

ナノ知識探索プロジェクト：実験記録からの知識発見 (第2報) 関連メタデータの追加とその活用

Knowledge Exploratory Project for Nano Device Design and Manufacturing: Knowledge Discovery from Experimental Records (2nd Report)
Utilization of Related Metadata

吉岡 真治*¹
Masaharu YOSHIOKA

Dieb Thaeer*¹
Dieb Thaeer

富岡 克広*²
Katsuhiko TOMIOKA

原 真二郎*¹
Shinjiro HARA

福井 孝志*¹
Takashi FUKUI

*¹北海道大学大学院情報科学研究科

Graduate School of Information Science and Technology, Hokkaido University

*²科学技術振興機構 さきがけ研究者

JST PRESTO

Development of nanodevices requires varieties of knowledge; some of this knowledge is tacit, based on the user's experience. Therefore, it is difficult to transfer knowledge of good engineers to novice ones. In order to accelerate this knowledge transfer process, a knowledge exploratory project for nanodevice design and manufacturing was organized based on collaboration between nanodevice engineers and computer science researchers. In last report, we have already developed experiment record management system with simple knowledge discovery techniques. However, based on the feedback from nanodevice researchers, it is necessary to include related information that are not described in the experimental record for meaningful analysis. In this paper, we propose a framework to describe information related to the experiment as metadata.

1. はじめに

近年、様々な機能を有する様々なナノ結晶デバイスが開発されている [Fukui 91, Noborisaka 05] が、その開発プロセスに関する知識は、十分に体系化されておらず、実際のデバイス開発のプロセスにおいては、主に経験に裏打ちされた多くの試行錯誤を必要としている。

本研究では、データマイニングやオントロジーといった人工知能の技術を応用する [山口 07] という立場に立ち、ナノ結晶デバイスの開発の裏側にある知識の明示化を支援する手法の構築を目的としている [吉岡 10]。前報では、ナノ結晶デバイスの実験記録を情報源として活用するために、実験記録シートのデータベース化とそのデータベースを活用したデータの検索・可視化を行う実験記録マネジメントシステムを提案し、ナノ結晶デバイス開発の技術者から一定の評価を得た。

一方、実際に作成したシステムを利用してもらったところ、実験記録シートに記述している情報以外の関連情報が存在し、その情報もあわせて処理をした方が良いという指摘があった。

本報では、前報で作成した実験記録データベースを拡張し、実験結果を分析するに役立つメタデータを付加する方法を提案する。

2. ナノ結晶デバイスの実験記録マネジメントシステムの構築

本研究では、北海道大学量子集積エレクトロニクスセンターにおける開発プロセスからの知識獲得を目指し、実験記録マネジメントシステムの構築を行っている [吉岡 10]。今回、分析の対象としているナノ結晶デバイスの構築手法は複数の材料を原子レベルで制御する有機金属気相成長法 (MOVPE法) [Fukui 91] と呼ばれる手法である。MOVPE法では、様々

なガスや金属を様々な条件 (温度・圧力・...) で反応させることにより、異なる化合物の層を積層し、様々な機能を持つデバイスを作成する。

この開発過程では、目的のナノ結晶構造を得るために、数多くの実験を行いながら、結晶を作るための物質の供給方法などのパラメータを決めていく必要がある。一つのナノ結晶構造を得るために決めなければならないパラメータ数は、非常に多いが、一回の実験に、多くの時間 (長いものでは、準備を入れると半日以上) とコスト (一回あたり数万円) がかかるため、そのパラメータは、主に、熟練者の経験に頼った試行錯誤によって決定されている。

この試行錯誤の過程は、図1に示すような実験記録シートに以下の情報を図や表などを用いて記載し、その記録の蓄積と紙ベースによる共有が行われてきた。

1. 実験関連情報

実験者、目的、日付、実験器具内での配置情報などで、主に、自由記述項目ならびに絵として記載される。

2. 実験パラメータ

ナノ結晶デバイスの作成実験で作成した積層構造ならびに、各層を作成したガスや金属の情報とそのパラメータ (温度・圧力・...) で、主に、表形式で記載される。

各研究者は、この実験記録シートを用いて、適切な類似事例などを見つけ出すことによって、新しいデバイス開発におけるパラメータ決定の参考にしている。しかし、初学者にとって、適切な類似事例を見つけ出すことは必ずしも容易ではなく、熟練した研究者から、類似事例を参照するようという指示に基づき、実際に実験を行いながら、経験的に知識を学ぶというのが現状であった。

そこで、本研究では、その第一段階として、実験記録シートの電子化を行い、以下のような機能を持つプロトタイプシステムを作成した。

連絡先: 吉岡真治, 北海道大学大学院情報科学研究科, 札幌市北区北14条西9丁目, 011-706-7107, yoshioka@ist.hokudai.ac.jp

2. 具体的なシステムの結果を見ることによって、さらなる要求を具体化することが可能となり、その要求を新しいシステムの試用として実現する。

3. 実験記録へのメタデータの追加

3.1 前報のシステムに関する評価と分析

前報で作成したシステムについての妥当性を検証するために、さらに、215 件のデータを追加し、実験結果の評価情報が入力可能であることを確認した。また、パラメトリック情報と頻度情報の可視化機能についても、熟練の研究者には、同様のイメージが頭の中にあるという評価を得た。

さらに、実際に実験結果に関する可視化を行った結果、同じ化合物の層を作成する際に利用するパラメータの組み合わせが複数存在することも、同様に確認できた。これらの違いを理解するためには、作成するデバイスが持つ機能の違いなどを考慮に入れる必要があると考えた。しかし、どのような情報が違いを理解するために必要なかが明確でない段階で、実験記録と関連づけるのは困難であると考えた。

このような違いを理解するために有用な情報について、デバイス開発の技術者に相談をしたところ、一連の実験の結果は論文という形でまとめられており、実験に関する背景情報などは、全て、その中に記載されているというコメントを得た。このコメントに対応する形で、実験記録と結果をまとめた論文を対応づけると共に、論文中の有用と考えられる情報のタグ付けを行うことにより、今後の分析に役立つと考えた。

3.2 ナノ結晶デバイス開発論文への情報のタグ付け

先に述べた考察に基づき、ナノ結晶デバイスの開発に関する論文に対して付与するタグの設計と具体的なタグ付け作業を行っている [Dieb 11]。

実際にドメインの研究者と相談することで、論文中の以下の情報についてタグを付与することとした。

材料 (SMat) : 実験に用いる素材や化合物 (As や InGaAs など)

材料の特性 (SMch) : 素材や化合物が持つ特性 (結晶の異方性に関する情報など)

実験パラメータ (ExP) : 実験でデバイス作成の制御のために用いるパラメータ (圧力や流量など)

実験パラメータの値 (ExPv) : 上記の実験パラメータに対応する値

評価パラメータ (EvP) : デバイスの性質を評価するためのパラメータ (ピークエネルギーや工学特性における半値幅など)

評価パラメータの値 (EvPv) : 上記の評価パラメータに対応する値

デバイスの作成手法 (Meth) 作成手法の名前 (SA-MOVPE など)

最終製品 (Prod) : 最終製品の名前 (半導体ナノワイヤなど)

しかし、上記のようなタグに関する大まかな規定では、タグ付け作業の一貫性が確保されないという問題がある。そのため、実際に、ドメインの知識を持つ北海道大学量子集積エレクトロニクスセンターに所属する修士課程の学生 2 名に、同一の

論文 1 件に対しタグ付け作業を行ってもらい、タグ付けに悩むような事例の収集を行った。また、それらの事例について、一貫した分類基準を与えることで、タグ付け基準とした。

図 3 にこのタグ付け基準でタグ付けをしたナノ結晶デバイス開発の論文 [Yoshimura 10] のアブストラクトを示す。

We fabricated <SMat>InGaAs</SMat> <Prod>nanowires (NWs)</Prod> in <SMat>SiO₂</SMat> mask openings on a <SMat>GaAs</SMat> <SMch>(111)B</SMch> substrate at <ExP>growth temperatures<ExP> of <ExPv>600 -700 </ExPv> using <Method>catalyst-free selective-area metal organic vapor phase epitaxy</Method>. At a <ExP>growth temperature</ExP> of <ExPv>600 </ExPv>, particle-like depositions occurred, but they decreased in number and density when the growth temperature was increased to <ExPv>650 </ExPv> and disappeared above <ExPv>675 </ExPv>. <EvP>The heights and growth rates</EvP> of the NWs increased when the <ExP>growth temperature</ExP> was increased and <ExP>the mask opening diameter</ExP> was decreased from <ExPv>300 to 50 nm</ExPv>. <EvP>Photoluminescence (PL) spectra</EvP> measured for the NWs indicated a blue shift in the peak from 0.95 to 1.3 eV as the growth temperature was increased from <ExPv>600 to 700 </ExPv>, indicating an increase in the <EvP><SMat>Ga</SMat> composition</EvP> from <EvPv>62 to 88% </EvPv> in the <SMat>InGaAs</SMat> NWs.

図 3: 論文アブストラクト [Yoshimura 10] へのタグ付け事例

このタグ付けの分類基準の一貫性を検証するために、タグ付け基準を作成した論文とは異なる同一論文に対して、先ほどと同じ二人の学生が同様のタグ付け作業を行い、その一致度をコーパスの分析で良く用いられる Kappa 係数 [Artstein 08] を用いて分析することとした。

このコーパスの一致度をはかる際に、一人目は、nanowires (NWs) に対して、最終製品 (Prod) のタグをつけ、もう一人が、nanowires のみにタグをつけるようなタグをつける語の単位の問題に起因する不一致が多いことが判明した。これは、単語の単位に関する基準が曖昧であったことに起因している。

まず、語の単位の基準が完全に一致しているものについて、二人の作業者のつけたタグの比較を行った結果を表 1 に示す。この表から、不一致の原因が主に、タグのタイプをつけることが難しい点にあるのではなく、タグをつける語の単位の認定も含めて、どのような範囲でつけるかどうか難しいということが読み取れる。また、唯一混同が見られた最終製品 (Prod) と材料の性質 (SMch) についてであるが、使用した材料の性質を持った最終製品が出来上がるような場合に、材料の性質として考えるのが最終製品の記述として考えるのかを決める必要があるということが分かった。

次に、語の単位の判定については、重なりがあれば対応関係があるとした語の部分一致を認めた場合の一致度を計算した結果を表 2 に示す。

タグ付けした語の単位の問題から、パラメータとその値などで一部、異なるタグ間の対応関係が見られるが基本的には、良く対応がとれており、[Artstein 08] によると比較的良好い一

	Smat	SMch	ExP	ExPv	EvP	EvPv	Meth	Prod	Other	Tot
Smat	91								11	102
SMch		30						4	11	45
ExP			32						28	60
ExPv			1	17					17	35
EvP					23	2			14	39
EvPv						4			34	38
Meth							9		4	13
Tart								44	3	47
Other	12	5	24	15	10	23	10	23		122
Tot	103	35	57	32	33	29	19	71	122	501

Kappa 係数 : 0.41

表 1: 完全一致による一一致度

	Smat	SMch	ExP	ExPv	EvP	EvPv	Meth	Tart	Other	Tot
Smat	105								1	106
SMch		36						10	4	50
ExP	1		53	2			8		1	65
ExPv			2	33		1			2	38
EvP					32	7			3	42
EvPv		1			1	24		2	12	40
Meth							14			14
Tart								47	1	48
Other	6	2	7	2	2		1	18		38
Tot	112	39	62	37	35	32	23	77	24	441

Kappa 係数 : 0.74

表 2: 部分一致による一一致度

致度であると判断できる。

これらの事から、タグ付け作業の一貫性をあげるためには、タグ付けをする際の語の単位に関する具体的な事例を収集し、その事例に基づいた新たな基準の作成が必要であるということが分かった。

今後は、上記の問題を踏まえた新しいタグ付け基準を作成し、実験に関連する論文へのタグ付け作業を行うと共に、これらのタグ付け作業で得られた情報が実験結果の違いを理解するのに有用であるかどうかについての検討を進めたいと考えている。

4. おわりに

本報では、前報で作成したナノ結晶デバイスの実験記録マネージメントシステムの評価結果を踏まえて、実験記録に関するメタデータを追加する方法を提案した。具体的には、関連論文中から有用と思われる情報にタグをつけることによって、実験記録に関する追加情報として利用する。また、実際のタグ付け作業を通して、今後のタグ付け作業の基準の改定を必要とするものの、高いレベルの一貫度を持つ安定的なタグ付けが可能であることが確認された。

今後の展望としては、本手法の有用性を確認するために、北海道大学量子集積エレクトロニクスセンターに所属する初学者(4年生)を対象にして、本システムの利用実験を行い、その結果について分析を行う予定である。

謝辞

本研究の一部は、北海道大学大学院情報科学研究科グローバル COE プログラム「知の創出を支える次世代 IT 基盤拠点(平成 19 年度~23 年度)」によるものである。

参考文献

[Artstein 08] Artstein, R. and Poesio, M.: Inter-coder agreement for computational linguistics, *Computational*

Linguistics, Vol. 34, pp. 555–596 (2008)

[Dieb 11] Dieb, T. M. and Yoshioka, M.: Construction of Tagged Corpus for Nanodevices Development Papers, 言語処理学会第 17 回年次大会発表論文集, pp. 272–275 (2011)

[Fukui 91] Fukui, T., Ando, S., Tokura, Y., and Toriyama, T.: GaAs Tetrahedral Quantum Dot Structures Fabricated using Selective Area Metalorganic Chemical Vapor-deposition, *APPLIED PHYSICS LETTERS*, Vol. 58, pp. 2018–2020 (1991)

[Noborisaka 05] Noborisaka, J., Motohisa, J., Hara, S., and Fukui, T.: Fabrication and characterization of freestanding GaAs/AlGaAs core-shell nanowires and AlGaAs nanotubes by using selective-area metalorganic vapor phase epitaxy, *APPLIED PHYSICS LETTERS*, Vol. 87, (2005)

[山口 07] 山口 高平: 知識マネジメントと AI 技術, 人工知能学会誌, Vol. 22, No. 4, pp. 461–466 (2007)

[Yoshimura 10] Yoshimura, M., Tomioka, K., Hiruma, K., Hara, S., Motohisa, J., and Fukui, T.: Growth and Characterization of InGaAs Nanowires Formed on GaAs(111)B by Selective-Area Metal Organic Vapor Phase Epitaxy, *Japanese Journal of Applied Physics*, Vol. 49, No. 4, pp. 04DH08–1–5 (2010)

[吉岡 10] 吉岡 真治, 富岡 克広, 原 真二郎, 福井 孝志: ナノ知識探索プロジェクト: 実験記録からの知識発見, 2010 年度人工知能学会全国大会(第 24 回) 論文集 (2010), CD-ROM 1B3-3