https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Discussions

INSPIRE standards as framework for artificial intelligence applications: a landslides example

Gioachino Roberti[1], Jacob McGregor[1], Sharon Lam[1], David Bigelow[1], Blake Boyko[1], Chris Ahern[1], Victoria Wang[1], Bryan Barnhart[1], Clinton Smyth[1], David Poole[1,2], and Stephen Richard[1]

[1]Minerva Intelligence Inc., 301 – 850 West Hastings Street, Vancouver, British Columbia, Canada V6C 1E1
[2]Computer Science department, University of British Columbia, Vancouver, V6T 1Z4, Canada

Correspondence: Gioachino Roberti (groberti@minervaintelligence.com)

Abstract. This study presents a landslide susceptibility map using an artificial intelligence (AI) approach that is based on standards set by the INSPIRE framework. We show how INSPIRE standards enhance the interoperability of geospatial data, and enable deeper knowledge development for their interpretation and explainability in AI applications. INSPIRE is a European Union Spatial Data Infrastructure (SDI) initiative to standardize spatial data across borders to ensure interoperability for management of cross-border infrastructure and environmental issues. Despite the theoretical effectiveness of the SDI, very few real-world applications make use of INSPIRE standards. We designed an ontology of landslides, embedded with INSPIRE vocabularies and then aligned geology, stream network and land cover data sets covering the Veneto region of Italy to the standards. INSPIRE was formally extended to include an extensive landslide type code list, a landslide size code list and the concept of landslide susceptibility to describe map application inputs and outputs. Using the terms in the ontology, we defined conceptual scientific models of slopes likely to generate landslides as well as map polygons representing real slopes. Both landslide models and map polygons were encoded as semantic networks and, by qualitative probabilistic comparison between the two, a similarity score was assigned. The score was then used as a proxy for landslide susceptibility and displayed in web map application. The use of INSPIRE-standardized vocabularies in ontologies that express scientific models promotes the adoption of the standards across the European Union and beyond. Further, this application facilitates the explainability of the generated results. We conclude that public and private organisations, within and outside the European Union, can enhance the value of their data by bringing them into INSPIRE-compliance for use in AI applications.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Discussions

1 Introduction

1.1 INSPIRE

Data accessibility and interoperability is key for multinational cross-border applications and fundamental for economic development (European Parliament and the Council, 2007). Different countries have different languages and data standards, hindering infrastructure planning, disaster risk reduction initiatives, and effective legislative implementation. To overcome these challenges, the European Union initiated INSPIRE (Infrastructure for Spatial Information in the European Community - Directive 2007/2/EC - (European Parliament and the Council, 2007)). INSPIRE is structured in 34 spatial data themes organized in three annexes. The themes span social (e. g. street addresses) and environmental domains (e. g. geology), and all EU countries are mandated by law to have implemented the data framework by 2021 (European Parliament and the Council, 2014). EU countries are aligning and serving INSPIRE data at a slow pace, and currently relatively few INSPIRE-compliant data sets are available across Europe (Cho and Crompvoets, 2019). Conferences and competitions are currently being organized to promote its implementation and to show the potential impact of real-world applications built on INSPIRE data sets (European Commission, 2019).

1.2 Artificial intelligence

Artificial Intelligence (AI) studies "the synthesis and analysis of computational agents that act intelligently" (Poole and Mackworth, 2017). Part of acting intelligently is building models of the world that make predictions. Probabilistic predictions are the most useful ones for subsequent decision making, and can be learned from data (Pearl, 1988). All models are based on human knowledge and data (observations of the world). For some problem domains, society has collected an overwhelming amount of data and still, useful human knowledge of the domain can be very vague. Machine learning has made great progress recently for such cases, particularly with deep learning (Goodfellow et al., 2016). However, for domains with relatively limited, but still very large in volume, data, human knowledge can complement the data to make useful predictions (Pearl, 1988). Many environmental problems do not have enough data to be solved by deep learning, but do have enough data to generate useful products when combined with human expertise.

1.3 The need for standards, ontologies, and taxonomies

Consistent vocabularies and data standards are essential in computer science applications, especially in AI. For the data to have any meaning, and for multiple datasets to be combined, we need consistent vocabulary that is well-defined. Deep learning techniques require meanings for the inputs and the outputs (often specified in standards such as JPEG, or WAV), but the internal representations do not have well-defined meanings, making the models very opaque (Marcus, 2018). For certain other representations, such as logical and probabilistic representations, the internal reasoning is done on symbols with well-defined meanings, which lend themselves to use in explanations (Marcus and Davis, 2019).

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Ontologies are "a specification of the meanings of the symbols in an information system" (Poole and Mackworth, 2017). In particular, an ontology stores the vocabulary used to define entities and relationships, and defines axioms controlling the use of the voabulary. Given these axioms, the vocabulary is unambiguously interpreted, and the implicit relationships between the entities can be inferred (Poole, 2009). Vocabularies should be, whenever possible, Aristotelean taxonomies. Aristotelean taxonomies are constructed by defining concepts from their relation to a more general parent concept (genus) and using differentiating properties (differentia) to distinguish concepts within the same genus (Aristotle, 350BC). Taxonomies based on Aristotelian definitions tend to be multi-hierarchical and can be used by computers to make logical inferences (Poole, 2009; Smith, 2003). The term 'multi-hierarchical' implies that there is more than one way to move through a taxonomy to arrive at a particular node or term. Knowledge stored in a domain-specific ontology (e.g. geohazards) can be accessed by computers, allowing for data investigation through various artificial intelligence (AI) techniques, including probabilistic matching as for this study.

Significant progress has been made in the development of taxonomies for geoscience information interchange by the IUGS CGI Geoscience Terminology Working Group which produced the GeoSciML standard along with the OGC (CGI, 2003). However, ontology applications in Earth Sciences are scarce. Notable exceptions are in economic geology (Smyth et al., 2007), geohazards (Jackson Jr et al., 2008), and disaster risk reduction domains (Phengsuwan et al., 2019; Sermet and Demir, 2019). The INSPIRE framework, through its standardised vocabularies ("Code Lists"), provides the necessary foundation upon which AI applications with explainable output can be constructed. INSPIRE application examples in landslide studies include the LAND-deFeND Italian landslide database structure (Napolitano et al., 2018) and a deep learning algorithm to map landslide susceptibility (Hajimoradlou et al., 2019). In this implementation of deep learning, training features were labelled with INSPIRE-compliant semantics to enable reproducibility of the experiment by other researchers.

In this study, we present an AI-based landslide susceptibility application using a natural hazard ontology. We do so by building from the ontology created by Jackson Jr et al. (2008), and by embedding INSPIRE code lists wherever possible and by alligning input and outpud data to the INSPIRE data standards.

1.4 Landslide susceptibility and hazard

Landslide susceptibility is defined as the relative spatial probability of occurrence for a landslide based on the intrinsic properties of a site (SafeLand, 2011). The concept of susceptibility differs from hazard in that the temporal probability of occurrence, the triggering factors, and the magnitude of the event are not considered in the definition of a susceptibility map (SafeLand, 2011; Van Den Eeckhaut and Hervás, 2012). Statistical methods, physical methods, and expert-based methods can be applied to produce landslide susceptibility maps (SafeLand, 2011). Statistical methods rely on the analysis of landslide databases and their relation to landscape properties (see review by Reichenbach et al., 2018); physical methods calculate the limit equilibrium between failure resisting forces and driving forces in slopes (e.g., Baum et al., 2008); and expert-based methods rely on expert opinion and the assumption that influencing factors are known and are specified in the models (Dai et al., 2002). The AI approach used in this study mimics the domain-expert reasoning, providing qualitative landslide susceptibility maps.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

2 Methods

2.1 INSPIRE extension

Technical guideline documents outline the data structure for each theme within the INSPIRE directive, its encoding rules, its metadata standards, and some of its use cases. Data structures are formalized as Unified Modeling Language (UML) and Extensible Markup Language (XML) application schemas. These application schemas model spatial objects as feature types (vector-based spatial data), describe properties of each feature type (often with standardised vocabularies) and characterize the relationships between different feature types. As such, each one of these data structures can be understood as an ontology (See Section 2.2 below), by defining various entities and the relationships between them.

INSPIRE data is generally encoded as Geography Markup Language (GML) and can also be provided as Open Geospatial Consortium (OGC) - compliant web services. Feature-type properties have value types (e.g. geometry for vector data sets) and code lists. Code lists store the terminology that can be used in each domain to specify the value attribute, and may incorporate vocabularies developed outside of INSPIRE (e.g. IUGS CGI rock type taxonomy, Natura 2000 And Emerald Bio-geographical Region Classification). Some code lists within INSPIRE are not extensible, some are extensible with narrower values, and some allow additional values at any level. Code list values, definitions and hierarchical structures are stored in the INSPIRE registry, making them accessible to and reusable by anyone. INSPIRE schemas can also be extended to include additional concepts and/or feature types. For this project, we worked with four INSPIRE themes: Geology, Land Cover, Hydrography and Natural Risk Zones. The Natural Risk Zone application schema was not fully adequate for this application as it lacked the 'landslide susceptibility' concept and 'landslide type' code lists (Tomas et al., 2015). We addressed this issue by formally extending the Natural Risk Zone schema and the Natural Hazards code list.

2.2 Ontological-grounded probabilistic matching

The method used to produce INSPIRE-based landslide susceptibility maps, is a probabilistic comparison system that mimics human expert reasoning, making qualitative predictions based on comparisons between models and instances (e.g., Sharma et al., 2010; Smyth et al., 2007; Poole and Smyth, 2005; Smyth and Poole, 2004). Models are expert-based conceptualized descriptions of a given phenomenon or entity (e.g. landslide susceptibility). The properties used in a model description are concepts stored in the ontology (e.g. soil slide – has slope – steep), along with frequency terms (e.g. soil slide – has slope – steep – always). Frequency terms used in this study included: "always", "usually", "sometimes", "rarely" and "never". These terms were chosen as they express experience-based judgements that geoscience practitioners use in field assessments. The term "never" allows the system to explicitly deal with negation (e.g. soil slide - has surficial material - bedrock - never). The properties and the frequency terms are encoded in semantic triple format and the resulting model is a semantic network. Real-world areas on the ground (map units – more generally revered to as "instances") are also described by semantic networks using the same properties stored in the ontology, but they are accompanied by true-or-false qualifiers (e.g. polygon – has slope – steep – true). Comparisons, referred to as matches, between instances and models is possible because models and instances all use the same structured terminology, as controlled by the ontology.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

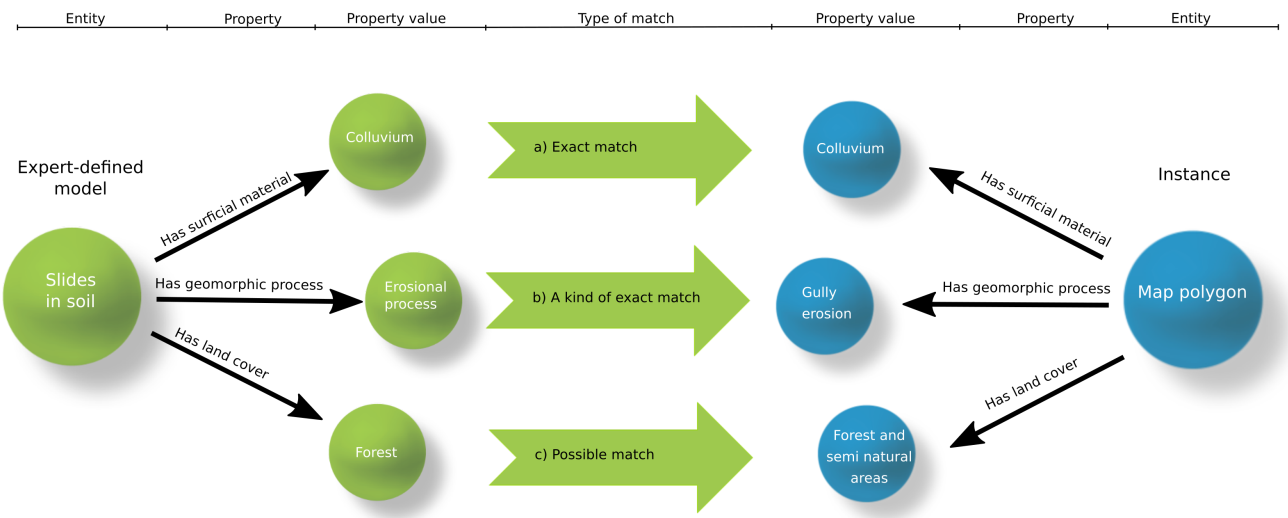


Figure 1. Graphical representation of the matching process between expert-defined models and map polygon instances. a) is an example of an Exact match between the property value "Colluvium"; b) is an example of a kind of (AKO) exact match, because "gully erosion" is a more specific kind of erosional process; c) is an example of a possible exact match because "Forest and semi natural areas" is a broader concept of "Forest". The vocabulary and the hierarchy are controlled by the ontology. Note that frequency terms for model properties are not shown in this figure.

Similarity scores are awarded based on the type of match between instance and model properties, the semantic distance in the taxonomy of compared property values and the model property frequency term (Figure 1). Match types include, Exact, A Kind Of (AKO) exact, and Possible. An Exact match indicates that the term used in the model is present in the instance ('a' in Figure 1), in which case full score is awarded for this component of the compared semantic networks. An AKO exact match indicates that the attribute found in the instance is a kind of the attribute found in the model ('b' in Figure 1), in which case a full score is also awarded. A "Possible" match occurs when the concept in the instance is broader than the concept in the model, based on the defined taxonomies, in which case the score is divided by the semantic distance between the two concepts. For example, 'forest' is a more specific type of 'forest and semi natural areas' ('c' in Figure 1) results in the score being divided by two. The score is lower because the instance is only possibly the kind of value that the model is looking for. In this study, an Exact match or an AKO exact match of a property with frequency "always" scores 10000, "usually" scores 9000, "sometimes" scores 1000, "rarely" scores "100" and "never" scores -10000. For an extensive review of the probabilistic comparison method, see Poole and Smyth (2005), Sharma et al. (2010), Smyth et al. (2007), and Smyth and Poole (2004). This approach has been successfully applied in economic geology to generate mineral deposit exploration targets (Smyth et al., 2007), and in geohazard mapping to produce landslide susceptibility maps (Jackson Jr et al., 2008).

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

2.2.1 Landslide models

This paper presents an AI expert-based landslide susceptibility map for three different landslide types: debris flows, slides in soil, and slides in rock (Hungr et al., 2014) for the Veneto region of Italy. These three landslide types are conceptualizations of landslide models defined using knowledge recorded in the scientific literature and the data available for the Veneto region. Here we briefly summarize the models, see Appendix C for a detailed explanation of each property-property value-frequency combination.

The ‘Debris Flow’ model describes the streams that may generate a debris flow. Debris flows are flow-like landslides generated when saturated sediments move down a stream. They can be originated when a slide in soil intersects a flowing body of water, or when saturated bed sediments are mobilized and begin flowing downstream. Debris flows are usually triggered by intense and persistent rainfall (Hungr et al., 2014).

The ‘Slides in Rock’ model describes slopes that may generate slides in rock. Slides in Rock form when steep rock slopes and cliffs fail under the influence of gravity, commonly triggered by intense rainfall or earthquakes. Slides in rock are usually very fast, and the failure can occur along planar, curved, and/or multiple surfaces. This model represents the collective class of landslides that have as material "rock" and movement type "slide", including rotational, planar, compound, wedge and irregular slides in rock (Hungr et al., 2014). Given the regional scale of this study, we do not have the data resolution to determine the possible failure plane geometry. For example, we cannot identify slopes more susceptible to planar rock slides rather than rotational rock slides.

The ‘Slides in Soil’ model describes slopes that may generate slides in soil. Slides in soil are downslope movements of soil under the influence of gravity, commonly triggered by intense rainfall or earthquakes. They can be slow or fast, and the failure can occur along one or many planar or curved surfaces (Hungr et al., 2014). With Slides in Soil, we refer to the collective class representing all landslides that have as material "soil" and movement type "slide”, including rotational, planar, and compound, clay, silt, sand, gravel, debris slides. Given the regional scale of this study, we do not have the data resolution to determine the possible failure plane geometry and the specific kind of soil that is involved in the failure.

In the presence of higher resolution information such as rock bedding orientation or shear geometry and stratigraphy in soil masses, specific kinds of rock slides (e.g. planar vs rotational) or different kinds of slides in soil (e. g. clay compound slide vs clay planar slide) susceptibility may be mapped.

2.2.2 Mapping unit and runout

The definition of the mapping unit is a critical step in any landslide susceptibility mapping application and there are many different approaches to subdividing the area of interest (see review by Guzzetti et al., 1999). For this study, we used slope units, which are a geomorphic representation of single slopes bounded by drainage and divide lines (Guzzetti et al., 1999), as mapping unit to identify areas susceptible to slides in soil or rock. We used the r.slopeunits software to objectively automate the slope unit delineation (Alvioli et al., 2016). We used stream line vector shapefiles provided by the Veneto Regional Government,

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Table 1. R.avaflow parameters for slides in soil, slides in rock and debris flows runout calculations

Variables (unit)	Slides in Soil	Slides in Rock	Debris Flow
Solid fraction (%)	60	70	60
Fluid fraction (%)	40	30	40
Solid fraction internal friction angle (degree)	18	18	5
Solid fraction basal friction angle (degree)	10	10	4
Fluid fraction internal friction angle (degree)	0	0	0
Fluid fraction basal friction angle (degree)	0	0	0
Solid fraction viscosity (m^2 s-1)	30	30	5
Fluid fraction viscosity (m^2 s-1)	3	3	3

buffered by a distance of 5 m as mapping units to map debris flow susceptibility. In total, the region of Veneto was subdivided into 93,262 polygons, including 9,302 stream buffer polygons and 83,960 slope-unit polygons.

Following the calculation of susceptibility, a first-order estimate of hazard is provided by calculating the likely extent of landslide runout for the most susceptible (highest scoring) instances for each model. Various physical methods have been developed to calculate landslide runout, given the physical properties of the material and the topography (see review by McDougall, 2016). To compute the runout extents, we applied the r.avaflow code (Mergili et al., 2017) which is an open source software package implementing the two-phase debris flow model developed by Pudasaini (2012). Physical model parameters for 'Slides in Rock' are inferred from the back-calculations of the recent Mt. Joffre landslide, in British Columbia, Canada (Friele et al., 2020), 'Slides in Soil' and 'Debris Flow' parameters use the default r.avaflow parameters for those landslide types (Table 1).

2.2.3 Web map

This study's landslide susceptibility maps and hypothetical landslide runouts for slides in soil, slides in rock and debris flows are delivered as an interactive web map based on OpenLayers (MetaCarta, 2005) and React (Facebook, 2013). Input layers are hosted through a Geoserver (The Open Planning Project, 2001) with a PostGIS (Refraction Researtch, 2001) backend database. INSPIRE-aligned layers are hosted on Hale Connect (WeTransform, 2014), a platform used to host and serve INSPIRE-compliant data.

3 Results

3.1 INSPIRE Natural Risk Zones extension

To develop an INSPIRE-compliant AI application to map of landslide susceptibility, we needed to extend the INSPIRE Risk Zones theme to include the concept of landslide susceptibility and the specific code list dealing with landslide terminology.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

The INSPIRE extensions developed in this project are documented and stored in the INSPIRE registry software (Minerva Intelligence, 2019a). The Minerva 'Re3gistry' is a version 1.3.1 instance of the INSPIRE registry based on the Re3gistry software (ISA, 2016). The registry service is packaged within a collection of Docker (Hykes, 2013) containers and hosted on a local server.

The Natural Risk Zone core (NZ-core) schema extension, which includes the Natural Risk Zone Susceptibility feature type was based on SafeLand recommendations (SafeLand, 2011). The 'Natural Hazard Classification' code list was extended to include a classification of various landslide types using the Updated Varnes Landslide Classification (Hungr et al., 2014), which has been widely adopted within the scientific community, the Natural Hazard Classification Landslide Extension' code list (Minerva Intelligence, 2019b), and a new code list of landslide size classes based on Jakob (2005) (see the 'Landslide Size Class' code list (Minerva Intelligence, 2019c)). This code list contains ten landslide size classes based on landslide volume and descriptions of approximate damage potential.

3.1.1 Code list extension

The Natural Hazard Classification code list extension for landslides was prepared using the open access ACE taxonomy editing software (Minerva Intelligence, 2019d). The classification tree considers material type and failure movement (Figure 2). Properties dealing with water content, depth of failure, rate of movement, loading state, channelized state, and failure plane geometry (see Appendix B) are used to further describe the different landslide types. The unique combination of these properties allows for unambiguous classification of the different landslide types into an Aristotelean taxonomy which is multi-hierarchical and amenable to reasoning by both humans and computers. Multi-hierarchical means that, for example, the concept "debris flow" is in different position in the taxonomy depending on the chosen top concept. It can be three level down from the top concept "flow-like landslide" (flow landslides>flows in soil>debris flow) or two level down from the top concept "fast landslides" (fast landslide>debris flow).

The formal extension registration process via the INSPIRE Registry software does not enable the representation of such multi-hierarchical classifications. Because of this we had to work with a single tree hierarchy, and consequently chose to first divide the classes on type of failure followed by a division based on type of movement (Figure 2).

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions

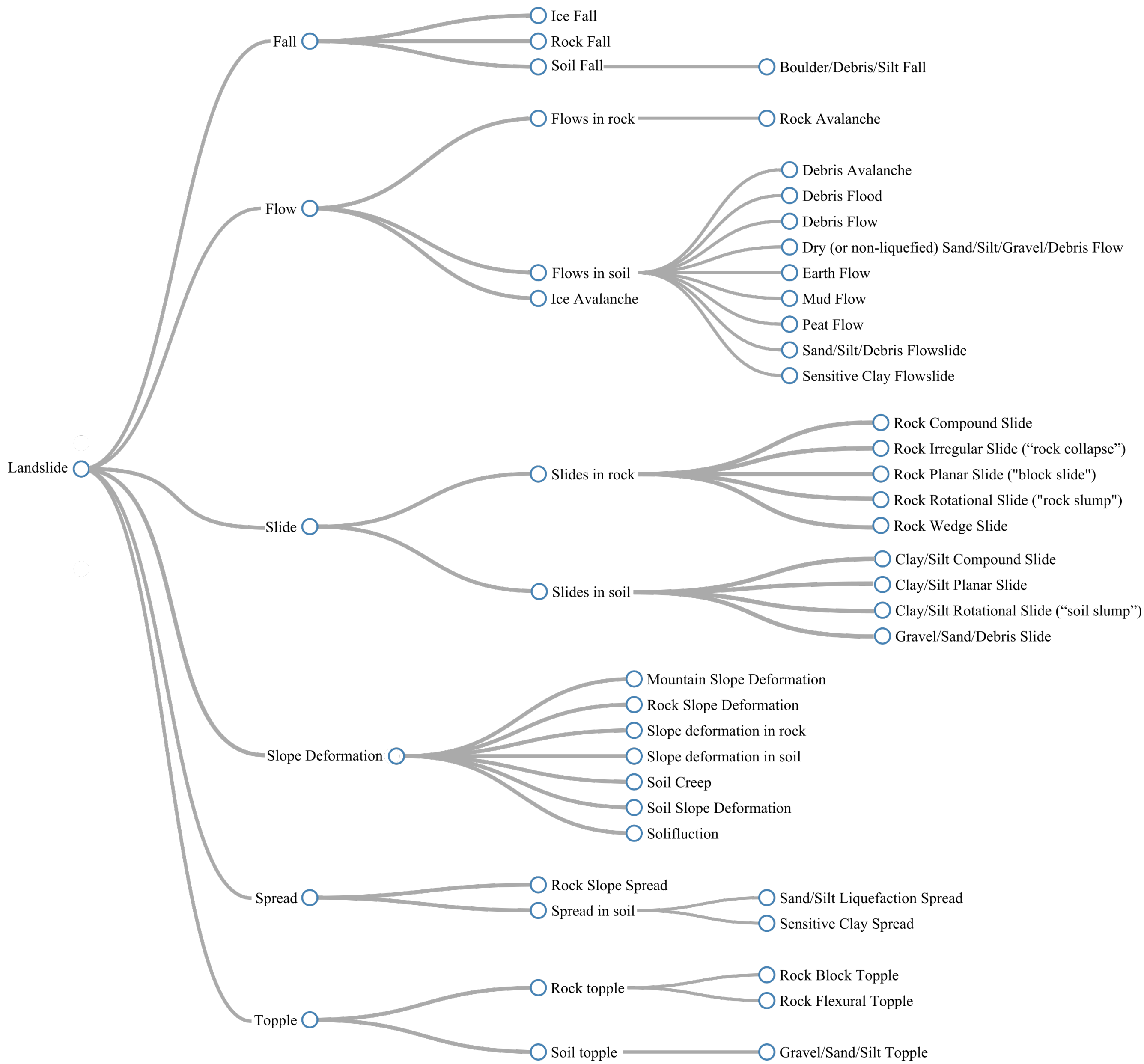


Figure 2. Natural Hazard Category code list extension for landslides

205 ### 3.1.2 Schema extension: susceptibility

The INSPIRE Natural Risk Zone schema includes hazard and risk feature types, but the concept of susceptibility as a feature type is missing. To overcome this problem, we extended the INSPIRE Natural Risk Zone core XML schema, adding a Natural Risk Zone Susceptibility schema (Minerva Intelligence, 2019e). The Natural Risk Zone Susceptibility schema includes

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions

Abstract Susceptibility Area (a in Figure 3) and Susceptibility Area feature types (b in Figure 3). The Susceptibility Area feature type is modelled following the structure of the Hazard Area and Risk Zone feature types in the Natural Risk Zone core schema. Susceptibility Area has three elements: Geometry, Influencing Factor and Relative Spatial Likelihood of Occurrence (b in Figure 3). Geometry, as with all INSPIRE vector datasets, is the geometric representation of the spatial feature. Influencing factors are defined as the intrinsic, preparatory variables which make an area susceptible to a hazard (SafeLand, 2011). Influencing factors are unbounded in multiplicity and can be defined qualitatively or quantitatively. Qualitative influencing factors are expressed as a string, while quantitative influencing factors are expressed as GML:MeasureType (c in Figure 3). Whether defined quantitatively or qualitatively, the influencing factor can also define a DataSetType attribute, such as slope or air quality. Influencing factors are used in the calculation of Relative Spatial Likelihood of Occurrence, which is an element that can be quantitatively or qualitatively defined (d in Figure 3). The relative spatial likelihood of occurrence refers to values that represent the spatial probability of occurrence of a specific hazard type, given the influencing factors present in the area (SafeLand, 2011). The Influencing Factor element allows end users of Susceptibility Area datasets to understand which known conditions of the specific area led to the resultant Relative Spatial Likelihood of Occurrence (susceptibility).

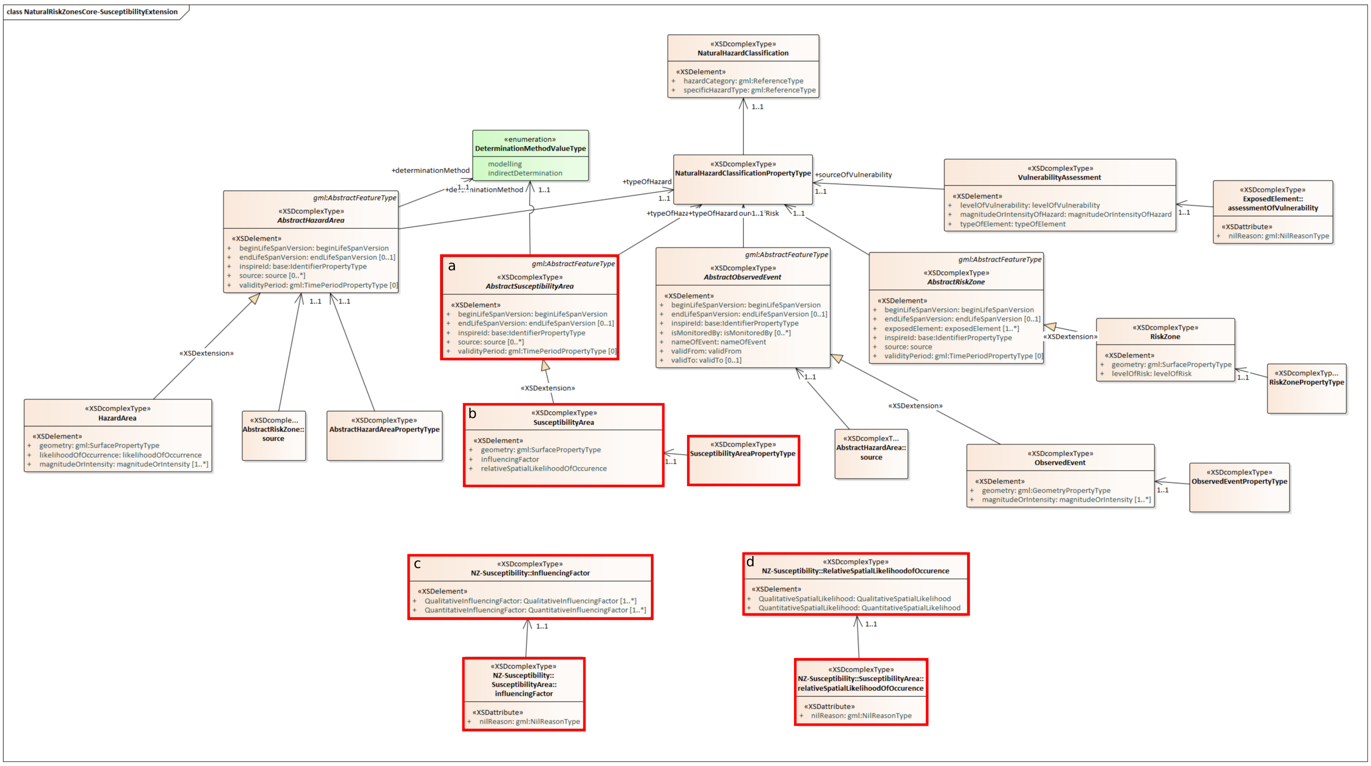


Figure 3. UML diagram showing the Natural Risk Zone Core Schema extended with Natural Risk Zone Susceptibility schema (red squares); a) Abstract susceptibility area; b) Susceptibility area; c) Influencing factor; d) Relative spatial likelihood of occurrence

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Discussions

3.2 Landslide susceptibility mapping in Veneto

3.2.1 Input data

For this study, we used open access datasets from the Veneto Region Geoportal and other sources (Table 2,3). We aligned 225 stream network, CORINE land cover, bedrock geology, and the Italian Landslide Inventory (IFFI) (Table 2) to INSPIRE standards using the software program Hale Studio (WeTransform, 2008). Datasets used that were not compliant with INSPIRE include lakes, watersheds, permafrost, fire, slope angle, faults, soil, roads and railways (Table 3).

Table 2. INSPIRE-compliant layers

Layer	Description	Source URL (last access: January 2020)
Streams	Hydrographic network in the Veneto region, including streams, rivers, and other inland flowing water bodies	https://idt2.regione.veneto.it
Land Cover (CORINE)	Land cover units in the Veneto region. The CORINE Land Cover (CLC) classification was used which includes 44 classes, and was last updated in 2018	https://land.copernicus.eu/pan-european corine-land-cover
Geology	Bedrock lithology in the Veneto region.	http://www.pcn.minambiente.it/mattm/en/wfs-service/
IFFI Landslide Points and Areas	Landslides that have been identified in the Veneto region as part of the IFFI project. The INSPIRE Natural Hazard Category code list was extended to include the updated Varnes landslide classification (Hungr et al., 2014), and the data were aligned to this standard	http://www.pcn.minambiente.it/mattm/en/wfs-service/

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Table 3. Layers not compliant with INSPIRE standards

Layer	Description	Source URL (last access: January 2020)
Lakes	Lakes in the Veneto region.	https://idt2.regione.veneto.it
Watersheds	Watersheds in the Veneto region, derived from a digital elevation model from the TINITALY project made available by the National Institute of Geophysics and Volcanology (INGV).	http://tinitaly.pi.ingv.it/
Permafrost	Permafrost derived from the Global Permafrost Zonation Index Map (Gruber, 2012)	http://www.geo.uzh.ch/microsite/cryodata/
Fires	Location and date of past forest fires in the Veneto region.	https://idt2.regione.veneto.it
Slope	The gradient of the slope in the Veneto region, derived from a digital elevation model from the TINITALY project made available by the National Institute of Geophysics and Volcanology.	http://tinitaly.pi.ingv.it/
Faults	Faults in the Veneto region, published as part of the Database of Individual Seismogenic Sources (DISS) provided by the National Institute of Geophysics and Volcanology (INGV).	http://diss.rm.ingv.it/diss/index.php/DISS321
Soils	Soil map of the Veneto region, including information about surficial deposit genesis, material, texture, thickness, geomorphic form and process.	https://idt2.regione.veneto.it
Railroads	Railroad network in the Veneto region.	https://idt2.regione.veneto.it
Roads	Road network in the Veneto region.	https://idt2.regione.veneto.it

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions

3.2.2 Mapping units and Spatial overlay

We used a spatial overlay analysis to aggregate data describing the physical properties of the mapping units, 83,960 slope units and 9,302 stream buffer polygons (Figure 4). The analysis was conducted using a custom QGIS script which aggregated the properties from all features that intersect the mapping units. For each property in an input layer, an aggregation type is specified as either: (a) list, whereby all of the intersecting properties are concatenated into the mapping unit (e.g. multiple rock types), or (b) boolean evaluation, which checks whether or not the mapping unit was intersected by a specific input feature (e.g. a fault). The end results are polygonal representations of the landscape which are attributed with all available data for landslide susceptibility mapping as required by our ontology.

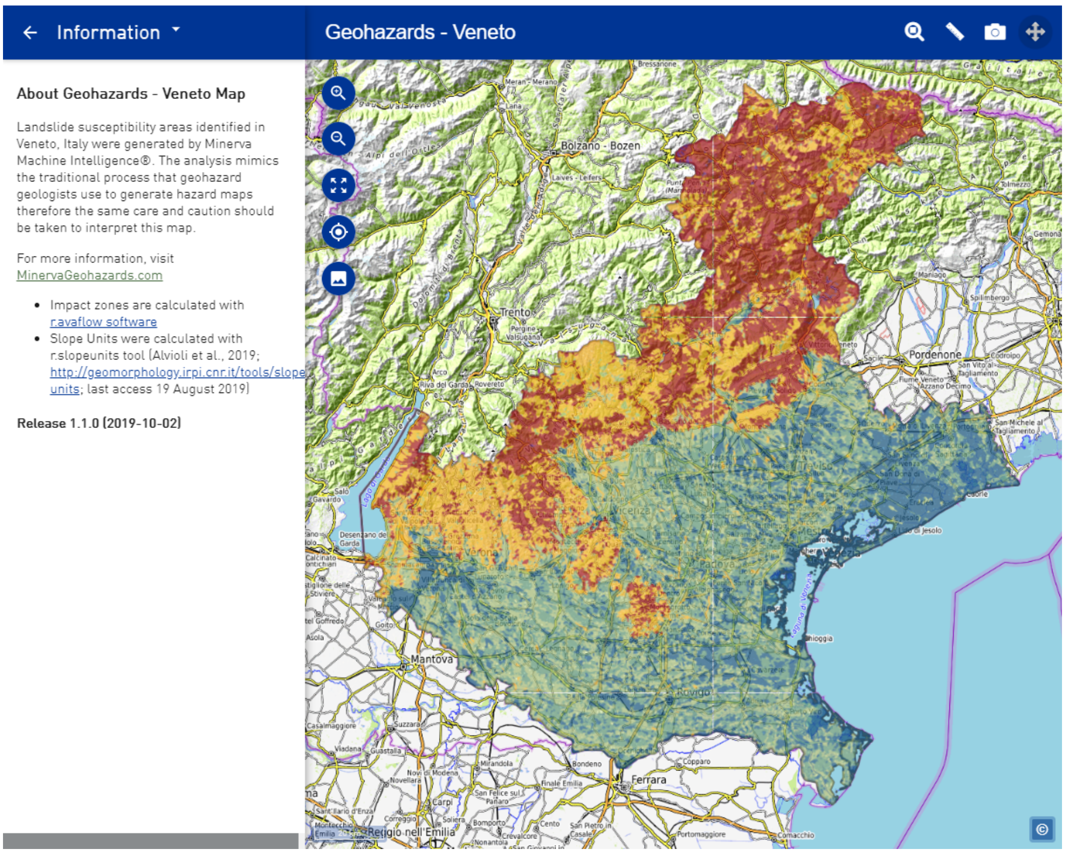


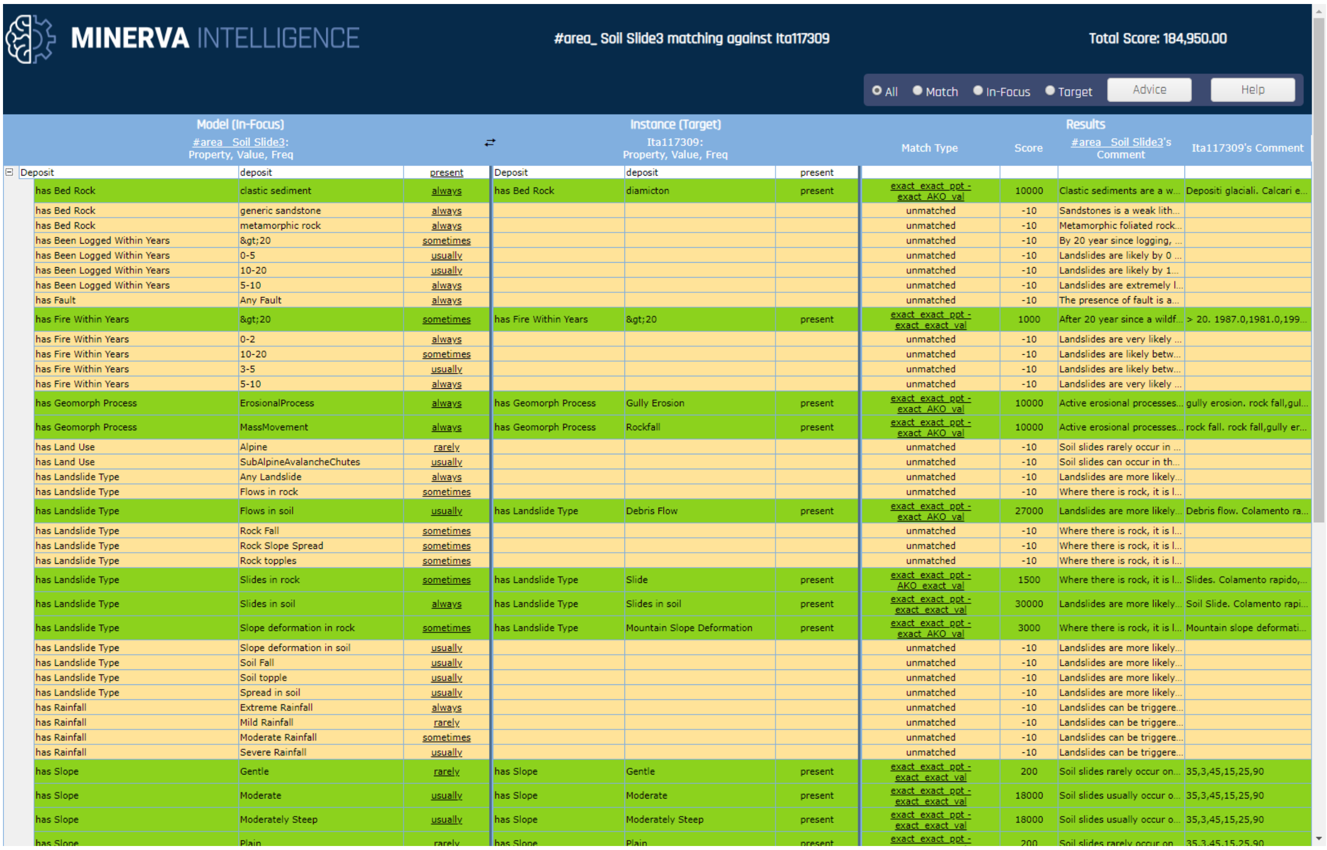
Figure 4. Web map interface portraying susceptibility to slides in soil in Veneto, Italy. Base map credit: © OpenTopoMap (CC-BY-SA)

3.2.3 Semantic network conversion, matching and impact zone modelling

The properties describing each mapping unit polygon were converted into semantic networks, one network for each polygon. This conversion allows for semantic reasoning to compare and rank, based on similarity, the mapping units against the expert-

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions



MINERVA INTELLIGENCE

#area_ Soil Slide3 matching against Ita117309

Total Score: 184,950.00

All Match In-Focus Target Advice Help

Model (In-Focus) #area_Soil Slide3: Property	Value	Freq	Instance (Target) Ita117309: Property	Value	Freq	Match Type	Score	#area_Soil Slide3's Comment	Ita117309's Comment
Deposit	deposit	present	Deposit	deposit	present				
has Bed Rock	clastic sediment	always	has Bed Rock	diamicton	present	exact_exact_ppt - exact_AKO_val	10000	Clastic sediments are a w...	Depositi glaciali. Calcari e...
has Bed Rock	generic sandstone	always				unmatched	-10	Sandstones is a weak lith...	
has Bed Rock	metamorphic rock	always				unmatched	-10	Metamorphic foliated rock...	
has Been Logged Within Years	>20	sometimes				unmatched	-10	By 20 year since logging, ...	
has Been Logged Within Years	0-5	usually				unmatched	-10	Landslides are likely by 0 ...	
has Been Logged Within Years	10-20	usually				unmatched	-10	Landslides are likely by 1...	
has Been Logged Within Years	5-10	always				unmatched	-10	Landslides are extremely l...	
has Fault	Any Fault	always				unmatched	-10	The presence of fault is a...	
has Fire Within Years	>20	sometimes	has Fire Within Years	>20	present	exact_exact_ppt - exact_exact_val	1000	After 20 year since a wildf...	> 20. 1987.0,1981.0,199...
has Fire Within Years	0-2	always				unmatched	-10	Landslides are very likely ...	
has Fire Within Years	10-20	sometimes				unmatched	-10	Landslides are likely betw...	
has Fire Within Years	3-5	usually				unmatched	-10	Landslides are likely betw...	
has Fire Within Years	5-10	always				unmatched	-10	Landslides are very likely ...	
has Geomorph Process	ErosionalProcess	always	has Geomorph Process	Gully Erosion	present	exact_exact_ppt - exact_AKO_val	10000	Active erosional processes...	gully erosion. rock fall,gul...
has Geomorph Process	MassMovement	always	has Geomorph Process	Rockfall	present	exact_exact_ppt - exact_AKO_val	10000	Active erosional processes...	rock fall. rock fall,gully er...
has Land Use	Alpine	rarely				unmatched	-10	Soil slides rarely occur in ...	
has Land Use	SubAlpineAvalancheChutes	usually				unmatched	-10	Soil slides can occur in th...	
has Landslide Type	Any Landslide	always				unmatched	-10	Landslides are more likely...	
has Landslide Type	Flows in rock	sometimes				unmatched	-10	Where there is rock, it is l...	
has Landslide Type	Flows in soil	usually	has Landslide Type	Debris Flow	present	exact_exact_ppt - exact_AKO_val	27000	Landslides are more likely...	Debris flow. Colamento ra...
has Landslide Type	Rock Fall	sometimes				unmatched	-10	Where there is rock, it is l...	
has Landslide Type	Rock Slope Spread	sometimes				unmatched	-10	Where there is rock, it is l...	
has Landslide Type	Rock topples	sometimes				unmatched	-10	Where there is rock, it is l...	
has Landslide Type	Slides in rock	sometimes	has Landslide Type	Slide	present	exact_exact_ppt - AKO_exact_val	1500	Where there is rock, it is l...	Slides. Colamento rapido,...
has Landslide Type	Slides in soil	always	has Landslide Type	Slides in soil	present	exact_exact_ppt - exact_exact_val	30000	Landslides are more likely...	Soil Slide. Colamento rapi...
has Landslide Type	Slope deformation in rock	sometimes	has Landslide Type	Mountain Slope Deformation	present	exact_exact_ppt - exact_AKO_val	3000	Where there is rock, it is l...	Mountain slope deformati...
has Landslide Type	Slope deformation in soil	usually				unmatched	-10	Landslides are more likely...	
has Landslide Type	Soil Fall	usually				unmatched	-10	Landslides are more likely...	
has Landslide Type	Soil topple	usually				unmatched	-10	Landslides are more likely...	
has Landslide Type	Spread in soil	usually				unmatched	-10	Landslides are more likely...	
has Rainfall	Extreme Rainfall	always				unmatched	-10	Landslides can be triggere...	
has Rainfall	Mild Rainfall	rarely				unmatched	-10	Landslides can be triggere...	
has Rainfall	Moderate Rainfall	sometimes				unmatched	-10	Landslides can be triggere...	
has Rainfall	Severe Rainfall	usually				unmatched	-10	Landslides can be triggere...	
has Slope	Gentle	rarely	has Slope	Gentle	present	exact_exact_ppt - exact_exact_val	200	Soil slides rarely occur on...	35,3,45,15,25,90
has Slope	Moderate	usually	has Slope	Moderate	present	exact_exact_ppt - exact_exact_val	18000	Soil slides usually occur o...	35,3,45,15,25,90
has Slope	Moderately Steep	usually	has Slope	Moderately Steep	present	exact_exact_ppt - exact_exact_val	18000	Soil slides usually occur o...	35,3,45,15,25,90
has Slope	Plain	rarely	has Slope	Plain	present	exact_exact_ppt -	200	Soil slides rarely occur on...	35,3,45,15,25,90

Figure 5. Sample Match Report showing polygon 117309 compared to slides in soil model

defined landslide models to evaluate landslide susceptibility. We deliver the similarity score between models and instances on the output maps. A higher similarity score between an individual mapping unit and a landslide susceptibility model signals a higher susceptibility to that type of landslide over other polygons. Explanations on how the scores were calculated are shown on the match report available for each spatial unit (Figure 5). The match report shows matching properties with green lines and unmatched properties in beige. Conflicting properties, if present, are coloured in red. Within the match report, there are hyperlinks to explanations describing how points are awarded for each property.

Landslide runout simulations were computed for all mapping units that fell within or above the 99.9th percentile of evaluated instances for each landslide type (Figure 6). Various landslide size classes were simulated for each instance, ranging from class 4 to class 6 (Jakob, 2005) for a total of 3696 landslide runout simulations. Classes 4 to 6 were chosen to provide a preliminary hazard assessment, where class 4 event may have an approximate return interval of hundreds of years and class 6 are very unlikely and extreme events with return intervals on the order of thousands of years (Jakob, 2005).

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

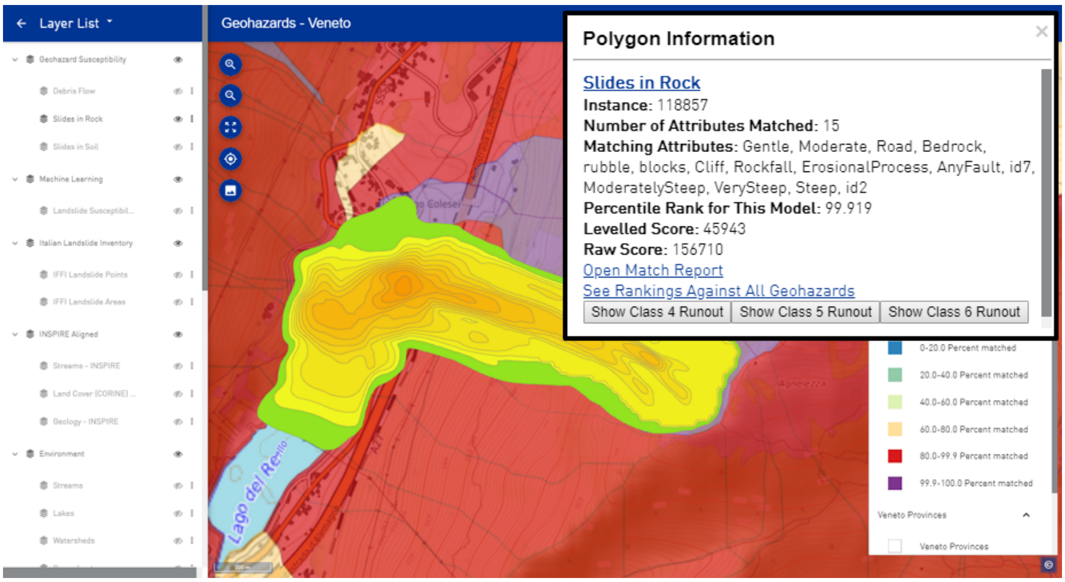


Figure 6. Landslide runout and information popup in Web Map. Base map credit: © OpenTopoMap (CC-BY-SA)

4 Discussion

4.1 INSPIRE as a framework for Explainable AI

Across society, the use of numerous complex and non-standardized earth science taxonomies results in interoperability problems, which hinder the widespread implementation of explainable AI solutions to natural hazard-related problems. This is evident in the landslide domain, where data layers for landslide susceptibility analysis, ranging from landslide databases (Van Den Eeckhaut et al., 2013) to geomorphology maps, vary across regions and countries. Consequently, despite the wealth of scientific literature on landslides in general, and landslide susceptibility in particular (Reichenbach et al., 2018), broad-scale operational landslide hazard management systems are scarce (Guzzetti et al., 2020), resulting in significant human and economic losses (Froude and Petley, 2018).

INSPIRE partially addresses this problem by providing standardised data structures for data-hosting and standard terminologies to use within those structures. As illustrated by this study, once it is INSPIRE-compliant, European data can be subjected to powerful AI analytical methods that can be efficiently and meaningfully applied to multiple other equivalent INSPIRE-compliant data sets. For example, the same landslide focused ontology that uses terminology and knowledge models based on INSPIRE code lists used in this project has been applied in South-Western British Columbia, Canada (Minerva Intelligence, 2019f).

By maintaining carefully curated standards, INSPIRE can play a critical role in AI applications that seek to be "explainable" (Gilpin et al., 2019). Its code lists can be imported into ontologies, enabling machines to make inferences based on data. The explainability in the application presented in this study is provided in the form of a comprehensive match report, which can be opened via an information popup for each slope instance. The match report provides the user with complete access to the logic that drives the AI reasoning engine, allowing interrogation of the results displayed on the map. By embedding explanations in a user-friendly interface, ontologically-based AI can improve the understanding of complex geospatial problems by decision-makers, insurance companies and the general public.

As shown in this study, INSPIRE compliance increases interoperability of data and enables AI applications to identify new insights from that data. Public and private organisations, within and outside the European Union, can significantly enhance the value of the data they collect and publish by using INSPIRE-compliant standards not only in natural hazard mapping but also in other domains. Quantification of this value has yet to be made, but calculations from (Craglia and Campagna, 2010) have shown that the development of an SDI for interoperable exchange of spatial data can save between 100-200 million Euro per annum in the Lombardia Region, Italy, alone.

4.2 INSPIRE extension and limitations

INSPIRE-compliant datasets are still rare across European countries in general, and in Italy in particular (Cetl et al., 2017; Mijić and Bartha, 2018; Cho and Crompvoets, 2019). Consequently, the authors were not able to find a jurisdiction in Europe with INSPIRE-compliant datasets for all the inputs necessary for this study. Therefore, instead of using already-compliant data, a region optimal for demonstrating the inter-relationship between INSPIRE and explainable AI was chosen, and the data

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions

for that region was transformed to INSPIRE compliance. In so doing, the study provides both a case study of dealing with non-INSPIRE-compliant data, and an illustration of the rewards achievable by bringing a coherent set of data into INSPIRE compliance.

The code lists and application schemas in the INSPIRE Natural Risk Zone theme lacked the level of detail necessary for this application. This is understandable as given the broad scope of the directive; schemas lack the necessary granularity for specific applications. INSPIRE is intended to be used as an overarching umbrella under which domain-specific applications can find their place by extending it where necessary. The Natural Risk Zone theme (Tomas et al., 2015) and the extension presented in this work is an example of using this extension facility. Within the Natural Risk Zone theme, the Natural Hazard Category Value code list includes geological/hydrological hazards, including 'flood' and 'landslide', but the different subclasses of floods and landslides are not specified. For this landslide evaluation, and other applications seeking the status of interoperable and explainable, the clear definition of landslide types, landslide size classes, and susceptibility is fundamental. For example, the hazard posed by a debris flow, which moves rapidly (tens of meters per second), and an earth flow, which may move slowly (meters per year) is very different. They can both destroy property but it is unlikely for an earth flow to result in fatalities while the opposite can be said of debris flows (Hungr et al., 2014). The definition of landslide sizes is also important: a size class 1 debris flow has a smaller impact area than a size class 6 event, but, by having a higher frequency, it may result in greater losses (Jakob, 2005). From a data perspective, INSPIRE code lists cannot currently host multi-hierarchical taxonomies. This limits the nature of reasoning that can be brought to bear on them. We understand the technical difficulties in handling multi-hierarchical taxonomies, but hope that future versions of the Registry software will be able to handle these complex knowledge representations.

The INSPIRE Natural Risk Zone theme also lacks the definition of susceptibility as a concept and feature type. The term susceptibility is not implemented as a feature type because for most hazards (e. g floods and earthquakes) the concept is embedded within the concept of hazard likelihood (Tomas et al., 2015). This does not apply in the landslide domain where susceptibility and hazard are distinct concepts (e.g. Van Den Eeckhaut and Hervás, 2012). In this study, we implemented the susceptibility feature type. Although we applied this feature type in the landslide domain, it will be useful for other natural hazard applications, when the spatial likelihood of hazard occurrence must be expressed separately from the general concept of hazard likelihood.

The extensibility of INSPIRE allows for domain-specific applications, like the approach presented in this paper, to fit within the INSPIRE framework. However, problems may also arise from the fact that INSPIRE is extensible. Extensibility allows greater precision in terminology and schema for a specific application but this allows different public and private institutions to implement separate, and eventually, incompatible extensions. For example, another landslide classification may be implemented by another institution: this implementation may not be interoperable with the one presented in this study, but will have the same INSPIRE compliance, leading to two conflicting standards. Much work remains at the level of thematic clusters to implement as many standardized vocabularies and schemas as possible. Our extension is open and free, and we hope that other entities will adopt it for other landslide applications.

4.3 AI-based probabilistic matching for landslide susceptibility mapping

The semantic AI system applied in this study aimed to replicate the reasoning with uncertainties typical of geological assessments, using the terminology that geological and geotechnical professionals use in their daily practice (Smyth et al., 2007). As they are based on expert-defined models, the landslide susceptibility maps produced in this study are comparable to qualitative heuristic assessments (SafeLand, 2011). The choice of using a qualitative method for landslide susceptibility assessment is in contrast with recent recommendations for the application of quantitative methods (Corominas et al., 2014). However, in current geological assessments, expert judgment is still widely applied (e.g., Association of Professional Engineers and Geoscientists of British Columbia, 2010), and quantitative (statistically and physically-based) methods rely on data that are not always available or of unknown quality. For example, landslide databases necessary for statistically-based susceptibility mapping are often incomplete, inaccurate, and geographically-limited (Guzzetti et al., 2012). Usually, the geotechnical parameters necessary for running physical models are approximated to carry out regional-scale studies (e.g., Mergili et al., 2014).

The semantic AI system applied in this study can be used in cases of data scarcity, and if coupled with numerical methods, can improve the explainability of predictions. For example, by embedding the ontology concepts related to statistical parameters (e.g. receiving operating curves, confidence intervals) or physical parameters (e. g. friction angles, viscosity), it will be possible for the numerical outputs of quantitative methods to be explained in natural language, helping to reduce the gap between scientists and decision-makers (Newman et al., 2017).

5 Conclusions

This study presents AI-based landslide susceptibility maps for debris flow, slides in soil and slides in rock for the province of Veneto, Italy, in the framework provided by the INSPIRE Natural Risk Zone theme. To produce the maps for specific landslide types, we extended the Natural Risk Zone theme to encompass both the concept of susceptibility and the different types of landslides. In particular, we registered a landslide classification extension of the Natural Hazard Category code list, a landslide size class code list, and Susceptibility Area and Abstract Susceptibility Area feature types schema extensions. After defining the extension, we aligned key input layers (geology, streams, and land cover) to INSPIRE and, by using a state-of-the-art ontologically-grounded probabilistic matching algorithm, we produced the landslide susceptibility layers. The processing outputs were mapped to the Natural Risk Zone Susceptibility schema extension. Then, impact zones of potential landslides for difference landslide classes were physically modelled for the most susceptible polygons. Finally, the results were embedded in a user-friendly interface, and made available online .

We have demonstrated the value of INSPIRE-compliance by showing how it enhances information and knowledge interoperability, and allows for explainability in AI applications by standardized interrogation of their inputs and outputs. Ontologies provide the formal structure for INSPIRE code lists to run algorithms similar to that applied here. The maps can explain the scientific results that they portray, and consequently improve the understanding of complex geospatial problems not only by domain experts but also by decision-makers and other non-specialized interested parties.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

This study also illustrates that, in their current state of development, the INSPIRE standards are not sufficiently expressive to support complex landslide susceptibility mapping. We provided an example of how INSPIRE's extension capabilities may be implemented to add the required expressivity. This extension framework ensures, through its Re3gistry register, that the expressivity extensions are documented and available to all interested parties for re-use. In so doing, it sets the context for the ongoing refinement of standards by the INSPIRE thematic committees.

6 Data availability

- The web application is available at: https://map.italy.minervageohazards.com/
- The schema extension is available at: https://github.com/minervaintelligence/INSPIRE-NZ-Susceptibility
- The code list extension is available at: http://minerva.codes/registry
- Data from the Italian National geoportal is available under "Attribution-NonCommercial-ShareAlike 3.0 Italy (CC BY-NC-SA 3.0 IT)" License, https://creativecommons.org/licenses/by-nc-sa/3.0/it/deed.en
- Data from the Veneto Geoportal are available under the "Italian Open Data License 2.0", https://www.dati.gov.it/content/italian-open-data-license-v20
- CORINE land cover data is available under EEA standard re-use policy: re-use of content on the EEA website for commercial or non-commercial purposes is permitted free of charge, provided that the source is acknowledged (http://www.eea.europa.eu/legal/copyright)
- Tinitaly DEM is available upon request by sending an email to simone.tarquini@ingv.it with the subject of TINITALY DEM. Terms and Conditions of Use: Data is provided for research purposes only. Data is provided solely to the person named on this application form and should not be given to third parties. Third parties who might need access to the same dataset are required to fill their own application forms http://tinitaly.pi.ingv.it/ Data from INGV is available under "Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)" license http://creativecommons.org/licenses/by-sa/4.0/
- The permafrost data is available under "Attribution 3.0 Unported (CC BY 3.0)" licence. http://www.geo.uzh.ch/microsite/cryodata/.

B Appendix A

B.1 Dictionary of Terms

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

Term	Description
Code list	Are vocabularies storing INSPIRE terms
CLC	CORINE land Cover
Feature type	Vector based spatial data
IFFI	Italian Landslide Inventory
Instance	A real-world entity, for this application an Instance is a spatial feature, either a slope unit polygon or a stream buffer polygon.
Model	Expert-defined conceptualization of some entity, in this case, landslides. Three landslide models were used in this project; debris flow, slides in soil and slides in rock.
Ontology	Definition of entities and of the rules describing the relation between the entities.
Semantic Network	A graph network of arcs and nodes storing data in semantic triple format.
Schema	The structure of the spatial data themes
Slope unit	Mapping unit polygon that is automatically derived from the terrain, based on hydrologic drainage and divide lines
Taxonomy	Hierarchical classification scheme based on shared characteristics between entities
Triple	A semantic triple is a subject-object-predicate expression, and it is the basic unit of a semantic network.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions

D Appendix B

D.1 Properties used for the landslide classification

Property	Property definition	Property value	Property value definition
Type of movement	Landslide movement types (Hungr et al., 2014)	Fall	A fall starts with the detachment of soil or rock from a steep slope along a surface on which little or no shear displacement takes place. The material then descends largely through the air by falling, saltation or rolling (Cruden and Couture, 2011)
		Topple	A topple is the forward rotation of material about a point or axis below the centre of gravity of the displaced mass. (Cruden and Couture, 2011)
		Slide	A slide is a downslope movement occurring dominantly on surfaces of rupture or relatively thin zones of intense shear strain (Cruden and Couture, 2011)
		Spread	Spread is an extension of mass combined with a general subsidence of a upper fractured mass of material into softer underlying material. (Cruden and Couture, 2011)
		Flow	A flow is a spatially continuous movement in which surfaces of shear are short-lived, closely spaced and not usually preserved (Cruden and Couture, 2011).
		Slope deformation	Slow, sometime unmeasurable, deformation of slopes (Hungr et al., 2014)
Material	Landslide-forming material types (Hungr et al., 2014)	Ice	Glacier ice or other solid water on steep slopes (Hungr et al., 2014)
		Rock	Intrusive, volcanic, metamorphic, strong sedimentary, (carbonatic or arenaceous) and weak sedimentary (argillaceous) (Hungr et al., 2014)

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences
Open Access
Discussions

	Strong		Rock broken with hammer (Hungr et al., 2014)
	Weak		Rock peeled with knife (Hungr et al., 2014)
Soil			Residual, colluvial, alluvial, lacustrine, marine, aeolian, glacial, volcanic, organic, random anthropogenic fills, engineered anthropogenic fills, mine tailings, and sanitary waste (Hungr et al., 2014).
	Peat		Organic material (Hungr et al., 2014).
	Debris		Low plasticity, unsorted and mixed material (Hungr et al., 2014).
	Silt, sand, gravel, and boulders		Nonplastic (or very low plasticicty), granular, sorted. Silt particles cannot be seen by eye. (Hungr et al., 2014).
		Partly saturated	GW, GP, and GM unified soil classes (Hungr et al., 2014)).
		saturated	SW, SP, and SM unified soil classes (Hungr et al., 2014).
		dry	ML unified soil class (Hungr et al., 2014).
	Mud		Plastic, unsorted, and close to Liquid Limit material. CL, CH, and CM unified soil classes (Hungr et al., 2014).
	Clay		Plastic, can be modeled into standard thread when moist, has dry strength. GC, SC, CL, MH, CH, OL, and OH unified soil classes (Hungr et al., 2014).
		Sensitive	Sensitive or quick clay is a special type of clay prone to sudden strength loss upon disturbance. From a relatively stiff material in the undisturbed condition, an imposed stress can turn such clay into a liquid gel (Geertsema, 2013).
		soft	Easily molded with fingers. Point of geologic pick easily pushed into shaft of handle. Easily penetrated several centimeters by thumb. (Hungr et al., 2014; USDA, 2012).
		stiff	Indented by thumb with great effort. Point of geologic pick can be pushed in up to 1 centimeter. Very difficult to mold with fingers. Just penetrated with hand spade (Hungr et al., 2014; USDA, 2012).

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Natural Hazards and Earth System Sciences Discussions

E Appendix C

E.1 Landslide models

- Debris flow model https://italy.minervageo.com/debris-flow-model/
- Slides in soil model https://italy.minervageo.com/slides-in-soil/
- Slides in rock model https://italy.minervageo.com/the-roberti-slides-in-rock-model/

Author contributions.

- Gioachino Roberti, Jake McGregor, Clinton Smyth and David Poole wrote the paper
- Gioachino Roberti conceptually designed the susceptibility schema and landslide extension, the expert-based landslide models and expanded the geohazard ontology.
- Jakob McGregor implemented the INSPIRE schema and code list extension and designed the web map application.
- David Poole and Clinton Smyth designed the qualitative probabilistic method used to calculate susceptibility.
- Sharon Lam and Blake Boyko implemented and maintain the web map.
- Victoria Wang implemented and maintained the geohazard ontology.
- Bryan and Chris implemented the qualitative probabilistic algorithm.
- Steve Richards supported the semantic implementations.
- David Bigelow helped in the redaction of the manuscript, reviewed the landslide models and the code list extensions.

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. This project was first presented at the Helsinki 2019 INSPIRE data challenge, where won the first price of the data challenge. The authors would like to acknowledge the conference organizer committee including the National Land Survey of Finland and Ministry of Agriculture, and the Joint Research Centre of the European Commission and Spatineo. The authors would also like to acknowledge M. Alvioli et al., for the availability of the r.slopeunit code, and M. Mergili and S. Pudasaini, for the r.avaflow code, and WeTransform GMBH for the HALE Connect and HALE Studio software licences. This research was funded by Minerva Intelligence Inc.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

References

Alvioli, M., Marchesini, I., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., and Guzzetti, F.: Automatic delineation of geomorphological slope units with r.slopeunits v1.0 and their optimization for landslide susceptibility modeling, Geoscientific Model Development, 9, 3975–3991, https://doi.org/10.5194/gmd-9-3975-2016, 2016.

Aristotle: The Categories, 350BC.

Association of Professional Engineers and Geoscientists of British Columbia: Guidelines for Legislated Landslide Assessments for proposed residencial development in BC, Tech. Rep. May, 2010.

Baum, R. L., Savage, W. Z., and Jonathan W., G.: Trigrs—A Fortran Program for Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability Analysis, Version 2.0, Tech. rep., 2008.

Cetl, V., Nunes De Lima, V., Tomas, R., Lutz, M. ., D 'eugenio, J., Nagy, A., and Robbrecht, J.: Summary Report on Status of implementation of the INSPIRE Directive in EU, https://doi.org/10.2760/162895, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC109035/jrc109035{_}jrc109035{_}jrc{_}inspire{_}eu{_}summaryreport{_}online.pdf, 2017.

CGI: GeoSciML, http://www.geosciml.org/, 2003.

Cho, G. and Crompvoets, J.: The INSPIRE directive: some observations on the legal framework and implementation, Survey Review, 51, 310–317, https://doi.org/10.1080/00396265.2018.1454686, https://doi.org/10.1080/00396265.2018.1454686, 2019.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M. G., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J., and Smith, J. T.: Recommendations for the quantitative analysis of landslide risk, Bulletin of Engineering Geology and the Environment, 73, 209–263, https://doi.org/10.1007/s10064-013-0538-8, 2014.

Craglia, M. and Campagna, M.: Advanced Regional SDI in Europe : Comparative cost-benefit evaluation and impact assessment perspectives, International Journal of Spatial Data Infrastructures Research, 5, 145–167, https://doi.org/10.2902/1725-0463.2010.05.art6, 2010.

Cruden, D. M. and Couture, R.: The Working Classification of Landslides: material matters, Pan-Am CGS Geotechnical Conference, p. 7, 2011.

Dai, F. C., Lee, C. F., and Ngai, Y. Y.: Landslide risk assessment and management: An overview, Engineering Geology, 64, 65–87, https://doi.org/10.1016/S0013-7952(01)00093-X, 2002.

European Commission: INSPIRE Helsinki 2019, https://www.inspire-helsinki-2019.fi/, 2019.

European Parliament and the Council: Directive 2007/2/EC of the European Parliament and of the council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Official Journal of the European Union, 50, 1–14, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF, 2007.

European Parliament and the Council: Commission Regulation (EU) No 1312/2014 of 10 December 2014 amending Regulation (EU) No 1089/2010 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data services, Official Journal of the European Union, pp. 13–34, https://doi.org/doi:10.3000/17252555.L_2010.323.eng, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:323:0011:0102:EN:PDF, 2014.

Facebook: React, https://reactjs.org/, 2013.

Friele, P., Millard, T. H., Mitchell, A., Allstadt, K. E., Menounos, B., Geertsema, M., and Clague, J. J.: Observations on the May 2019 Joffre Peak landslides, British Columbia, Landslides, https://doi.org/10.1007/s10346-019-01332-2, 2020.

Froude, M. J. and Petley, D. N.: Global fatal landslide occurrence from 2004 to 2016, Natural Hazards and Earth System Sciences, 18, 2161–2181, https://doi.org/10.5194/nhess-18-2161-2018, https://www.nat-hazards-earth-syst-sci.net/18/2161/2018/, 2018.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

Geertsema, M.: Quick Clay, pp. 803–804, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-1-4020-4399-4_282, https://doi.org/10.1007/978-1-4020-4399-4{_}282, 2013.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., and Kagal, L.: Explaining explanations: An overview of interpretability of machine learning, Proceedings - 2018 IEEE 5th International Conference on Data Science and Advanced Analytics, DSAA 2018, pp. 80–89, https://doi.org/10.1109/DSAA.2018.00018, 2019.

Goodfellow, I., Bengio, Y., and Aaron, C.: Deep Learning, MIT press., 2016.

Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, Cryosphere, 6, 221–233, https://doi.org/10.5194/tc-6-221-2012, 2012.

Guzzetti, F., Carrara, A., Cardinali, M., and Reichenbach, P.: Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy, Geomorphology, 13, 1995, https://doi.org/10.1016/S0169-555X(99)00078-1, 1999.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., and Chang, K. T.: Landslide inventory maps: New tools for an old problem, Earth-Science Reviews, 112, 42–66, https://doi.org/10.1016/j.earscirev.2012.02.001, http://dx.doi.org/10.1016/j.earscirev.2012.02.001, 2012.

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., and Melillo, M.: Geographical landslide early warning systems, Earth-Science Reviews, 200, 102 973, https://doi.org/10.1016/j.earscirev.2019.102973, https://doi.org/10.1016/j.earscirev.2019.102973, 2020.

Hajimoradlou, A., Roberti, G., and Poole, D.: A Probabilistic Approach for Predicting Landslides by Learning a Self-Aligned Deep Convolutional Model, http://arxiv.org/abs/1911.04651, 2019.

Hungr, O., Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update, Landslides, 11, 167–194, https://doi.org/10.1007/s10346-013-0436-y, 2014.

Hykes, S.: Docker, https://www.docker.com/, 2013.

ISA: Re3gistry software 1.3, https://ec.europa.eu/isa2/solutions/re3gistry{_}en, 2016.

Jackson Jr, L., Smyth, C., and Poole, D.: Hazardmatch: an application of artificial intelligence to landslide susceptibility mapping, Howe Sound Area, Bristish Columbia, 4th Canadian Conference on Geohazards: From Causes to Management, p. 594, 2008.

Jakob, M.: A size classification for debris flows, Engineering Geology, 79, 151–161, https://doi.org/10.1016/j.enggeo.2005.01.006, 2005.

Marcus, G.: Deep Learning: A Critical Appraisal, pp. 1–27, http://arxiv.org/abs/1801.00631, 2018.

Marcus, G. and Davis, E.: Rebooting AI: Building Artificial Intelligence We Can Trust, Knopf Doubleday Publishing Group, 2019.

McDougall, S.: 2014 Canadian Geotechnical Colloquium: Landslide runout analysis — current practice and challenges, Canadian Geotechnical Journal, 54, 605–620, https://doi.org/10.1139/cgj-2016-0104, 2016.

Mergili, M., Marchesini, I., Alvioli, M., Metz, M., Schneider-Muntau, B., Rossi, M., and Guzzetti, F.: A strategy for GIS-based 3-D slope stability modelling over large areas, Geoscientific Model Development, 7, 2969–2982, https://doi.org/10.5194/gmd-7-2969-2014, 2014.

Mergili, M., Fischer, J. T., Krenn, J., and Pudasaini, S. P.: R.avaflow v1, an advanced open-source computational framework for the propagation and interaction of two-phase mass flows, Geoscientific Model Development, 10, 553–569, https://doi.org/10.5194/gmd-10-553-2017, 2017.

MetaCarta: OpensLayer, https://openlayers.org/, 2005.

Mijić, N. and Bartha, G.: Infrastructure for Spatial Information in European Community (INSPIRE) Through the Time from 2007. Until 2017, pp. 34–42, https://doi.org/10.1007/978-3-030-02577-9_5, 2018.

Minerva Intelligence: Minerva INSPIRE registry, http://minerva.codes/registry, 2019a.

Minerva Intelligence: Natural Hazard Category Landslide Extension, http://minerva.codes/codelist/NaturalHazardCategoryLandslideExtension, 2019b.

Minerva Intelligence: Landslide Size Class, http://minerva.codes/codelist/LandslideSizeClass, 2019c.

Minerva Intelligence: Aristotelean Class Editor - ACE, https://minervaintelligence.com/project/ace/, 2019d.

Minerva Intelligence: INSPIRE Natural Risk Zone Schema Extension for Susceptibility Area, https://github.com/minervaintelligence/INSPIRE-NZ-Susceptibility, 2019e.

Minerva Intelligence: Sea to Sky hazards, https://www.minervageohazards.com/, 2019f.

Napolitano, E., Marchesini, I., Salvati, P., Donnini, M., Bianchi, C., and Guzzetti, F.: LAND-deFeND – An innovative database structure for landslides and floods and their consequences, Journal of Environmental Management, 207, 203–218, https://doi.org/10.1016/j.jenvman.2017.11.022, https://doi.org/10.1016/j.jenvman.2017.11.022, 2018.

Newman, J. P., Maier, H. R., Riddell, G. A., Zecchin, A. C., Daniell, J. E., Schaefer, A. M., van Delden, H., Khazai, B., O'Flaherty, M. J., and Newland, C. P.: Review of literature on decision support systems for natural hazard risk reduction: Current status and future research directions, Environmental Modelling and Software, 96, 378–409, https://doi.org/10.1016/j.envsoft.2017.06.042, http://dx.doi.org/10.1016/j.envsoft.2017.06.042, 2017.

Pearl, J.: Probabilistic reasoning in intelligent systems: networks of plausible inference, Morgan Kaufmann, 1988.

Phengsuwan, J., Shah, T., James, P., Thakker, D., Barr, S., and Ranjan, R.: Ontology-based discovery of time-series data sources for landslide early warning system, Computing, https://doi.org/10.1007/s00607-019-00730-7, https://doi.org/10.1007/s00607-019-00730-7, 2019.

Poole, D.: Ontology Design for Scientific Theories That Make Probabilistic Predictions, Semantic scientific knowledge integraion, 2009.

Poole, D. and Mackworth, A.: Artificial Intelligence: foundations of computational agents, Cambridge University Press, second edn., 2017.

Poole, D. and Smyth, C.: Type uncertainty in ontologically-grounded qualitative probabilistic matching, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 3571 LNAI, 763–774, 2005.

Pudasaini, S. P.: A general two-phase debris flow model, Journal of Geophysical Research: Earth Surface, 117, 1–28, https://doi.org/10.1029/2011JF002186, 2012.

Refraction Researtch: PostGIS, https://postgis.net/, 2001.

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., and Guzzetti, F.: A review of statistically-based landslide susceptibility models, Earth-Science Reviews, 180, 60–91, https://doi.org/10.1016/j.earscirev.2018.03.001, https://doi.org/10.1016/j.earscirev.2018.03.001, 2018.

SafeLand: Recommended procedures for validating landslide hazard and risk models and maps, Living with landslide risk in Europe: Assessment, effects of global change, and risk management strategies, p. 162, 2011.

Sermet, Y. and Demir, I.: Towards an information centric flood ontology for information management and communication, Earth Science Informatics, 12, 541–551, https://doi.org/10.1007/s12145-019-00398-9, https://doi.org/10.1007/s12145-019-00398-9, 2019.

Sharma, R., Poole, D., and Smyth, C.: A framework for ontologically-grounded probabilistic matching q, International Journal of Approximate Reasoning, 51, 240–262, https://doi.org/10.1016/j.ijar.2009.05.007, http://dx.doi.org/10.1016/j.ijar.2009.05.007, 2010.

Smith, B.: Ontology, in: Blackwell Guide to the Philosophy of Computing and Information, edited by Luciano Floridi, chap. 11, pp. 155–166, Blackwell, Oxford, 2003.

Smyth, C. and Poole, D.: Qualitative Probabilistic Matching with Hierarchical Descriptions, {KR}2004: Principles of Knowledge Representation and Reasoning, pp. 479–486, 2004.

Smyth, C., Poole, D., and Sharma, R.: Semantic e-Science and Geology, Association for the Advancement of Artificial Intelligence, 2007.

https://doi.org/10.5194/nhess-2020-134
Preprint. Discussion started: 27 May 2020

The Open Planning Project: Geoserver, http://geoserver.org/, 2001.

Tomas, R., Harrison, M., Barredo, J. I., Thomas, F., Llorente Isidro, M., Pfeiffer, M., and Čerba, O.: Towards a cross-domain interoperable framework for natural hazards and disaster risk reduction information, Natural Hazards, 78, 1545–1563, https://doi.org/10.1007/s11069-015-1786-7, 2015.

USDA: Chapter 3 Engineering Classification of Earth Materials, Part 631 National Engineering Handbook, p. 35, 2012.

Van Den Eeckhaut, M. and Hervás, J.: State of the art of national landslide databases in Europe and their potential for assessing landslide susceptibility, hazard and risk, Geomorphology, 139-140, 545–558, https://doi.org/10.1016/j.geomorph.2011.12.006, http://dx.doi.org/10.1016/j.geomorph.2011.12.006, 2012.

Van Den Eeckhaut, M., Hervas, J., and Montanarella, L.: Landslide Databases in Europe: Analysis and Recommendations for Interoperability and Harmonisation, in: Landslide Science and Practice: Landslide Inventory and Susceptibility and Hazard Zoning, vol. 1, pp. 243–247, https://doi.org/10.1007/978-3-642-31325-7, http://link.springer.com/10.1007/978-3-642-31325-7{%}0Ahttp://www.springer.com/earth+sciences+and+geography/natural+hazards/book/978-3-642-31324-0{%}0Ahttp://grass-italia.como.polimi.it/meeting2012/09marchesini.pdf, 2013.

WeTransform: Hale Studio, https://www.wetransform.to/products/halestudio/, 2008.

WeTransform: Hale Connect, https://www.wetransform.to/products/haleconnect/, 2014.