Chapter 1

Introduction

The RAKIP Markup Language (RakML) is an XML-based format for the description of model metadata.

...

Chapter 2

Technical specification

2.1 Primitive data types

The primitive data types used in RAKIP-ML are taken from the XML Schema 1.0 including: **string**, **boolean**, **int** and **date**.

2.2 General structure

Every RAKIP model involves four main metadata components: general information, scope, data background and model math. A RAKIP-ML document has one model with these components.

2.3 Common types

...

2.3.1 Assay

Element	Type	Min. Ocurrences	Max. Ocurrences
Name	string	1	1
Description	string	0	1
MoisturePercentage	string 0	1	
FatPercentage	string	0	1
DetectionLimit	string	0	1
QuantificationLimit	string	0	1
LeftCensoredData	string	0	1
ContaminationRange	string	0	1
UncertaintyValue	string	0	1

Name A name given to the assay.

Description General description of the assay. Corresponds to the Protocol REF in ISA.

MoisturePercentage Percentage of moisture in the original sample.

FatPercentage Percentage of fat in the original sample.

DetectionLimit Limit of detection reported in the unit specified by the variable "Hazard Unit".

QuantificationLimit Limit of quantification reported in the unit specified by the variable "Hazard Unit".

LeftCensoredData Percentage of measures equal to LOQ and/or LOD.

ContaminationRange Range of result of the analytical measure reported in the unit specified by the variable "Hazard unit".

UncertaintyValue Indicate the expanded uncertainty (usually 95% confidence interval) value associated with the measurement expressed in the unit reported in the field "Hazard unit".

Listing 2.1: Example of Assay

```
<Name>Bradford protein assay</Name>
<Description>spectroscopic analytical procedure used to measure
    the concentration of protein in a solution. It is subjective,
    i.e., dependent on the amino acid composition of the
    measured protein.
</Description>
<DetectionLimit>30-300</DetectionLimit>
<QuantificationLimit>5000 - 8000</QuantificationLimit>
<ContaminationRange>500-4000</ContaminationRange>
```

2.3.2 Contact

Element	Type	Min. Ocurrences	Max. Ocurrences
Title	string	0	1
FamilyName	string	0	1
GivenName	string	0	1
Email	string	1	1
Telephone	string	0	1
StreetAddress	string	0	1
Country	string	0	1
City	string	0	1
ZipCode	string	0	1
Region	string	0	1
TimeZone	string	0	1
Gender	string	0	1
Note	string	0	1
Organization	string	0	1

Listing 2.2: Example of Contact

```
<Title>Dr.</Title>
<FamilyName>Romanov</FamilyName>
<GivenName>Natalia</GivenName>
<Email>black_widow@marvel.com</Email>
<Telephone>030 12345</Telephone>
<StreetAddress>Nahmitzer Damm 40</StreetAddress>
<Country>Russian Federation</Country>
<City>Berlin</City>
<Region>Berlin-Brandenburg</Region>
<Organization>SHIELD</Organization>
```

2.3.3 Hazard

Element	Type	Min. Occurrences	Max. Occurrences
Type	string	0	1
Name	string	1	1
Description	string	0	1
Unit	string	0	1
AdverseEffect	string	0	1
SourceOfContamination	string	0	1
BenchmarkDose	string	0	1
MaximumResidueLimit	string	0	1
NoObservedAdverseAffectLevel	string	0	1
AcceptableOperatorExposureLevel	string	0	1
AcuteReferenceDose	string	0	1
AcceptableDailyIntake	string	0	1
IndSum	string	0	1

Type General classification of the hazard for which the model or data applies.

Name Name of the hazard for which the model or data applies.

Description Description of the hazard for which the model or data applies.

Unit Unit of the hazard for which the model or data applies.

AdverseEffect Morbidity, mortality, origin.

SourceOfContamination Source of contamination, origin.

BenchmarkDose A dose or concentration that produces a predetermined change in response rate of an adverse effect (called the benchmark response or BMR) compared to background.

MaximumResidueLimit International regulations and permissible maximum residue levels in food and drinking water.

NoObservedAdverseAffectLevel Level of exposure of an organism, found by experiment or observation, at which there is no biologically or statistically significant increase in the frequency or severity of any adverse effects in the exposed population when compared to its appropriate control.

LowestObservedAdverseAffectLevel Lowest concentration or amount of a substance found by experiment or observation that causes an adverse alteration of morphology, function, capacity, growth, development, or lifespan of a target organism distinguished from normal organisms of the same species under defined conditions of exposure.

AcceptableOperatorExposureLevel Maximum amount of active substance to which the operator may be exposed without any adverse health effects. The AOEL is expressed as milligrams of the chemical per kilogram body weight of the operator.

AcuteReferenceDose An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure for an acute duration (24 hours or less) to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.

AcceptableDailyIntake Measure of amount of a specific substance in food or in drinking water tahta can be ingested (orally) on a daily basis over a lifetime without an appreciable health risk.

IndSum Define if the parameter reported is an individual residue/analyte, a summed residue definition or part of a sum a summed residue definition.

Listing 2.3: Example of Hazard

```
<Type>Organic contaminants</Type>
<Name>Norovirus (Norwalk-like virus)</Name>
<Description>norovirus is described as nast and hard to get rid
of</Description>
<Unit>CFU</Unit>
<AdverseEffect>morbitity</AdverseEffect>
<SourceOfContamination>sewage</SourceOfContamination>
<MaximumResidueLimit>0.01 mg/kg</MaximumResidueLimit>
<NoObservedAdverseAffectLevel>10 mg</NoObservedAdverseAffectLevel>
<LowestObservedAdverseAffectLevel>40 mg</LowestObservedAdverseAffectLevel>
<AcceptableOperatorExposureLevel>50 mg</AcceptableOperatorExposureLevel>
<AcuteReferenceDose>80 mg</AcuteReferenceDose>
<AcceptableDailyIntake>20 mg</AcceptableDailyIntake>
```

2.3.4 Laboratory

Element	Type	Min. Occurrences	Max. Occurrences
Accreditation	string	0	1
Name	string	0	1
Country	string	0	1

Accreditation The laboratory accreditation to ISO/IEC 17025.

Name Laboratory code (National laboratory code if available) or Laboratory name

Country Country where the laboratory is placed. (ISO 3166-1-alpha-2).

Listing 2.4: Example of Laboratory

```
<Accreditation>Accredited</Accreditation>
<Name>National High Magnetic Field Laboratory</Name>
<Country>United States</Country>
```

2.3.5 ModelCategory

Element	Type	Min. Occurrences	Max. Occurrences
ModelClass	string	1	1
ModelSubClass	string	0	1
ModelClassComment	string	0	1
BasicProcess	string	0	1

ModelClass Type of model used to build-up the risk assessment structure.

ModelSubClass Sub-cassification of the model given the Model Class

BasicProcess Defines the impact of the specific process on the hazard

Listing 2.5: Example of ModelCategory

```
<ModelClass>Dose-response model</ModelClass>
<ModelClassComment>This Model Class is very special</ModelClassComment>
```

2.4 GeneralInformation

Element	Type	Min. Ocurrences	Max. Ocurrences
Name	string	1	1
Source	string	0	1
Identifier	string	1	1
Author	Contact	0	unbounded
Creator	Contact	1	1
CreationDate	date	1	1
ModificationDate	date	0	unbounded
Rights*	string	1	1
Available	string	0	1
Format	string	0	1
Reference	Reference	1	unbounded
Language	string	0	1
Software	string	0	1
LanguageWrittenIn	string	0	1
ModelCategory	ModelCategory	0	1
Status	string	0	1
Objective	string	0	1
Description	string	0	1

Name Name given to the model or data.

Source A related resource from which the described resources is derived.

Identifier An unambiguous ID given to the model or data.

Author Person who generated the model code or generated the data set originally.

Creator The person responsible for creating the model file in the present form or the person responsible for creating the data file in the present form.

CreationDate Temporal information on the model creation date.

ModificationDate Temporal information on the last modification of the model.

Rights Information on rights held in and over the resource.

Available Availability of data or model.

Format Form of model or data (file extension).

Reference

Language Language of the resource.

ABST	CHAP	DICT	GEN	MANSCPT	PCOMM	VIDEO
ADVS	CHART	EBOOK	GOVDOC	MAP	RPRT	
AGGR	CLSWK	ECHAP	GRANT	MGZN	SER	
ANCIENT	COMP	EDBOOK	HEAR	MPCT	SLIDE	
ART	CONF	EDJOUR	ICOMM	MULTI	SOUND	
BILL	CPAPER	ELECT	INPR	MUSIC	STAND	
BLOG	CTLG	ENCYC	JOUR	NEW	STAT	
BOOK	DATA	EQUA	JFULL	PAMP	THES	
CASE	DBASE	FIGURE	LEGAL	PAT	UNPB	

Table 2.1: Publication types

Software Program in which the model has been implemented.

LanguageWrittenIn Language used to write the model, e.g. R or Matlab.

ModelCategory

Status Curation status of the model.

Objective Objective of the model or data.

Description General description of the study, data or model.

2.4.1 Reference

Element	Type	Min. Ocurrences	Max. Ocurrences
IsReferenceDescription*	boolean	1	1
Type	string	0	1
Date	string	0	1
Pmid	string	0	1
Doi	string	0	1
AuthorList	string	0	1
Title	string	1	1
Abstract	string	0	1
Journal	string	0	1
Volume	int	0	1
Issue	int	0	1
Status	string	0	1
Website	string	0	1
Comment	string	0	1

IsReferenceDescription Indicates whether the publication serves as the reference description for the model.

Type Type of the publication. Takes a value from the reserved words listed at 2.1.

Year Temporal information on the publication date.

Pmid The PubMed ID related to this publication.

Doi The DOI related to this publication.

AuthorList Name and surname of the authors who contributed to this publication.

Title Title of the publication in which the model or the data has been described.

Abstract Abstract of the publication in which the model or the data has been described.

Journal Publication journal.

Volume Publication volume.

Issue Publication issue.

Status Publication status.

Website Publication website.

Comment Publication comment.

Listing 2.6: Example of Reference

```
<Reference>
    <IsReferenceDescription>true</IsReferenceDescription>
    <Type>PAMP</Type>
    <Date>3805-07-02</Date>
    <Doi>10.1111/risa.12758</Doi>
    <AuthorList>Jack Bauer, Kiefer Sutherland</AuthorList>
    <Title>Quantitative Risk Assessment of Norovirus Transmission
        in Food Establishments: Evaluating the Impact of
        Intervention Strategies and Food Employee Behavior on the
        Risk Associated with Norovirus in Foods.
    </Title>
    <Abstract>
        This research looks at the work of Margaret C. Anderson,
        the editor of the Little Review. The review published
        first works by Sherwood Anderson, James Joyce, Wyndham
        Lewis, and Ezra Pound. This research draws upon mostly
        primary sources including memoirs, published letters, and
        a complete collection of the Little Review. Most prior
        research on Anderson focuses on her connection to the
        famous writers and personalities that she published and
        associated with. This focus undermines her role as the
        dominant creative force behind one of the most influential
        little magazines published in the 20th Century. This case
        example shows how little magazine publishing is arguably a
```

```
        literary art.
    </Abstract>
    <Status>Accepted</Status>
    <Website>https://nature.com</Website>
    <Comment>publisher demands edits</Comment>
</Reference>
```