

データの整理・分類のためのワークシート

○宇宙航空研究開発機構(JAXA)

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	JAXA	公開用Web	公開用Webコンテンツ(静的)	非構造	1回/週程度の更新	千差万別	元データがFirewall内サーバに有り	コンテンツ作成者	単一サーバを参照	なし	なし	不明(数GB程度?)	ディスク	不明	コンテンツ作成者がそれぞれの分担を管理	過去のコンテンツの保管はこのシステム上ではなく、コンテンツ作成者が各自のローカルな環境にバックアップする。
2	JAXA	数値シミュレーションデータ	スパコンの入出力データ	非構造	千差万別	千差万別	仕組みは無し	ジョブの所有者	スパコンシステム内では共有可能	なし	なし	不明	HSM(ディスク+テープ×2)で保存。上記に加えて、ローカル環境にコピーする場合もある。	300TB	ジョブの所有者が管理	HSMのポリシー管理はスパコン管理者が実施。
3	JAXA	実験データ	風洞試験データ、飛行試験データ、観測データ等	非構造	1回/週程度か?(厳密には不明)	不明(今までの実験データは削除せずテープにアーカイブして保管している)	仕組みは無し	実験実施者	実験計測システムとアーカイブシステムはファイルシステムのクロスマウントにより共有を実現。	なし	無し(但し、市販アプリケーションのインターフェースを通じて管理)	30TB程度	HSM(ディスク+テープ×2)で保存。実験測定直後は、実験計測システムの高速度RADIディスクに保管している。	100TB程度	実験実施者	ネットワークファイルシステムがローカルの高速ディスク並みに高速であれば、直接書き込みも考えられるが、現時点では、実験計測システムのデータ出力速度に耐えられるネットワークファイルシステムが見当たらない。
4	JAXA	電子メール	業務遂行上のメール	非構造	12000通/日程度	千差万別	仕組みは無し	メールの受信者/メールサーバ管理者	無し IMAPによるメール処理は共有と言う?	なし	なし	不明	ハードディスク。ローカル環境にバッチ処理でコピーして保管。	300GB以上	各人で管理	

○国立天文台

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	国立天文台 ハワイ観測所	すばる望遠鏡観測データ	観測データおよび望遠鏡エンジニアリングデータ(STARS)、環境・気象データ	コンテンツ管理	常時	永久	バックアップ有、一部のエンジニアリングデータはバックアップなし	計算機・データ管理部門	Webベースでデータ検索/要求	ORACLE	UIは独自開発、ミドルウェアとしてSafeFile/Global(PetaServe)を使用	10TB	Disk+Tape(PetaSite)	60TB	消去せず、Read-Only	バックアップは山頂でテープ保存および三鷹MASTARSが該当する。
2	国立天文台 (三鷹)	すばる望遠鏡観測データ	観測データおよび望遠鏡エンジニアリングデータの一部(MASTARS)	コンテンツ管理	常時	永久	バックアップ有	計算機・データ管理部門	Webベースでデータ検索/要求	ORACLE	UIは独自開発、ミドルウェアとしてSafeFile/Global(PetaServe)を使用	5TB	Disk+Tape(PetaSite)	30TB	消去せず、Read-Only	バックアップは山頂でテープ保存およびハワイ観測所STARSが該当する。
3	国立天文台 野辺山 宇宙電波観測所	野辺山45m観測データ	観測データ	コンテンツ管理	常時	永久	バックアップ無	計算機管理部門	複数サーバから共有	ORACLE	独自にAPIを提供	2.0TB	Disk+Tape(STKライブラリ)	3.7TB	消去せず、Read-Only	
4	国立天文台 野辺山 宇宙電波観測所	野辺山10m観測データ	観測データ	コンテンツ管理	常時	永久	バックアップ無	計算機管理部門	複数サーバから共有	ORACLE	独自にAPIを提供	1.0TB	Disk+Tape(STKライブラリ)	4.7TB	消去せず、Read-Only	

○中京大学

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	中京大学	個人用データ	個人用ホーム領域(学部学生)	非構造	適宜, 特に講義時	4年	あり	NW管理者	複数サーバ間で共有	ディスククォータによる使用上限設定	なし	合計 1400GB	Disk+Tape (LTOライブラリ)	合計 510GB	4年で消去	
2	中京大学	個人用データ	個人用ホーム領域(教員・教員研究用・大学院生用)	非構造	適宜	退職・修了まで	あり	NW管理者	複数サーバ間で共有	なし	なし	160GB	Disk+Tape (LTOライブラリ)	合計 83GB	退職後特に依頼が無ければ1年で消去	
3	中京大学	講義課題用データ	講義用課題提示・レポート提出用領域	非構造	適宜, 特に講義・レポート作成時	1年	あり	NW管理者	複数サーバ間で共有	ディスククォータによる使用上限設定	なし		Disk+Tape (LTOライブラリ)		1年で消去	個人用ホーム領域(教員, 教員研究用, 大学院生用)と共用 高可用性が必要
4	中京大学	個人用データ	個人用Profile領域(学生・教員共用)	非構造	適宜, 特に講義時	退職・卒業まで	なし	NW管理者	クライアントへ共有	なし	なし	100GB	Disk	60GB	学生: 4年で消去 教員: 退職後消去	
5	中京大学	テンポラリー領域	テンポラリー領域	非構造	適宜, 特に講義時	次回PowerONまで	なし	ユーザ	なし	なし	なし	7.8GBx390 +30GBx95	Disk	(7.8GBx390 +30GBx95)	各ユーザが管理	各PCに用意されているテンポラリーディスク(Dドライブ). 講義中に発生する中間データ一時保存用
6	中京大学	メールスプール	メールスプール(学生・教員共用)	非構造	適宜 2000通~5000通/日程度	退職・卒業まで	なし	NW管理者	複数サーバ間で共有	なし	なし	55GB	Disk	8GB	学生: 4年で消去 教員: 退職後特に依頼が無ければ1年で消去	

○北陸先端科学技術大学院大学(JAIST)

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	JAIST	ホーム領域	教員・学生用ホーム 事務職員用ホーム	非構造	随時	所有者アカウントが存在する期間	バックアップ有	情報科学センター	WS, TS, 並列計算機等から共有 事務用TSから共有	特になし	特になし	6.5TB	Disk	30.5TB	所有者アカウントが存在する間は保存, アカウントがなくなると圧縮して保管	[fs3, fs8: 統合ストレージシステム] バックアップは約3ヶ月分
2	JAIST	テンポラリー領域	申請により作成される研究用データ領域/申請なしで利用できるデータ領域	非構造	随時	申請期間/不定	バックアップ有/無	情報科学センター	WS, 並列計算機等から共有	特になし	特になし	7TB	Disk	11TB	申請期間は保存	[fs1, fs3: 統合ストレージシステム] 必要に応じ人手で容量追加
3	JAIST	アプリケーション領域	WS用アプリケーション領域	非構造	ほぼ更新無し	長期間	バックアップ有	情報科学センター	WSから共有	特になし	特になし	0.1TB	Disk+Tape	0.5TB	ほとんど更新無し	[fs9] バックアップは約3ヶ月分
4	JAIST	ワーク領域	計算データ	非構造	随時	不定	バックアップ有	情報科学センター	なし	特になし	特になし	3TB	Disk+Tape	9TB	あふれそうになったらお知らせ	[個別Disk]
5	JAIST	サーバ領域	外部公開用Web, FTPサーバ	非構造	随時	所有者アカウントが存在する期間	バックアップ有	情報科学センター	なし	特になし	特になし	7TB	Disk+Tape	8TB	あふれそうになったらお知らせ	[個別Disk]
6	JAIST	サーバ領域	メールスプール領域	非構造	随時	不定	バックアップ有	情報科学センター	なし	特になし	特になし	0.1TB	Disk+Tape	0.5TB		[個別Disk]
7	JAIST	ログデータ領域	各種サーバから収集したログ	非構造	随時	不定	バックアップ有	情報科学センター	なし	特になし	特になし	0.5TB	Disk+Tape	0.5TB		[個別Disk]
8	JAIST	サーバ領域	講義アーカイブデータ	非構造	約10件/日	長期間	バックアップ有	遠隔教育研究センター	複数配信サーバから共有	特になし	特になし	1TB	Disk	5TB	1年後に別ディスクに移動	[個別Disk] 移動は人手

○理化学研究所

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	理研	ライフサイエンス	公開されている遺伝子・タンパク質データ	コンテンツ管理	適宜	1世代位	一部バックアップあり	サーバ管理者	なし	独自	なし	数TB	Disk	8TB	新規DBが公開されるまで保持 消去しない	バックアップは基本的にしない
2	理研	実験データ	高エネ物理実験データ	コンテンツ管理	適宜	永久	本データは元データのレプリカ	サーバ管理者	複製	なし	なし	約300TB	Disk+Tape	1PB		元データはUS BNLに存在する
3	理研	シミュレーションデータ	スパコンユーザの利用領域	非構造	適宜	不明	ホーム領域のみ	ユーザ	なし	なし	なし	2TB	Disk+Tape	約37TB	ユーザ次第	テープへのバックアップはホーム領域のみ。容量の内訳はディスクが約7TB、テープが約30TB