

スキーマ日本語化ファイル の見方

はじめに

Open mHealthスキーマ日本語化のためのファイルを見られる方へ

2024年度Open mHealth班報告書の添付資料であるスキーマ日本語化ファイルはプログラムの知識がない方でもスキーマの内容に対してある程度の議論ができるようにと作成したのですが、残念ながら現在のバージョンはJSON Schemaの知識がないと理解困難です。

そこで、JSON Schemaの知識がない人でも何となく理解できるように本資料を作成しました。わかりやすさを優先して厳密さを犠牲にしていますので、そこに関してはご理解いただけますと幸いです。

文責：山下 暁士

ファイルの概要

1つのファイルにはそれに関連した複数のスキーマが記載されています。
1シートに1スキーマが記載されています（下記は1例です）。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Sachema	http://json-schema.org/draft-07/schema#												
2	Sid	https://w3id.org/ieee/ieee-1752-schema/physical-activity.json												
3	title	Physical Activity												
4	type	object												
5	description	このスキーマは身体活動のエピソードを表す												
6														
7	references	value		description			url			コメント				
8														
9														
10	definitions				Sref				コメント					
11	length_unit_value				length-unit-value-1.0.json				propertyは単位と値。Unitには["fm","pm","nm","um","mm","cm","m","km","in","ft","yd","mi"]から用いるものを選び					
12	unit_value				unit-value-1.0.json				propertyは単位と値。					
13	duration_unit_value				duration-unit-value-1.0.json				propertyは単位と値。単位には["ps","ns","us","ms","sec","min","h","d","wk","Mo","yr"]から用いるものを選び					
14	kcal_unit_value				kcal-unit-value-1.0.json				propertyは単位と値。単位には"kcal"を用いる					
15	speed_unit_value				speed-unit-value-1.0.json				propertyは単位と値。単位には["m/s","km/h"]から用いる単位を選び					
16	time_frame				time-frame-1.0.json				時間枠は日時と時間区間の2タイプが許可されているが、ここでは時間区間のみが用いられる。時間区間は開始時間と持続時間、終了時間と持続時間、開始時間と終了時間の3種類から選べる					
17	descriptive_statistic				descriptive-statistic-1.0.json				使用できる統計量は 平均、総数、最大値、メジアン、最小値、標準偏差、合計、分散、20パーセンタイル値、80パーセンタイル値、下位四分位値、上位四分位値、四分位偏差、第1(2・3・4)五分位値					
18	descriptive_statistic_denominator				descriptive-statistic-denominator-1.0.json				測定が定期的な期間行われる場合の記述統計量の分母。["d","wk","Mo","episode","meal","session"]から選択する。これが"d"だと、1日当たりの平均・・・、episodeだと1エピソード当たりの最大・・・となる					
19														
20	Sref													
21														
22	properties	必須		description		Sref	references/description		references/url		type	format	データ内容	コメント
23	activity_name	○		その人が従事している身体活動の名称。これは座って行うものでもよい。活動名は、標準エネルギー消費量(METs)へのマッピングを容易にするための、最新の CDC ガイドラインまたは他のピアレビューされた情報源から引用することが推奨される。			標準エネルギー消費量 (METs) に関するCDCガイドライン	https://www.cdc.gov/nccdphp/dnpa/physical/pdf/pa_intensity_table_2_1.pdf		string				
24							2011年身体活動大全	https://www.ncbi.nlm.nih.gov/pubmed/21681120						
25							身体活動オントロジー	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658272/						
	base_movement_quantity			身体活動の基本動作の繰り返し回数（該当する場合）（例えば、活動が歩行の場合、base_movement_quantityは歩数となる）。列挙された単位のいずれにも当てはまらない場合は、単位を明示的に指定する必要がある。		#/definitions/unit_value								

ヘッダー

	A	B	C	D	E	F
1	Schema		http://json-schema.org/draft-07/schema#			
2	Sid		https://w3id.org/ieee/ieee-1752-schema/physical-activity.json			
3	title		Physical Activity			
4	type		object			
5	description		このスキーマは身体活動のエピソードを表す			

この部分はスキーマのヘッダーです。

2行目はスキーマの**UUID**(この世で一意的IDがURL形式で記載されている)

3行目はこのスキーマの**タイトル**

4行目はこのスキーマの**データ型**(ほとんどがオブジェクト型ですが)

5行目が言葉による簡単なスキーマの説明

外部参照

7	references	value	description	url	コ
8					

この部分はこのスキーマ全体に適用される外部参照を記載するところです。

Open mHealth規格では、スキーマ全体に関係する別の標準（例えばSNOMEDなど）に関するURLとその簡単な説明(description)が記載されていることが多いです。スキーマによっては、スキーマの戻り値(value)が何を示すのかを説明(description)する場合があります

7	references	value	description	url
8			SNOMEDコードは頻度（定量的概念）を表す	http://purl.bioontology.org/ontology/SNOMEDCT/260864003

7	references	value	description	url
8		nth percentile	順番に並んだスコアを均等に100分割するポイントの中のn番目のポイント	
9		lower quartile	データの下半分の中央値	
10		upper quartile	データの上半分の中央値	
11		quartile deviation	分布において、上位四分位値と下位四分位値の差の半分（半四分位範囲）。	
12		1st quintile	それぞれが母集団全体の5分の1を含むように全体を5つに分ける4つの値のうち最低値	
13		2nd quintile	それぞれが母集団全体の5分の1を含むように全体を5つに分ける4つの値のうち低い方から2番目の値	
14		3rd quintile	それぞれが母集団全体の5分の1を含むように全体を5つに分ける4つの値のうち低い方から3番目の値	
15		4th quintile	それぞれが母集団全体の5分の1を含むように全体を5つに分ける4つの値のうち最高値	

定義

10	definitions	\$ref	コメント
11	length_unit_value	length-unit-value-1.0.json	propertyは単位と値。Unitには["fm","pm","nm","um","mm","cm","m","km","in","ft","yd","mi"]から用いるものを選択
12	unit_value	unit-value-1.0.json	propertyは単位と値。
13	duration_unit_value	duration-unit-value-1.0.json	propertyは単位と値。単位には["ps","ns","us","ms","sec","min","h","d","wk","Mo","yr"]から用いるものを選択
14	kcal_unit_value	kcal-unit-value-1.0.json	propertyは単位と値。単位には"kcal"を用いる
15	speed_unit_value	speed-unit-value-1.0.json	propertyは単位と値。単位には["m/s","km/h"]から用いる単位を選択
16	time_frame	time-frame-1.0.json	時間枠は日時と時間区間の2タイプが許容されているが、ここでは時間区間のみが用いられる。時間区間は開始時間と持続時間、終了時間と持続時間、開始時間と終了時間の3種類から選べる

この部分はこのスキーマ全体に適用される定義を記載するところです。

definitionsで定義される語がどのスキーマを参照するものかをここに記載することで、繰り返しスキーマファイルの名称を記載する必要がなくなります

→ 結果として、そのスキーマファイルのバージョンが上がった際に定義の部分さえ書き換えればスキーマ本体を修正する必要がなくなり、書き換え忘れによるエラーを防止することができます。

\$refに参照するスキーマファイルが記載されています。コメントは日本語化の担当者が付けた参照先のスキーマファイルの簡単な説明です。

参照

20	\$ref	
----	-------	--

\$refが定義やプロパティではなく、スキーマにじかに記載されている場合は、そのスキーマ自体が参照するスキーマを表します。それを記載するのがこの部分です。簡単に言えば、そのスキーマが規定するJSONの戻り値(?)のデータ型がどのスキーマで決まっているかを表現したものとなります。

referencesがこのスキーマが参照する（主に）ドキュメント情報を記載するものとするれば、この\$refはこのスキーマが参照するプログラムの情報を記載するものと思っています（わかりにくいですかね）。

プロパティ 1 —総論—

22	properties	必須	description	\$ref	references/description	references/url	type	format	データ内容	コ.
23	activity_name	○	その人が従事している身体活動の名称。これは座って行うものでもよい。活動名は、標準エネルギー消費量（METs）へのマッピングを容易にするため、最新の CDC ガイドラインまたは他のピアレビューされた情報源から引用することが推奨される。		標準エネルギー消費量（METs）に関する CDC ガイドライン	https://www.cdc.gov/nccdphp/dnpa/physical/pdf/pa_intensity_table_2_1.pdf	string			
24					2011年身体活動大全	https://www.ncbi.nlm.nih.gov/pubmed/21681120				
25					身体活動オントロジー	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658272/				
	base_movement_quantity		身体活動の基本動作の繰り返し回数（該当する場合）（例えば、活動が歩行の場合、base_movement_quantityは歩数となる）。列挙された単位のうちいずれにも当てはまらない場合は、その列挙された単位は、列挙された単位とは異なる単位である。	#/definitions/unit_value						

プロパティは各スキーマが定義しているJSONの状態を格納する変数みたいなものだと思います。

日本語化ファイルには必ずしもそのJSONスキーマのすべてのプロパティが記載されているわけではありません。スキーマそのもの、definitions、プロパティなどの\$refで参照されたスキーマで定義されているプロパティが適用されるためです。

実際にデータを利用する際には、Metadataスキーマで定義されているData-pointスキーマかData-seriesスキーマのいずれかが用いられることになっているため、実際のデータには必ずHeaderスキーマで定義されたプロパティ（たとえば、（スキーマ名).header.modalityなど）がデータ本体（各スキーマで定義される）とともに存在することになります。

プロパティ 2 — 各論 —

++	properties	必須	description	Sref	references/description	references/url	type	format	データ内容	コ.
23	activity_name	○	その人が従事している身体活動の名称。これは座って行うものでもよい。活動名は、標準エネルギー消費量 (METs) へのマッピングを容易にするため、最新の CDC ガイドラインまたは他のピアレビューされた情報源から引用することが推奨される。		標準エネルギー消費量 (METs) に関する CDC ガイドライン	https://www.cdc.gov/nccdphp/dnpa/physical/pdf/pa_intensity_table_2_1.pdf	string			
24					2011年身体活動大全	https://www.ncbi.nlm.nih.gov/pubmed/21681120				
25					身体活動オントロジー	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658272/				
	base_movement_quantity		身体活動の基本動作の繰り返し回数 (該当する場合) (例えば、活動が歩行の場合、base_movement_quantityは歩数となる)。列挙された単位のうちいずれにも当てはまらない場合は、空欄の状態で記述すること。列挙されていない単位も除く。	#/definitions/unit_value						

プロパティ名。子プロパティは親プロパティの下に段落下げした形で記載

必須プロパティかどうか。必須とはプロパティが必ず値を持っているということとってください

プロパティの簡単な説明が記載されています

プロパティ 2 — 各論 —

properties	必須	description	Sref	references/description	references/url	type	format	データ内容	コ.
activity_name	○	その人が従事している身体活動の名称。これは座って行うものでもよい。活動名は、標準エネルギー消費量 (METs) へのマッピングを容易にするため、最新の CDC ガイドラインまたは他のピアレビューされた情報源から引用することが推奨される。		標準エネルギー消費量 (METs) に関する CDC ガイドライン 2011年身体活動大全 身体活動オントロジー	https://www.cdc.gov/nccdphp/dnpa/physical/pdf/pa_intensity_table_2_1.pdf https://www.ncbi.nlm.nih.gov/pubmed/21681120 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658272/	string			
base_movement_quantity		身体活動の基本動作の繰り返し回数 (該当する場合) (例えば、活動が歩行の場合、base_movement_quantityは歩数となる)。列挙された単位のうちいずれにも当てはまらない場合は、空欄の値を指定する。列挙された単位は、列挙されていない単位を除く。	#/definitions/unit_value						

そのプロパティが参照するJSONスキーマ。参照するスキーマの定義に従い、子プロパティができる場合もある

プロパティが参照するURLに関する簡単な説明

プロパティを使用するにあたって参照すべき情報が記載されているURL
主に外部の標準に関する情報が記載されている

スキーマのデータ型。参照スキーマがない場合に規定される

データのフォーマット。
YYYYMMDDなど

取りうる値が決まっていそれを列挙するときなどに記載

その他

50	値	
51	format	
52		

プロパティのformatとデータの内容（=値）と同じもの。そのスキーマの値を示すプロパティ（value）の値のデータフォーマットと取りうる値が記載されている

値	["sitting","lying down","standing","semi-recumbent"]
format	

値の例(体位)

値	
format	date-time

データフォーマットの例(日時)

下記のように勉強会で議論されたことや担当者の意見、注意すべき点などの記載欄をいくつか用意していますので、それらもご参照ください。

Note

議論内容

日本語化する方法に関する質問が大部分（日本語化に関する説明を一緒に実施したため）

主な質問

Q：このExcelファイルをどのような形で提供する形にするのかを事前に決めておいた方がいいのでは→相互参照をExcelで指定する形にするか、個々のファイルの関係を別のドキュメントに記載するかするしなければ、全体像を確認することができない

A：あるべき姿を今すぐ決めること自体がなれていないので難しい→当面は日本語化を進めてみてその経験をもとにどのような方法を採用するのか決める形にするべきか。Excelファイル自体のブラッシュアップも継続的にすべき

担当者意見

このスキーマは身体活動のデータを網羅的に保存するためのもので、実際に使用する際にはウォーキング、ランニング、縄跳び、ダンスなど運動ごとに実装のための約束事を決める必要がある可能性が高い

基本的には運動名 → METS変換 → 測定値保存で済むと思われるが、強度の分類などは参考文献（referenceで定義されている）を見ながら決めざるを得ないところがありそう