

【データモデリング・アンソロジー】 論理/物理データモデル活用さまざま - 「マスタ = リソース」発想だけでは、 無理を生む?! ... などなど

説明日: 2021年1月29日

インフォラボ遊悠 (ゆうゆう) PB

中岡 実

minoru.nakaoka@infolabyouyou.com

https://info.infolabyouyou.com

「インフォラボ遊悠」およびロゴマークは、当代表者の登録商標です。

Eyes Wide Open



Eyes Wide Open

【今日の話】 データモデリングを柱に様々な視点から

- (話題1) データモデリングの位置付け/目的/意義を再考する
- (話題2) データモデリング・アプローチの視点から考える
- (話題3) マスタデータ = リソースデータの発想を見直す
- (話題4) リレーショナルスキーマ以外のモデルへの適用を考察する
- (話題5) データモデリングの適用範囲の更なる拡大性を考える
- (話題6) 第10分科会で出てきた「疑問」を取り上げる

その前に インフォラボ遊悠 (ゆうゆう) - ご紹介

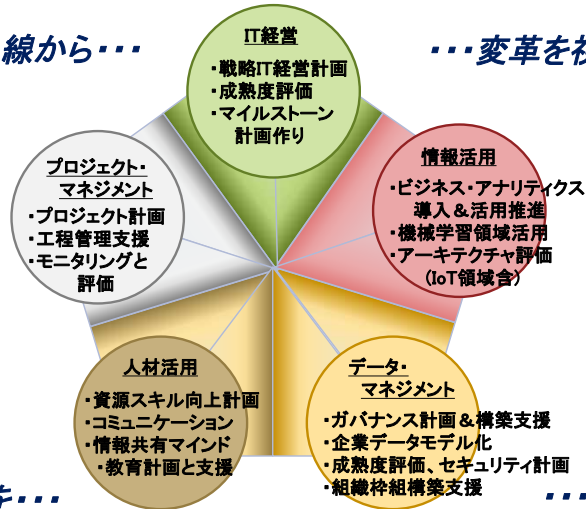
目標: 「情報力」、「人間力」の拡大に尽力
ビジネスとITの活用・融合領域のサービス&ご支援範囲(概要):



「インフォラボ遊悠」およびロゴは、当代表者による登録商標です。

IT経営の目線から...

...変革を視野に入れ



望ましい姿を...

...より俊敏に

Eyes Wide Open

Eyes Wide Open

<代表者ご紹介>



監修者
2003年(翔泳社)



EDW2020 San Diego
オンライン講演実施
(Video packageにて
限定公開)

名前: 中岡 実(なかおか みのる)
役割: インフォラボ遊悠 代表/所長 兼 チーフコンサルタント
背景など:

- 情報セキュリティ・マネジメント(経済産業省)
- ITコーディネータ(ITCA認定)
- PMP(プロジェクトマネジメント国際資格、米国PMI会員)
- 認定心理士(日本心理学会)
- CBA(Certification of Business Administration: グロービスマネジメントスクール)
- データ解析・統計関連知識など((財)実務教育研究所認定他)
- 学術修士(Master of Arts)
- 東商2級カラーコーディネーター認定

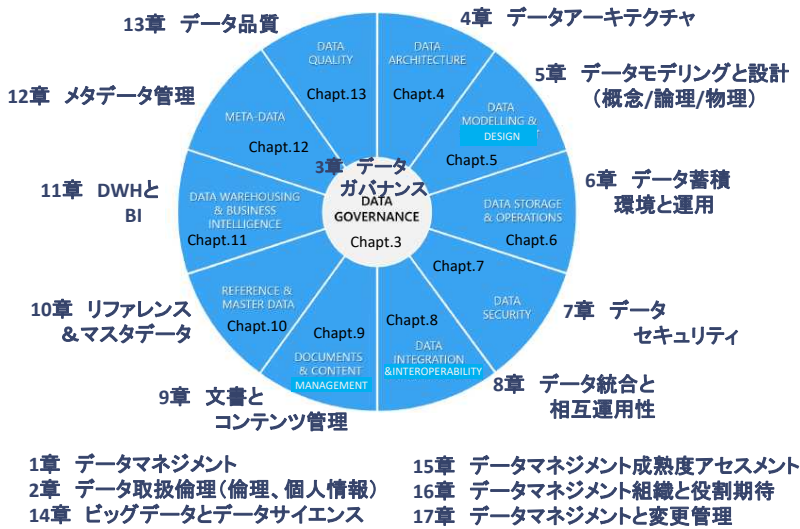
活動略歴:

- 2014年まで米国系ビジネスインテリジェンスベンダーにて、コンサルティング、ビジネスインテリジェンス/DWH系プロジェクト 参画経験多数(金融、保険、通信、製造、流通他、各種業界大手企業の情報システム計画、構築コンサル&プロジェクト管理を経験)
- 日本データマネジメント・コンソーシアム(JDMC)設立・運営に参画
- データマネジメント協会(DAMA)日本支部会員...分科会活動参画中
- データモデリング関連活動(JEMUGメンバ) ...JEMUG賞受賞等
- その他、国内大手ベンダーにて、コンピュータOS開発経験など

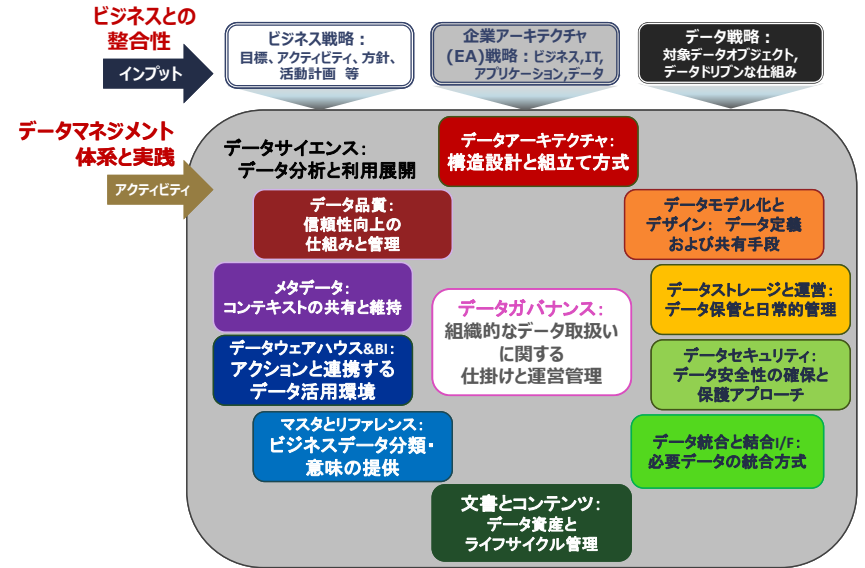
他文化系活動:

- 元NPO中国茶文化協会認定中国茶コーディネータ
- 元NPO日本健康太極拳協会 師範(2020年3月退会)
- 易学研究者(周易、五行易) ...「易学研究」誌執筆(筆名:普 人士) 等
- 趣味 二胡演奏

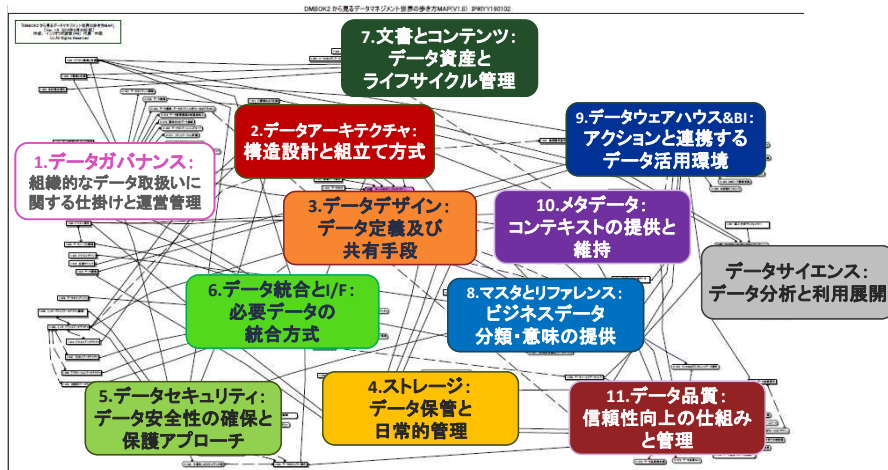
データマネジメント知識体系ガイド(DMBOK 第2版)の内容構成(知識領域)



DMBoK2 データマネジメント歩き方マップから捉えるポイント



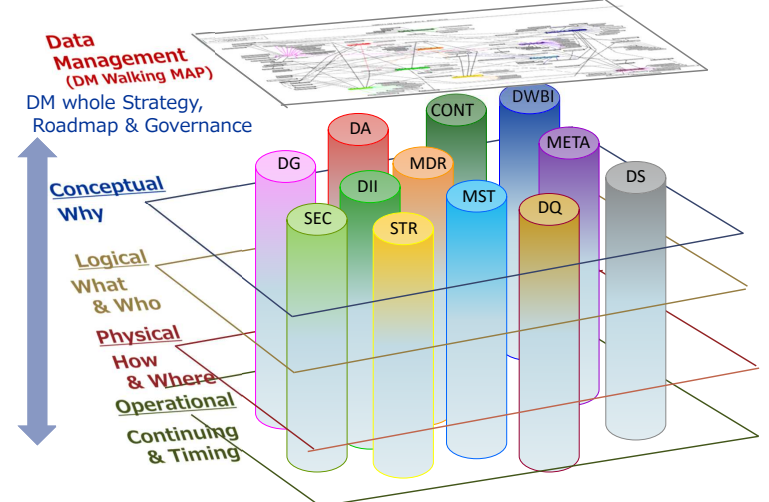
【参考】DMBOK2 歩き方マップ Ver. 1.8 (日本語版、英語版)



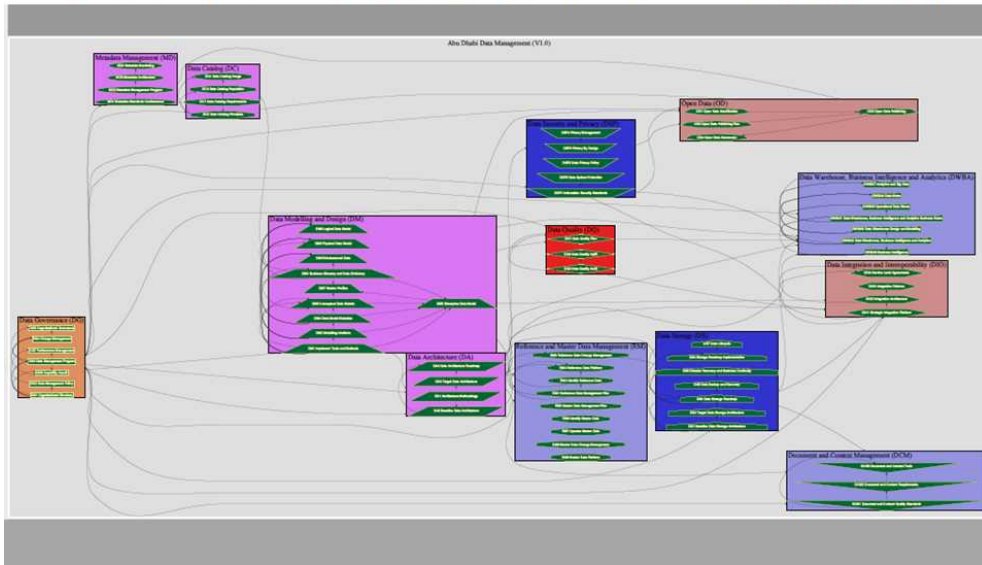
ビジネス活動には、俯瞰的視点からの有機的な「データマネジメント発想」が重要である

Hierarchical view(3D+Time) of Data Management

> Our targets for Data Management reside in a hierarchical (3D) and time sensitive structure. We live in a real DM world of complex networks.



【参考】Abu Dhabi データマネジメント標準 (1.0) 統制要素の関連概要図 (2020/07)



©2020.インフォラボ遊悠, All Rights Reserved

9

Eyes Wide Open

【今日の話】 データモデリングを柱に様々な視点から

- (話題1) データモデリングの位置付け／目的／意義を再考する
- (話題2) データモデリング・アプローチの視点から考える
- (話題3) マスタデータ = リソースデータの発想を見直す
- (話題4) リレーショナルスキーマ以外のモデルへの適用を考察する
- (話題5) データモデリングの適用範囲の更なる拡大性を考える
- (話題6) 第10分科会で出てきた「疑問」を取り上げる

©2020.インフォラボ遊悠, All Rights Reserved

10

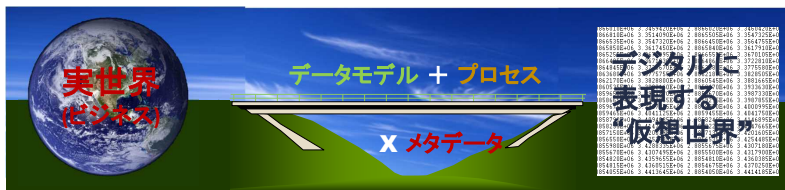
Eyes Wide Open

【一つの考え方】 データモデルとは？ - 世の中を図式的に表現する方法

デジタルな世界: データ(信号)をやり取りし、論理で結びついた一連の処理を実行する。

現実世界の模倣(モデル化):

⇒ [モデルで表わされる“仮想的現実世界”構造] + [プロセス] x [メタデータ(文脈)]



特に、関係者(人・コンピュータ)が“意味を交換する”には、
([データモデル] + [プロセスモデル]) x [メタデータ]
による意図の共有・連携が欠かせない。..... コミュニケーションの基礎

©2020.インフォラボ遊悠, All Rights Reserved

11

Eyes Wide Open

データ活用に向けて: 論理と物理モデル、物理実装 (データ理解) を支援

●今更の、クラウド連携時代のデータ活用における、データモデルの位置付け考察

【これまでのデータモデル記述】

- 論理データモデル図
 - ・ER図、その他記法を利用
- 物理データモデル図
 - ・ER図+DBスキーマなどを利用

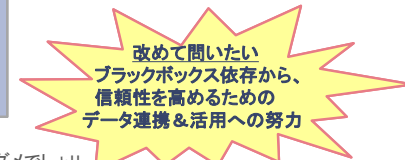


【データモデル記述への期待】

- 論理/物理実装 ネットワーク構造活用
- データモデル対象の多様性
 - ・多層的/多次的表現の採用
- ビッグデータ/事象発生表現を拡大
 - ・インスタンスの考え方をどう捉え、表すか
 - ・個別的意味表現の利用拡大

【期待役割の変化】

- ①記述厳密性 ⇒ コンテキスト整合性重視へ
 - 意味境界記述のドメイン世界
 - ・論理記述は概念整理寄り
 - ・物理は実装整理寄り
- ②RDBを意識した実装
 - ⇒ 実装の多様性と混在環境表現
 - 記述標準の選択の必要性
- ③外部データ/クラウドサービス(Web)連携の必要性の拡大
 - 論理コンテキストの整理&付与を期待
 - 俯瞰的アーキテクチャ指向の必要性



閑話休題: 「技術的にできる」だけではダメでしょ!!

例: とても迷惑なAmazon広告 ...一旦検索すると、どこへ行っても表示される商品の存在

©2020.インフォラボ遊悠, All Rights Reserved

12

Eyes Wide Open

データモデリングの位置付けに関する「誤解」(意識レベルでの過少・過大評価)

●「データモデル」、「データモデリング」の可能性について再考する

▶ 様々な「誤解」発言の存在(例)

- マネージャAさんの誤解
 - 「データモデルは役に立たない??」
 - ⇒ 使い方、利用範囲・応用について認知されていない
- 開発者Bさんの誤解
 - 「アジャイル時代」「JSONを使えばアプリ開発できているのでデータモデルなど不要。モデル屋さんが面倒みていれば良い」
 - ⇒ データモデルが物理実装のためのものだけと捉えている(開発者は単なる利用者意識)
- ベンダー技術者Cさんの誤解
 - 「自分が理解できるデータモデルこそ最高のモデル」
 - ⇒ データモデルが関係者間のコミュニケーション技術・手段と捉えていない(?)
- CIO/CDO Dさんの誤解
 - 「データモデルはできた。これで安心。大切に金庫に保管しておこう」
 - ⇒ 維持・利用継続を広める重要性の自覚に欠ける(TCO、資産化視点欠如)
- 経営者Eさんの誤解
 - 「ツール、AIの導入が済んだ。後は営業努力で利益が出せるはず」
 - ⇒ データ経営のための基盤的意義を十分認識できていない(資産化意識の不足)
- データモデラーFさんの誤解
 - 「自分のモデルは正しく描けている。理解できないあの人たちに問題がある」
 - ⇒ それで良い? / あきらめ?? / 被害意識???

◆データモデル、データモデリングは使い道・工夫次第で金の卵となりえる!?

◆但し、“One size does not fit ALL.” に注意を払うこと!!

©2020,インフォラボ遊悠, All Rights Reserved

Eyes Wide Open

13

【今日の話】 データモデリングを柱に様々な視点から

- (話題1) データモデリングの位置付け/目的/意義を再考する
- (話題2) データモデリング・アプローチの視点から考える
- (話題3) マスタデータ = リソースデータの発想を見直す
- (話題4) リレーショナルスキーマ以外のモデルへの適用を考察する
- (話題5) データモデリングの適用範囲の更なる拡大性を考える
- (話題6) 第10分科会で出てきた「疑問」を取り上げる

Eyes Wide Open

©2020,インフォラボ遊悠, All Rights Reserved

14

【DMBoK2】 5.1 イントロ 3. 基本概念 3.4 データモデリング・スキーマ

●データモデリングスキーマの利用種類は、実装するDBMSによって適切に使い分けことが期待される。データモデルのレベルと利用スキーマ、DBMSの種類との関係イメージ (Table 10)

スキーマ種類	DBMS種類						
	RDBMS	多次元DB (MDBMS)	オブジェクトデータベース	ドキュメントDB	カラムデータベース	グラフデータベース	キーベースデータベース
リレーショナル (Relational)	CDM LDM PDM	CDM LDM	CDM LDM	CDM LDM	CDM LDM	CDM LDM	CDM LDM
ディメンショナル (Dimensional)	CDM LDM PDM	CDM LDM PDM					
オブジェクト指向 (Object-Oriented)	CDM LDM PDM		CDM LDM PDM				
ファクトベース (Fact-Based)	CDM LDM PDM	CDM LDM	CDM LDM	CDM LDM	CDM LDM	CDM LDM	CDM LDM
タイムベース (Time-Based)	PDM						
NoSQL			PDM	PDM	PDM	PDM	PDM

備考: CDM (概念データモデル)、LDM (論理データモデル)、PDM (物理データモデル)

中間コメント: ここでLDMを考えた場合、グラフデータベースの場合のLDMでは、「コンテキスト」を表現するためのモデル図という意味合いを含むものとする(以前のJEMUGミーティングで説明)
 ・PDMについては、RDBMSにおけるLDM-PDM関係の表現はかなり異なったものになると考えられる

出典: DMBoK 2nd, P.136~137, Table 10 を元に中間訳 & 整理

©2020,インフォラボ遊悠, All Rights Reserved

Eyes Wide Open

15

トップダウンvs.ボトムアップのアプローチ・ギャップの認識典型例(1)

- ▶ [視点] **トップダウンアプローチ**では、ビジネスユーザとのコミュニケーション、スコーピング、用語定義、関係者調整等を重視
 - ✓ 経験的に、部門横断的に整備し、最新化に向けて定期メンテナンスが実施された用語集やコード定義資料が、その場ですぐに提供されることはない
 - ✓ 関係するビジネス部門が多いほど、データ項目に関する意味および利用法の調整時間は、非線形的に増大する
 - ✓ うまく進められれば、ビジネスメタデータ整備の良い機会・動機付けとなる
 - ✓ ITユーザを当初から巻き込めれば、後工程を進めやすくなることが期待できる
- ▶ [視点] **ボトムアップアプローチ**では、ITユーザ(システム関係者)とのコミュニケーション、スコーピング、コード確認、関係者調整等を重視
 - ✓ モデル定義のために調査対象となるシステムの識別、選択に手間を要することが少なくない(放っておくと膨大な数のファイル/テーブルが対象となる可能性)
 - ✓ 調査対象となるデータに関して、システム内およびシステム間横断的に、データ項目がどのように管理運用できているかを説明できる要員は大変限られていることが多い
 - ✓ ビジネスユーザの関与度が低いほど、現行システムの置き換えが主目的になる可能性が高い
 - ✓ うまく進められれば、データのリネージ関係を整理するきっかけとできる可能性が高い

Eyes Wide Open

©2020,インフォラボ遊悠, All Rights Reserved

16

トップダウンvs.ボトムアップのアプローチ・ギャップの認識典型例(2)

➤[視点]テーブルやファイルの論理/物理名称が「**XXXマスタ**」となっていることが少なくない(ボトムアップアプローチ)

✓主にレガシー系システムにおけるデータベースおよびシステムでの傾向

✓分類上や会話で「**XXXマスタ**」として取り上げられるならば、ギャップ問題は余り大きくならないと考えられる(コミュニケーションが取り易い)。

✓なぜなら「**XXXマスタ**」と論理/物理名称を付けているということは、エンティティを意識したデータモデル発想が欠けていることを意味するから(反論あるかもしれないが)

✓エンティティ構造を持つしっかりしたデータモデル設計発想からは、例えば「顧客^(※)」、「商品」がエンティティ名であり、「顧客マスタ」、「商品マスタ」とする考え方には行き着きにくいということ



データモデリング発想を組織的に広げるにはこの辺りのIT関係者認識変革が必要となることも!!

注※ ・ここでは分かり易さのために取り上げたが、説明者は「顧客」というエンティティ名利用を必ずしも推奨している訳ではない。

©2020,インフォラボ遊悠, All Rights Reserved

17

Eyes Wide Open

【今日の話】 データモデリングを柱に様々な視点から

(話題1) データモデリングの位置付け/目的/意義を再考する

(話題2) データモデリング・アプローチの視点から考える

(話題3) マスタデータ = リソースデータの発想を見直す

(話題4) リレーショナルスキーマ以外のモデルへの適用を考察する

(話題5) データモデリングの適用範囲の更なる拡大性を考える

(話題6) 第10分科会で出てきた「疑問」を取り上げる

©2020,インフォラボ遊悠, All Rights Reserved

18

Eyes Wide Open

マスタデータおよびリファレンスデータ管理の必要性

➤マスタデータ(広義)で扱われるリファレンスデータは、データに対する自企業の見方・共通視点を定義するもの

- ✓正確性に欠ければ、処理して出てくる情報も信頼できない
- ✓意味合いが共有されず、曖昧に使われれば、結果にも大きなブレが生まれる
- ✓漏れがあれば、見えるものが見えなくなる



・取扱いには手間が掛かるが、うまく共有できれば力強い味方になる(統制の肝)

➤マスタデータ(狭義)は、内容の一つ一つが具体的な対象物を表現している(主に、インスタンス視点に着目した利用の仕方)

- ✓例えば、顧客マスタは、ビジネスでの収益源泉の特性を表す
- ✓例えば、取引先マスタは、ビジネスパートナーの特性を表す
- ✓例えば、製品マスタ/商品マスタは、扱っている「モノ」が何かを表す



・マスタデータの内容・正確性(項目)を充実させて、より具体的な活動への手助けにする
・「データ資産」として捉えられるマスタデータは主にこちらの領域

データモデリングを通じた管理を重要視する理由

©2020,インフォラボ遊悠, All Rights Reserved

19

Eyes Wide Open

【今日の話】 データモデリングを柱に様々な視点から

(話題1) データモデリングの位置付け/目的/意義を再考する

(話題2) データモデリング・アプローチの視点から考える

(話題3) マスタデータ = リソースデータの発想を見直す

(話題4) リレーショナルスキーマ以外のモデルへの適用を考察する

(話題5) データモデリングの適用範囲の更なる拡大性を考える

(話題6) 第10分科会で出てきた「疑問」を取り上げる

©2020,インフォラボ遊悠, All Rights Reserved

20

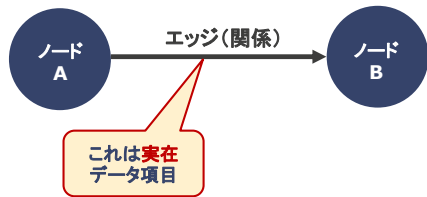
Eyes Wide Open

リレーショナル表現(ER図)を用いた論理データモデル表現の試み

【データ構造スコープの相違】 ここでは有向グラフを対象に

(I) グラフでの基本構造(物理)

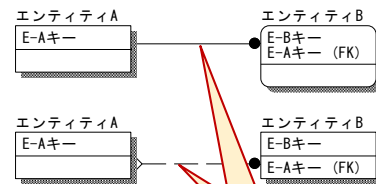
- ・インスタンス(ノード)同士の関係性表現
- ・エッジは実在するデータ要素



- 論理モデルでの表現をするためには、
1. ノード群(インスタンス)の抽象化が必要
(ノード群のエンティティ表現を行うということ)
 2. エッジのエンティティ表現が必要
 3. グラフでのエッジ関係でのグループ化を通じた属性表現が必要

(II) ER図での基本構造(論理/物理)

- ・エンティティ同士の関係
- ・エンティティキーを用いて関係存在を表記



これは仮想の関係線

Eyes Wide Open

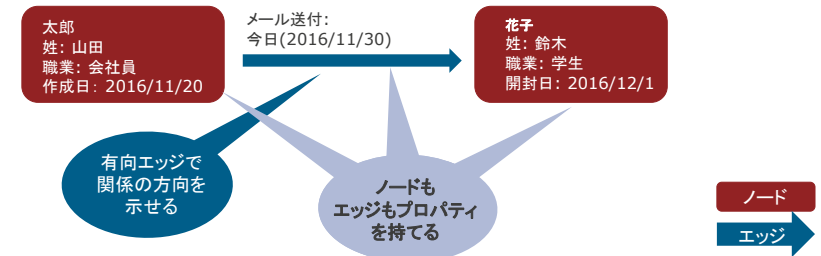
©2020,インフォラボ遊悠, All Rights Reserved

21

ネットワークグラフの「ノード」と「エッジ」

- ネットワークグラフは「ノード」(頂点)と「エッジ」(辺)で関係性の存在を图示表現する
- ノード間の関係を表すエッジには、方向性を持たないもの(無向)と持つもの(有向)がある
- 技術的実装の考え方においては、ネットワークグラフ・データベース仕様は二種類ある
- 「プロパティグラフ・モデル仕様」では、ノードもエッジもプロパティ(属性項目)を持つことができる
- ここでの表現は、インスタンスレベルでのデータ構造の表現図と考えると分かりやすい

例1: プロパティグラフ ... モデル管理には、プロパティの属性整理にはある程度正規化的な考え方を持つことが重要
(論理モデル管理連携の必要性)



Eyes Wide Open

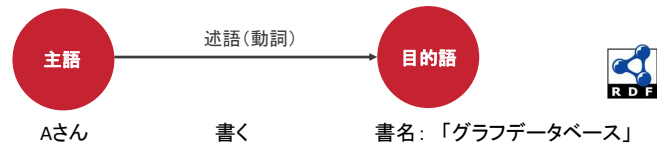
©2020,インフォラボ遊悠, All Rights Reserved

22

ER図を用いた論理データモデル表現2 (RDF系グラフ)

- 第二のグラフ表現でのセマンティックWebの領域の話題としてRDF(Resource Description Framework)の考え方がある。
- 「主語 (Subject)」、「述語 (Predicate)」、「目的語 (Object)」の三つ組(トリプル)でWeb上にあるリソースを記述しようという考え方から始まった。意味を表すメタデータのための利用として当初規定されており、RSS(RDF Site Summary)などでの利用ケースがある。
- 現在は「リンクトデータ」の表現としても提案されており、オープンデータの流れにより、様々な利用が提唱されている。
- この形式のデータ構造枠組みをサポートするグラフデータベースも提供されている

例4: RDFのトリプル(triple)表現



- ✓このタイプのグラフデータを論理モデル表現するには、特に注意が必要
⇒ ノードを単純にインスタンスの表現として捉えられない点があり、モデル表現上の工夫が必要
- ・オントロジー的発想
 - ・ノードとエッジの「エンティティ・グループ」的対応付けと管理

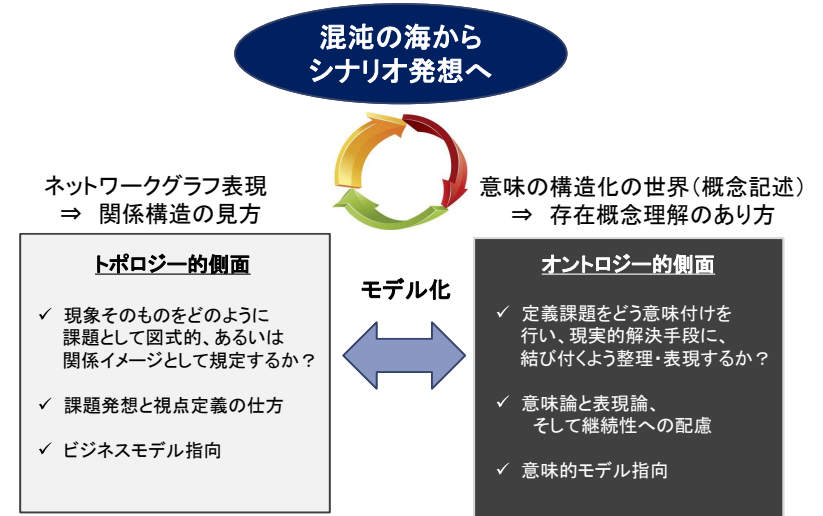
備考: RDFにおけるノードの識別は、URI参照またはリテラルで示される

Eyes Wide Open

©2020,インフォラボ遊悠, All Rights Reserved

23

【参考】ネットワークグラフ活用への重要視点 -トポロジーとオントロジー-



Eyes Wide Open

©2020,インフォラボ遊悠, All Rights Reserved

24

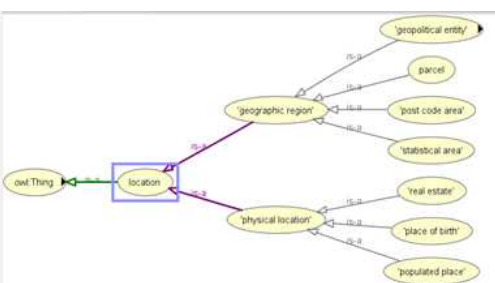
【参考】RDF/OWLで記述されたオントロジーの例(FIBO)

- EDMC(※)が提唱している金融業界の体系モデルとしてFIBO(Financial Industry Business Ontology)がある。
- FIBOは、金融に係る用語・概念を定義した用語集(Vocabulary)、RDFによるオントロジー定義、概念図等のデータの集合体を総称している。「金融業界で共有利用可能な概念・用語」を標準化したものという位置付けである。
- そこで定義されている内容の一部を参考として取り上げる(例 a), b))。

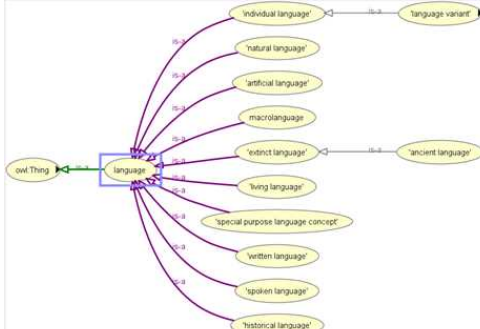
<https://spec.edmcouncil.org/fibo/>



例 a) FIBO: FunctionalEntities > Location



例 b) FIBO: FunctionalEntities > Language



※ EDMC(The Enterprise data management council)は、米国金融機関、証券会社等が中心となって設立した団体。2005年に金融機関におけるデータ管理実務の発展を目的にニューヨークで設立された。実質的な業界標準である「データ管理評価モデル:Data Management Capability Assessment Model(DCAM)」や「金融機関の体系モデル:Financial Industry Business Ontology(FIBO)」を提唱。

©2020,インフォラボ遊悠, All Rights Reserved

25

Eyes
Wide
Open

【今日の話】 データモデリングを柱に様々な視点から

- (話題1) データモデリングの位置付け／目的／意義を再考する
- (話題2) データモデリング・アプローチの視点から考える
- (話題3) マスタデータ = リソースデータの発想を見直す
- (話題4) リレーショナルスキーマ以外のモデルへの適用を考察する
- (話題5) データモデリングの適用範囲の更なる拡大性を考える
- (話題6) 第10分科会で出てきた「疑問」を取り上げる

次回話題

©2020,インフォラボ遊悠, All Rights Reserved

26

Eyes
Wide
Open



©2020,インフォラボ遊悠, All Rights Reserved

27

Eyes
Wide
Open