XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXX Aluna: Gina Szajnbok Harari

XXX TIA: 72008075 - CURSO MESTRADO EM CIÊNCIAS DA COMPUTAÇÃO

XXX Universidade Presbiteriana Mackenzie

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

**Relatório das sub-Tarefas para o Projeto de Final de Curso**

**De 11/06 a 18/06**

**- Planejamento da metodologia para a elaboração da Ontologia**

**- Utilização do Prótegé**

**De 19/06 a 25/06**

**- Execução da formatação dos textos de apoio.**

**De 26/06 a 28/06**

**- Inclusão de Data Proprieties e Execução da formatação dos textos de apoio.**

**De 29/06**

**- Execução da formatação dos textos de apoio.**

**De 30/06 a 07/07**

**-Revisão Geral da Ontologia.**

**-Verificação de Erros.**

**-Aguardando suporte da OWL para proceder a solução dos erros com o *Reasoner*.**

**-Inclusão de mais algumas *Data Proprieties* e *Individuals*.**

**-Revisão da Estratégia 5W4H.**

**Planejamento da metodologia para a elaboração da Ontologia**

1. **Motivação Histórica.**

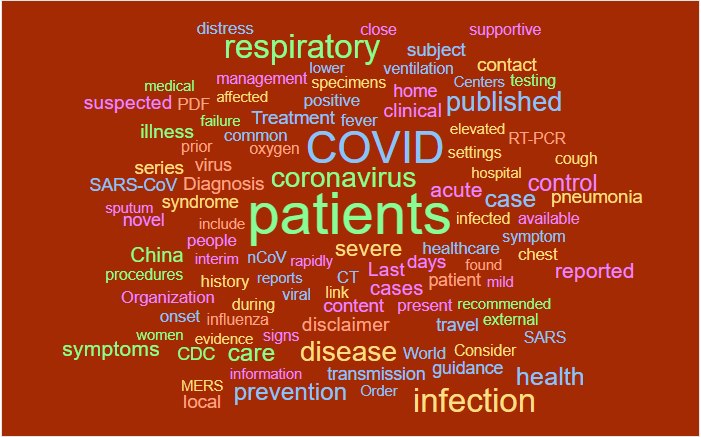
Selecionar e verificar a existência de Ontologistas ao longo da história.<Ontologistas.docx> Ex. Spinoza, Russel.

1. **Seleção do texto básico sobre COVID-19para produzir os Objetos, Classes, e Propriedades da Ontologia:**

BMJ best Practice COVID-19. **BMJ Publishing Group Ltd. 2020**, London. Disponível em:

<https://bestpractice.bmj.com/topics/en-gb/3000168/pdf/3000168/Coronavirus%20disease%202019%20%28COVID-19%29.pdf>. Acesso em: 17 de jun. de 2020.

Resultado: 98 palavras.



acute affected available care case CDC Centers chest China clinical close common Consider contact content control coronavirus cough COVID CT days Diagnosis disclaimer disease distress during elevated evidence external failure fever found guidance health healthcare history home hospital illness include infected infection influenza information interim Last link local lower management medical MERS mild nCoV novel onset Order Organization oxygen patient patients PDF people positive present prevention prior procedures published rapidly recommended reported reports respiratory RT-PCR SARS-CoV series settings severe signs specimens sputum subject supportive suspected symptom syndrome testing transmission travel Treatment ventilation viral virus women World

1. **Seleção de textos de apoio1.**

Os textos de apoio serão retirados das seguintes fontes:

1. Referências bibliográficas da apresentação do Pré-Projeto da Dissertação da Disciplina de Metodologia de Trabalho Científico. (O Pré-Projeto possui 123 referências, o que tornou inviável para esta entrega. Na apresentação do Pré-Projeto foram selecionados os textos mais atuais e relevantes). <Apresentação_projeto_V14.pdf>, quais sejam:

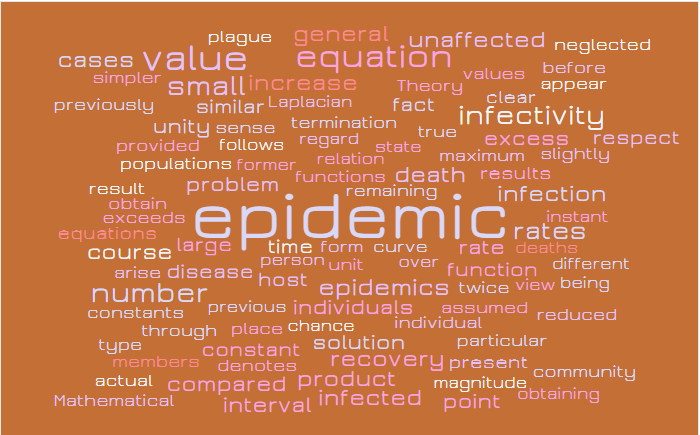
\_

**1 Observação importante**: Percebeu-se, por causa dos textos serem científicos, que eles possuem *stopwords* próprios. Após alguns testes com os textos selecionados, percebeu-se que alguns deles, embora científicos, não serviram como base para a criação da Ontologia, mas tão somente como texto de apoio, pois o objetivo do texto não era conceitual e sim procedural, com execuções e resultados. Percebeu-se também que algumas *stopwords* eram válidas ou não, dependendo de textos específicos, como por exemplo, a palavra “**others**” **não é** *stopword* para o texto de FERRAZ e MONTEIRO (2019). Isto porque o texto trata de “morte por outras causas”. No entanto ela é *stopword* para o texto de CHANG (2020), pois é um pronome comumente utilizado. Isto foi analisado em todos os textos de apoio para que não fosse imposto um viés na API do WordItOut. Percebeu-se também a necessidade de inclusão de símbolos matemáticos nas *stopwords.* As letras gregas, comumente utilizadas nos textos matemáticos também foram consideradas *stopwords.*

**- Trabalhos seminais em epidemiologia**

KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character, The Royal Society London, v. 115, n. 772, p.700–721, 1927.

Resultado: 89 palavras

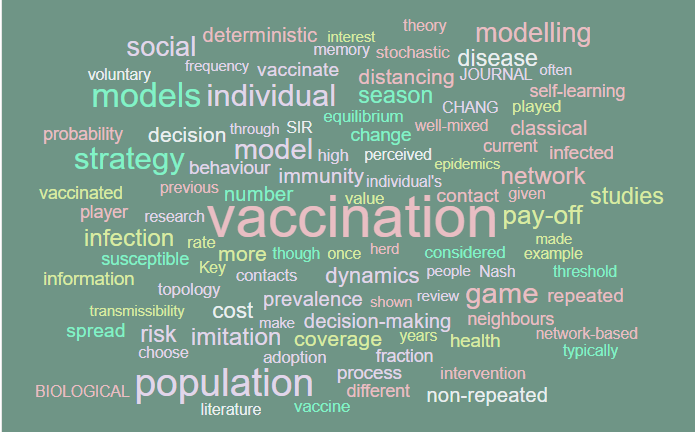


actual appear arise assumed before being cases chance clear community compared constant constants course curve death denotes different disease epidemic epidemics equation equations exceeds excess fact follows form former function general host increase individual infected infection infectivity instant interval Laplacian large magnitude Mathematical maximum members neglected number obtain obtaining over particular person place plague point populations present previous previously problem product provided rate recovery reduced regard relation remaining respect result sense similar simpler slightly small solution state termination Theory through time true twice type unaffected unit unity value view

**- Trabalhos recentes em epidemiologia, Sistemas Dinâmicos, *OSN (Online Social Network),* Teoria dos Jogos (Revisão Sistemática), Base de Dados**

CHANG, Sheryl L. et al. Game theoretic modelling of infectious disease dynamics and intervention methods: a review. Journal of Biological Dynamics, v. 14, n. 1, p. 57-89, 2020

Resultado: 97 palavras



adoption behaviour BIOLOGICAL CHANG change choose classical considered contact contacts cost coverage current decision decision-making deterministic different disease distancing dynamics epidemics equilibrium example fraction frequency game given health herd high imitation immunity individual individual's infected infection information interest intervention JOURNAL Key literature made make memory model modelling models more Nash neighbours network network-based non-repeated number often once pay-off people perceived played player population prevalence previous probability process rate repeated research review risk season self-learning shown SIR social spread stochastic strategy studies susceptible theory though threshold through topology transmissibility typically vaccinate vaccinated vaccination vaccine value voluntary well-mixed years

FAN, D. et al. Novel fake news spreading model with similarity on PSO - based networks. Physica A: Statistical Mechanics and its Applications, Elsevier, p. 124319, 2020.

Resultado: 83 palavras



analysis angular average Based clustering coefficient connected constant contact correlation critical decreases degrees denoted denotes densities density different disk distance distribution dynamics each effect epidemic Eq evolution exponent fake final follows function generated heterogeneous hyperbolic ignorants increases individual individuals interacts larger maximum model modeling models Monte-Carlo new news node obtained other pair paper Pearson Physica popular popularity power-law probability product proposed PSO-based radius rate real-world represent results set show simulation SIR smaller social Song spread spreaders standard stifler subsection three threshold time while

MISTRY, Dina et al. Inferring high-resolution human mixing patterns for disease modeling. arXiv preprint arXiv:2003.01214, 2020.

Resultado: 100 palavras



age approach attack available average based both Canada case census characteristics child children China community composition contact contacts countries country data demographic derived detailed different disease diseases distance distributions dynamics economic employment epidemic estimated framework frequency general geographical household households human increased India individual individuals infectious influenza Japan level linear location locations macro matrices matrix members Methods micro mixing model modeling models multiple national number other patterns population populations Province rate rates relevant reported resulting results Russia sample school schools setting settings show similar size social socio specific States statistics structure structured study sub-national survey synthetic transmission United well world

**- Trabalhos recentes efetuados na UPM**

FERRAZ, D. F.; MONTEIRO, L. H. A. The impact of imported cases on the persistence of contagious diseases. Ecological Complexity, v. 40, p. 100788, 2019.

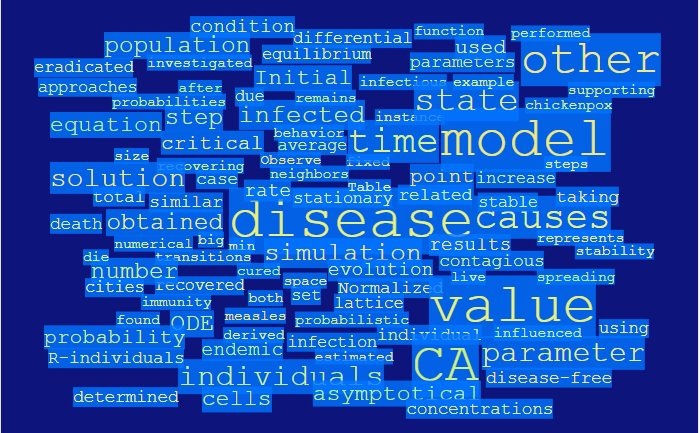
Resultado: 79 palavras



account affected analytical analytically approaches approximation Assume asymptotic asymptotically automaton balanced basic bifurcation computed confirmed constant contact contagious control death derived disease edges effective eigenvalue estimated found geographical gives health high host I-individual immune impact individual infection infectious instance lattice line mean-field measles model neighbors numerical occurs ODE parameter PCA per point proposed rare rate region related reproduction result rule secondary section sick single solid solution stable state stationary step steps susceptible system taking time transcritical vaccinated vaccination value

MONTEIRO, L. H. A.; CHIMARA, H. D. B.; BERLINCK, JG Chaui. Big cities: shelters for contagious diseases. Ecological Modelling, v. 197, n. 1-2, p. 258-262, 2006.

Resultado: 98 palavras



after approaches asymptotical average behavior big both CA case causes cells chickenpox cities concentrations condition contagious critical cured death derived determined die differential disease disease-free due endemic equation equilibrium eradicated estimated evolution example fixed found function immunity increase individual individuals infected infection infectious influenced Initial instance investigated lattice live measles min model neighbors Normalized number numerical Observe obtained ODE other parameter parameters performed point population probabilistic probabilities probability rate recovered recovering related remains represents results R-individuals set similar simulation size solution space spreading stability stable state stationary step steps supporting Table taking time total transitions used using value

MORAES, A. L. S.; MONTEIRO, L. H. A. On considering the influence of recovered individuals in disease propagations. Communications in Nonlinear Science and Numerical Simulation, v. 34, p. 224-230, 2016. – utilizado como consulta.

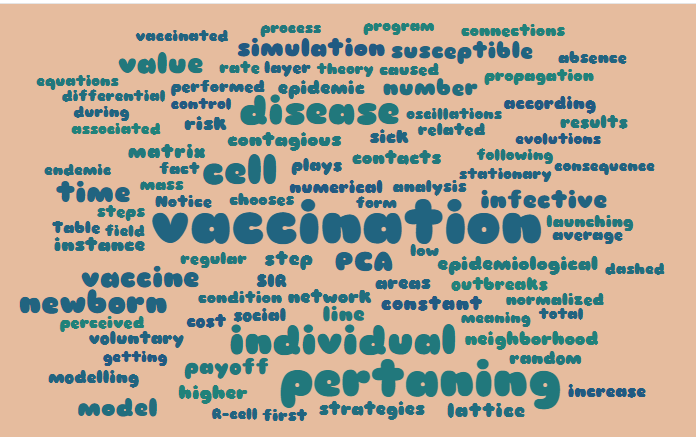
Resultado: 95 palavras



affect annual attractor becomes Belgium bifurcation case childhood children comparing condition constant contagious corresponds countries cycle dashed data day death diagram disease disease-free eigenvalue either endemic epidemic found free Germany given global Hopf host immune incidence increases increasing infections infective influence initial Italy limit line locally model normalized Note number numerically Observe obtained one opposite other parameter people point positive predictions proposed rate real records recovered recovery reduces related results R-individuals sick simulations S-individuals Solid solution space spread stability stable state stationary steady steady-state susceptible system terms time total transcritical unstable vaccination varicella varying year

SCHIMIT, P. H. T.; MONTEIRO, L. H. A. A vaccination game based on public health actions and personal decisions. Ecological Modelling, v. 222, n. 9, p. 1651-1655, 2011.

Resultado: 90 palavras

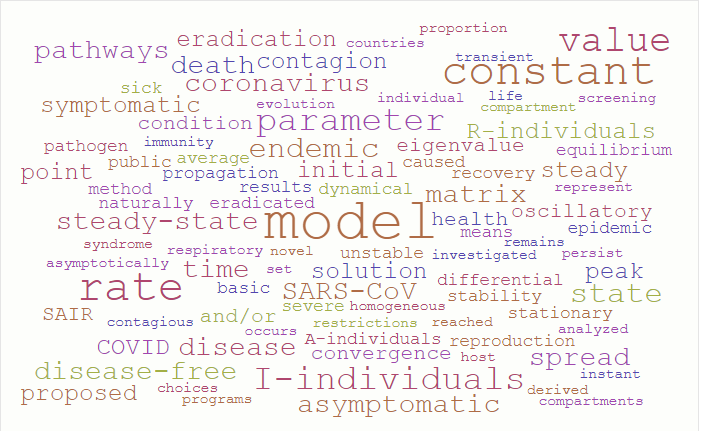
****

absence according analysis areas associated average caused cell chooses condition connections consequence constant contacts contagious control cost dashed differential disease during endemic epidemic epidemiological equations evolutions fact field first following form getting higher increase individual infective instance lattice launching layer line low mass matrix meaning model modelling neighborhood network newborn normalized Notice number numerical oscillations outbreaks payoff PCA perceived performed pertaning plays process program propagation random rate R-cell regular related results risk sick simulation SIR social stationary step steps strategies susceptible Table theory time total vaccinated vaccination vaccine value voluntary

1. Será incluído na análise o seguinte texto recém publicado:

MONTEIRO, L. H. A. An epidemiological model for SARS-CoV-2. Ecological Complexity, p. 100836, 2020.

Resultado: 90 palavras



A-individuals analyzed and/or asymptomatic asymptotically average basic caused choices compartment compartments condition constant contagion contagious convergence coronavirus countries COVID death derived differential disease disease-free dynamical eigenvalue endemic epidemic equilibrium eradicated eradication evolution health homogeneous host I-individuals immunity individual initial instant investigated life matrix means method model naturally novel occurs oscillatory parameter pathogen pathways peak persist point programs propagation proportion proposed public rate reached recovery remains represent reproduction respiratory restrictions results R-individuals SAIR SARS-CoV screening set severe sick solution spread stability state stationary steady steady-state symptomatic syndrome time transient unstable value

1. **Metodologia para a formalização das classes, subclasses, membros e propriedades**

**4.1 Contagem de palavras**

Uma API 2 será utilizada para contagem das palavras mais utilizadas em todos os textos compactados no arquivo [Seleção de todas as palavras dos textos - cada texto eh uma coluna.xlsx](Seleção%20de%20todas%20as%20palavras%20dos%20textos%20-%20cada%20texto%20eh%20uma%20coluna.xlsx) e posteriormente essas palavras serão pesquisadas no texto básico e se transformarão em classes e as relações serão efetuadas manualmente via análise das classes e subclasses.

Serão utilizados os seguintes **limites máximos** como característica da Ontologia:

**52** classes e subclasses,

**33** propriedades para cada classe/subclasse 3.

* 1. **Ferramenta 5W4H**

Será aplicada a ferramenta **5W4H** para cada elemento elencado pelo software, usando a seguinte transformação:

– WHAT -> o quê – “Classes, subclasses e membros” (Objeto)

– WHY -> Por quê – “Propriedades” (Justificativa)

– HOW -> Como fazer - “Propriedades” (Estratégia)

– WHEN -> Quando - “Propriedades” (Unidade temporal)

– WHERE -> Onde - “Propriedades” (Localização Geográfica)

– HOW FEEL -> Para quê - “Propriedades” (Resultados esperados)

– WHO -> Para quem - “Classes, subclasses e membros” (Público Alvo)

– HOW MUCH -> Quanto - “Propriedades” (Unidade de valor)

– HOW MEASURE -> Como medir - “Propriedades” (Monitoramento, controle, parametrizações)

2 Worditout (<https://worditout.com/word-cloud/create>)

3 Número passível de modificação, conforme o desenvolvimento da tarefa.

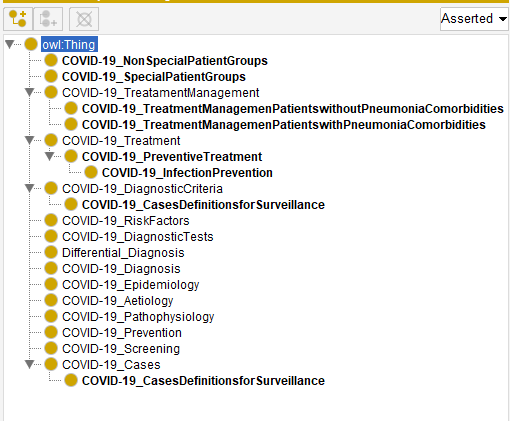
* 1. **Exemplos**

Exemplos gerais do 5W4H:

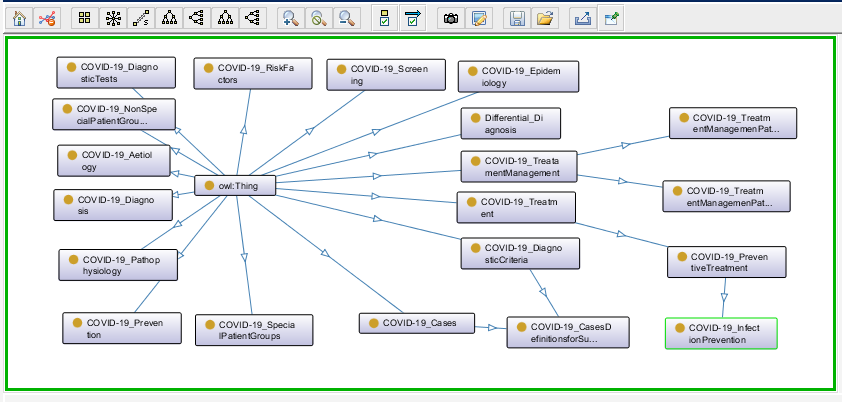
1. (WHAT) Coronavirus disease 2019 (COVID-19) - is a - Severe acute respiratory infection **[Subclasses]**
2. (HOW) Coronavirus disease 2019 (COVID-19) – was caused by - severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) **[proprieties]**
3. (WHERE) Coronavirus disease 2019 (COVID-19) – was started - in Wuhan City, Hubei Province, China **[proprieties]**
4. (WHEN) Coronavirus disease 2019 (COVID-19) – was started - in December 2019 **[proprieties]**
5. (WHAT) Middle East respiratory syndrome (MERS) – is a - Respiratory infection **[Subclasses]**
6. HOW MUCH) - Respiratory infection – has as a cause – a mild common cold-like illness **[proprieties]**
7. (HOW MUCH) - Respiratory infection – has as a cause - a severe viral pneumonia that is potentially fatal **[proprieties]**
8. (WHO) SARS-CoV-2 - is the name of - the virus owing to the virus's genetic similarity to the SARS-CoV vírus disease spectrum **[proprieties]**
9. (HOW MEASURE) Disease spectrum - is diferent than - SARS-CoV vírus disease **[Subclasses disjuntas]**
10. (HOW MEASURE) Disease transmission – is diferent than - SARS-CoV vírus disease **[Subclasses disjuntas]**

**Utilização do Prótegé**

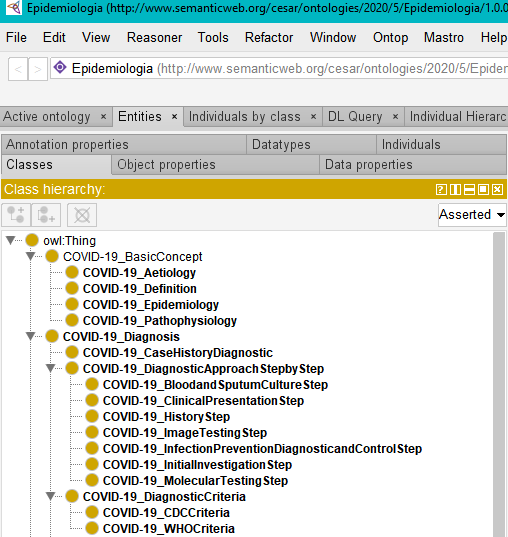
1. **Criação de Classes, Subclasses, Propriedades, Subpropriedades, membros e anotações no Software Protégé.**

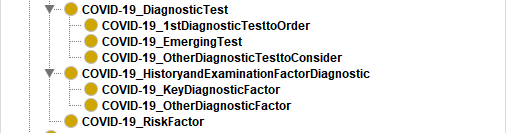
****

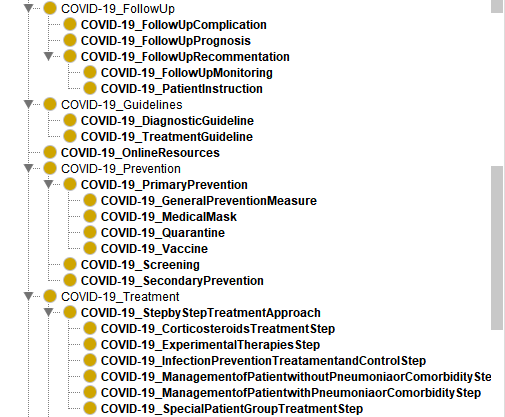
1. **Execução de testes iniciais (somente com algumas classes).**

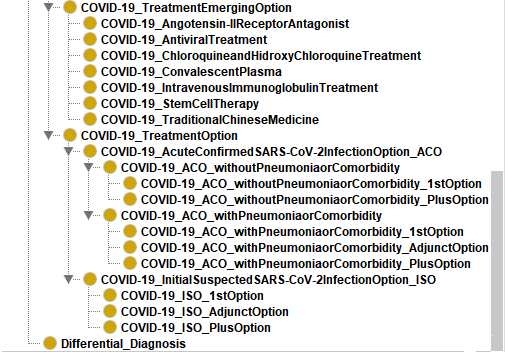
****

1. **Execução dos testes finais (com todas as classes, Propriedade dos Objetos, e Propriedade dos dados).**

****

****

****

****

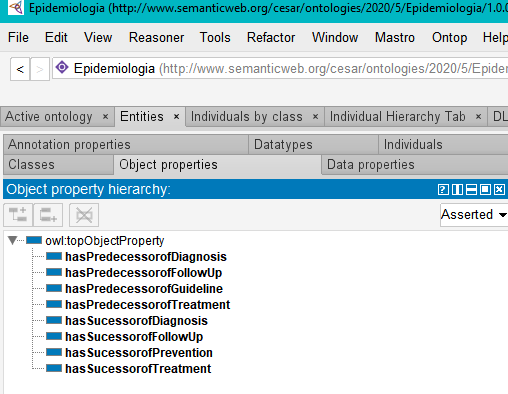
**Não há relação de Predecessor e Sucessor entre a classe BasicConcept e qualquer outra classe, no entanto, entre as outras classes isso ocorre porque elas fazem parte do seguinte fluxo:**

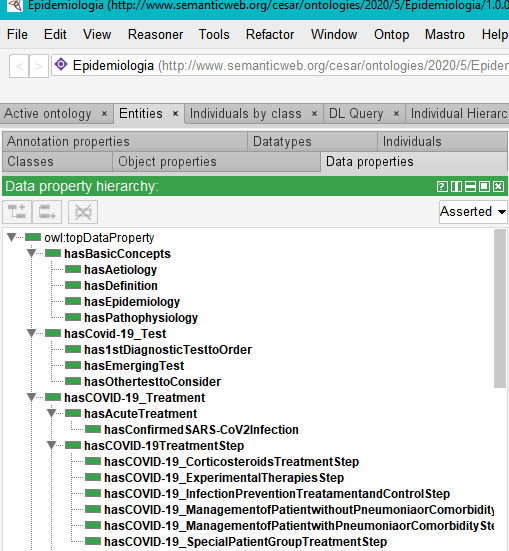
**- Predecessor**

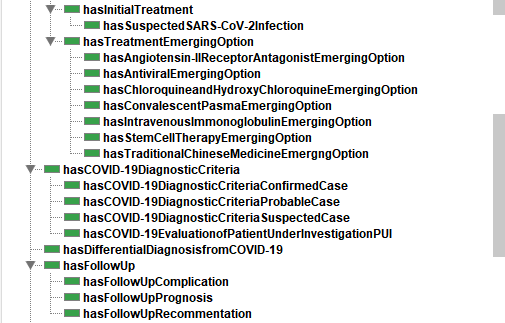
**Prevention -> Diagnosis -> Treatment -> Follow up -> Guidelines**

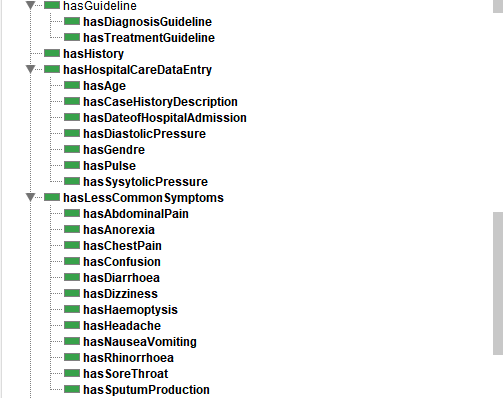
**- Sucessor**

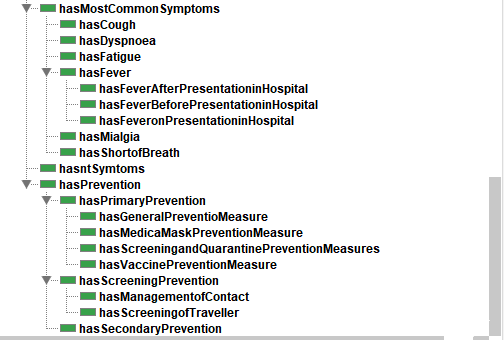
**Guidelines -> Follow up -> Treatment -> Diagnosis -> Prevention**

****

****

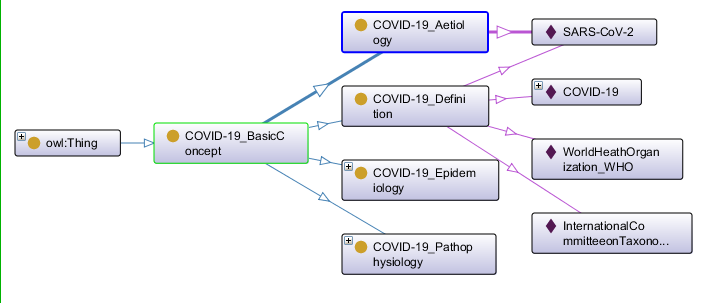
****

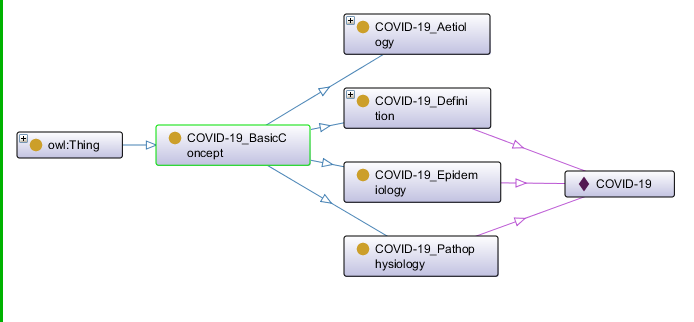
****

****

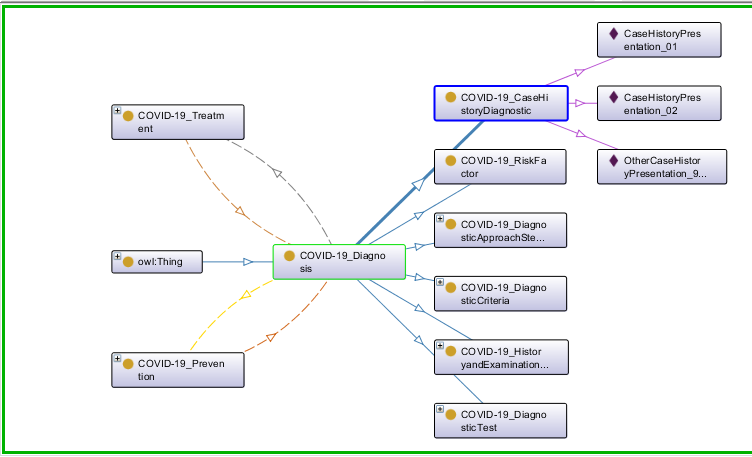
**Windows > Views > Class views > Ontograf**

**BasicConcept**

****

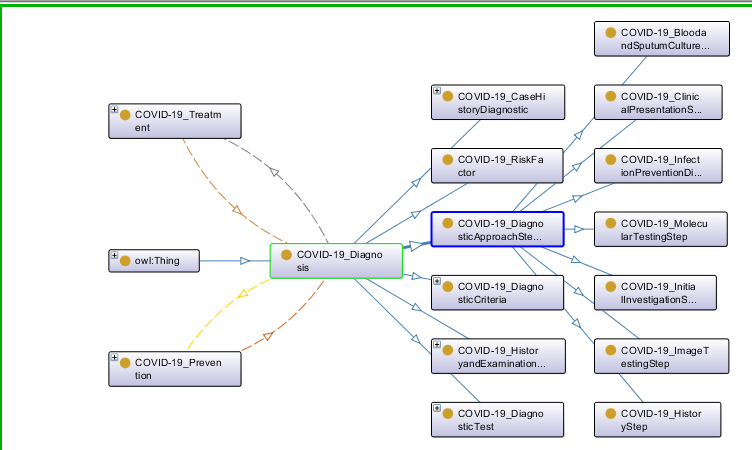
****

**Diagnosis**

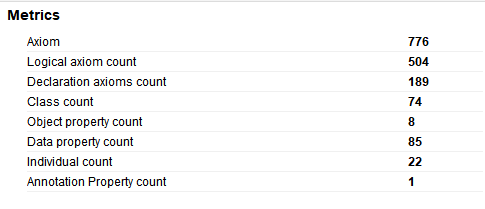
****

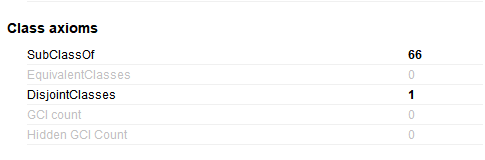
**DiagnosticApproachStep**

**Essa foi a Estratégia adotada para que procedimentos com os pacientes se relacionassem com os DiagnosisSteps. A solução não foi trivial. Foi alcançada graças às características do Protege, uma delas é de ser mais “maleável” sob este aspecto. A palavra “Step” acabou servindo como um “Complemento Nominal” que uniu tudo.**

****

Os **limites máximos** não foram mantidos, para que a qualidade da Ontologia fosse satisfeita.

****

****

1. **Github**

Projeto Final da Ontologia:

<https://github.com/Gina-Szajnbok-Harari/MESTRADO_MACK_2020_ISMAR/blob/master/Ontology_project_GSH_72008075>

Arquivo:

Phase%205%20-%20Final%20Project\_Epidemiology\_RDF\_XML\_V5.owl

Atividades (Última versão):

Atividades do dia 30 junho a 07 julho.docx

Texto Base:

2020\_BMJ%20best%20Practice%20COVID-19.%20BMJ%20Publishing%20Group%20Ltd.%202020,%20London\_COVID-1.pdf

Demais arquivos: Textos auxiliares, arquivos de fases de testes e elaboração.