Bioschemas Tools Specification  
Version: 0.1

Note: For the background and rationale of this specification see the [Bioschemas website](http://bioschemas.org/). See also [this Github issue](https://github.com/schemaorg/schemaorg/issues/1431) and the [field mappings](https://docs.google.com/spreadsheets/d/1s342mMwfXAGl5uwgTCcwsVLQztxRIDwTi92zEouP33o/edit#gid=1063339077) from tools providers.

**Schema**

Legend:

*M: minimum; R: recommended; O: optional*

*CN: Cardinality (one, many)*

*CV: Controlled Vocabulary*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Property | Expected Type | Description | CN | CV |
| [name](https://schema.org/name) | [Text](https://schema.org/Text) | The name of the item. | One |  |
| [description](https://schema.org/description) | [Text](https://schema.org/Text) | A short description of the item. | One |  |
| [url](https://schema.org/url) | [URL](https://schema.org/URL) | URL of the item.This property can be used on a page listing many software tools, to indicate each individual tool’s page. | One |  |
| featureList | Text | Features or modules provided by this application (and possibly required by other applications). | Many | x |
| softwareVersion | Text | Version of the software instance. | One |  |
| publisher | Person or Organization | Use for contacts or ‘credit’. | Many |  |
| Citation | CreativeWork or Text | Publication about this software | Many |  |
| license | Text | The applicable software license | Many | ? |
| applicationCategory | Text | Type of software e.g. a database, tool, service. | Many | x |
| keywords | Text | Keywords or tags used to describe this content. Multiple entries in a keywords list are typically delimited by commas. | Many | x |
| potentialAction (object of a ControlAction)\* | URL | The input format of the data. Must be one of the [EDAM Data](http://edamontology.org/data_0006) concept labels or one of its synonyms.  See [Appendix I](#_99393w8t45f5). | Many | x |
| potentialAction (result of a ControlAction)\* | URL | The output format of the data. Must be one of the [EDAM Data](http://edamontology.org/data_0006) concept labels or one of its synonyms. | Many | x |
| offers | Offer | An offer to provide this item—for example, an offer to sell a product, rent the DVD of a movie, perform a service, or give away tickets to an event. | Many |  |
| softwareRequirements | Text or URL | Component dependency requirements for application. This includes runtime environments and shared libraries that are not included in the application distribution package, but required to run the application (Examples: DirectX, Java or .NET runtime). | Many |  |
| dateCreated | Date or DateTime | The date on which the CreativeWork was created or the item was added to a DataFeed. |  |  |
| dateModified | Date or DateTime | The date on which the CreativeWork was most recently modified or when the item's entry was modified within a DataFeed. |  |  |
| softwareHelp | CreativeWork | A documentation for the tool | Many |  |
| hasPart | SoftwareApplication | Used to describe tools included into suites/bundles/workflows. | Many |  |
| operatingSystem | Text | Operating systems on which the app runs | Many |  |

You can also use other properties from [SoftwareApplication](https://schema.org/SoftwareApplication), [CreativeWork](https://schema.org/CreativeWork) and [Thing](https://schema.org/Thing).

\* potentialAction is used to describe input and output formats. See [Appendix I](#_99393w8t45f5) for details about this.

# **Examples**

Schema.org [suggests](http://schema.org/docs/gs.html) implementing metadata using JSON-LD, RDFa or Microdata. JSON-LD is the recommended format by Google, but any of these formats can be used for embedding information about tools in a web page or other online resource.

|  |
| --- |
| Example 1. Using all minimum fields |
| JSON-LD  <script type="application/ld+json">  {  "@context": "http://schema.org/",  "@type": "SoftwareApplication",  "name": "ProMateus",  "description": "ProMateus extends ProMate, a protein binding site predictor, and allows the user to suggest new features that might improve the prediction of the binding site locations.",  "url": "http://bioportal.weizmann.ac.il/promate/",  "featureList": "Peak calling, Protein feature detection, Transcription factor binding site prediction, Protein binding site prediction, Protein sequence feature detection",  "softwareVersion": "None(Latest)"  }  </script>  RDFa  <div vocab="http://schema.org/" typeof="SoftwareApplication">  <div property="name">ProMateus</div>  <div property="description">ProMateus extends ProMate, a protein binding site predictor, and allows the user to suggest new features that might improve the prediction of the binding site locations.</div>  <div property="url"> http://bioportal.weizmann.ac.il/promate/  </div>  <div>Feature list:  <span property="featureList">Peak calling, Protein feature detection, Transcription factor binding site prediction, Protein binding site prediction, Protein sequence feature detection</span>  </div>  <div>Version:  <span property="softwareVersion">None(Latest)</span>  </div>  </div>  Microdata  <div itemscope itemtype="http://schema.org/SoftwareApplication">  <div itemprop="name">ProMateus</div>  <div itemprop="description">ProMateus extends ProMate, a protein binding site predictor, and allows the user to suggest new features that might improve the prediction of the binding site locations.</div>  <div itemprop="url"> http://bioportal.weizmann.ac.il/promate/  </div>  <div>Feature list:  <span itemprop="featureList">Peak calling, Protein feature detection, Transcription factor binding site prediction, Protein binding site prediction, Protein sequence feature detection</span>  </div>  <div>Version:  <span itemprop="softwareVersion">None(Latest)</span>  </div>  </div> |

|  |
| --- |
| Example 2. Showing a property with single cardinality (name) and one with multiple cardinality (applicationCategory). |
| JSON-LD  <script type="application/ld+json">  {  "@context": "http://schema.org/",  "@type": "SoftwareApplication",  "name": "ExpaRNA",  "applicationCategory":["Command-line tool","Script"]  }  </script>  RDFa  <div vocab="http://schema.org/" typeof="SoftwareApplication">  <div property="name">ExpaRNA</div>  <div property="applicationCategory">Command-line tool</div>  <div property="applicationCategory">Script</div>  </div>  Microdata  <div itemscope itemtype="http://schema.org/SoftwareApplication">  <div><meta itemprop="name">ExpaRNA</div>  <div>Type:  <span itemprop="applicationCategory">Command-line tool</span>,  <span itemprop="applicationCategory">Script</span>  </div>  </div> |

|  |
| --- |
| Example 3. Showing the use of controlled vocabulary for applicationCategory |
| JSON-LD  <script type="application/ld+json">  {  "@context": "http://schema.org/",  "@type": "SoftwareApplication",  "name": "Pladipus",  "description": "The [CompOmics](http://www.compomics.com/) Pladipus system is an affordable cluster computing framework allowing for distributed processing, management and analyses for bio-informatics pipelines.",  "applicationCategory": ["Command line","Desktop GUI"],  "keywords": "Proteomics",  "softwareVersion": "0.4.1"  }  </script>  RDFa  <div vocab="http://schema.org/" typeof="SoftwareApplication">  <div property="name">Pladipus</div>  <div property="description">The [CompOmics](http://www.compomics.com/) Pladipus system is an affordable cluster computing framework allowing for distributed processing, management and analyses for bio-informatics pipelines.</div>  <div property="applicationCategory">Command-line tool</div>  <div property="applicationCategory">Desktop GUI</div>  <div property="keywords">Proteomics</div>  <div property="softwareVersion">0.4.1</div>  </div>  Microdata  <div itemscope itemtype="http://schema.org/SoftwareApplication">  <div><meta itemprop="name">Pladipus</div>  <div itemprop="description">The CompOmics Pladipus system is an affordable cluster computing framework allowing for distributed processing, management and analyses for bio-informatics pipelines</div>  <div>Type:  <span itemprop="applicationCategory">Command-line tool</span>,  <span itemprop="applicationCategory">Desktop GUI</span>  </div>  <div itemprop="keywords"> Proteomics </div>  <div itemprop="softwareVersion">0.4.1</div>  </div> |

|  |
| --- |
| Example 4. Extended example using JSON-LD, RDFa and microdata |
| JSON-LD:  <script type="application/ld+json">  {  "@context": "http://schema.org",  "@type": "SoftwareApplication",  "name": "SNPs and GO",  "description": "SNPs&GO is a server for the prediction of single point protein mutations likely to be involved in the insurgence of diseases in humans.",  "url": "http://snps-and-go.biocomp.unibo.it/snps-and-go/index.html",  "applicationCategory": "Tool",  "potentialAction": {  "@type": "ControlAction",  "object": {  "@type": "Dataset",  "additionalType": "http://edamontology.org/data\_3021"  }, "result": {  "@type": "Dataset",  "additionalType": "http://edamontology.org/data\_0896" }  },  "Keywords": ["DNA polymorphism","Bioinformatics","Protein properties"],  "softwareVersion": "1.0",  "featureList": "Variant classification",  "publisher": [  {  "@type": "Person",  "name": "Rita Casadio",  "email": "casadio@biocomp.unibo.it"  },  {  "@type": "Person",  "name": "Pier Luigi Martelli",  "email": "gigi@biocomp.unibo.it"  }],  "operatingSystem": "All"  }  </script>  RDFa:  <div vocab="http://schema.org/" typeof="SoftwareApplication">  <div property="name">SNPs and GO</div>  <div property="description">SNPs &amp; GO is a server for the prediction of single point protein mutations likely to be involved in the insurgence of diseases in humans.</div>  <div>Url:  <a property="url" href="http://snps-and-go.biocomp.unibo.it/snps-and-go/index.html"> http://snps-and-go.biocomp.unibo.it/snps-and-go/index.html </a>  </div>  <div>Type:  <span property="applicationCategory">Tool</span>  </div>  <div>Input:  <div property="potentialAction">  <div typeof="<https://schema.org/ControlAction>">  <div property="object">  <div typeof="<https://schema.org/Dataset>">  <span property="additionalType" content=”<http://edamontology.org/data_3021>”>  UniProt accession  </span>  </div>  </div>  <div property="result">  <div typeof="<https://schema.org/Dataset>">  <span property="additionalType" content="<http://edamontology.org/data_0896>">  Protein report  </span>  </div>  </div>  </div>  </div>  </div>  <div>Keywords:  <span property="keywords">DNA polymorphism, Bioinformatics, Protein properties</span>  </div>  <div>Version:  <span property="softwareVersion">1.0</span>  </div>  <div>Function:  <span property="featureList">Variant classification</span>  </div>  <div>Contacts:  <div property="publisher">  <div typeof="<http://schema.org/Person>">  <div property="name">Rita Casadio</div>  <div property="email">casadio@biocomp.unibo.it</div>  </div>  <div typeof="<http://schema.org/Person>">  <div property="name">Pier Luigi Martelli</div>  <div property="email">gigi@biocomp.unibo.it</div>  </div>  </div>  </div>  <!-- The following is to appease Google's Structured Data  Testing Tool, which requires at least two properties from these four: operatingSystem, aggregateRating, applicationCategory, offers. -->  <meta property="operatingSystem" content="All" />  </div>  Microdata:  <div itemscope itemtype="http://schema.org/SoftwareApplication">  <div itemprop="name">SNPs and GO</div>  <div itemprop="description">SNPs&GO is a server for the prediction of single point protein mutations likely to be involved in the insurgence of diseases in humans.</div>  <div>Url:  <a itemprop="url" href="http://snps-and-go.biocomp.unibo.it/snps-and-go/index.html"> http://snps-and-go.biocomp.unibo.it/snps-and-go/index.html </a>  </div>  <div>Type:  <span itemprop="applicationCategory">Tool</span>  </div>  <div>Input:  <div itemprop="potentialAction">  <div itemtype="<https://schema.org/ControlAction>">  <div itemprop="object">  <div itemtype="<https://schema.org/Dataset>">  <span itemprop="additionalType" content=”<http://edamontology.org/data_3021>”>  UniProt accession  </span>  </div>  </div>  <div itemprop="result">  <div itemtype="<https://schema.org/Dataset>">  <span itemprop="additionalType" content="<http://edamontology.org/data_0896>">  Protein report  </span>  </div>  </div>  </div>  </div>  </div>  <div>Keywords:  <span itemprop="keywords">DNA polymorphism, Bioinformatics, Protein properties</span>  </div>  <div>Version:  <span itemprop="softwareVersion">1.0</span>  </div>  <div>Function:  <span itemprop="featureList">Variant classification</span>  </div>  <div>Contacts:  <div itemprop="publisher">  <div itemscope itemtype="<http://schema.org/Person>">  <div itemprop="name">Rita Casadio</div>  <div itemprop="email">casadio@biocomp.unibo.it</div>  </div>  <div itemscope itemtype="<http://schema.org/Person>">  <div itemprop="name">Pier Luigi Martelli</div>  <div itemprop="email">gigi@biocomp.unibo.it</div>  </div>  </div>  </div>  <!-- The following is to appease Google's Structured Data  Testing Tool, which requires at least two properties from these four: operatingSystem, aggregateRating, applicationCategory, offers. -->  <meta itemprop="operatingSystem" content="All" />  </div> |

# **Appendix I: Describing input and output formats**

Schema.org does not have the concept of ‘input’ and ‘output’. You can’t easily describe how a process changes something. Here we want to describe the input data format of some software and the output data format.

By ‘Data format’ we don’t mean file type (XML, JSON, CSV etc) but a descriptive keyword or phrase like ‘protein sequence’ or ‘gene name’, or a link to such a term in the [EDAM Data ontology](http://bioportal.bioontology.org/ontologies/EDAM?p=classes&conceptid=data_0006). We use the potentialAction attribute of Thing to describe input and output. The object is the input format and result is the output format. See the [GitHub issue](https://github.com/schemaorg/schemaorg/issues/1431) for a discussion about this.

|  |
| --- |
| <script type="application/ld+json"> {  "@context": "<http://schema.org>",  "@type": "SoftwareApplication",  "name": "SignalP",  …  "potentialAction": {  "@type": "ControlAction",  "object": {  "@type": "Dataset",  "additionalType": "http://edamontology.org/data\_2044"  }, "result": {  "@type": "Dataset",  "additionalType": "http://edamontology.org/data\_2544" }  },  …  }  </script> |

# 