MODEL 3 (GEO):

pip install sense2vec

prodigy sense2vec.teach geo3_method /Users/sashaqanderson/Dropbox/USGS/NER_Work/s2v_old --seeds "water_flow"

prodigy sense2vec.teach geo3_method /Users/sashaqanderson/Dropbox/USGS/NER_Work/s2v_old --seeds "heat_transfer"

prodigy sense2vec.teach geo3_method /Users/sashaqanderson/Dropbox/USGS/NER_Work/s2v_old --seeds "subduction_zone"

prodigy sense2vec.teach geo3_method /Users/sashaqanderson/Dropbox/USGS/NER_Work/s2v_old --seeds "mass_transfer"

prodigy sense2vec.teach geo3_method /Users/sashaqanderson/Dropbox/USGS/NER_Work/s2v_old --seeds "soil_erosion"

prodigy sense2vec.teach geo3_method /Users/sashaqanderson/Dropbox/USGS/NER_Work/s2v_old --seeds "geothermal_energy" bad one! Don't use again.

prodigy terms.to-patterns geo3_method --label METHOD

{"label":"METHOD","pattern":[{"lower":"water_flow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"reservoir|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"flow_rate|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_level|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"waterflow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_table|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_movement|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"evaporation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"drainage|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"reservoirs|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"surface_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"aquifer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sediment|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ground_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_surface|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_levels|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"groundwater|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"aquifers|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"seawater|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"waste_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sediments|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_column|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"salinity|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fresh_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"pollutants|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"particulates|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"dissolved_oxygen|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_tables|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"salt_water|NOUN"}]}

{"label":"METHOD","pattern":[{"lower":"rain_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"evaporated_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"silt|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"contaminants|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_system|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"saltwater|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"compaction|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rainfall|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"nutrient_solution|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"gas_exchange|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"particulate_matter|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"river_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rainwater|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"surface_waters|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"wastewater|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"dissolved_gasses|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"underground_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_cycle|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"particulate|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"basins|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"oceans|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"snowmelt|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_floor|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"thermal_mass|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"source_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"temperature_gradient|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"basin|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"evaporation_rate|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"algae_blooms|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"phytoplankton|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"algal_blooms|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"freshwater|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"organic_matter|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ozone|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"waterways|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"lake_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"contaminates|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sedimentation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sewage_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_waters|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"seafloor|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanic_ash|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_floor|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_ways|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"eutrophication|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_currents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_surface|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"saline_water|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"turbidity|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"greenhouse_gasses|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"topsoil|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"microorganisms|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"acidification|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mass_transfer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"heat_transfer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"heat_exchange|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fluid_flow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"thermal_energy|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"Evaporation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"heat_equation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"Thermodynamics|ORG"}]}
{"label":"METHOD","pattern":[{"lower":"finite_element_analysis|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"radioactive_decay|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"energy_transfer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"gas_laws|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"heat_flow|NOUN"}]}

{"label":"METHOD","pattern":[{"lower":"thermal_radiation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"boundary_layer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ionization|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"atmospheric_pressure|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"temperature_gradients|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"laminar_flow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"chemical_reactions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"nuclear_decay|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mass_flow_rate|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"magnetic_field|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mass_flow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"chemical_reaction|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"magnetic_fields|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"magnetic_flux|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"biological_systems|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"solar_radiation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"pressure_gradient|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"heat_radiation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"convection_currents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"gravitational_fields|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"EM_radiation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"electromagnetic_radiation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"EM_waves|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"atmospheric_composition|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"upper_atmosphere|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"electromagnetic_waves|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"EM_fields|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"subduction_zone|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"Pacific_ocean|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fault_lines|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ice_sheets|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fault_line|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"continental_shelf|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coastal_areas|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"subduction_zones|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"jet_stream|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"subsidence|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tectonic_plate|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tectonic_plates|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"continental_crust|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coastline|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"subduction|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"caldera|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"glaciers|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coastlines|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"oceanic_crust|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rising_seas|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coastal_regions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"gulf_stream|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rising_sea_levels|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanic_activity|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"subducted|VERB"}]}
{"label":"METHOD","pattern":[{"lower":"wave_action|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ice_shelf|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rain_shadow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_level_rise|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"floodplain|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coast_lines|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanism|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"marshland|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"polar_regions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"waterspouts|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"glacier|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tectonic_activity|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mountain_range|NOUN"}]}

{"label":"METHOD","pattern":[{"lower":"quakes|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"boreal_forests|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_ice|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"wind_patterns|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_level|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_levels|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanoes|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mountain_ranges|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"shorelines|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"seabed|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"continental_plates|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"jet_streams|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coast_line|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"plate_tectonics|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tropics|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"polar_caps|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"deserts|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"lakes|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"temperate_zones|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"polar_ice_caps|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"lagoons|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"wetlands|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"weather_patterns|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanic_eruptions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"glaciation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rivers|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"reefs|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"currents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"dust_storms|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"lava_flow|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"polar_ice|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"earthquakes|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"grasslands|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanic_eruption|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"volcanos|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tectonic_movement|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"liquefaction|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_life|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"hydrothermal_vents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"coral_reefs|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ice_caps|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"monsoons|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tsunamis|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"vegetation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"precipitation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"geysers|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rain_forests|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"desertification|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"climates|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rainforests|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea-level|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ice_age|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ice_cap|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rock_formations|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"snowpack|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_levels|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"droughts|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"weather_systems|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mountains|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"marshlands|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"hurricanes|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"seismic_activity|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"wind_currents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"water_currents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"marine_life|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ice_ages|NOUN"}]}

{"label":"METHOD","pattern":[{"lower":"rain_forest|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"icecaps|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"flash_floods|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"watersheds|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"dams|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"wildfires|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"marshes|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"icebergs|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"geothermal_heat|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"meteor_impacts|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"forests|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"eruptions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_walls|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"air_currents|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"tropical_rainforest|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"rainforest|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"sea_bed|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"land_bridges|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"prairies|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ocean_acidification|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"geothermal_energy|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"soil_erosion|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"heat_transfer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"thermal_transfer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"mass_transfer|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"temperature_differential|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"surface_tension|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"pressure_differential|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"pressure_differentials|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"soil_erosion|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"deforestation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"biodiversity|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"overfishing|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ecological_impact|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"erosion|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"pollution|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"overgrazing|VERB"}]}
{"label":"METHOD","pattern":[{"lower":"reforestation|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"clear_cutting|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"acid_rain|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"carbon_sequestration|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"wildlife_habitat|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fertilizers|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"crops|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"oil_spills|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"greenhouse_gases|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"ecosystems|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"agriculture|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"green_house_gases|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"habitats|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"food_supply|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"habitat|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fossil_fuels|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"invasive_species|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"greenhouse_gas_emissions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"air_pollution|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"CO2_emissions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fracking|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"native_species|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fossil_fuel_emissions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"carbon_emissions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"biosphere|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"fossil_fuel|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"methane_emissions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"green_house_gasses|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"overpopulation|NOUN"}]}

{"label":"METHOD","pattern":[{"lower":"greenhouse_emissions|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"carbon_emission|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"biomass|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"polution|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"geothermal_energy|NOUN"}]}
{"label":"METHOD","pattern":[{"lower":"geothermal|ADJ"}]}
{"label":"METHOD","pattern":[{"lower":"solar_energy|NOUN"}]}

prodigy db-out geo3_method > /Users/sashaqanderson/Dropbox/USGS/NER_Work/geo3method_patterns.jsonl

Prodigy ner.manual geo3method_data blank:en /Users/sashaqanderson/Dropbox/USGS/NER_Work/ner_text_train40.jsonl --label METHOD

prodigy db-out geo3method_data > /Users/sashaqanderson/Dropbox/USGS/geo3method_data_model.jsonl

python -m spacy pretrain /Users/sashaqanderson/Dropbox/USGS/NER_Work/ner_text_train40.jsonl en_vectors_web_lg /Users/sashaqanderson/Dropbox/USGS/NER_Work/pretrain_geo3model/ --use-vectors

Prodigy ner.batch-train geo3method_data en_vectors_web_lg --init-tok2vec /Users/sashaqanderson/Dropbox/USGS/NER_Work/pretrain_geo3model/model999.bin --output geo3_model --eval-split 0.2 --label METHOD

53% Accuracy (oops – I didn't pretrain it first)

Prodigy ner.make-gold geo3method_data_correct ./geo3_model /Users/sashaqanderson/Dropbox/USGS/NER_Work/ner_text_train40.jsonl --label GEO

Prodigy ner.batch-train geo3method_data_correct en_vectors_web_lg --init-tok2vec /Users/sashaqanderson/Dropbox/USGS/NER_Work/pretrain_geo3model/model999.bin --output geo3_model --eval-split 0.2 --label GEO --n-iter 20

69% (pretraining + make-gold)

Prodigy ner.make-gold geo3method_data_correct ./geo3_model /Users/sashaqanderson/Dropbox/USGS/NER_Work/ner_text.jsonl --label GEO

Prodigy ner.print-stream geo3_model /Users/sashaqanderson/Dropbox/USGS/NER_Work/ner_text_train40.jsonl --label GEO

Ran into a problem with model 2 (NER labeled as METHOD and GEO)… start again.