

Van: 5.1.2e, 5.1.2e, (5.1.2e)

Verzonden: woensdag 27 januari 2021 08:48

Aan: 5.1.2e, 5.1.2e <5.1.2e@rivm.nl>; 5.1.2e, 5.1.2e <5.1.2e@rivm.nl>; 5.1.2e, 5.1.2e, 5.1.2e <5.1.2e@rivm.nl>
 <5.1.2e@rivm.nl>
CC: 5.1.2e, 5.1.2e, 5.1.2e (5.1.2e) <5.1.2e@minvws.nl>; 5.1.2e, 5.1.2e, (5.1.2e) <5.1.2e@minvws.nl>; 5.1.2e, 5.1.2e <5.1.2e@AethiQs.nl>

Onderwerp: RE: RIVM vaccinatie open data

Hey 5.1.2e

Bedankt voor je reactie! Om het alvast zo goed mogelijk op weg te helpen doe ik dan bij dezen alvast zelf een voorstel qua formaat/kolommen. Zodra de RIVM programmeurs er dan wel aan toe zijn, dan hebben ze alvast een begin en gelijk iets om te bespreken. @ 5.1.2e Kan jij t.z.t. bij hen toetsen of dit werkbaar is, en wat er nog aan moet veranderen, wat de nice to have's zijn etc.?

In de bijlage een voorbeeld CSV bestand (met verzonden waarden), en hieronder een technische omschrijving. Ik heb zoveel mogelijk de RIVM open data conventies proberen te volgen.

Bestandsnaam: COVID-19_gezette_vaccinaties.csv

In de CSV-file wordt als scheidingsteken een ";" puntkomma gebruikt. Er staan geen lege regels in het bestand. Hieronder de kolomnamen en de typen waarden in de CSV:

- **Version:** Bestaande uit vrije tekst (maar geen puntkomma). Is voor elke rij altijd gevuld. Voorbeeld: '1'. Deze kolom is een voorstel van 5.1.2e om wijzigingen in het bestand bij te houden en is bv. ook doorgevoerd in https://data.rivm.nl/covid-19/COVID-19_ziekenhuisopnames.csv.
- **Date_of_report:** Geschreven in formaat JJJJ-MM-DD HH:MM. Is voor elke rij altijd gevuld. Voorbeeld: '2020-10-29 10:00:00'. Deze kolom zit ook in veel andere open data bestanden en het geeft weer wanneer het bestand voor het laatst is bijgewerkt. Heeft dezelfde waarde voor alle rijen.
- **Date_of_statistics:** Geschreven in formaat JJJJ-MM-DD. Is voor elke rij altijd gevuld. Voorbeeld: '2020-02-27'. Deze kolom zit ook in veel andere open data bestanden en het geeft de dag weer waar de betreffende rij over gaat.
- **Region_code:** Code van de regio. Nederland als geheel is meegenomen onder de code 'NL00'. Een veiligheidsregio wordt aangeduid door 'VR' gevolgd door twee cijfers: bijvoorbeeld 'VR01'. Een gemeente met 'GM' gevolgd door vier cijfers: bijvoorbeeld 'GM1945'. Dit in overeenstemming met de CBS notatie. Deze kolom is voor elke rij altijd gevuld. Deze werkwijze wordt bv. gehanteerd in https://data.rivm.nl/covid-19/COVID-19_gedrag.csv. Deze kolom zal vermoedelijk altijd NL00 bevatten op het begin, maar hopelijk kunnen de gegevens uiteindelijk ook op veiligheidsregio en/of gemeente niveau gepubliceerd kunnen worden. Door deze kolom op te nemen hoeft er dan niet iets aan de structuur van het bestand te worden.
- **Region_name:** Bestaande uit vrije tekst string (maar geen puntkomma). Is voor elke rij altijd gevuld. Voorbeelden: 'Nederland', of 'Midden- en West-Brabant' of 'Venlo'. Deze kolom biedt wat extra info over de Region_code. Deze werkwijze wordt bv. gehanteerd in https://data.rivm.nl/covid-19/COVID-19_gedrag.csv.
- **Vaccine_name:** Bestaande uit vrije tekst string (maar geen puntkomma). Is voor elke rij altijd gevuld. Deze kolom geeft aan welk vaccin type gezet is. Voorbeelden: 'BioNTech/Pfizer' of 'Moderna'. Vraag: Moeten we hier ook voorzien in de optie 'Unknown' indien dit onbekend is, of weten we dit in principe altijd?
- **Vaccine_round:** Bestaande uit enkel hele getallen (integer). Is voor elke rij altijd gevuld. Voorbeeld: '1'. Deze kolom geeft aan of iemand een eerste of tweede prik van het vaccin krijgt. Vraag: Moeten we hier ook voorzien in de optie 'Unknown' indien dit onbekend is, of weten we dit in principe altijd?
- **Vaccine_administered:** Bestaande uit enkel hele getallen (integer). Is voor elke rij altijd gevuld. Voorbeeld: '4512'. Deze kolom geeft aan hoeveel van deze vaccins op een dag gezet zijn.

Verder zijn er nog een aantal kolommen die ook zinvol kunnen zijn, maar ik weet niet of die ook aangeleverd kunnen worden bij de eerste livegang. Het is uiteraard prima mogelijk om later nog extra kolommen toe te voegen aan het bestand, naarmate meer gegevens beschikbaar komen. Het heeft onze voorkeur om eerst met een minimale set aan info (zoals hierboven beschreven) z.s.m. live te gaan qua open data. Zie hieronder een paar suggesties voor extra info:

- Doelgroep (bv. acute zorgmedewerkers, bewoners van gehandicaptenzorginstellingen, huisartsen).
- Door wie is de prik gezet (bv. GGD, ziekenhuizen, huisartsen)
- Eventueel los nog wat de bron van de data is, mocht dit afwijken van de partij wie de prik zet.

Daarnaast is een andere optie om het bestand op een andere manier in te richten. Het voorstel hierboven is namelijk op geaggregeerde basis, zoals https://data.rivm.nl/covid-19/COVID-19_verpleeghuizen.csv. Echter je zou er ook voor kunnen kiezen om het op casus basis (dus per prik een regel) in te richten, zoals https://data.rivm.nl/covid-19/COVID-19_casus_landelijk.csv. Dan kan je makkelijker informatie kwijt op persoonsniveau zoals leeftijd, geslacht. Dit is wel weer privacy gevoeliger, en hangt denk ik af of een persoon toestemming geeft tot verwerking van zijn/haar gegevens.

Met vriendelijke groet,

5.1.2e 5.1.2e
5.1.2e
Ministerie van Volksgezondheid, Welzijn en Sport
5.1.2e
5.1.2e

Van: 5.1.2e 5.1.2e <5.1.2e @rivm.nl>

Verzonden: dinsdag 26 januari 2021 18:15

Aan: 5.1.2e, 5.1.2e, (5.1.2e) <5.1.2e @minvws.nl>; 5.1.2e, 5.1.2e <5.1.2e @rivm.nl>

CC: 5.1.2e, 5.1.2e, 5.1.2e <5.1.2e @rivm.nl>; 5.1.2e, 5.1.2e, 5.1.2e (5.1.2e) <5.1.2e @minvws.nl>; 5.1.2e, 5.1.2e, (5.1.2e) <5.1.2e @minvws.nl>; 5.1.2e, 5.1.2e <5.1.2e @AethiQs.nl>

Onderwerp: RE: RIVM vaccinatie open data

Beste 5.1.2e,

We zijn telefonisch inderdaad niet goed te bereiken ivm de veelheid aan noodzakelijke overleggen.

We zijn ons bewust van het belang van de open data en welke stappen daarbij komen kijken; de hulp van 5.1.2e is daarbij ook zeer welkom.

Zoals je weet zijn we nu nog in de fase om data te vergaren, omdat CIMS nog niet operationeel is. We zullen op een opportuun moment onze acties op het gebied van open data starten.

Groet, 5.1.2e

From: 5.1.2e, 5.1.2e (5.1.2e) <5.1.2e @minvws.nl>

Sent: dinsdag 26 januari 2021 13:15

To: 5.1.2e, 5.1.2e <5.1.2e @rivm.nl>; 5.1.2e, 5.1.2e <5.1.2e @rivm.nl>

Cc: 5.1.2e, 5.1.2e, 5.1.2e <5.1.2e @rivm.nl>; 5.1.2e, 5.1.2e, 5.1.2e (5.1.2e) <5.1.2e @minvws.nl>; 5.1.2e, 5.1.2e, (5.1.2e) <5.1.2e @minvws.nl>; 5.1.2e, 5.1.2e <5.1.2e @AethiQs.nl>

Subject: RIVM vaccinatie open data

Hey 5.1.2e

Telefonisch krijg ik jullie helaas niet te pakken, dus vandaar maar weer even via de mail. Zoals jij 5.1.2e gisteren al aangaf, is het RIVM bezig om z.s.m. het gezette aantal vaccinaties (althans in ieder geval die via de GGD) als dagelijkse open data te publiceren. Eerder had ik al met 5.1.2e besproken dat haar team daartoe met 5.1.2e, 5.1.2e, 5.1.2e van het RIVM contact zou gaan opnemen om dit voor te bereiden. Ik sprak 5.1.2e echter vandaag en die is tot nu toe nog niet benaderd hierover.

Er zijn namelijk best al wat dingen die, ongeachte of de data nu al betrouwbaar is of niet, al besproken en uitgewerkt kunnen worden. Bijvoorbeeld de structuur van het open data bestand, de metadata pagina en hoe de scripts moeten ingesteld zijn om de data consistent en tijdig te publiceren. 5.1.2e heeft hier al veel ervaring mee en is technisch betrokken geweest bij vrijwel alle bestanden die staan op <https://data.rivm.nl/covid-19/>. Ik snap dat bepaalde details nu nog niet voor 100% zekerheid vast te leggen zijn, maar het lijkt mij wel verstandig om dit proces alvast serieus op gang te brengen. Door verder uit te werken hoe je de data gaat publiceren kan je namelijk ook tegen zaken lopen waar je aanvankelijk niet meteen aan denkt, en die impact hebben op hoe je data vergaart.

Enfin, ik hoop dat jullie 5.1.2e er snel bij kunnen betrekken. Hij heeft korte lijnen met ons kan jullie helpen om details met ons af te stemmen.

Met vriendelijke groet,

5.1.2e 5.1.2e
5.1.2e
Ministerie van Volksgezondheid, Welzijn en Sport
5.1.2e
5.1.2e

verband houdt met risico's verbonden aan het elektronisch verzenden van berichten.
www.rivm.nl *De zorg voor morgen begint vandaag*

This message may contain information that is not intended for you. If you are not the addressee or if this message was sent to you by mistake, you are requested to inform the sender and delete the message. RIVM accepts no liability for damage of any kind resulting from the risks inherent in the electronic transmission of messages.

www.rivm.nl/en *Committed to health and sustainability*