

4.1 日本 IBM 支援による開発促進事業

4.1.1 ユーザーニーズ調査

I. 主旨・目的

「みちのく震録伝」では、収集した情報を公開するだけでなく、情報を研究者が活用することで防災・減災に関する2次的情報を社会展開することを目的としている。そのため、情報を利用する一般の利用者、行政・自治体の防災担当者、企業、研究者から幅広く利用者のニーズを調査する必要がある。

II. アンケート／ヒアリング実施要領について

「みちのく震録伝」システムの構築に向けて、研究者を対象に、ニーズ調査を実施した。現状のデータの登録や検索の方法、関連する課題を把握するための、アンケート調査と、1時間程度のヒアリングを行なっている。

1) アンケート調査及びヒアリング実績

下記対象者へのアンケート調査及びヒアリングを実施した。

表 4.1-1 アンケート調査、及びヒアリング実施記録(下)

【アンケート調査、およびヒアリング対象者】

No.	ヒアリング日			専門分野
1	アンケート	今村先生	教授	津波工学研究
2	アンケート	真野先生	教授	津波工学研究
3	アンケート	杉浦先生	准教授	防災ポテンシャル研究
4	アンケート	佐藤(源)先生	教授	脳科学
5	アンケート	箕浦先生	教授	電波応用工学
6	8/10(水)	吉田先生	教授	福祉経済設計学
7	8/12(金)	磯田先生	准教授	人文地理学, 応用経済学
8	8/12(金)	邑本先生	教授	認知心理学
9	8/12(金)	柴山先生	助教	建築防災, 地域防災, 地震工学
10	8/12(金)	佐藤(翔)先生	助教	災害危機情報解析
11	8/12(金)	今井先生	助教	津波工学
12	8/17(水)	石渡先生	教授	地質学, 岩石学
13	8/17(水)	原先生	准教授	科学哲学
14	8/17(水)	増田先生	教授	地域経済開発
15	8/19(金)	石井先生	助教	量子エネルギー工学
16	8/19(金)	島田先生	教授	都市法, 都市政策, 景観計画, 災害対策法
17	8/19(金)	菅原先生	研究支援者	堆積学
18	8/22(月)	上原先生	教授	国際保健学
19	8/22(月)	佐藤(大)先生	助教	日本近世史, 歴史資料保全学
20	8/23(月)	田所先生	教授	ロボット工学, レスキュー
21	8/23(月)	阿部先生	教授	心理学
22	8/24(水)	奥村先生	教授	土木計画学, 交通計画学
23	8/24(水)	藤本先生	教授	海底地殻変動観測
24	8/26(金)	大野先生	准教授	地震工学
25	8/26(金)	越村先生	准教授	自然災害科学, 津波工学
26	8/31(水)	本江先生	准教授	都市・建築デザイン

2) アンケート調査内容

アンケート調査では、大きく以下の3つのパートについての現状を質問し、回答を得た。

- ・ パート0. 基本情報としてご担当の研究について
- ・ パートI. 研究データの保存および取り出しに関して
- ・ パートII. 現在のシステム環境について

パートIに当たる部分では、研究業務の中で、データや資料などの情報を、電子的に保存す

る場面（質問 2～8），および保存した情報を後に取り出す場面（次ページ質問 9～15）について、確認している。

【データを保存する場面】

2. 研究に必要なデータ（研究業務で扱っているデータの種類と内容を確認するため）
3. 保存手順（研究業務で扱っているデータの保存の手順を確認するため）
4. データの保存ルール（研究業務上の確立したルールが有るかを確認するため）
5. データの位置情報の管理（データの位置情報の利用状況確認するため）
6. 外部には開示できないデータ（開示出来ないデータの例などを確認するため）
7. アクセス制限のレベル分け（アクセス制限に関する要件を確認するため）
8. データ保存で苦労している点（データ保管に関する要件を確認するため）

【データを取り出す場面】

9. 保存したデータの探しかた（データの検索に関する要件を確認するため）
10. 研究室等でのデータの保存と共有（データの共有に関する要件を確認するため）
11. データの分析（データの分析手順に関する傾向を確認するため）
12. データの取り出しで苦労している点（データの検索などに関する要件を確認するため）
13. 保存したデータの利用状況（データの検索などに関する要件を確認するため）
14. 今後のアーカイブ（データ）の利用方法（アーカイブに関する要件を確認するため）
15. 今必要なデータ（アーカイブに関する要件を確認するため）

1. 結果の要約・考察

1) 研究データについて

研究テーマが多岐に渡るのと同様に、研究に利用するデータについても多岐に渡っている。研究時には、複数のデータを組み合わせて解析・分析することが多く、インターネットなどを利用して、基礎データを取得するためのワークロードがかかっている現状が明らかとなった。

これらの基礎データは整備（自身、他含む）するとともに、それらの属性（権利など）についても管理する必要がある。また、復興の過程を記録する必要があり、データ容量も増加するため、伸び率を考えると共に、長期保管の仕組みを考える必要がある。

2) データの保存・取出ルールについて

データの保存手順や、ルールについては、特に統一した基準を設けて運用しているケースは少なく、ファイル名、フォルダ名に分かりやすい名前をつける、日付別に整理するなど、簡易的な方法により運用しているケースがほとんどであった。そのため、取り出し方も、保存と同様に、ファイル名、フォルダ名、日付などを基軸にしている。

データ登録に関する標準を設けることと、検索時においても、ユーザー使用性に優れた仕組みが望まれる。

3) データの登録について

東日本大震災に関わる新聞記事や、各種調査伝票など、アナログデータを大量に保管してい

るケースもあり、ワークロードの観点から、電子化、分類化できていない現状もあきらかとなった。

また、全ての登録データには、タグ付けをしている訳ではないので、データの登録時のタグ付けを簡素化できる仕組みが必要。

4) 公開できる情報とできない情報

基本的には、論文など、整理／加工後の情報については、広く公開して良いが、その前の段階では、公開できないデータが多い。理由は、データの権利の関係や、データが一人歩きする危険性などによるものである。

また、アンケート調査用紙や、外部からの入手情報など、一般向けには公開はできないケースが多いが、オリジナルデータの重要性は認知されており、研究者向けだけに公開するような仕組みも必要と考えられる。

5) データの帰属（著作権）について

データを活用する際、著作権の確認、データの信憑性の確認、オリジナルはどこか、使用しても良いか確認することに時間を要する現状が明らかとなった。

アーカイブでは、これらの情報を管理する仕組みを構築し、利用者（主に研究者）に伝えることのできるインターフェースが必要である。また、使用許諾なども、簡単にできる仕組みが望まれる。

6) ツールの活用

GIS 表現の重要性が認識されている。実際に研究データを GIS と連携させ、研究を行っている回答もあった。アーカイブデータ活用のアプリケーションの一つとして、GIS 機能をアーカイブ側で提供することが良いと考えられるが、そのためには、研究データとの重ね合わせを容易に実現できる GIS インターフェースの提供が必要となる。

また、SPSS を利用している回答も多く、広く利用されていた。

その他、分析系の機能のニーズも多く、研究者が情報を活用するためのアプリケーションとして、このような機能をアーカイブ側に持たせる検討も必要である。

II. 「みちのく震録伝」に対する期待

期待する内容については、以下のような回答であった。

- ・ 地震がなぜ起きたかの解明を行い、災害の軽減のためにどうすれば良いかを一般の人に知らせることが必要。
- ・ アーカイブにあるデータを検索した時、オリジナルのデータの所在や所有者が明確であり、利用手続きが簡単に行えることが必要であり、データの所有者には閲覧者の情報などのフィードバックが必要。
- ・ データを長期に蓄積できることが必要。
- ・ 蓄積されたデータは他の人から容易に検索されて、利用されるような仕組みがあることが必要。
- ・ アーカイブは蓄積されて、利用されることが重要であり、そのためには想定される利用者にどのように利用されるのかを予め考慮して、良い利用方法が作り込まれる事が必要。アーカイブに蓄積される情報は、想定利用者にとって利用する価値の有るものでなくてはならず、一般の人向けには防災のためのガイドなどが必要。
- ・ 情報を蓄積する側も、その情報を使う側も双方にメリットが必要。
- ・ アーカイブでは検索された情報が直接公開されていなくても、利用するための糸口となる情報が得られることが必要。
- ・ アーカイブは情報を長期に亘って蓄積・保存し、将来の災害の防止の対策に役立つ必要がある。
- ・ 建物被害のオリジナルデータや、公の調査のオリジナルデータなどが利用できる仕組みが必要。また、検索されたデータの利用許可の手続きがアーカイブの機能の一部に含まれているなど、データが利用しやすいシステムである必要がある。データの検索のためにはタグ付けが便利であるが、分類学的な階層構造も必要。芸術・歴史分野ではファイル形式が様々で研究への利用が進まなかった。データの分析やガイドの利用などの活用の場面を想定したファイル形式の管理が必要。
- ・ アーカイブは研究の材料となるデータを提供するだけでなく、研究を促進するために研究者の協業のきっかけを容易に作る事ができる仕組みの提供などが必要。
- ・ 研究者に対してはアクセス権限管理と使いやすい検索機能、一般の人にとっては理解しやすい防災マニュアルなど想定利用者向けの機能やコンテンツが必要。
- ・ 想定利用者の要求や行動に合わせた、必要な情報を必要なときに得られるシステムを設計することが必要。
- ・ 研究の手順や進展に合わせてデータを取り込むことができるような、柔軟なデータの構造が必要。研究や分析の手順を観察し分析作業の効率化・省力化を図れる機能を開発する必要がある。
- ・ アーカイブに蓄積された情報は、一般の人にも興味が沸き、情報探求の入口となる面白い表現方法で提供することも重要。(情報の視覚化)

III. 設問ごとの回答サマリー

1) 設問 2 : 研究に必要なデータ

アンケート調査の結果から、現在の研究業務で必要としているデータの種類を集計すると、計測データは6割近く、写真とテキストは8割近くの回答者が必要としていた。

計測データは主に理学・工学系の研究者が必要と回答しており、写真などのデータは幅広い分野の研究者が必要と回答していた。

また、外部の Web サイトで公開されているデータの収集に、多くの時間が割かれているケースも見られた。

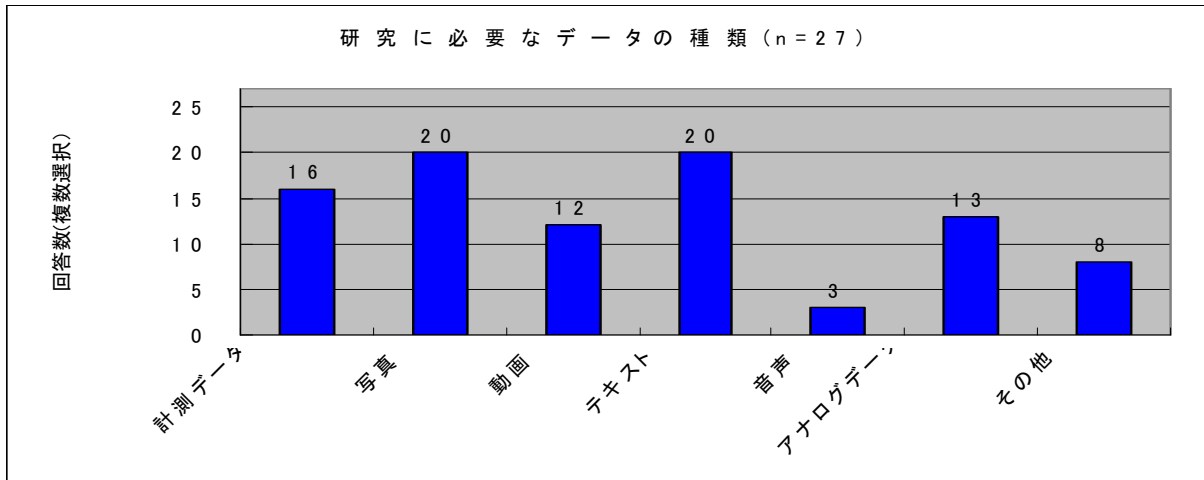


図 4.1-1 研究に必要なデータの種類の種類

表 4.1-2 研究データの内訳(下)

(参考)研究データの内訳例

地震観測記録	海底地殻変動観測データ	地層・堆積学	地盤	津波	災害空間	社会・経済	パニックやデマ被害に付随する諸般データ	医学・生理
地震記録	海底GPS観測	津波痕跡データ	地中レーダ	津波遡上高に関する測定データ	数十m四方程度の3次元点データ	死亡者リスト	経済指数	脳活動データ
気象庁	海底圧力観測	浸水データ	地表設置合成開口レーダ	家屋被害に関する調査データ	地図データ	人口分布データ	犯罪発生件数等	
K-net		GPSデータ	航空機・衛星リモートセンシングデータ	衛星画像等		男女平等度合		
Kiki-net		浸水深		津波波形データ				
震度情報NW		年代測定データ		津波痕跡データ				
浸水深		流向データ		海岸林諸元データ				
浸水高		地球化学的データ						
		古生物学的データ						
		地下構造データ						
		GPS測量データ						
		地震前後の5mDEM						
		DEMの解析で作成した浸食・堆積マップ						

2) 設問 3 : 保存手順

写真については、日付別や、場所別、イベント別にフォルダを作成し、その中に写真データを保存する方式がほとんどであった。GPS 付のカメラを使用するようになってからは、フォルダで分けられない保存方法を利用しているケースもある。個々の写真については、タグ付けをしているケースは、ほとんどなかった。

テキストや、アナログデータは、電子化やタグ付けの手間が多く、保存に時間がかかる、保存時に十分な整理が出来ない、或いは電子化されていない大量のデータがあるなど、未だ十分に活用出来ていないとの回答が多かった。

写真を日付や場所で整理する、それらをキーワードとして検索する運用では、今後データ増加に対応出来なくなるとの回答もあった。

3) 設問 4：データの保存ルール

27 人の回答者中、データの保存ルールを決めていないとの回答が半数以上であった。この中には写真データも含まれている。

またルールを決めているとの回答でも、日付や場所をフォルダ名にして保存しているケースが多く、写真を保存しているが、有効に利用できていないとの回答もあった。

4) 設問 5：データの位置情報の管理

回答者の 25%程度がGPS 付のカメラを利用しているが、位置情報を積極的に利用しているとの回答は少なかった。

5) 設問 6：外部に開示できないデータ

個人情報を適切に扱うために容易には公開できないと考えている、または、論文発表前であるためデータを公開できないという回答が多く、データのアクセス制限については関心が高いことが明らかになった。

適切な情報にのみアクセスできるよう、公開情報に際してデータの選別が必要であるとの意見が多かった。

6) 設問 7：アクセス制限のレベル分け

一般向けには報告書は公開したいが基になるデータなどは出せない、研究者同士は引用ルールが正しく守られれば共有しても良いと感じる、基のデータはオーナーのみがアクセスできる事が望ましい等の意見があった。

細かい制限が必要になる半面、公開する側の設定の負担も軽くなるよう、にシンプルな手順が望まれる。

7) 設問 8：データ保存で苦労している点

保存に関しては、特に困っていないとの回答が多い半面、広範囲なデータを収集し、保存するために困っているとの回答もあった。特に、アナログデータの保存や、加工前のデータに関しては、収集は行っているものの、リソース（時間・人手）不足が原因で、データ保存まで達していないとの回答が多かった。

8) 設問 9 : 保存したデータの探し方

データを作成した日付や、保存時のフォルダ名など保存者の記憶を基にしてデータを探すケースが大半であった。

そのため、情報の属性（ファイル名、日付・・・）では検索できなくなるケースや、管理者が変わると、全て探し出せているかどうか分からないなど、この方式の限界に関する回答が多く、データの再検索に満足できていない状況が伺える。

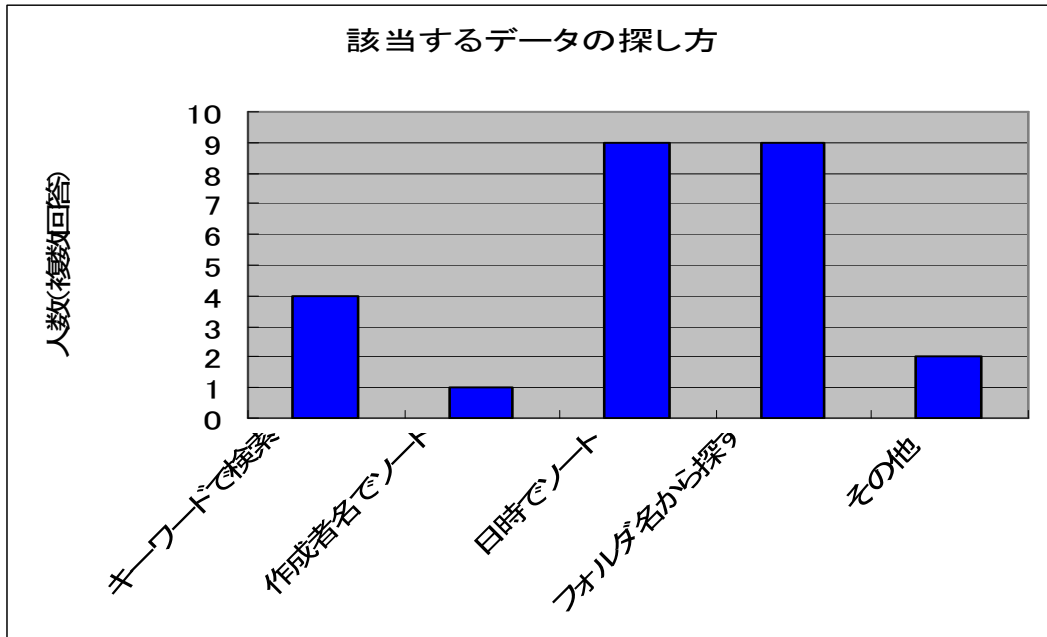


図 4.1-2 該当するデータの探し方

9) 設問 10 : 研究室等でのデータの保存と共有

データの保存と共有について手順ありと回答されたものの内訳は、フォルダや、ファイルの共有がほとんどで、ネーミングルールに基づいて運用しているケースは、少ない。

属性による検索や、データの管理担当者に聞いて、データを取り出している状況が明らかとなった。

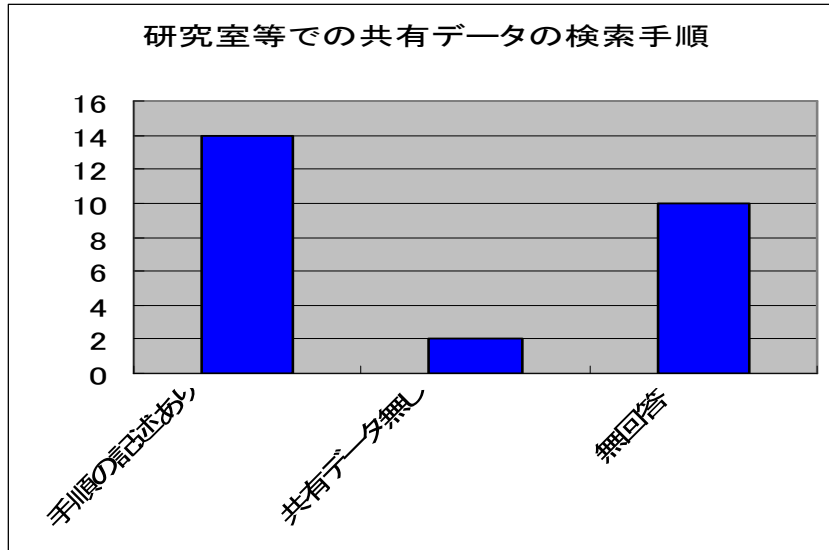


図 4.1-3 共有データの検索手順

10) 設問 11 : データの分析

分析については，以下のような項目に分類される．

手持ちの様々な情報と，地理情報（GIS）とを組み合わせた分析に対するニーズが高い．

- ・ 地理情報を活用している（したい）
- ・ 定性分析を行っている（行いたい）
- ・ 時系列での分析を行っている（行いたい）
- ・ 統計パッケージ SW を使っている（使いたい）
- ・ 独自の数値解析を行っている
- ・ グラフでの視覚化を行っている（行いたい）
- ・ 初心者向けの GUI だけでなく，熟練者・高度利用者向けの要件を取り入れた UI で効率よく分析したい
- ・ 数値データ公開示には背景情報の提供が必要
- ・ 外部の公開データの更新情報（差分情報）が必要
- ・ データ収集中で分析に至っていない

11) 設問 12 : データの取り出しで苦労している点分析

様々な課題について回答があったが，なかには，ツールの不備・不足に関する意見もあり，環境整備への要求が伺えた．

データを取り出しのほか，抽出した後の分析段階でも，同様な課題があることが明らかになった．

表 4.1-3 データを取り出す際の課題(下)

データを取り出す際の課題	事例
地理情報と組み合わせた分析のためのツールの不足	地図とのリンク。テキストデータからの位置情報の抽出が出来ない
定性情報のデータ化が出来ない	経済的・心理的ダメージなど、写真だけでは記録できないものがある
タグ付けなどデータの編集ツールの不足	時空間のタグ付けが簡単に行えない
データ整理編集の効率化・人的リソースの不足	データが膨大なので整理に手間がかかる
個人向けのツールが研究用途では機能不足	google マップもピンが多くなると扱いにくい
データの背景情報が現在では得られない	コンタクト者情報もほしい
共通の業務効率化のツールが不足	名刺管理が出来ていない 研究者紹介の更新もままならない
音声データを自動認識しテキスト化できない	テキスト化できれば斜め読み、表記や表現の統一も出来る
データを動的・時系列で視覚化出来るツールがない	動的・時系列での表現や分析が出来ない。
データの内容で検索できる環境やシステムがない	簡単にデータが引き出せる環境が欲しい。
データごとの著作権情報が未整備	著作権の所在がはっきりしていないとデータの使用許諾の連絡がすばやくできない
データファイルの管理機能がない	以前作成したデータのファイル名が分からなくなり、同じ作業を繰り返すことがある
研究者を対象としたSNSなどのツールがない	自分のデータを参照する人とコミュニケーションしたい
図書館所蔵の文献を内容で検索できない	古いDATAの文献が見つからない、東海・東南海関連や関東大震災関連も見つからない

12) 設問 13 : 保存したデータの利用状況

蓄積および保存目的のほか、具体的には以下のような目的でデータを利用することが明らかになった。共有や協業を行っているケースは僅かであった。

自己の研究

- ・ ガイドラインや評価手法の検討の参考
- ・ 事例調査と表現手法の構想
- ・ 分析、テキストマイニングに利用
- ・ プレゼン準備
- ・ データの解析、論文・発表資料の作成
- ・ データを再度分析、人に頼まれて探す

共有

- ・ 一部 HP で公表

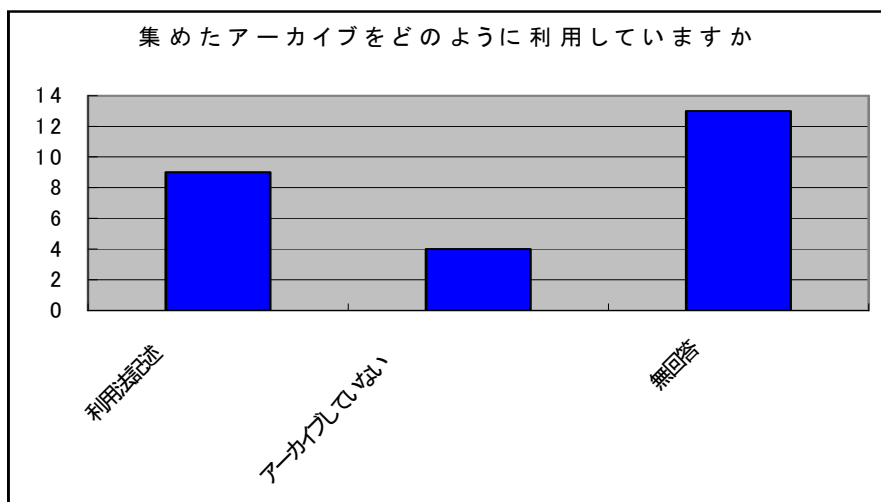


図 4.1-4 データの利用

13) 設問 14 : 今後のアーカイブ（データ）の利用

アーカイブ（データ）を自己の研究に利用にするとの回答のほか、他研究分野のデータを組

み合わせた活用，アンケートデータをより多面的に活用など，データ利用の拡張を志向する意見も見られた。

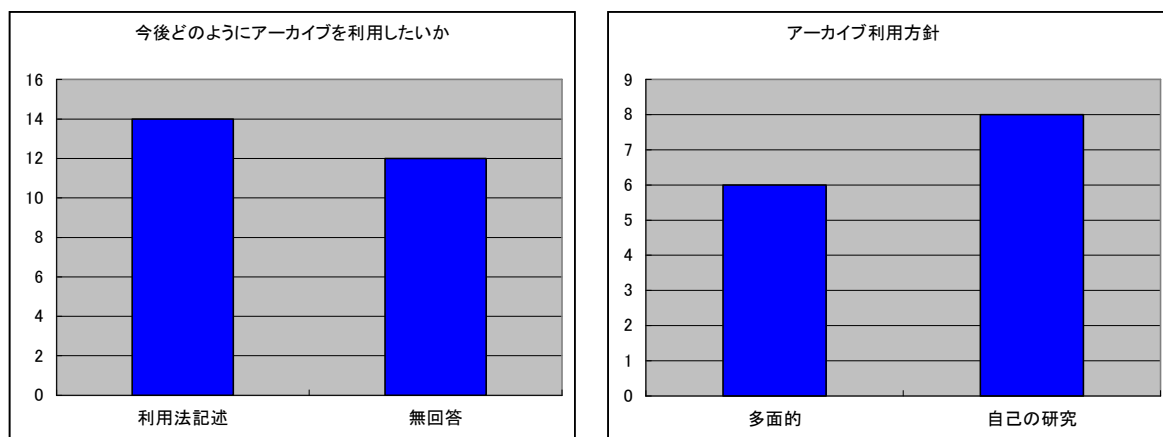


図 4.1-5 アーカイブデータの利用ニーズ

14) 設問 15 : 今必要なデータ

以下のようなデータが研究に必要と回答されている。

自己の研究領域以外のデータ

- ・行政・住民，住民間での復興への合意のプロセス，政策の効果を見るデータ

自己の研究領域のデータ

- ・海底および陸上の地形データ
- ・被害に関する統計情報
- ・被災後の特異的体験（普段と違うどのような行動が起きたのか）
- ・防災授業で使用するための，災害画像データ・動画データ
- ・東日本大震災に関する研究成果のうち，地図の形式で表現された図版
- ・東日本大震災に関するなるべく多くのテキストデータ
- ・ビデオ，写真，体験談など
- ・同様のデータを今年度末に同一被験者群から再び採取
- ・詳細な空中写真（被災前および後），津波遡上過程の動画
- ・地震観測データ，被災建物写真，津波・震災時の室内の動画・・・
- ・地形データ（津波前・津波後の測量データ，海底地形データ）
- ・かつての土地利用のメッシュあるいはポリゴンデータ
- ・浸水区域の詳しい情報がほしい
- ・原発避難者の個人所在データなど年齢・性別・場所で集計されたもの

15) 設問 16 : 現在使用しているシステム環境

Windows 系の OS が多いが，研究目的の端末については，一般企業のように OS についての統制をかけるようなことはないため，Mac，および Unix の OS など，バージョンまで含めると様々な種類の OS が使用されている。

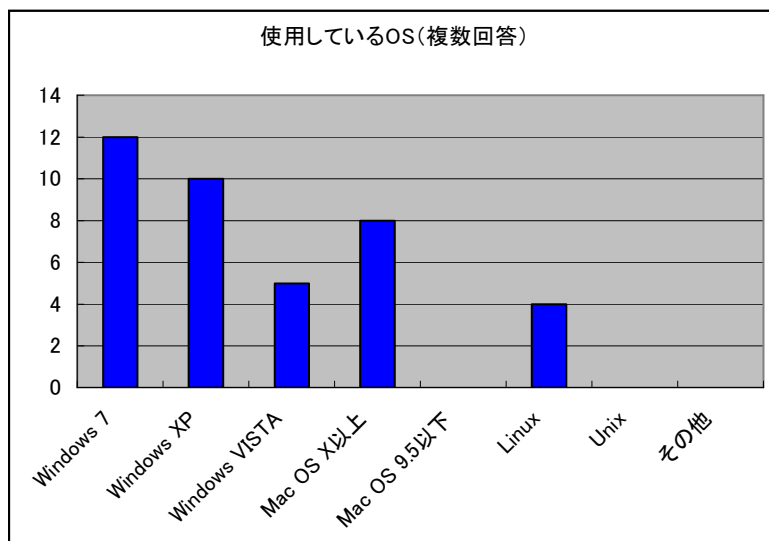


図 4.1-6 使用している OS

16) 設問 17 : データの総量

写真のデータ量が多く、総量では、数十 TB を超えるデータが格納される予定である。また、報告書などのテキスト情報等、今後の伸び率を含めた詳細な調査が必要である。

情報として、電子化前の新聞・雑誌の大量データがあることも考慮しなければならない。

表 4.1-4 データ総量(下)

	数値	写真	動画	テキスト	音声	アナログ	その他
写真 1千枚程度 1TB程度		1TB					
詳細不明							
衛星画像, 航空写真:300GB		303.5 GB					
現地調査写真:3.5GB							
新聞紙と雑誌4か月分, 日記やその他のデータ 1TB程度						4ヶ月	1TB
新聞紙(河北新報)5ヶ月分(3月～現在まで)						5ヶ月	防災教育関連資料数点
防災教育関連資料数点							
現在収集中のため、明示できる段階にありません							
ニュースデータ:300MB推定 Twitter:不明						ニュースデータ:300MB推定	
不明(数えたことがない)							
数値 DATA 3GB	3GB						640MB
EXCEL+WORD 合計で640 MB 以 脳画像データ全体で200GB、うち開 示可能分20GB		200GB					
50TB							50TB
被災地写真 5000枚 2TB, 動画		2TB					
写真 100GB		100GB					
建物の写真 1万5000枚 30GB		30GB、					
360度カメラ画像 20TB		20TB					
3.11の地震に関するデータを納めたフォルダ(震災後に作成)の容量	48GB					フィールドノートの記載数 10ページ	堆積物試料 数 kg
48GB							
堆積物試料 数 kg							
ムービー/写真データ 数TB	数TB	数TB	数TB				
計測データ 数TB							

4.1.2 プロトタイプシステム

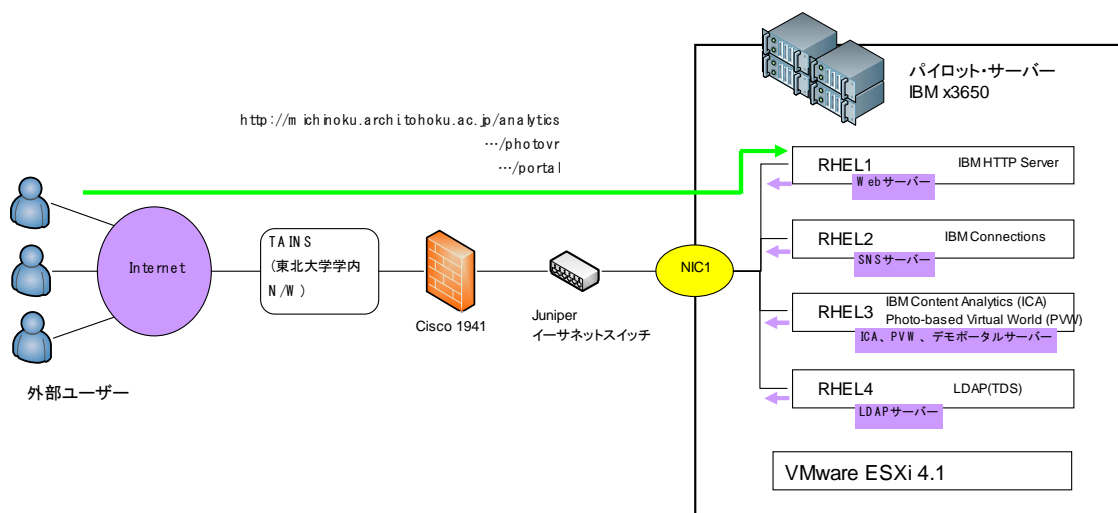
I. 主旨・目的

本システムを構築する前に、収集したデータの格納，収集データの検索結果表現の試行などを目的としてプロトタイプシステムを構築した。主に，開発環境，デモ環境，また，賛同協力機関とのコミュニケーションサイト運営環境として活用している。

II. システム構成

プロトタイプシステムは，IBM System x3650 M3 PLX 上に VMware ESXi を導入し，サーバー仮想化を行っている。4つの仮想サーバーを構築し，Red Hat Enterprise Linux をそれぞれ導入し，必要なソフトウェアの導入・セットアップを行った。

図 4.1-7 システム構成(下)



1) システムのハードウェア諸元

今回のシステム構築に当たり，導入したハードウェアは以下の通りである。
(ラックや電源などの設備は記載を割愛)

表 4.1-5 システムのハードウェア諸元(下)

種別	製品名	数量	備考
サーバーハードウェア	IBM System x3650 M3 PLX		
CPU	Intel Xeon E5649(6core) 2.53GHz	12	全体 12core 24thread
Memory	4GB(1x4GB.1Rx4.1.35V)PC3L-10600 CL9 DDR3 1333MHz LP RDIMM	32	全体 32GB
HDD	146GB 15K SAS 2.5 型 SFF Slim-HS HDD	2	RAID-1

			実効 136GB
	900GB 10K 6Gbps SAS 2.5 型 SFF HS HDD	3	RAID-0 実効 2.45TB
	1TB 7200rpm 6Gbps NL 2.5 型 SAS SFF HS HDD	7	RAID-5 実効 4.5TB
ネットワークインターフェース	全二重 Ethernet 10BASE-T/100BASE-TX/1000BASE-T (Broadcom 5709)	2	
ルーター	Cisco 1941 Security Bundle	1	ファイアウォール NAT 変換
ネットワークスイッチ	Juniper 24ポート 1Gb EX2200 イーサネットスイッチ	1	

システムは、VMware ESXi により、サーバー仮想化を行っているが、各仮想サーバーに割り当てている資源は表 2 のとおりである。

表 4.1-6 デモシステムのシステム資源の割り当て(下)

サーバー名	vCPU 数 ※1	Memory (GB) ※2	Disk 合計 (GB) ※3
RHEL1 (Web サーバー)	4	2	210
RHEL2 (SNS サーバー)	4	16	1500
RHEL3 (ICA,PVW サーバー)	4	11	525
RHEL4 (LDAP サーバー)	4	3	115
Ubuntu (東北大サーバー)	2	1	100
割当合計	18	33	2450
全体	24	32	4600

※1 システムの vCPU 数は、6 コア (12 スレッド) × 2 = 24vCPU. VMware ESXi 4.1 Standard の制限により、1つの仮想サーバーへの割り当て最大数は“4”

※2 Memory は、オーバーコミット (搭載量 < 割当量) させている

※3 Disk 合計には、VMware 使用域、バックアップ専用領域は含まれていない。

2) システム導入ソフトウェア

今回のシステム構築に当たり、導入した商用ソフトウェア及び主なオープンソースソフトウェアとそのバージョンは以下の通り。

(PVW は、 WebSphere Application Server 上で稼働するアプリケーションとして導入)

表 4.1-7 デモシステムの導入ソフトウェア一覧(下)

サーバー名	種別※1	ソフトウェア名称・バージョン
RHEL1	OS	Red Hat Enterprise Linux V5.6 64bit
	Web	IBM HTTP Server V7.0 fp 19 (+ Plugin)
	セキュリティ	Trend Micro ServerProtect™
RHEL2	OS	Red Hat Enterprise Linux V5.6 64bit
	DB	DB2 ESE V9.7.2
	AP	WebSphere Application Server ND 64bit V7.0 fp 19
	AP	IBM Connections V3.0.1
	AP	IBM Tivoli Directory Integrator V6.x
	セキュリティ	Trend Micro ServerProtect™
RHEL3	OS	Red Hat Enterprise Linux V5.6 64bit
	AP	WebSphere Application Server ND 64bit V7.0 fp 19
	Web	IBM HTTP Server V7.0 fp 19 (+ Plugin)
	AP	IBM Content Analytics V2.2
	セキュリティ	Trend Micro ServerProtect™
RHEL4	OS	Red Hat Enterprise Linux V5.6 64bit
	DB	DB2 ESE V9.7.2
	AP	WebSphere Application Server Base 64bit V7.0 fp 19
	AP	IBM Tivoli Directory Server V6.3
	セキュリティ	Trend Micro ServerProtect™
Ubuntu	OS	Ubuntu Server 10.10
	DB	PostgreSQL 9.1 + PostGIS 1.5.3-1
	AP	Tomcat 7.0.21
	WEB	Apache 2.2.20
VMware	仮想化	VMware ESXi 4.1 Standard

3) システムの運用

システム全体としての運用を記載する。

(a) サービス提供時間

24 時間/365 日。ただし、毎週金曜日 18:00-22:00 のシステムバックアップ他、メンテナンスのため停止する場合を除く。

(b) システム起動・停止

システム管理者が手動にて、起動・停止運用を行う。オペレーションは、サーバールーム内の管理用ローカルネットワーク経由で実施する。

(c) システムバックアップ

毎週金曜日 18:00-22:00 に、システム管理者によってバックアップを取得する。

VMware のユーティリティを利用し、仮想サーバー全体のバックアップを取得する。

現在、バックアップの対象となっているのは、RHEL1(Web サーバー)、RHEL2(SNS サーバー)、RHEL4(LDAP サーバー)であり、RHEL3(ICA,PVW サーバー)は、バックアップの対象外としている。

現在の所要時間は、約 20 分程度である。(バックアップ取得前後の起動・停止時間を除く)バックアップデータは、1 世代のみの保管。

(d) ユーザー登録

システム「RHEL4」の LDAP サーバー上のユーザー情報での認証を行う。同じ LDAP サーバー上の同エントリーで SNS サービスの認証も行っている。ユーザー情報の登録、修正、削除は、ユーザーからの依頼に基づいて、システム管理者が手動で実施する。

III. 今後必要となる対応

大量のデータを蓄積・活用するシステムとして、今後、以下のような対応が必要となる。

1) ユーザーインターフェースの整備

ユーザーに対しては、単一のアクセスポイント(操作画面)を提供し、その画面上で直観的にすべての操作が行えるようにする必要がある。そのようなインターフェースデザイン的前提として、どのユーザーにどのような機能を使用させるか、また、各要素技術をどのように連携させるかの検討は欠かせない。

2) システム利用ルールの整備

外部に公開することを目的として、システムを構築する場合、システム利用ルールの整備は必須である。検討すべき項目は多岐にわたるが、特にユーザー視点からのルール作りが必要である。

どのユーザーが、どこまでのデータに、どのようにアクセスができるか、利用ユーザーや各ユーザーが登録するデータはどのようにチェックするか等、検討項目は多い。

ルールの整備が完了したら、ルール順守を補助する機能を組み込み、長期にわたり、ルールが形骸化しないようにする。

3) 登録データの管理・保護

膨大なデータを蓄積・分類・提供し、そのデータを基にした二次利用、三次利用を通じて、新たな知の創出につなげることを目的とした場合、登録データの管理や保護といった観点での検討は必須である。

蓄積されたデータは著作権やその他の法律などによって保護されている場合も多く、データの提供者、利用者ともに合意できる仕組みづくりが必要である。

また、膨大なデータの中に個人情報など、開示が適切でない情報が紛れ込む可能性もある。それらの適切でない情報の検出・排除方法などの検討も必要となる。

4) データバックアップの確立

テキストデータのみならず、画像、音声、動画を活用する当システムには、数テラバイト～数百テラバイトまでのデータが格納される。貴重な学術データの長期保管も前提となるため、定期的なデータバックアップは必須である。

バックアップ対象データ量が膨大になることが想定されるため、バックアップの取得タイミング、頻度、保管世代数、利用技術の検討は慎重に行う必要がある。また、蓄積データの希少性が高まるほど、分散保管の検討も必要になるであろう。

5) パフォーマンス管理

長期間にわたり、データ量の増大が想定されるシステムであるため、システムの定期的なパフォーマンス管理は重要である。管理項目は、提供するサービスのレスポンスタイムやCPU、メモリーなどのシステム資源、ディスク容量などである。ある一定の閾値を判断基準として、システム資源の再配分や増強を検討することになるであろう。判断のよりどころとなるパフォーマンスデータを蓄積することはかなりのワークロードを要する。

管理ワークロードの軽減のために、システムのパフォーマンス管理を行うツールの検討も一考に値する。

4.1.3 ホームページ

I. 主旨・目的

時間の経過と共に、東日本大震災関連情報も、被害状況や避難状況を伝える内容から、復旧・復興情報へ変化がみられるようになった。「みちのく震録伝」においても、震災時の厳しい状況を伝える情報から、被災地の「今」を伝えるホームページへの変更が必要となったことから、ホームページを再作成することとした。

コンテンツ管理システム（CMS）のひとつである WordPress を採用し、今後のメンテナンスや記事の追加に対する利便性を考慮したホームページの作成を行なっている。

II. ホームページの構成

サイトマップに示すように、ホームページは大きく 6 個のカテゴリに分類する。

- ・ トップ
 - メッセージ性の強いコラムや、プレスリリースなどをまとめたカテゴリ。
 - 特にトップページは、図にあるように分割表示している。
- ・ 「みちのく震録伝」
 - プロジェクトの概要や活動について紹介するカテゴリ。
- ・ シンポジウム・報告会
 - シンポジウムや報告会など、研究成果報告を紹介するカテゴリ。
- ・ コンテンツ
 - 研究者や企業とのコラボレーション企画など、情報にアクセスするためのコンテンツを紹介するカテゴリ。
- ・ 賛同・協力／パートナー
 - 賛同・協力機関、アーカイブに関するパートナー機関の紹介カテゴリ。
- ・ みちのく・いまをつたえ隊
 - 現地情報収集活動の報告カテゴリ。

サイトマップ

- ▶ トップ
 - ▶ コラム
 - ▶ プレスリリース
 - ▶ 報道
- ▶ みちのく震録伝
 - ▶ プロジェクト概要
 - ▶ ロゴ・パートナー
 - ▶ ワーキンググループ
 - ▶ 情報の収集・協力のお題①
 - ▶ 寄付金のお題①
 - ▶ プロジェクトメンバー
- ▶ シンポジウム・報告会
 - ▶ 東日本大震災アーカイブ国際合同シンポジウム 開催要項
 - ▶ 東日本大震災アーカイブ国際合同シンポジウム 実施
- ▶ コンテンツ
- ▶ 賛同・協力／パートナー
 - ▶ 賛同・協力機関の一覧
 - ▶ 賛同・協力機関の紹介
 - ▶ 賛同・協力機関の活動
 - ▶ パートナー
- ▶ みちのく・いまをつたえ隊
 - ▶ 調査地域
 - ▶ 現在の活動状況

図 4.1-8 サイトマップ構成

これらのカテゴリには、それぞれサブカテゴリがあり、個別のページとなっている。WordPress では、これらの内容を固定ページもしくは投稿ページとして作成している。投稿ページへの項目追加は、投稿の新規追加により、簡単に行うことができる。

トップページは、以下のように重要な情報の集約された内容となっている。

The screenshot shows the homepage of 'Mitsuru no Kuraiki' (Mitsuru's Diary) with several callout boxes highlighting key features:

- タブによるカテゴリ表記**: Points to the top navigation tabs (Home, Mitsuru's Diary, Shinjokawa New Year Meeting, Contents, Mitsuru's Diary/Partners, Mitsuru's Diary/Partners).
- コラムによる注目度の高い情報の表示**: Points to the featured article section on the right side of the page.
- コンテンツの紹介**: Points to the 'Contents' section on the left side of the page.
- 新着情報の表示**: Points to the 'New Arrivals' section on the right side of the page.
- イベントのカレンダー表示**: Points to the 'Calendar' section at the bottom right of the page.

The main content area includes sections for 'Contents', 'New Arrivals', and 'Calendar'. The 'Calendar' section shows a grid for March 2012 with dates 1 through 31.

図 4.1-9 トップページ構成

また、サブカテゴリのページでは、左フレームにサブカテゴリの目次を表記し、操作性を向上させている。



図 4.1-10 サブカテゴリ構成

III. システム環境

ホームページ作成においては、下記ソフトウェア環境で実施した。

- Wordpress V3.1.4
- PHP V5.3.9
- mySQL V5.1

また、これらの稼働については、ブラウザ製品およびそのバージョンによって表示のずれや振る舞いの違いが発生する。今回のホームページ作成において、前提としたブラウザは、下記の通りである。

- Microsoft IE V7, V8
- Firefox V10
- Google Chrome V18