FAIR in times of Crisis

Virus Outbreak Data Network (VODAN)

FAIR Data Points as a service for data-driven research
(COVID-19 pressure-cooker use case)

Barend Mons

The VODAN Implementation Network is one of the joint activities carried out by CODATA, RDA, WDS, and GO FAIR (Link to the Data Together Statement).

Read the full statement on Data Together COVID-19 Appeal and Actions.

Active GO FAIR Implementation Network

The spread of the virus causing the COVID-19 outbreak is far from over. During this epidemic and in earlier occasions, we have seen severely suboptimal data management and data reuse. Moreover, access to the immensely valuable data of past and current epidemics is not always equally accessible for different affected populations and countries. For instance, the data from the past Ebola epidemics are very difficult to find, to access, and if accessible, they are not interoperable, let alone reusable. Under the urgent need to harness machine-learning and future AI approaches to discover meaningful patterns in epidemic outbreaks, we need to do better and ensure that data are FAIR (in this sense also meaning Federated, AI-Ready).

Purpose of the Implementation Network
Dear GO FAIR community,

This issue of the Newsletter will have a significantly longer editorial than what you are used to. The reason is that we are at a natural inflection point with GO FAIR, where we need to collectively reflect on our achievement and future ambitions.

GO FAIR was originally conceived as a temporary initiative, with the aim to ‘kick-start’ early developments towards the EOSC, while taking on an additional global perspective. After a little over two years on a roller coaster of amazing community building, we can conclude that GO FAIR achieved a visibility which is much higher than we expected. The main reason is that the instrument we chose, Open Implementation Networks (INs), clearly answered an unmet need.
The VODAN IN: support of a FAIR-based infrastructure for COVID-19

Barend Mons

European Journal of Human Genetics (2020) | Cite this article

189 Accesses | 2 Altmetric | Metrics

The VODAN core consortium

Origin and first weeks

The Virus Outbreak Data Network (VODAN) Implementation Network (IN) was conceived to kick-start a ‘community of communities’ that could design and rapidly build a truly international and interoperable, distributed data network infrastructure that supports evidence-based responses to the viral outbreak. The IN has a longer-term goal to reuse the resulting data and service infrastructure, also for future outbreaks.
Minimal Valuable “EOSC”: 1: basic principles (generic)

- For H2M, M2H and M2M
- Based on FAIR GP
- Architecture according to FIF
- All elements machine-actionable where possible
- As open as possible, as closed as necessary
- As distributed as possible, as centralised as necessary
Minimal Valuable “EOSC”: 2: Basic Implementation Choices

- Based on FAIR GP
- Implementation is FAIR implementation Framework (FIF)
- Architecture FIF based on FD(o)F
- All elements have FAIR metadata, some data are FAIR
- All elements become ‘inter’-actionable via FAIR implementation profiles (FIP)
Minimal Valuable “EOSC”: 3: Corona demonstration use case (0.1 alpha)

- Driving use case: Corona > COVID-19 outbreaks
- Implementation data sources: FAIR data points with Corona relevant data (RWO and EK)
- Architecture for FD(o)F routing: Data Biology
- All elements have FAIR implementation profiles associated
- Funded, done in a small GO FAIR ad hoc Implementation network: VODAN

Q2-3 2020
VODAN adds visitable FAIR RWO data Across borders
The VODAN-IN approach: distributed analytics over FAIR data

- FAIR reference (EK) data stations
- FAIR RWO data stations
- Trains - FAIR algorithms

→ COVID-19 FAIR data model (extensible)
→ FDP-in-a-box
→ PHT-infrastructure
The machine knows what we mean.…
The end of data travel……
Federated AI Ready…..

‘Why did we not already have this’?

Is this a dream
(as some advocate, also in COVID-19 crisis)?
**Disease Modelling Workflow (COVID-19)**

**A1**
- FAIR abstract Paper 1
- FAIR abstract Paper 2
- FAIR abstract Paper 3

**A2**
- Case Report Forms/EPD
- Clinical measurements
- Self Reporting (apps)

**B**
- SARS-CoV-2/human proteome
- ACE2/ACE
- Cytokine Storm (CRS)
- Renin-angiotensin system
- Further sub-hypotheses

**C**
- Greatest Common Denominator (GDC)
- ‘Connectome’

**D**
- Interactive disease Model(s)

**E**
- Expert introduced concept
- Drug, mechanism, cell type etc.

**F**
- Case Report
- Interactive disease Model(s)

**G**
- Expert Curation, annotation and Hypothesis discussion
- Disease phase alignment

**H**
- Community annotation options

**New Publications**
- Real World Observations
- Clinical and Self reporting

**AI-ready Established Knowledge**
- Plus selected connectors Subhypotheses

**FAIR abstract**
- Paper 1
- Paper 2
- Paper 3

**Systematic in silico Rationalisation**

**Systematic in silico Rationalisation**
<table>
<thead>
<tr>
<th>CT characteristics</th>
<th>Normal</th>
<th>Normal</th>
<th>Perfusion defects/GGO/hyperemia</th>
<th>Infarction/GGO (increase)</th>
<th>Interlobular septa, Pulmonary embolism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase</td>
<td>Phase 1 (mild symptoms)</td>
<td>Phase II (onset Cytokine storm)</td>
<td>Phase III Early endothelial/vascular disorders</td>
<td>Phase IV Vascular and thrombotic disorders (severe)</td>
<td>Phase V Severe, life threatening manifestations</td>
</tr>
<tr>
<td>Lung</td>
<td>Cytokine release syndrome, shortness of breath</td>
<td>Endotheliitis/Vasculitis,</td>
<td>Diffuse alveolar damage, microvascular obstruction, Hyperemia, obstruction, Capillary leak syndrome.</td>
<td>Venous thrombosis/permeability disorders, extravasation</td>
<td>Organised pneumonia, pulmonary embolism.</td>
</tr>
<tr>
<td>Heart</td>
<td>Arrhythmia, tachycardia</td>
<td></td>
<td></td>
<td>Myocardial dysfunction</td>
<td>Heart failure</td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td></td>
<td></td>
<td>Stroke</td>
<td></td>
</tr>
<tr>
<td>Kidney</td>
<td></td>
<td></td>
<td></td>
<td>Acute renal failure</td>
<td></td>
</tr>
<tr>
<td>General symptoms/manifestations</td>
<td>Nausea, shortness of breath, dyspnea, headache, tachycardia, exanthema, coughing, fever, hypotension</td>
<td>Macrophage activation syndrome, cell injury, Histiocytosis, Hematophagocytic, reactive hematophagocytic syndrome, Lymphohistiocytosis, hemo...</td>
<td>Acute inflammatory disease, systemic inflammatory response</td>
<td>Thrombus, blood coagulation disorders, hemorrhages</td>
<td>Septicemia, hemodynamic instability, DIC, diffuse hemorrhages, septic shock, toxic shock syndrome, anaphylaxis, multiple organ failure</td>
</tr>
<tr>
<td>Direct interaction SARS-COV-2 proteome-human proteome</td>
<td>GGCX, SLC9A3R1, HDCA2</td>
<td>GNG5, RHOD, FGFR1OP, CNTRL, PRKACA, ITGB1, HMOX1, IL17RA, PABPC1, LOX</td>
<td>PTPN11, CD14, CD28, TLR4, MPO, C5, IFNG, TLR3</td>
<td>CCL3, CSF3, CXCL10, GPX1, KNG1, IKK1, HMOX1, PPARA, VEGFA, TGFB1, APP</td>
<td></td>
</tr>
<tr>
<td>Genes affected in human proteome</td>
<td>CRP, CSF2, CD19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processes</td>
<td>Viral induction?</td>
<td>Cytokine release syndrome</td>
<td>Transition</td>
<td>Onset cytokine storm</td>
<td>Downstream effects renal-angiotensin disturbance</td>
</tr>
<tr>
<td>Small molecules</td>
<td>Glutathione</td>
<td>Myeloperoxidase</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ACE2/ACE2
GGCX
HDAC2
CSF2
CRP
RAS
Cytokine release syndrome
Mild symptoms
Vasculitis inflammation
Intermediate actors
Leakage
Thrombosis
47 genes (mostly immune actors)
Multiple Organ failure
Intermediate actors
10 downstream genes
2 downstream genes
IL6
IC-life threatening & interventions
Drug a-g
Drug g-m
Drug m-t
Hospital admission?
Hospital discharge?
Disease progression
Mild symptoms
Vasculitis inflammation
Cytokine release syndrome
Leakage
Thrombosis
47 genes (mostly immune actors)
Multiple Organ failure
Intermediate actors
Drug a-g
Drug g-m
Drug m-t
IC-life threatening & interventions
Recovery
Lung
Heart
Brain
Kidney
Liver
Fatal outcome
Organ Specific
Where needed
Dinoprostone connectome with disease model: filter: genes only
Dinoprostone connectome with disease model: filter: diseases only

- dyspnea
- exanthema
- cardiac arrhythmia
- tachycardia
- nausea
- headache
- dry cough
- hypotension
- fever

- vasculitis
- macrophage activation syndrome
- cell injury
- histiocytosis

- hyperemia
- diffuse alveolar damage
- capillary leak syndrome

- acute inflammatory disease
- systemic inflammatory response

- myocardial dysfunction
- blood coagulation disorder
- hemmorhage
- exsanguination

- venous thrombosis
- extravasation

- pulmonary embolism
- organized pneumonia
- heart failure
- cerebrovascular accident
- septicemia
- haemodynamic instability
- septic shock
- toxic shock syndrome
- anaphylaxis
- multiple organ failure

- cytokine release syndrome
- cytokine storm

- dinoprostone
The VODAN-IN approach: distributed analytics over FAIR data

- FAIR reference (EK) data stations
- FAIR RWO data stations
- Trains - FAIR algorithms