Le altre iniziative regionali	22
Regione Abruzzo	22
Regione Basilicata	22
Provincia di Bolzano	22
Regione Calabria	22
Regione Campania	22
Regione Emilia Romagna	22
Regione Friuli Venezia Giulia	23
Regione Lazio	23
Regione Liguria	23
Regione Lombardia	23
Regione Marche	23
Regione Molise	23
Regione Piemonte	23

Regione Puglia	24
Regione Sardegna	24
Regione Sicilia	24
Regione Toscana	24
Provincia Autonoma di Trento	24
Regione Valle d'Aosta	25
Regione Veneto	25
Specifiche di comunicazione - API	25
Panoramica dello schema	26
Raccordo con le raccomandazioni del GDL 7	27
Gruppo di lavoro	28
Riferimenti bibliografici	29
ALLEGATO I - Specifiche di comunicazione API - Casi d'uso ed esempi	30
Osservazioni quotidiane regionali	30
Osservazioni in un intervallo mensile	32
Osservazioni in una Unità Amministrativa e nelle sottounità	33

Introduzione

Nell'ambito delle attività della Task Force Data Driven COVID19, il gruppo di lavoro 2 "Data collection and infrastructure" ha deciso di realizzare un progetto pilota per la definizione di un modello dati e specifica API per la condivisione e pubblicazione degli indicatori, e delle relative osservazioni che concorrono al loro calcolo, presenti nel [decreto emanato in data 30 aprile al fine di monitorare la Fase 2](#). Il progetto vuole essere una risposta concreta alle attività che il gruppo di lavoro 2 è stato chiamato a compiere, nel contribuire alla diffusione di buone pratiche per favorire interoperabilità tra i vari flussi di dati sanitari e amministrativi e la creazione di un ecosistema di scambio virtuoso.

Durante il progetto di Interoperabilità indicato nel Piano Triennale per l'Informatica si è stabilito un rapporto di collaborazione costante con una serie di Regioni e con il - [Centro Interregionale per i Sistemi Informatici Geografici e Statistici](#) (CISIS)

Il CISIS gestisce il progetto "Interoperabilità e cooperazione applicativa delle regioni - Azione di transizione ai nuovi standard" a cui attualmente partecipano 11 tra Regioni e Province Autonome.

Già prima dell'emergenza era emerso l'interesse comune di contribuire alla definizione di una API standard tra i partecipanti alle attività del CISIS in materia sanitaria, per standardizzare la fruizione dei dati associati e permettere il riuso non solo dei dati ma anche di prodotti associati quali mobile app o dashboard. Inoltre, la cooperazione tra erogatori per definire API comuni è anche indicata nelle linee d'azione del Piano Triennale in redazione 2020-2023.

Durante l'emergenza COVID-19, sulla spinta del lavoro fatto per le dashboard della Protezione Civile e da alcune regioni come ad esempio [l'Umbria](#), [Toscana](#) e [Provincia Autonoma di Trento](#) (si veda il dettaglio presentato in sezione [Open Data 19 - Iniziative regionali](#)), è emersa la disponibilità da parte del CISIS e dei rappresentanti delle regioni che collaborano anche all'interno della task force data-driven, a sperimentare un progetto pilota per la a e standardizzazione di un'API utile a fronteggiare l'emergenza covid.

L'obiettivo è quindi individuare un modello dati e specifica per l'interfaccia API, che tutti gli erogatori possano implementare per pubblicare i dati utili alla gestione dell'emergenza.

Infine, nel contesto dei lavori portati avanti dai vari sottogruppi di lavoro della Task Force, il progetto si integra con le raccomandazioni contenute nel documento di Follow up prodotto dal sottogruppo di lavoro 7 pubblicato al [seguito link](#), precisamente alla sezione 3 "Indicazioni tecniche per l'implementazione del flusso descritto nel decreto del Ministero della Salute del 30/4/202015"

Obiettivi

I dati principali relativi al sistema di Sorveglianza integrata COVID-19 in Italia (ordinanza del n. 640 del 27 febbraio 2020) sono realizzati dall'Istituto Superiore di Sanità (ISS) integrando i dati microbiologici ed epidemiologici forniti dalle Regioni/PPAA e dal Laboratorio nazionale di riferimento per SARS-CoV-2 dell'ISS. Oltre all'ISS sono coinvolti nella raccolta dei dati anche la Protezione Civile ed il Ministero della Salute.

Adottare un modello dati comune e API standardizzate per esporre i dati delle varie fonti collegate all'emergenza COVID-19, da utilizzare al di sopra dei sistemi informativi di produzione del dato, favorisce l'interoperabilità, armonizza ad esempio, la creazione di dashboard e rapporti

automatizzati e contribuisce a costruire una visione unitaria dell'evolversi del fenomeno pandemico sul territorio, individuando eventuali criticità.

Dati, Governance e Standardizzazione

Il sistema emergenziale deve identificare i fenomeni da osservare, l'insieme di dati da utilizzare, il livello di dettaglio di cui si ha bisogno e quali sono le relazioni utili per prendere decisioni (per questo motivo, nel seguito è riportato il tracciato COVID-19 pubblicato dal dipartimento della protezione civile e il tracciato pubblicato dalla Regione Umbria, che hanno ispirato il lavoro proposto nel presente documento).

Per superare eventuali vincoli sull'utilizzo dei dati è necessario un accordo tra tutte le parti interessate in modo da creare e utilizzare API standardizzate per tale sistema. Attualmente il Decreto del Ministro della salute non dettaglia le specifiche di una API per esportare questi dati, l'iniziativa non si sovrappone quindi ad altre attività in corso.

Costruendo le API secondo gli standard concordati (vedi [LG ModI-AglD](#)) gli attori assicurano:

- una governance del servizio adeguata alle necessità ed alla tipologia dei dati - che comunque non includono dati personali - compreso il monitoraggio dei rischi per la sicurezza e gli eventuali requisiti di autenticazione e autorizzazione;
- una evoluzione non distruttiva del servizio ma capace di adattare i dataset all'evoluzione delle esigenze;
- il rispetto dei contratti tecnologici dichiarati nell'interfaccia e degli ulteriori accordi utili all'erogazione e alla fruizione dei dati.

Sarà inoltre necessario supportare i vari stakeholder nell'uso delle API e nell'integrazione con i propri dati o servizi nel sistema.

Per velocizzare l'evoluzione e coinvolgere le comunità di interessati, il progetto sin dalle prime fasi pubblica le specifiche tramite una piattaforma di code-hosting e può far leva sui servizi messi a disposizione da Developers Italia (slack, github, forum).

DATASET COVID-19

La sorveglianza comprende tutti i casi di COVID-19 diagnosticati dai laboratori di riferimento regionali, come indicato dalla Circolare del Ministero della Salute n. 0005443 del 22 febbraio 2020. I dati individuali, aventi un livello di dettaglio maggiore rispetto a quello previsto in altri flussi esistenti, al momento, vengono aggiornati da ciascuna Regione/PPAA con cadenza giornaliera.

Disponibilità del DATO e indicazioni nazionali di riferimento

Nelle prime fasi del presente (progetto) pilota è necessario accettare che i dati talvolta possono non essere immediatamente disponibili e/o mancare nei vari livelli di aggregazione territoriale (regione, provincia, comune, ...); questo processo di pubblicazione aiuta a: i) far emergere eventuali difficoltà, ii) individuare punti di intervento e gli ambiti di ulteriori investimenti e iii) promuovere una governance più allargata.

Flusso informatico a livello nazionale - NOTIFICA DELLE MALATTIE INFETTIVE

Il Servizio definito all'interno di ogni Regione (Sistema Sanitario Regionale) è impegnato nel debito informativo quotidiano sui dati della sorveglianza COVID-19 al Ministero della Salute, alla

Protezione Civile ed all'Istituto Superiore di Sanità. Nel documento già pubblicato su "[Analisi dei flussi e mappatura delle banche dati di interesse per la task force dati per l'emergenza COVID-19](#)" sono stati descritti i flussi dati COVID-19 provenienti dal livello locale durante le prime fasi dell'emergenza. In quei flussi, il dato è stato raccolto attraverso fogli elettronici. Dal 25 giugno 2020, il flusso si è modificato prevedendo due direttrici:

- il flusso dei dati aggregati inviati giornalmente dalle Regioni. Questo flusso è coordinato dal Ministero della Salute non solo con il supporto della Protezione Civile, come accadeva in precedenza, ma anche con il supporto dell'Istituto Superiore della Sanità. Il flusso consente di raccogliere informazioni sul numero totale di test positivi, decessi, ricoveri in ospedale e ricoveri in terapia intensiva in ogni Provincia d'Italia.
- il flusso dei dati individuali inviati giornalmente dalle Regioni all'Istituto Superiore di Sanità (Sorveglianza integrata Covid-19, secondo quanto indicato nell'ordinanza 640 della Protezione Civile del 27/2/2020). In questo flusso sono compresi anche i dati demografici, le comorbidità (in ambito sanitario il termine comorbidità indica la coesistenza di più patologie diverse in uno stesso individuo), lo stato clinico e la sua evoluzione nel tempo, per un'analisi più accurata.

Dal 25 giugno 2020 la scheda con l'aggiornamento quotidiano dei dati (pubblicata sotto forma di tabella PDF sul sito del Ministero della Salute) è stata integrata con i "casi identificati dal sospetto diagnostico" (casi positivi al tampone emersi da attività clinica) e "casi identificati da attività di screening" (indagini e test, pianificati a livello nazionale o regionale, che diagnosticano casi positivi al tampone).

Da giugno 2020 quindi, il dato non è più raccolto attraverso fogli elettronici, ma compilato tramite un servizio di raccolta dati messo a disposizione dell'Istituto Superiore di Sanità che, a seguito di approvazione dei dati da parte del Ministero della Salute, invia le informazioni al Dipartimento della Protezione Civile. Quest'ultima processa automaticamente le informazioni, eroga i dati aperti e popola la dashboard ufficiale nazionale. Ad oggi, a parte il popolamento dei dati e l'approvazione, tutto il processo è automatizzato.

La Figura 1 rappresenta i nuovi flussi sopra descritti.

ricoverati_con_sintomi	Ricoverati con sintomi	Numero	3
terapia_intensiva	Ricoverati in terapia intensiva	Numero	3
totale_ospedalizzati	Totale ospedalizzati	Numero	3
isolamento_domiciliare	Persone in isolamento domiciliare	Numero	3
totale_positivi	Totale attualmente positivi (ospedalizzati + isolamento domiciliare)	Numero	3
variazione_totale_positivi	Variazione del totale positivi (totale_positivi giorno corrente - totale_positivi giorno precedente)	Numero	3
nuovi_positivi	Nuovi attualmente positivi (totale_casi giorno corrente - totale_casi giorno precedente)	Numero	3
dimessi_guariti	Persone dimesse guarite	Numero	3
deceduti	Persone decedute	Numero	3
casi_da_sospetto_diagnostico	Casi positivi al tampone emersi da attività clinica	Numero	3

casi_da_screening	Casi positivi emersi da indagini e test, pianificati a livello nazionale o regionale	Numero	3
totale_casi	Totale casi positivi	Numero	3
tamponi	Totale tamponi	Numero	3
casi_testati	Totale dei soggetti sottoposti al test	Numero	3
note	Note sui dati forniti	Testo	Lorem ipsum...

I dati vengono erogati nei formati csv, json, shp e geojson oltre che tramite API e GraphQL

In questo dataset l'area geografica è rappresentata dalla Regione, a livello provinciale vengono raccolti soltanto i casi totali, ma allo stesso modello è possibile indicare, utilizzando i codici istat anche Province e Comuni. Nei CSV e Json nei campi lat (latitudine) e long (longitudine) vengono riportate le coordinate del capoluogo di Regione, analogamente, negli shape e nei geojson sono riportate le aree geografiche del territorio.



Criteri relativi alle attività di monitoraggio del rischio sanitario

Principi del Decreto del 30 Aprile 2020

Ai fini del monitoraggio e della necessità di classificare tempestivamente il livello di rischio, con l'obiettivo di valutare la necessità di modulazioni nelle attività di risposta all'epidemia, sono stati disegnati alcuni indicatori con valori di soglia e di allerta che dovranno essere monitorati, attraverso sistemi di sorveglianza coordinati a livello nazionale, al fine di ottenere dati aggregati nazionali, regionali e locali. Tali indicatori non sono finalizzati ad una valutazione di efficienza/efficacia dei servizi ma ad una raccolta del dato e ad una migliore comprensione della qualità dello stesso, al fine di realizzare nel modo più corretto possibile una classificazione rapida del rischio di concerto con l'Istituto Superiore di Sanità e le Regioni/PPAA. Alcuni indicatori, definiti opzionali, sono relativi a flussi di sorveglianza non attualmente attivi che potranno essere istituiti in alcune Regioni/PPAA in base alla fattibilità ed opportunità. Tali indicatori verranno considerati nella classificazione del rischio solo qualora la Regione/PPAA raccolga il dato a seguito dell'attivazione del relativo flusso informativo.

Il monitoraggio comprenderà i seguenti indicatori:

- indicatori di processo sulla capacità di monitoraggio;
- indicatori di processo sulla capacità di accertamento diagnostico, indagine e gestione dei contatti;
- indicatori di risultato relativi a stabilità di trasmissione e alla tenuta dei servizi sanitari.

Considerazioni su un possibile modello comune dei dati per gli indicatori

Il modello, presentato di seguito si caratterizza per:

- adattabilità: essere un modello generico che si adatta a diversi tipi di indicatori e non necessariamente a quelli COVID-19;

- standardizzazione: essere usato nella serializzazione secondo lo standard OpenAPI (si veda sezione [Specifiche di comunicazione - API](#)) ma anche per essere serializzato in un formato tabellare di riferimento, eventualmente da predisporre;
- modellazione: aver utilizzato pattern di modellazione inseriti nell'ambito di alcuni standard di riferimento ([INSPIRE](#) e [dominio sanitario](#))
- modularità: si presta all'utilizzo sia da parte degli organi locali (regioni) sia centrali (enti centrali competenti nella raccolta dati e monitoraggio) dell'intero processo di gestione dei dati sugli indicatori.

Definizione dell'ontologia degli indicatori

La definizione della [prima bozza del modello degli indicatori](#), che ha fornito la base concettuale per definire la [specificazione di comunicazione API](#), è iniziata analizzando lo stato dell'arte. In particolare si sono cercati:

1. standard disponibili in letteratura (e.g., standard INSPIRE Observation & Measure, HL7/FHIR, paper scientifici, sotto riportati, sul tema dei KPIs, degli indicatori di risultato nel contesto scientifico e di qualità del software) e altre proposte di standard per il caso specifico COVID quali [CovidJSON](#);
2. si sono studiate iniziative analoghe a livello europeo per la definizione degli indicatori e per la loro rappresentazione concettuale.

A fronte di alcuni confronti sul tema con interlocutori della Commissione Europea (uno nel contesto SEMIC (SEMantic Interoperability Community) e uno nel contesto Open Science) è emerso che non esiste ancora un modello comunitario di rappresentazione degli indicatori che possa considerare anche indicatori per il monitoraggio del COVID-19. Tuttavia esistono lavori avviati come <https://www.covid19dataportal.org/> per facilitare la condivisione dei dati e l'analisi su aspetti di ricerca scientifica inerenti il coronavirus, accelerando così i risultati. Tale piattaforma consiste di:

- un portale di dati;
- un data hub, con nodi (per ora è in atto una sperimentazione con un nodo spagnolo) che pubblicano metadati quali gli identificatori di monitoraggio del campionamento, il tempo di campionamento, la posizione geografica, il metodo di campionamento, che vengono successivamente raccolti attraverso meccanismi automatici (harvesting).

Inoltre, per integrarsi nel contesto nazionale sul tema modellazione e dati in ambito statistico, dare più ampio respiro al modello e non limitarlo al contesto COVID-19, è stata avviata una collaborazione con ISTAT da cui sono nate alcune riflessioni in parte recepite nel modello descritto.

Sulla base di questo scenario si è elaborato un modello generale (ontologia) per Osservazioni (già presente nella rete nazionale di ontologie e vocabolari controllati dalla pubblica amministrazione richiamata dal piano triennale ICT attualmente in vigore) e Indicatori applicabile anche ad altri contesti, oltre COVID-19.

Il modello è rappresentato nello schema grafico ontologico seguente:

- gli elementi in giallo sono Classi/concetti;
- le frecce blu scure che collegano i box gialli sono le relazioni tra le classi;
- gli elementi verde acqua sono valori (stringhe o tipi primitivi come date);
- le frecce verde acqua collegano le classi ai valori;
- le frecce nere sono predicati (subClassOf → relazione isA).

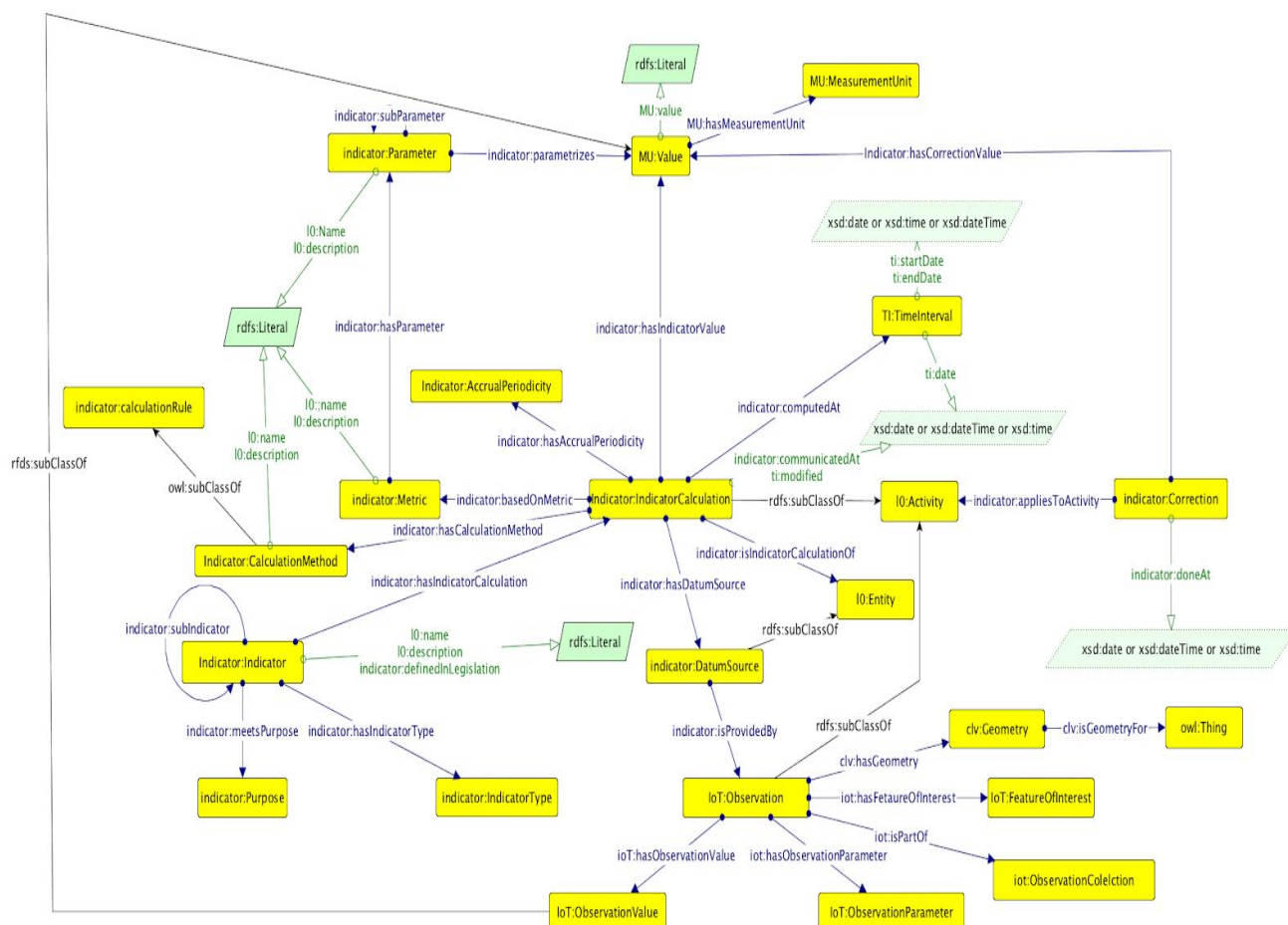


Figura 1: Ontologia degli indicatori

Nel modello si possono distinguere tre moduli principali:

1. modulo degli **Indicatori**: modulo non disponibile a livello di ontologie nazionali. Il modulo è stato pensato per essere il più generale possibile e non vincolato al contesto degli indicatori COVID-19, che rappresentano invece una specializzazione del modello in Figura 1. Concetti rappresentati in questa parte del modello quali per esempio Purpose derivano dal modello [Measure](#) dello standard in ambito sanitario HL7/FHIR; il modulo si caratterizza dalla distinzione tra indicatore e calcolo dell'indicatore. Questo consente per esempio di definire vocabolari controllati sui diversi possibili indicatori separando tale definizione dal calcolo dell'indicatore (dall'osservazione dell'indicatore stesso) che ha un valore ed è associato a una sorgente dati).
2. modulo dei **Valori e Misure** già modellato a livello nazionale attraverso la relativa ontologia il cui prefisso è MU. Questo modulo rappresenta valori (di qualunque tipo) e unità di misura.
3. modulo delle **Osservazioni** già modellato a livello nazionale attraverso l'ontologia della sensoristica IoT. Quest'ultimo modello segue due standard disponibili a livello internazionale ed europeo: lo standard [INSPIRE Observation & Measurements](#) e lo standard Web della [Semantic Sensor Network Ontology](#)¹. Inoltre, nel modello HL7/FHIR

¹ Si noti che il modulo della sensoristica rappresenta elementi che non sono completamente pertinenti per gli scopi dell'ontologia degli indicatori. Infatti, di tutto quel modello, la parte relativa alle reti e piattaforme di sensori e

specifico del mondo sanitario le [osservazioni](#) sono concetti che si ritrovano nel modulo della [diagnostica](#).

Il modulo sugli indicatori è stato predisposto tenendo conto del decreto del ministero della salute del 30 aprile 2020 e visionando una serie di ontologie analoghe già presenti allo stato dell'arte, anche scientifico [1, 2, 3, 4, 5, 6, 7].

Il razionale dietro al modello può essere così riassunto: gli indicatori sono calcolati in base a diverse metriche elementari. Il calcolo ha un valore ottenuto sulla base di un certo metodo/processo. Il processo che sottende il calcolo presuppone la raccolta di una serie di osservazioni su determinate caratteristiche di interesse e proprietà che complessivamente alimentano una sorgente dati usata per il calcolo degli indicatori. Analizzando le tabelle degli indicatori del decreto del Ministero della Salute del 30 aprile 2020 emerge come gli stessi siano calcolati sulla base di una certa fonte dati (nel modello in figura 1 questo concetto è chiamato DatumSource, che può essere il sistema di sorveglianza integrata COVID-19 dell'ISS, rilevazioni dirette delle regioni) Tali fonti dati si pensa siano diverse, alcune già formate nella fase 1 sulla base di flussi già attivi verso il Ministero della Salute/Protezione Civile o verso l'ISS e altre da gestire mediante altri flussi da definire e alimentare.

Elementi di calcolo degli indicatori

Il modello definisce due importanti concetti: l'indicatore e il calcolo dell'indicatore.

Gli **indicatori** possono essere indicatori composti (super indicatori) e semplici. In Figura 1 questo è catturato con la proprietà subIndicator (e la sua inversa super indicator) che consente di modellare gerarchie di indicatori. Gli indicatori (e.g., il rapporto tra il numero di casi sintomatici notificati per mese con una data di inizio sintomi sul totale) sono calcolati (calcolo dell'indicatore) sulla base di una o più metriche. Le metriche sono misure più atomiche rispetto agli indicatori che ne consentono il calcolo; un esempio di metrica rispetto all'indicatore sopra riportato è il numero di casi sintomatici con data inizio sintomi valorizzata.

La metrica può essere associata a uno o più parametri o caratteristiche (o unità di osservazione del calcolo) dell'indicatore) che possono parametrizzare un valore. La relazione tra il parametro e il valore consente di modellare casi in cui i parametri sono definiti nel contesto di uno specifico range o regione in uno spazio dimensionale (e.g., un parametro che indica la qualità di qualcosa che può essere definita attraverso un range che va da [molto buona, buona, sufficiente, pessima]).

L'indicatore ha uno scopo/goal (e.g., monitorare la stabilità di trasmissione e la tenuta dei servizi sanitari), ha un calcolo dello stesso associato che a sua volta ha un valore che può avere un'unità di misura.

Il calcolo (o l'osservazione) dell'indicatore è comunicato in una certa data, è rilevato in un certo intervallo temporale (che può essere un istante o un vero e proprio intervallo).

L'indicatore è calcolato in base a un metodo di calcolo (tipicamente matematico - elemento opzionale), può essere di diversi tipi (ad esempio, di processo, di risultato e di struttura), ha una frequenza di aggiornamento (ad esempio gli indicatori 1.1 -1.4 del decreto Ministero salute

attuatori non è utilizzata. L'uso che si fa in questo ambito è relativo al solo pattern Osservazioni / Collezione di osservazioni / Parametri delle osservazioni / Valori delle osservazioni / Caratteristiche di interesse, particolarmente utile e flessibile in un contesto dove i parametri delle osservazioni che concorrono a costruire le basi di dati su cui calcolare gli indicatori, potrebbero variare anche nel breve periodo e rispetto ai diversi livelli amministrativi, per rispondere più efficacemente al monitoraggio di determinati elementi di un fenomeno.

hanno una frequenza di aggiornamento mensile)), ha un nome, una descrizione e uno o più riferimenti normativi (non essendo ancora in grado di modellare le norme, per ora i possibili riferimenti normativi sono semplicemente stringhe che riportano a un decreto, altra legge o uno specifico articolo e/o comma in una legge).

Elementi amministrativi e/o territoriali

Il calcolo dell'indicatore è associato a una generica entità (IO:Entity nella figura): La relazione tra il calcolo dell'indicatore e l'entità è utilizzata per poter associare la misura dell'indicatore a elementi amministrativi (è il calcolo dell'indicatore per una certa pubblica amministrazione) e/o territoriali (è il calcolo dell'indicatore per un certo territorio). Entità infatti è una classe generica al di sotto della quale si ritrovano sia agenti come organizzazioni, ma anche location e unità amministrative territoriali.

Elementi raw (osservazioni e proprietà)

Il calcolo dell'indicatore avviene utilizzando tipicamente una sorgente dati (nelle tabelle del decreto 30 aprile del Ministero della Salute, viene infatti indicata una fonte dati per ciascun indicatore). La sorgente dati è formata mediante la raccolta di dati raw (micro e/o macro), ossia di singole osservazioni su caratteristiche di interesse e specifici parametri. Questa parte del modello, già esistente a livello nazionale, richiama sia CovidJson prima menzionato, sia lo standard INSPIRE sulle osservazioni e misure, ma include anche le osservazioni del modello HL7/FHIR. In sostanza, singole osservazioni alimentano una base di dati usata poi per il calcolo degli indicatori.

Una rappresentazione semplificata UML dell'ontologia proposta è fornita nelle seguenti Figure 2 e 3 dove si è voluto separare la parte di modellazione degli indicatori, potenzialmente utile nel caso dei flussi COVID-19 per gli organismi centrali coinvolti nel processo, dalla parte di modellazione delle osservazioni che può essere per esempio, nel caso COVID-19, adottata da un livello locale per l'invio dei dati richiesti per l'alimentazione di basi dati centrali.

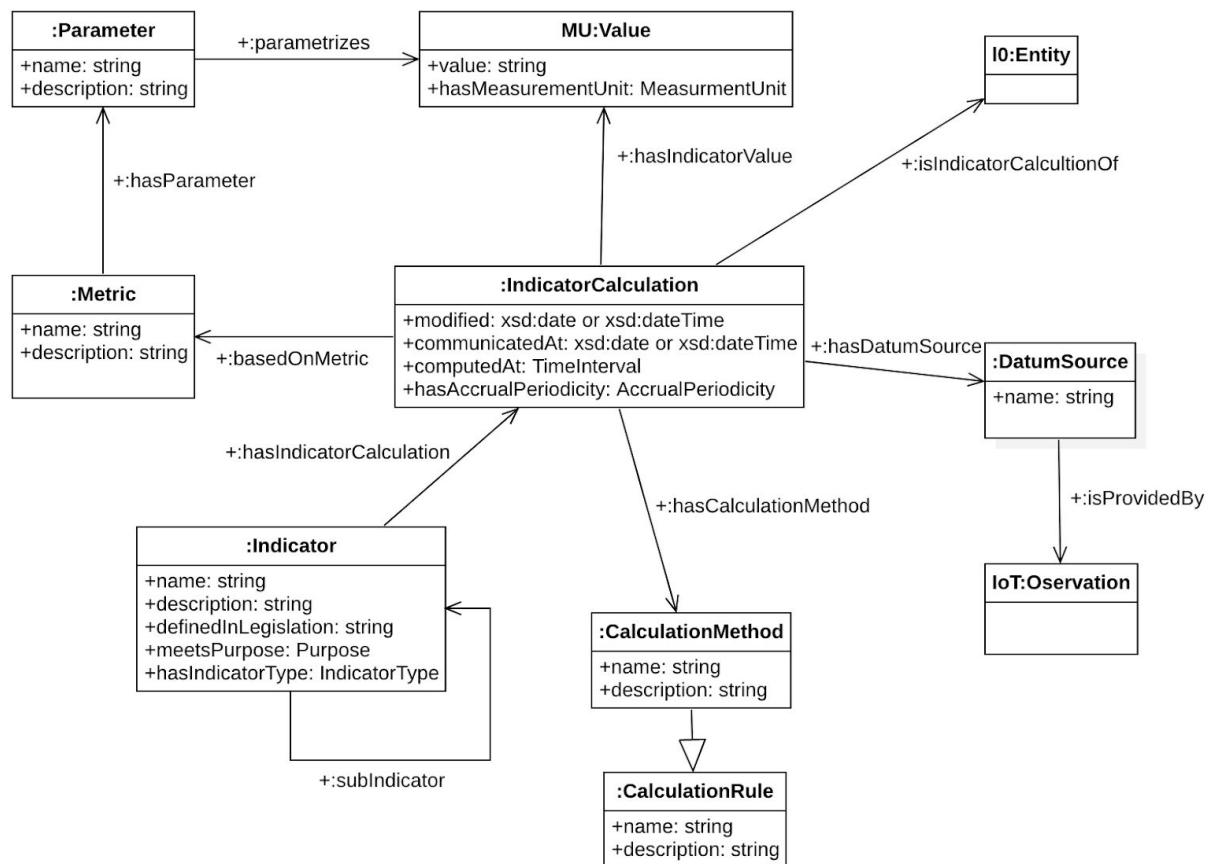


Figura 2: Rappresentazione UML della parte del modello sugli indicatori

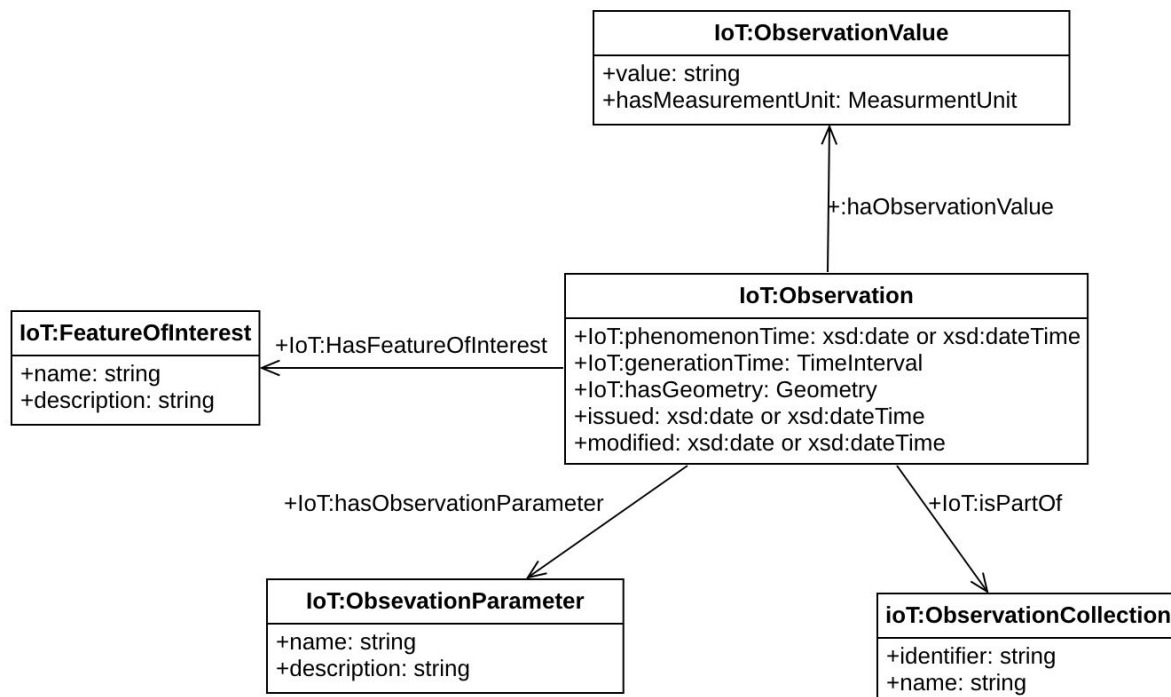


Figura 3: Rappresentazione UMI della parte del modello sulle osservazioni

Il modello proposto consente di rispondere a domande di competenza (competency question) come le seguenti; si noti che non è una lista esaustiva ma è un primo elenco a titolo d'esempio:

ID domanda di competenza	Domanda di competenza
CQ1	Quali sono gli indicatori di tipo processo?
CQ2	Quali sono gli indicatori calcolati per il Comune XXX?
CQ3	Qual è la fonte dati che consente di calcolare l'indicatore XXX?
CQ4	Qual è la frequenza di aggiornamento per il calcolo dell'indicatore XXX?
CQ5	Qual è il valore di un certo indicatore?
CQ6	Qual è la metrica/metriche usata per calcolare un certo indicatore?
CQ9	Quali sono le proprietà che forniscono i dati per il calcolo di un certo indicatore?

Insieme a questo modello è stata impostata una prima [bozza di vocabolario controllato degli indicatori della fase 2A COVID-19](#). Grazie agli esperti di dominio del Ministero della Salute e della Protezione Civile, i primi 6 indicatori (tabella 1 del decreto) del vocabolario controllato sono stati specificati nelle loro *label* (etichette) ufficiali in maniera puntuale tenendo conto anche delle soglie individuate nel decreto.

Lo stesso lavoro di raffinamento dell'etichetta (label) ufficiale associata all'indicatore potrà essere fatta insieme agli esperti di dominio del Ministero della Salute e della Protezione Civile anche per gli indicatori delle tabelle 2 e 3.

Il modello e il vocabolario sono utilizzati per la definizione dello schema di API. Nel caso di regioni non completamente attrezzate con sistemi automatizzati, il modello e il vocabolario potrebbero essere anche efficacemente utilizzati per la creazione di un template di dataset in formato CSV da mettere a disposizione per la compilazione e l'invio dei dati agli organi competenti centrali.

Open Data COVID-19 - Iniziative regionali

Questa sezione evidenzia le varie iniziative regionali di dialogo e promozione della trasparenza con i cittadini sui dati e le statistiche relative al fenomeno COVID-19.

Dall'analisi emerge un panorama eterogeneo sia sulle modalità di erogazione (frequenza, formato) sia sui contenuti informativi che vengono comunicati.

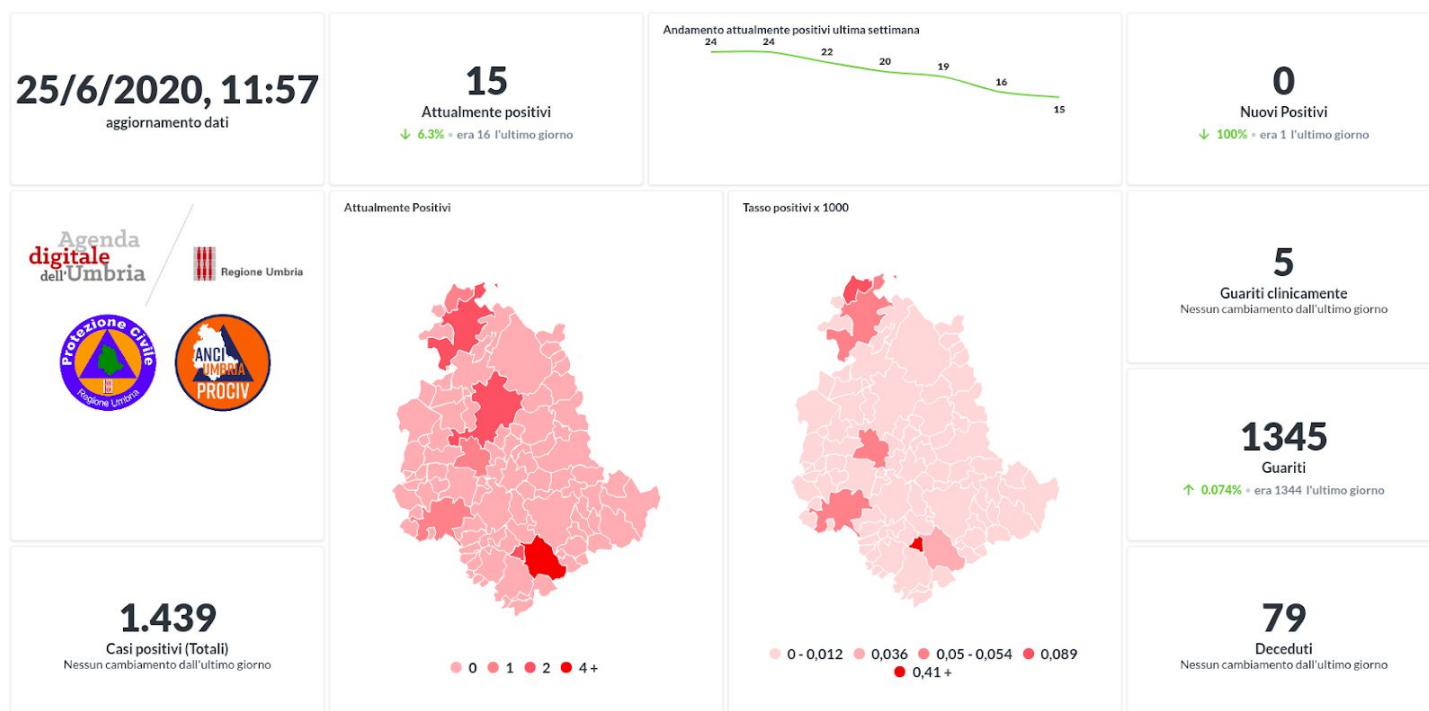
In particolare, alcune regioni diffondono autonomamente i propri dati una e addirittura più volte al giorno, altre condividono informazioni molto granulari (casi, tamponi, ricoveri, ricoveri in terapia intensiva, decessi), altre ancora invece si limitano a comunicare solo i casi e i decessi. Inoltre, ogni regione utilizza metodi e modalità di rappresentazione dei dati differenti: la maggior parte non pubblica i dati in modalità raw e in formato aperto, e utilizza informazioni testuali e tabelle spesso distribuite in formato PDF, che limita il riutilizzo dei dati. Sono pochissime le realtà che seguono in parte le linee guida sulla valorizzazione del patrimonio pubblico, offrendo ad esempio dati aperti e servizi esposti mediante API con relative dashboard di analisi territoriale.

Rispetto all'analisi effettuata all'inizio maggio, si rileva positivamente, che diverse regioni hanno ampliato l'offerta di informazioni, servizi e dati pubblicati (anche attraverso dashboard e/o infografiche) promuovendo quindi più trasparenza.

Si rinvia allo [spazio unico dati e risorse COVID-19](#), realizzata nell'ambito della Task force, per la consultazione delle iniziative regionali sotto elencate.

Buone pratiche dal territorio: il caso della Regione Umbria

Anche sull'esempio del Dipartimento per la Protezione Civile, alcune Regioni hanno pubblicato delle dashboard informative con un alto livello di dettaglio a livello territoriale. E' il caso della Regione Umbria, che fornisce ad oggi una delle dashboard più dettagliate presenti sul panorama nazionale.



I dati presenti nella dashboard vengono distribuiti mediante API; essi seguono la struttura riportata nella seguente tabella di raccordo che mappa i valori osservati dalla Regione Umbria

rapportandoli a quelli nazionali. Alcuni dei valori raccolti a livello regionale non sono presenti nel tracciato del DPC perché non tutte le Regioni sono riuscite a reperirli. Questo tipo di tabella può fungere da ausilio per aiutare i vari erogatori ad allineare le osservazioni fatte a livello locale con quanto stabilito a livello nazionale.

Nome campo	Descrizione	Disp.Naz. (equivalente campo DPC)
data	Data dell'informazione	data
stato	Stato di riferimento	stato
codice_regione	Codice della Regione (ISTAT)	codice_regione
denominazione_regione	Denominazione della Regione	denominazione_regione
codice_geo	Codice area geografica (ISTAT per i comuni) a cui si riferiscono i dati	
denominazione_geo	Denominazione dell'area o struttura a cui si riferiscono i dati	
tipo_geo	Tipo di area geografica o struttura a cui si riferiscono i dati (valorizzato "extra" per i dati sui non residenti)	
lat_geo	Latitudine (centroide)	lat
long_geo	Longitudine (centroide)	long
totale_residenti	Totale residenti da ISTAT 2019	
fascia_eta	Ove applicabile al dato, indica l'intervallo di età, compreso tra un valore minimo e uno massimo, nel modo seguente: A. Minore di 6 anni; B. da 6 anni a 13 anni; C. da 14 a 17 anni; D. da 18 a 39 anni; E. da 40 a 64 anni; F. da 65 a 79 anni; G. maggiore di 80 anni	
sex	Ove applicabile al dato, indica il sesso delle persone a cui si riferisce il dato.	
casi_positivi	Totale complessivo delle persone ad oggi risultate positive ad almeno un tampone oro-faringeo; anche uguale a: attualmente_positivi + guariti + deceduti	totale_casi

tasso_casi_positivi_x1000	Prevalenza del totale dei casi positivi ogni 1000 abitanti residenti, dall'inizio dell'emergenza; calcolo: $\text{casi_positivi} / \text{residenti} * 1000$	
nuovi_positivi	Nuovi casi positivi dal giorno precedente; calcolo: $\text{casi_positivi} - \text{casi_positivi del giorno prima}$	nuovi_positivi
tasso_nuovi_settimana_x1000	Incidenza dei nuovi casi positivi ogni 1000 abitanti residenti, negli ultimi 7 giorni; calcolo: $\text{nuovi_positivi degli ultimi 7 giorni} / \text{residenti} * 1000$	
isolamento	Totale complessivo delle persone sottoposte alle due modalità di isolamento possibile, fiduciario o contumaciale. Quindi non per forza testate positive; calcolo: $\text{attualmente_positivi} + \text{isolamento_volontario}$	
isolamento_fiduciario	Attuale numero di persone sottoposte ad isolamento fiduciario in quanto hanno avuto contatti stretti con casi positivi, loro familiari o casi sospetti definiti dall'ISP. Non sono testate positive e sono in isolamento in casa o altra struttura non ospedaliera	
isolamento_contumaciale	Attuale numero di casi positivi che sono tenuti a restare in isolamento contumaciale, senza ricovero	
isolamento_positivi	Attuale numero di casi positivi che sono tenuti a restare in isolamento contumaciale, tolti i guariti clinici; calcolo: $\text{isolamento_positivi_sintomatici} + \text{isolamento_positivi_asintomatici}$; oppure anche: $\text{isolamento_contumaciale} - \text{guariti_clinici}$	isolamento_domiciliar e
isolamento_positivi_sintomatici	Attuale numero di casi positivi in isolamento contumaciale, con sintomi	
isolamento_positivi_asintomatici	Attuale numero di casi positivi in isolamento contumaciale, senza sintomi	
usciti_da_isolamento	Totale numero di casi positivi usciti dall'isolamento contumaciale (non più positivi al tampone)	
ricoverati_totale	Attuale numero dei casi positivi che sono ricoverati in ospedale	totale_ospedalizzati

di_cui_ricoverati_con_sintomi	Attuale numero di casi positivi che sono ricoverati in reparti diversi dalla terapia intensiva; calcolo: $\text{ricoverati_totale} - \text{di_cui_ricoverati_in_terapia_intensiva}$	ricoverati_con_sintomi
di_cui_ricoverati_in_terapia_intensiva	Attuale numero di casi positivi ricoverati in terapia intensiva	terapia_intensiva
terapia_intensiva_attuali_posti	Attuale numero di posti in terapia intensiva	
in_terapia_intensiva_attuali_degenti	Attuale numero di ricoverati (anche non positivi) in terapia intensiva	
totale_positivi	Attuale numero di casi positivi tolti i guariti clinici ; calcolo: $\text{attualmente_positivi} - \text{guariti_clinici}$; oppure anche: $\text{isolamento_positivi} + \text{ricoverati_totale}$	totale_positivi
attualmente_positivi	Attuale numero di casi positivi; calcolo: $\text{ricoverati_totale} + \text{in_isolamento_contumaciale}$	
tasso_positivi_x1000	Tasso degli attuali casi positivi ogni 1000 abitanti residenti; calcolo: $\text{attualmente_positivi} / \text{residenti} * 1000$	
sign_positivi_x1000	Significatività degli attuali casi positivi ogni 1000 abitanti residenti rispetto al tasso medio regionale; calcolo: 1 se tasso_positivi_x1000 superiore del 5% alla media, oppure -1 se tasso_positivi_x1000 inferiore del 5% alla media, altrimenti 0 se intorno alla media	
dimessi_guariti	Totale dei guariti e dei guariti clinici; = calcolo: $\text{guariti} + \text{guariti_clinici}$	dimessi_guariti
guariti	Totale dei casi positivi che risolvono i sintomi dell'infezione da Covid-19 e che risultano negativi in due test consecutivi effettuati a distanza di 24 ore uno dall'altro	
guariti_clinici	Totale dei casi positivi che risultano clinicamente guariti, anche se ancora positivi al test. Pur non essendo più necessario il ricovero, la persona non può ritornare alla vita di comunità perché ancora con una carica virale elevata.	

deceduti	Totale dei casi positivi che sono deceduti, anche con diagnosi post-mortem; La conferma che la causa del decesso è attribuibile esclusivamente al SARS-CoV-2 verrà dichiarata dall'Istituto Superiore di Sanità	deceduti
indice_letalita	Percentuale di deceduti sul totale dei casi positivi; calcolo: $\text{deceduti} * 100 / \text{casi_positivi}$	
tamponi_eseguiti	Totale dei tamponi (test) effettuati, un soggetto può essere sottoposto a più tamponi quindi non è indicativo delle persone controllate	tamponi
nuovi_tamponi_eseguiti	Nuovi tamponi eseguiti nel singolo giorno; calcolo: $\text{tamponi_eseguiti} - \text{tamponi_eseguiti del giorno prima}$	
tamponi_positivi	Totale dei tamponi (test) effettuati con esito positivo, un soggetto può essere sottoposto a più tamponi	
casi_testati	Totale dei soggetti sottoposti a test	
nuovi_casi_testati	Nuovi soggetti su cui si è eseguito il test nel singolo giorno; calcolo: $\text{casi_testati} - \text{casi_testati del giorno prima}$	
testo_acc	Testo sostitutivo del dato a fini di accessibilità.	
note	Eventuali note sul dato raccolto. In questo campo viene indicata anche l'eventuale revisione di aggiornamento del dato.	
status	Se dato ufficiale confermato = "pub"	

Inoltre, dal 3 aprile si è sviluppata la campagna #daCasaPuoi sulle pagine facebook dei DigiPASS regionali (punti di accesso assistito di Assisi, Città di Castello, Foligno, Gubbio, Media Valle del Tevere, Narni, Orvieto, Perugia, Spoleto) attraverso schede grafiche o video per offrire informazioni e suggerimenti utili sui servizi digitali disponibili online per svolgere attività quotidiane senza uscire di casa. Altra campagna denominata #emerGemma è stata portata avanti dal progetto pubblico-privato finanziato nell'avviso #OpenUmbria.

Le altre iniziative regionali

Si rinvia allo [spazio unico dati e risorse COVID-19](#), realizzata nell'ambito della Task force, per la consultazione delle iniziative regionali sotto elencate.

Regione Abruzzo

La Regione Abruzzo ha una pagina web dedicata su COVID-19 con molte informazioni, compresi i numeri di telefono dell'assistenza COVID per ciascuna provincia. Lo stesso per il portale regionale della salute.

La Regione ha pubblicato sul suo sito web un bollettino giornaliero (con slide riepilogative) con dati su casi, decessi per ciascuna provincia / ASL e dati aggregati su terapia intensiva e guarigioni.

Regione Basilicata

La Regione pubblica un'infografica relativa al bollettino epidemiologico da COVID-19. La regione ha condiviso il bollettino quotidiano con i giornalisti.

Provincia di Bolzano

La Provincia Autonoma di Bolzano ha una pagina web specifica su COVID-19: è stato anche prodotto materiale informativo di vario genere. Lo stesso sulla pagina web dell'Autorità Provinciale della Sanità.

Per quanto riguarda la condivisione dei dati, la provincia di Bolzano pubblica quotidianamente dati aperti in formato Excel con casi positivi giornalieri per comune e ASL. I casi di ricoveri ospedalieri e terapia intensiva sono invece riportati (solo testo) nel comunicato stampa giornaliero.

Regione Calabria

E' stato creato un portale per l'emergenza coronavirus (in collaborazione con la protezione civile) in cui viene pubblicato il bollettino giornaliero e sono presenti le mappe ed i grafici dei contagi.

Regione Campania

La Regione ha una pagina web dedicata a COVID-19, e pubblica ogni giorno un'immagine di sintesi con i dati positivi e negativi quotidiani, con alcuni dettagli sull'ospedale in cui si trovano, ma tutti in prosa, senza tabelle.

Regione Emilia Romagna

Il portale della Regione Emilia Romagna ha in Home page una sezione dedicata a COVID-19, dove sono pubblicati i comunicati stampa. La Regione ha anche prodotto propri materiali informativi e video, anche con testimonianze di VIP che vivono nella regione (tutte nella home page). Inoltre, anche il portale Salute ER fa lo stesso.

A livello regionale, un comunicato stampa viene pubblicato quotidianamente tra le notizie sulla homepage con dati di prosa per provincia su casi, ricoveri, decessi e dimissioni, ma con dettagli provinciali, non comunali o ASL.

Regione Friuli Venezia Giulia

Il portale web ha una sezione specifica dedicata a COVID-19. Esiste inoltre, il portale della Protezione civile FVG con la mappa interattiva dei casi per comune (positivi, guariti, decessi), con dati scaricabili in csv.

Regione Lazio

La Regione Lazio ha un portale regionale dedicato a COVID-19: dove i collegamenti a: ordinanze, news, FAQ, materiale informativo, comprese infografiche regionali, supporto psicologico. Il portale Salute Lazio fa lo stesso. È stata lanciata una app regionale per parlare con il proprio medico.

Per quanto riguarda la diffusione dei dati, il dipartimento di Epidemiologia regionale ha sviluppato una mappa interattiva per ciascun comune.

Regione Liguria

Il portale web della Regione Liguria ha un link alla home page alla pagina dedicata a COVID-19 con collegamenti ai materiali del Ministero della Salute.

Dal sito web della regione è possibile trovare il bollettino giornaliero (solo testo) con i dettagli a livello di ASL. I dati su casi positivi, persone in quarantena e ricoveri per ASL sono diffusi quotidianamente su Facebook e Twitter.

Regione Lombardia

Il portale regionale ha una parte dedicata a COVID-19 chiaramente visibile con un ChatBot sul coronavirus sulla homepage. È stata lanciata la App AllertaLOM: già utilizzata per gli avvisi di protezione civile, sin dall'inizio dell'emergenza di Coronavirus che ha consentito a oltre 50 mila utenti di ricevere notifiche e informazioni. Gli utenti compilano un questionario integrato all'interno dell'app che consente di raccogliere dati, in forma anonima, e rendere disponibili informazioni complete e strutturate sulla diffusione dell'infezione all'Unità di crisi regionale e specialisti territorio lombardo.

Per quanto riguarda i dati, il riferimento è il portale Lombardia News dove è presente una mappa interattiva con il numero di casi in ciascuna provincia.

Regione Marche

La Regione condivide (attraverso il GORES – Gruppo Operativo Regionale per l’Emergenza Sanitaria) ogni giorno dati completi (in pdf) sul numero di casi, test, ricoveri, terapia intensiva, per provincia, ASL e ospedale.

Regione Molise

Il sito Web della Regione Molise ha aperto una pagina Web specifica su COVID-19 con collegamenti a tutti i Decreti e gli aggiornamenti sulla pandemia. La Regione Molise pubblica ogni giorno una nota con dati anche su ricoveri e cure intensive con dettagli per l'ospedale, anche in formato grafici interattivi (). È stato anche prodotto materiale informativo sulle buone pratiche.

Regione Piemonte

La Regione Piemonte ha aperto una pagina Web su COVID-19 sul sito Web regionale. E' presente una Mappa dei contagi suddivisa per Comune di domicilio (totale positivi e positivi per

1000 abitanti) e anche una sezione sul sito web della Giunta Regionale, chiamato "Piemonte Informa" con i dati aggiornati (solo testo) per ogni provincia.

Regione Puglia

La Regione Puglia ha una pagina web dedicata su COVID-19 all'interno del sito Web regionale. Sono stati prodotti alcuni materiali regionali ad esempio video su chiedere all'esperto, un utile glossario sulla pandemia e materiali di comunicazione in 6 lingue per persone migranti (inglese, francese, arabo, bambara, pasthu e wolof).

Per quanto riguarda la condivisione dei dati, la regione pubblica quotidianamente il bollettino giornaliero in pdf con molti dati e grafici, solo sui casi positivi e sui decessi

Regione Sardegna

La Regione Sardegna ha una pagina web su COVID-19 all'interno del sito Web regionale dove viene pubblicato il bollettino quotidiano con tutti i dati (solo testo) su casi e ricoveri tra le province

La regione pubblica anche un documento chiamato "open data room" con grafici anche sulle persone che hanno viaggiato dall'Isola all'Italia.

Regione Sicilia

La Regione Sicilia utilizza il portale "Costruire Salute" per condividere dati e informazioni ai cittadini (solo testo). Lo stesso su Coronavirus Sicilia.

È stata lanciata la Web App Covid-19 – SiciliaSiCura, un servizio creato dalla Regione Siciliana in collaborazione con la Protezione Civile regionale, dedicato ai cittadini che sono tornati in Sicilia a partire dal 1 marzo 2020 e che si sono registrati sul sito www.siciliacoronavirus.it. Gli utenti registrati o coloro che si registrano sul sito www.siciliacoronavirus.it riceveranno via SMS ed e-mail, un nome utente e un collegamento per la creazione della password. Per utilizzare il servizio, l'utente, utilizzando il proprio smartphone, deve essere fisicamente geolocalizzato nel territorio della Regione Siciliana.

Regione Toscana

La Regione Toscana ha aperto una pagina Web su COVID-19 all'interno del sito Web regionale. L'Autorità Sanitaria Regionale (ARS Toscana) ha pubblicato una dashboard interattiva (con set di dati), ma solo con dati su casi e morti per province e città ().

Provincia Autonoma di Trento

La Provincia autonoma di Trento ha una pagina web su COVID-19 sul sito web principale che contiene video, consigli, materiali informativi. Anche una parte della pagina web dell'autorità sanitaria è dedicata a COVID19.

Dal 16 marzo è disponibile l'app #TreCovid19 (sviluppata da PA Trento, Fondazione Bruno Kessler, Servizio sanitario provinciale). Le nuove funzionalità consentiranno il monitoraggio remoto delle persone in quarantena e dei pazienti positivi al virus trattati a casa. Grazie all'app, saranno in grado di inserire informazioni sullo stato di salute (automonitoraggio) e gli operatori sanitari potranno monitorare i dati da remoto.

Per quanto riguarda i dati, il Servizio sanitario provinciale e Trentino Digitale hanno aperto una dashboard interattiva con dati e mappe aperti per ogni comune ed è possibile scaricare i dati in formato csv.

Regione Valle d'Aosta

La Regione Valle d'Aosta ha una pagina Web su COVID-19 sul sito Web regionale principale e sono stati prodotti materiali di comunicazione locali (infografiche) sia in italiano che in francese . La Regione pubblica quotidianamente un bollettino in pdf con la somma regionale del numero di test, casi e casi isolati, oltre al totale ospedalizzato, con i dati dei casi di ciascun comune.

Regione Veneto

La Regione Veneto ha aperto una pagina Web sul principale sito Web regionale con materiali di comunicazione dedicati (), e anche l'Autorità Regionale Sanitaria (Azienda Zero) ha una pagina web su COVID-19.

Per quanto riguarda i dati, fino a metà aprile l'Azienda Zero ha pubblicato due volte al giorno un bollettino con dati su casi, decessi, ricoveri e cure intensive di ciascun ospedale e provincia. Da metà aprile si è deciso di pubblicare il bollettino solo al mattino, per essere più omogenei con i dati della Protezione Civile. Dall'inizio hanno pubblicato due tabelle: uno dei casi COVID19 e un altro sul dettaglio dei ricoveri (età, sesso, ospedale). Attualmente viene generata una infografica settimanale.

Specifiche di comunicazione - API

Questa sezione descrive la specifica di interscambio per le osservazioni utilizzando una API REST conforme alle indicazioni del Modello di Interoperabilità per la Pubblica amministrazione. La specifica proposta permette di pubblicare delle osservazioni (Observations) di una serie di parametri semanticamente ben definiti (observation-parameters) - eg. morti_totali - associate ad una determinata unità amministrativa (comune, unità territoriale sovracomunale, regione). L'unità amministrativa è identificata dal codice istat.

L'API è stata progettata con l'intento di supportare in primis le seguenti funzionalità:

- i parametri esposti possono variare nel tempo;
- ad ogni osservazione è possibile associare sia l'intervallo temporale (data o timestamp) dell'evento (phenomenon_time) sia l'eventuale istante di pubblicazione (generation_time) o di una modifica (modified);
- si può recuperare la definizione dei singoli parametri in formato machine-readable;
- si possono ritornare le definizioni dei parametri osservati insieme alle osservazioni;
- i dati possono essere ottenuti in diversi formati (JSON, GeoJSON, CSV) attraverso il [meccanismo della Content-Negotiation](#) implementato dal protocollo HTTP. Nel caso del GeoJSON è possibile restituire anche i confini delle unità amministrative di riferimento ottenendo la collezione di osservazioni come una property di un geojson.

Ecco alcune limitazioni deliberatamente introdotte nel design utili a semplificare - in una prima fase - l'implementazione del servizio:

- gli URL associati alle varie tipologie di unità amministrative sono separati perché non è detto che la somma, eg. dei dati delle province, corrisponda al totale dei dati regionali a causa di possibili difficoltà nel campionamento o di problematiche amministrative (eg. trasferimenti, disallineamenti dovuti alla durata temporale dei campionamenti, errori umani, ...);
- le response in formato CSV possono contenere dati mancanti quando non ci sono dati associati a determinati observation-parameter;

- la possibilità di fare ricerche è opzionale.

Panoramica dello schema

La specifica dell'API è reperibile [nella bozza di specifica](#) nel [repository api-pilota-osservazioni](#).

I path (URL) principali sono due:

- `observation-parameters/` che permette di recuperare la definizione dei parametri di osservazione;
- `observations/` dove è possibile scaricare i dati veri e propri relativi ad una unità amministrativa come indicato di seguito

```
/observations/administrative-units/regioni/{codice_regione}  
/observations/administrative-units/province/{codice_provincia}  
/observations/administrative-units/comuni/{codice_comune}
```

E' possibile estendere il modello ad una granularità diversa (es. per singola struttura sanitaria, quartiere della città, ...) aggiungendo nuovi path.

Lo schema di un `ObservationParameter` è il seguente, ed ha associati un URI ed una definizione che ne sancisce semanticamente ed universalmente il significato. Col tempo è possibile creare nuovi `ObservationParameter` utili a rispondere meglio ai requisiti informativi: se il significato di un parametro cambia, è fondamentale cambiare anche il suo identificativo (campi `@id` ed `id`) in modo che sia sempre possibile comparare nel tempo le osservazioni in maniera corretta. Con l'evolversi delle fasi della pandemia ad esempio si potrebbe definire un nuovo `observation-parameters` chiamato `positives_total_2021`.

```
id: positives_total  
'@id': https://w3id.org/italia/data/observation-parameter/positives_total  
name: casi sintomatici con data inizio sintomi
```

Lo schema di una `Observation` è estensibile e supporta questi parametri minimi: data di validità, identificativo del parametro e valore.

```
value: 4  
parameter_id: intensive_care  
phenomenon_time: "2020-05-20"
```

E' possibile restituire anche dati più dettagliati, che includono i metadati del parametro e dell'oggetto della misurazione, l'istante di generazione o di modifica del dato e la caratteristica di interesse dell'osservazione (e.g., il paziente).

```
generation_time: "2020-05-22"  
feature_of_interest_id: patient  
value_measurement_unit: people  
_embedded:  
  parameter:  
    '@id': https://w3id.org/italia/data/observation-parameter/positives_total  
    name: casi sintomatici con data inizio sintomi  
  feature_of_interest:
```

```
name: Patient
'id': https://w3id.org/italia/data/feature-of-interest/patient
```

Ad ogni Observation possiamo associare una serie di informazioni, come la struttura che le ha raccolte o delle informazioni geografiche. La nostra API riporta - per compattezza - le informazioni raggruppate per Unità Amministrativa all'interno di un oggetto di tipo Observations, che contiene:

- una lista di Observation (il concetto di Collezione di Osservazioni dell'ontologia delle osservazioni [collegata con l'ontologia degli indicatori](#))
- un campo **_meta** con i metadati associati: nel nostro caso ci sono le informazioni sull'unità amministrativa a cui sono associate le informazioni

```
Observations:
  items:
  - parameter_id: intensive_care
    phenomenon_time: '2020-05-20'
    value: 4
  - parameter_id: intensive_care
    phenomenon_time: '2020-05-21'
    value: 3
  - parameter_id: intensive_care
    phenomenon_time: '2020-05-20'
    value: 5
  - modified: '2020-05-21'
    parameter_id: positives_total
    phenomenon_time: '2020-05-21'
    value: 11
  - parameter: positives_total
    phenomenon_time: '2020-05-22'
    value: 8
  _meta:
    administrative_unit:
      city_id: 054039
      city_name: Perugia
      country_id: ITA
      country_name: Italia
      province_acronym: PG
      province_id: '054'
      province_name: Perugia
      region_id: '10'
      region_name: Umbria
    source:
      IPA_code: r_umbria
      description: Regione Umbria
      url: https://www.indicepa.gov.it/ricerca/n-dettaglioservfe.php?cod_amm=r_umbria
```

Per poter ritornare Observations relative ad unità amministrative diverse, l'API ritorna un ObservationDataset che contiene una lista di Observations.

Si consulti l'allegato A in cui vengono illustrati alcuni esempi di richieste e risposte in formato HTTP che illustrano la modalità di richiesta dei dati e i contenuti inviati dal server.

Raccordo con le raccomandazioni del GDL 7

Con riferimento al documento "[Follow-up report: Analisi e raccomandazioni per la fase di riapertura](#)" a cura del gruppo di lavoro #7 della task force, nel modello dati e nella specifica API qui proposti si evidenziano sia elementi di allineamento ai principi descritti dal GDL 7 sia elementi di supporto tecnico per la realizzazione di quanto da loro proposto.

In particolare, il modello dati e la specifica API mirano a fornire un ecosistema condiviso e comune di interscambio di dati tra i diversi livelli amministrativi coinvolti, contribuendo pertanto a implementare alcuni degli aspetti tecnici e organizzativi chiave menzionati dal GDL 7 nel contesto dell'architettura dell'ecosistema integrato data-driven per il monitoraggio. Infatti, sia il modello che la specifica API possono essere efficacemente adottati per governare i flussi di raccolta dei dati epidemiologici e di impatto sul sistema sanitario e abilitare un controllo sulla qualità dei dati nonché sulla loro possibile integrazione. Nello specifico, il modello semantico, reso disponibile anche attraverso standard del Web Semantico (e.g., RDF/OWL) nel contesto più ampio della rete di ontologie e vocabolari controllati per la pubblica amministrazione, può facilitare l'integrazione dei dati tra un livello centrale e locale, supportare l'analisi dati, la creazione di cruscotti di navigazione e i modelli predittivi, grazie anche alla possibile inferenza di nuova conoscenza che può essere abilitata.

Sempre nell'ambito del lavoro del GDL 7 relativo alle *"indicazioni tecniche per l'implementazione del flusso descritto nel decreto del Ministero della Salute del 30/04/2020"*, si nota come alcune delle raccomandazioni operative suggerite possono essere rese operative dal lavoro descritto nel presente documento.

Nello specifico, il modello e la specifica API consentono alle regioni di inviare i dati in forma esclusivamente elettronica, con un formato predefinito che segue appunto il modello proposto. La specifica API permette di avere risposte in diversi formati, tra cui anche il CSV. Inoltre, il modello delle osservazioni, ossia i dati puntuali osservati dal livello locale che contribuiscono a formare le banche dati su cui poi calcolare gli indicatori, supporta l'organizzazione del dato coppia chiave-valore, come suggerita dal GDL 7. Infatti ragionando in questi termini, la chiave potrebbe essere quello che nel modello proposto per la parte osservazioni è il *parametro dell'osservazione* e il valore potrebbe essere il *valore dell'osservazione* stessa.

In conclusione, gli strumenti introdotti in questo documento sono allineati ai principi che anche altri gruppi di lavoro della task force raccomandano e costituiscono un primo importante tassello per l'implementazione di alcune raccomandazioni della task force.

Gruppo di lavoro

Hanno contribuito alla realizzazione di questo documento:

Massimo Bernaschi, CNR - Istituto per le Applicazioni del Calcolo (componente del gruppo di lavoro 7)

Maria Claudia Bodino, Dipartimento per la Trasformazione Digitale

Pierluigi Cara, Dipartimento della Protezione Civile

Giorgio Consol, Regione Piemonte Direzione Competitività del sistema regionale, Settore Sistema informativo regionale

Giovanni Gentili, Regione Umbria

Fabio Godorecci, Centro Interregionale per I Sistemi Informatici Geografici e Statistici

Giorgia Lodi, Agenzia per l'Italia Digitale

Roberto Polli, Dipartimento per la Trasformazione Digitale

Umberto Rosini, Dipartimento della Protezione Civile

Riferimenti bibliografici

- [1] P. Ciancarini, A. G. Nuzzolese, V. Presutti, D. Russo, *SQuAP-Ont: an Ontology of Software Quality Relational Factors from Financial Systems* in Semantic Web Journal - IOS Press, 2019 - <http://www.semantic-web-journal.net/system/files/swj2256.pdf>
- [2] Y. Li, R. Garcia-Castro, N. Mihindukulasooriya, J. O Donnell, S. Sánchez. *Enhancing energy management at district and building levels via an EM-KPI ontology*. in Automation in Construction. 99. 152-167. 10.1016/j.autcon.2018.12.010. - <http://energy.linkeddata.es/em-kpi/ontology/index-en.html>
- [3] A. G. Nuzzolese, V. Presutti, A. Gangemi, P. Ciancarini *Extending ScholarlyData with research impact indicators* in: Lecture Notes in Computer Science Vol 10959 Springer Verlag, atti del 3rd International Workshop on Semantics, Analytics, Visualization, SAVE-SD 2017, SAVE-SD 2018 - DOI: https://doi.org/10.1007/978-3-030-01379-0_4 disponibile anche in , https://save-sd.github.io/2018/accepted/nuzzolese/index.html#endnote_15
- [4] M. S. Fox *ISO 37120 indicator URIs*, University of Toronto - <http://ontology.eil.utoronto.ca/ISO37120.html>
- [5] M. Fox, "A Foundation Ontology for Global City Indicators", University of Toronto - Agosto 2013 - https://www.researchgate.net/publication/262675058_A_Foundation_Ontology_for_Global_City_Indicators
- [6] O. D. Beyan, N. Baykal, *A Knowledge Based Search Tool for Performance Measures in Health Care Systems*, in Journal of Medical Systems 36, 201-221 (2012). DOI:<https://doi.org/10.1007/s10916-010-9459-2> - <https://link.springer.com/article/10.1007/s10916-010-9459-2#article-info>
- [7] L. Olsina, M. De Los Angeles Martín, *Ontology for Software Metrics and Indicators* in Journal of Web Engineering (JWE) 2(4):262-281 · Gennaio 2004 - <https://www.semanticscholar.org/paper/Ontology-for-Software-Metrics-and-Indicators-Olsina-Mart%C3%ADn/047e4cce849f91abdfa1c5c699521f572fe65cf5?citationIntent=result#citing-papers>

ALLEGATO I - Specifiche di comunicazione API

- Casi d'uso ed esempi

Nei prossimi paragrafi troviamo degli esempi di richieste e risposte in formato HTTP che illustrano la modalità di richiesta dei dati e i contenuti inviati dal server.

Osservazioni quotidiane regionali

Come CITTADINO voglio conoscere LO STATO ODIERNO DEI CONTAGI NELLA REGIONE Piemonte

Esempio di richiesta HTTP

```
GET
/observations/administrative-units/regioni/{codice_regioni}?phenomenon_time_start=2020-05-20&phenomenon_time_end=2020-05-20
```

RISPOSTA in formato CSV MINIMO

```
phenomenon_time,administrative_unit,administrative_unit_id,positives_total,intensive_care,...
2020-05-20,regione,01, 10, 4,...
```

RISPOSTA in formato CSV ESTESO

```
phenomenon_time,administrative_unit,administrative_unit_id,positives_total,intensive_care,...,region_id,region_name
2020-05-20,regione,01, 10, 4,...,01,Piemonte
```

RISPOSTA in formato JSON MINIMO

```
{
  "items": [
    {
      "items": [
        {
          "value": 3,
          "parameter_id": "intensive_care",
          "phenomenon_time": "2020-05-22"
        },
        {
          "value": 8,
          "parameter_id": "positives_total",
          "phenomenon_time": "2020-05-22"
        },
        ...
        {
          "value": 8,
          "parameter_id": "deceased",
          "phenomenon_time": "2020-05-22"
        }
      ],
      "_meta": {
        "administrative_unit": {
          "country_id": "ITA",
          "region_id": "01",
          "region_name": "Piemonte"
        },
        "source": {
          "IPA_code": "r_piemonte",
          "description": "Regione Piemonte",
          "url": "https://www.indicepa.gov.it/ricerca/n-dettaglioservfe.php?cod_amm=r_piemonte"
        }
      }
    }
  ]
}
```

RISPOSTA in formato JSON ESTESO

```

{
  "items": [
    {
      "items": [
        {
          "phenomenon_time": "2020-05-01",
          "parameter_id": "intensive_care",
          "value": 10,
          "generation_time": "2020-05-22",
          "feature_of_interest_id": "patient",
          "value_measurement_unit": "people",
          "_embedded": {
            "parameter": {
              "@id": "https://w3id.org/italia/data/observation-parameter/intensive_care",
              "name": "Pazienti in terapia intensiva alla data indicata"
            },
            "feature_of_interest": {
              "name": "Patient",
              "@id": "https://w3id.org/italia/data/feature-of-interest/patient"
            }
          }
        },
        {
          "phenomenon_time": "2020-05-01",
          "parameter_id": "positives_total",
          "value": 10,
          "generation_time": "2020-05-22",
          "feature_of_interest_id": "patient",
          "value_measurement_unit": "people",
          "_embedded": {
            "parameter": {
              "@id": "https://w3id.org/italia/data/observation-parameter/positives_total",
              "name": "Casi sintomatici e positivi al tampone alla data indicata"
            },
            "feature_of_interest": {
              "name": "Patient",
              "@id": "https://w3id.org/italia/data/feature-of-interest/patient"
            }
          }
        }
      ],
      ...
    },
    {
      "phenomenon_time": "2020-05-31",
      "parameter_id": "intensive_care",
      "value": 10,
      "generation_time": "2020-05-22",
      "feature_of_interest_id": "patient",
      "value_measurement_unit": "people",
      "_embedded": {
        "parameter": {
          "@id": "https://w3id.org/italia/data/observation-parameter/intensive_care",
          "name": "Pazienti in terapia intensiva alla data indicata"
        },
        "feature_of_interest": {
          "name": "Patient",
          "@id": "https://w3id.org/italia/data/feature-of-interest/patient"
        }
      }
    },
    {
      "phenomenon_time": "2020-05-31",
      "parameter_id": "positives_total",
      "value": 10,
      "generation_time": "2020-05-22",
      "feature_of_interest_id": "patient",
      "value_measurement_unit": "people",
      "_embedded": {
        "parameter": {
          "@id": "https://w3id.org/italia/data/observation-parameter/positives_total",
          "name": "Casi sintomatici e positivi al tampone alla data indicata"
        },
        "feature_of_interest": {
          "name": "Patient",

```

```

        "@id": "https://w3id.org/italia/data/feature-of-interest/patient"
      }
    }
  ],
  "_meta": {
    "administrative_unit": {
      "country_id": "ITA",
      "country_name": "Italia",
      "region_id": "01",
      "region_name": "Piemonte",
    },
    "source": {
      "IPA_code": "r_piemonte",
      "description": "Regione Piemonte",
      "url": "https://www.indicepa.gov.it/ricerca/n-dettaglioservfe.php?cod_amm=r_umbria"
    }
  }
}
]
}

```

Osservazioni in un intervallo mensile

Come CITTADINO voglio conoscere LO STATO DEI CONTAGI NELLA PROVINCIA di Pisa tra il 2020-05-01 e il 2020-05-31

RISPOSTA in formato CSV MINIMO

```

phenomenon_time,administrative_unit,administrative_unit_id,positives_total,intensive_care,...
2020-05-01,provincia,050, 10, 4,...
2020-05-02,provincia,050, 10, 4,...
..
2020-05-31,provincia,050, 10, 4,...

```

RISPOSTA in formato CSV ESTESO

```

phenomenon_time,administrative_unit,administrative_unit_id,positives_total,intensive_care,...,region_id,region_name,province_id,province_name,province_acronym
2020-05-01,provincia,050, 10, 4,...,10,Toscana,050,Pisa,PI
2020-05-02,provincia,050, 10, 4,...,10,Toscana,050,Pisa,PI
..
2020-05-31,provincia,050, 10, 4,...,10,Toscana,050,Pisa,PI

```

RISPOSTA in formato JSON MINIMO

```

{"items": [{
  "items": [
    {"value": 3, "parameter_id": "intensive_care", "phenomenon_time": "2020-05-01"},
    {"value": 8, "parameter_id": "positives_total", "phenomenon_time": "2020-05-01"},
    ...,
    {"value": 8, "parameter_id": "deceased", "phenomenon_time": "2020-05-01"},
    ...,
    {"value": 3, "parameter_id": "intensive_care", "phenomenon_time": "2020-05-31"},
    {"value": 8, "parameter_id": "positives_total", "phenomenon_time": "2020-05-31"},
    ...,
    {"value": 8, "parameter_id": "deceased", "phenomenon_time": "2020-05-31"}
  ],
  "_meta": {
    "administrative_unit": {
      "country_id": "ITA",
      "province_id": "50",
      "province_name": "Pisa",
    },
    "source": {

```



```

        "IPA_code": "r_toscana",
        "description": "Regione Toscana",
        "url": "https://www.indicepa.gov.it/ricerca/n-dettaglioservfe.php?cod_amm=r_toscana"
    }
}
]
}

```

Osservazioni in una Unità Amministrativa e nelle sotto-unità

Come CITTADINO voglio conoscere LO STATO ODIERNO DEI CONTAGI NEI COMUNI di una provincia.

```

/observations/administrative-units/province/{codice_provincia}?sub=true&phenomenon_time_begin=2020-05-01&phenomenon_time_end=2020-05-01

```

RISPOSTA in formato CSV MINIMO

```

phenomenon_time,administrative_unit,administrative_unit_id,positives_total,intensive_care,...
2020-05-01,comune,054001, 10, 4,...
2020-05-01,comune,054002, 10, 4,...
..
2020-05-31,comune,054050, 10, 4,...

```

RISPOSTA in formato CSV ESTESO

```

phenomenon_time,administrative_unit,administrative_unit_id,positives_total,intensive_care,...,region_id,region_name,province_id,province_name,province_acronym
2020-05-01,comune,054001, 10, 4,...,10,Umbria,54,Perugia,PG,054001,Assisi
2020-05-01,comune,054002, 10, 4,...,10,Umbria,54,Perugia,PG,054002,Bastia Umbra
..
2020-05-01,comune,054050, 10, 4,...,10,Umbria,54,Perugia,PG,054050,Spello

```

RISPOSTA in formato JSON MINIMO

```

{"items": [{
  "items": [
    {"value": 3, "parameter_id": "intensive_care", "phenomenon_time": "2020-05-01"},
    {"value": 8, "parameter_id": "positives_total", "phenomenon_time": "2020-05-01"},
    ...,
    {"value": 8, "parameter_id": "deceased", "phenomenon_time": "2020-05-01"},
    ...
    {"value": 3, "parameter_id": "intensive_care", "phenomenon_time": "2020-05-31"},
    {"value": 8, "parameter_id": "positives_total", "phenomenon_time": "2020-05-31"},
    ...,
    {"value": 8, "parameter_id": "deceased", "phenomenon_time": "2020-05-31"}
  ],
  "_meta": {
    "administrative_unit": {
      "country_id": "ITA",
      "city_id": "054001",
      "city_name": "Assisi",
    },
    "source": {
      "IPA_code": "r_umbria",
    }
  }
}]

```

```

        "description": "Regione Umbria",
        "url": "https://www.indicepa.gov.it/ricerca/n-dettaglioservfe.php?cod_amm=r_umbria"
    }
}
],
..
{
    "items": [
        {"value": 3, "parameter_id": "intensive_care", "phenomenon_time": "2020-05-01"},
        {"value": 8, "parameter_id": "positives_total", "phenomenon_time": "2020-05-01"},
        ...,
        {"value": 8, "parameter_id": "deceased", "phenomenon_time": "2020-05-01"},
        ...,
        {"value": 3, "parameter_id": "intensive_care", "phenomenon_time": "2020-05-31"},
        {"value": 8, "parameter_id": "positives_total", "phenomenon_time": "2020-05-31"},
        ...,
        {"value": 8, "parameter_id": "deceased", "phenomenon_time": "2020-05-31"}
    ],
    "_meta": {
        "administrative_unit": {
            "country_id": "ITA",
            "city_id": "054050",
            "city_name": "Spello",
        },
        "source": {
            "IPA_code": "r_umbria",
            "description": "Regione Umbria",
            "url": "https://www.indicepa.gov.it/ricerca/n-dettaglioservfe.php?cod_amm=r_umbria"
        }
    }
}
]
}

```