

# 技術研究会報告集データベース構築への取り組み

澤 昌孝、水谷文保

分子科学研究所 技術課

## 概要

「技術研究会報告集データベース化」について提案および検討を行い、一昨年[1]および昨年[2]の技術研究会において報告を行ってきた。過去に分子研で開催した技術研究会の報告書は、印刷物でのみ存在しているため、インターネットを通じて簡単に参照できる様に電子化を進めてきた。ようやくこの作業が完了したため、ここ数年に開催された技術研究会において、電子ファイルで提出された報告書と合わせてデータベース化を開始した。その進捗状況について報告する。

## 1 はじめに

技術研究会は、過去 28 年間に 31 回開催されており、総数 2031 演題が報告されてきた[3]。この膨大な技術情報は、残念ながら一覧できないのが現状である。

平成 10 年度の技術研究会で報告された「技術情報ネットワーク」構想[4]をもとに、過去 2 回（平成 10、11 年度）の技術研究会で「インターネット討論会」が行われ、技術者間の交流と技術情報の蓄積および交換をインターネット上で展開するアイデアについて議論がなされた。そこでは組織または個人で技術情報を蓄積することが、「技術情報ネットワーク」実現への第一歩であることが確認された。しかし日常業務の合間に個人のレベルで技術情報をコツコツ作り上げていく、という作業は、個人のメリットが明確でない以上ボランティア要素が強く、また日々の業務に忙殺されている状況下では非常に困難である。さらにたとえ技術情報の蓄積が行われ始めたとしても、あるまとまった量として蓄積されるには、今後数十年という年月が必要であると思われる。国立大学および共同利用機関の法人化まで残り数ヶ月と迫った段階にあって、技官体制の将来を危ぶむ声が多い中、この遠大な計画が実現されるのは極めて困難であると思われる。

以上を踏まえ、すでに膨大な技術情報として存在している技術研究会報告集の内容をデータベース化することについて検討してきた。分子研で 16 回の技術研究会を開催していること、また過去 7 年間ににおいて、電子化された技術研究会報告は 4 回あり[5][6][7][8]、電子化されていない 3 回中 2 回は分子研開催分であることより、まず分子研開催分の技術研究会報告を電子化できれば、不完全ながら実用性を認識するのに十分な技術研究会報告データベース（以下、暫定版データベース、とする）が構築できることが想像できる。

データベース化に向けての動きは、平成 12 年度より開始している。所内の合意の後、平成 13 年 4 月に高エネ研、核融合研の両技術部長に報告書のデータベースに向けて著作権の観点からの打診を行っており、好感触を得たことから本格的に検討を開始した。検討内容については、平成 13 年度、14 年度の技術研究会において、技術研究会報告のデータベース化の可能性と試みについて報告してきた[1][2]。また技術の伝承について検討している低温技術者の研究会でも報告を行った[9]。平成 15 年度、分子研開催分の全報告集の電子化作業が完了した。これを元に早急に検索機能を整備して暫定版データベースの試験的運用を行いたいと考えているが、一般的に公開するには克服しなければならない問題がたくさん残っている。今回の報告では、改めて問題点を整理するとともに、暫定版データベースの開発の進捗状況について報告を行う。

## 2 技術情報データベース構築の問題点

技術研究会報告集のデータベース構築とその公開に向けて以下の様々な問題を克服しなければならない。

### 2.1 著作権の問題

技術研究会報告という技術情報において、各報告書の著作権は誰が所有しているか、という点はインターネットを通じて正式公開する上で明確にしておかなければならない。これまでのいろいろな場面で議論をさせて頂いたなかでは、個々の技術情報の著作権は発表者に帰属している、という点を前提に考えることで問題なさそうである。さらに技術研究会報告集を作成する段階で、発表者は報告集発刊を前提にして著作権の一部を開催機関に承諾していると考えられる点で、合意が得られそうな感触を得ている。ただしどの権利についてどの範囲までが承諾済みであるか、が明確ではない。出版する以上、出版した機関には公開や再配布の権利が承諾されていることには疑問をはさむ余地は無いが、出版物を閲覧する側の利用が、著作権を侵害しない範囲で紳士的に行われる、という点も今後は明確にしなければならない。

2003年5月に高エネ研の呼びかけによって、過去の技術研究会開催機関の代表者を集めて、データベース化することに対する意見交換を行った。議論の中で、インターネット上で公開し検索環境を整備するところまで著作権承諾範囲を拡大解釈してよいか、という点については意見が分かれた。結局個人に対して再度了解を取るべき、という見解で一致をみた。この作業を過去28年にさかのぼって実現することは困難であるが、打開策を検討して克服しなければならないと考えている。

### 2.2 技術面の問題

出版物しか残っていない過去の技術情報をデータベース化するため、イメージスキャニングによって画像情報としてPDF化する方向で作業を行っている。その後、全文検索のためにOCRすることを検討しているが、そのためにはスキャニングの解像度が300dpi程度必要である。スキャニング後のOCRは、PDFから直接操作できるイー・アイ・ソフト株式会社の「読ん de!! ココ Ver.8」を使用している。平成13年度開催分は、既に電子化されているが、スキャニング解像度がOCRに不十分なものがあり、再度電子化作業が必要と思われる。また分子研開催分については、一部手書き原稿のオフセット印刷を中心としたものがあり、この情報のOCRは不可能なため、何らかの方法でテキスト情報を作成する必要がある。以上の様に、過去の情報収集には、多くの問題があり1つ1つ克服して行かなければならない状況である。

新規の技術情報に関しては、データベース化に必要な情報収集に対して追加作業を不要とするために、画一的な規格を作って徹底させる必要がある。電子出版について報告[10]を行っている高エネ研が今回の主催機関であることより、電子情報の回収ガイドライン整備について提案されると思われる。

### 2.3 労力面の問題

電子化およびOCR作業など、膨大な情報を処理するには、相当の労力が必要である。これをボランティアで処理するには限界があるため、組織だった体制や予算措置が必要である。

今後回収される技術情報は、電子情報を回収することで労力が少なく済むが、データベース収録という労力は継続的に発生する。これについては、どのような体制で行っていくかを検討していかなければならない。この点は、暫定版データベースの完成し、その効果を検証した後に検討を進めていきたい。

### 2.4 継続性の問題

今後の技術研究会の技術情報も収録していくためには、継続性を確保しなければならない。これに関して、技術面、労力面の問題を認識した上で、今後の開催機関に徹底する体制が必要となる。特に技術研究会は分散した機関で開催されること、報告集の詳細までとりまとめを行う機関が存在していないことなど、自動的

には継続性が維持できない状況であることを認識する必要がある。この点については、2003年5月の意見交換の場で、高エネ研からの提案が出された。この提案は、報告集編集作業を次回開催機関の担当予定者も手伝えることで、方法論の伝承を行う、というものである。今回の技術研究会では、この方法を試験的に行ってみる、という合意の元で作業が行われている。

## 2.5 収録範囲の問題

技術研究会報告集をデータベース化することがスタートラインであるが、分析技術研究会や生理学技術研究会など継続して開催されている研究会がいくつか存在しており、技術情報という点では同等な性格を持っている。しかし開催母体が異なるため、前出の4点の問題が同様に発生することは想像に難くない。将来的には収録を前提として検討を進めたい。

## 3 進捗状況と今後の計画

過去28年に開催された31回分の技術研究会報告の電子化状況を表1に示す。今回分子研開催分の電子化が完了したことから、20回分の技術研究会における60%弱の技術情報が閲覧可能になった。平成10年度、12年度、14年度の分については、電子ファイルの回収が行われているので、テキスト情報が完全に収録されている。その他の開催分については、OCRによるテキスト変換を行うため、不完全なテキスト情報となる。

今後は、電子化されていない分についての演題と発表者の収録を進め、全開催分の演題および発表者の検索を可能にしたい。その後、組織・発表者毎に整理した情報を元にして、著作権における電子的な再配布などの確認を行い、一般公開を行う準備を行いたい。さらに予算確保を行い、全開催分の電子化を実現させ完全版データベースの構築を目指したい。

## 4 さいごに

技官の交流の場を提供してきた技術研究会は、技官の約30年の活動足跡を残してきたと言える。実際それ以前の足跡をたどることは、非常に困難である。法人化後の技官の位置づけを考えていく中でも、この足跡が十分有益な情報を提供してくれるものと確信している。技術研究会への参加者および機関は、年々増大しているが、過去の技術報告が再利用できない状況のままでは、イベントにすぎない。再利用できる環境整備ができてこそ、技術者のよりどころとしての技術研究

表 1. 技術研究会開催一覧

年度	開催機関	発表数	電子済
50	分子研	5	5
51	分子研	6	6
	分子研	6	6
52	分子研	7	7
	分子研	9	9
53	分子研	8	8
54	分子研	11	11
	分子研	11	11
55	分子研	23	23
56	分子研	28	28
57	高エネ研	45	
58	分子研	35	35
59	プラズマ研	65	
60	高エネ研	69	
61	分子研	49	49
62	プラズマ研	57	
63	高エネ研	91	
1	分子研	61	61
2	核融合研	79	
3	高エネ研	86	
4	分子研	86	86
5	核融合研	78	
6	高エネ研	69	
7	天文台、名大、北大	122	
8	分子研	74	74
9	核融合研	78	
10	高エネ研	115	115
11	分子研	112	112
12	東北大	164	164
13	核融合研	136	136
14	東大本郷	246	246
	合計	2031	1192

会の性格が明確になると思われる。技術情報作成の意義がデータベース構築によって明確化されることを期待している。それによって技術研究会以外においても、定常的な技術情報作成への呼び水となると思われる。

分子研開催分の報告集バックナンバーは、すでに存在せず、過去の技術情報が風化しかかっていた。今回電子化したことによって、風化を食い止めることができた。この意味からも、データベース構築を促進する価値は十分あると言える。技術研究会は、数年前から大学でも行われる様になった。広範囲で分散開催することは、技術情報としてまとまりがつきにくくなること意味している。早急に技術情報を集中管理する体制を整備する必要がある。

データベースの常時閲覧が可能になれば、技術者間の交流も常時行われる様になることは想像に難くない。日常的に情報交換が可能となれば、意識も変化して技術習得の効率向上につながるであろう。インターネットの普及によって、従来このような効果が期待されてきたが、インターネットの存在だけでは、なかなか効果が発揮されなかった。インターネット討論会においても、メーリングリストの整備が提案されたが、メーリングリストの存在だけでは解決にならないことは示唆されていた。誰でも質問に即座に答えてくれる環境を切望するが、答える側のボランティアになってしまって長続きせず、返答が無いメーリングリストは魅力が無いのですたれてしまう、という構図が見えてしまうのである。インターネットの有効性を発揮するには、有益なコンテンツの存在が不可欠である。よりどころとなる場が存在して、道となるインターネットが生かされる。この点からも、技術情報に人情報が付加された技術研究会報告集をデータベース化することは、大変意味があると言える。完全版データベース構築と平行して、技術情報ネットワークの実現に向けて、コミュニケーション環境整備もぜひ行っていきたい。

## 参考文献

- [1] 水谷文保, “技術研究会報告データベース化の検討”, 平成 13 年度核融合科学研究所技術研究会報告, March 14-15, 2002, p.5\_19, URL [http://etwww.nifs.ac.jp/kenkyu01/publish/no5/5\\_19.pdf](http://etwww.nifs.ac.jp/kenkyu01/publish/no5/5_19.pdf)
- [2] 水谷文保, “技術研究会報告データベース化の試み”, 平成 14 年度東京大学総合技術研究会報告, March 6-7, 2003, p.44-47, URL <http://www.ut-tech.iis.u-tokyo.ac.jp/uttech /5/05-16.pdf>
- [3] 核融合科学研究所技術部, “技術研究会の記録”, URL <http://etmikan.lhd.nifs.ac.jp/giken/kiroku.html>
- [4] 伊藤篤, “大学および研究所の技術部におけるネットワーク環境とその利用状況について”, 平成 10 年度高エネルギー加速器研究機構技術研究会 March 4-5, 1999, URL <http://www-eng.kek.jp/giken/procedng/paper/met 102.pdf>
- [5] 平成 10 年度高エネルギー加速器研究機構技術研究会報告 March 4-5, 1999, URL <http://www-eng.kek.jp/giken/procedng/index.pdf>
- [6] 平成 12 年度東北大学技術研究会報告 March 1-2, 2001, URL <http://platypus.lns.tohoku.ac.jp/Program.htm>
- [7] 平成 13 年度核融合科学研究所技術研究会報告 March 14-15, 2002, URL <http://etwww.nifs.ac.jp/kenkyu01/publish.html>
- [8] 平成 14 年度東京大学総合技術研究会報告 March 6-7, 2003, URL <http://www.ut-tech.iis.u-tokyo.ac.jp/g.index.html>
- [9] 水谷文保, “技術研究会報告書のデータベース化がもたらすもの”, 東京大学物性研究所低温液化室研究会報告書, October 8-10, p.24-25
- [10] 片桐広明他, “国際会議等における電子出版”, 平成 13 年度核融合科学研究所技術研究会報告, March 14-15, 2002, URL [http://etwww.nifs.ac.jp/kenkyu01/publish/poster/pos\\_42.pdf](http://etwww.nifs.ac.jp/kenkyu01/publish/poster/pos_42.pdf)