



2018年度 JDSFストレージ要素技術 (SET)部会活動成果報告 「アーカイブシステムを考える時のポイント」

部会長：齊藤

副部会長：村竹

副部会長：須賀田

版数	日付	内容
V1.0	2019/7/10	初版作成

注記：

本資料に記載されている資料は、公開されている情報を参考としています。

検討内容に関しては、JDSF要素技術部会において議論・検討された内容であり、データ・アーカイブを全体的にとらえて考えた内容となっています。個々の要件や条件における最適な方法を提供する目的では記載されておりませんので、ご留意お願い致します。

尚、本内容に関してはJDSFにお問い合わせください。

アーカイブシステムを考える時のポイント

- 本活動の目的
- 背景
- データ特性の調査・検討
- ILM調査・検討
- メディア特性調査
- まとめ
- 活動メンバ

- 本活動の目的
- 背景
- データ特性の調査・検討
- ILM調査・検討
- メディア特性調査
- まとめ
- 活動メンバ

■ 本活動として以下のような内容を提案したい

- ✓ データ利活用を支えるデータのアーカイブを実現するストレージ・システムを、どの様に考えれば良いかを示す
 - データにはデータ種類ごとの価値（特性）があることを調査から示す
 - アーカイブをインフォメーションライフサイクルマネジメント(ILM)の視点からとらえ、どの様な軸で考えるのが良いかを示す
 - ストレージメディアの特長
 - アーカイブを実現するストレージ・システムを作成する際のポイント

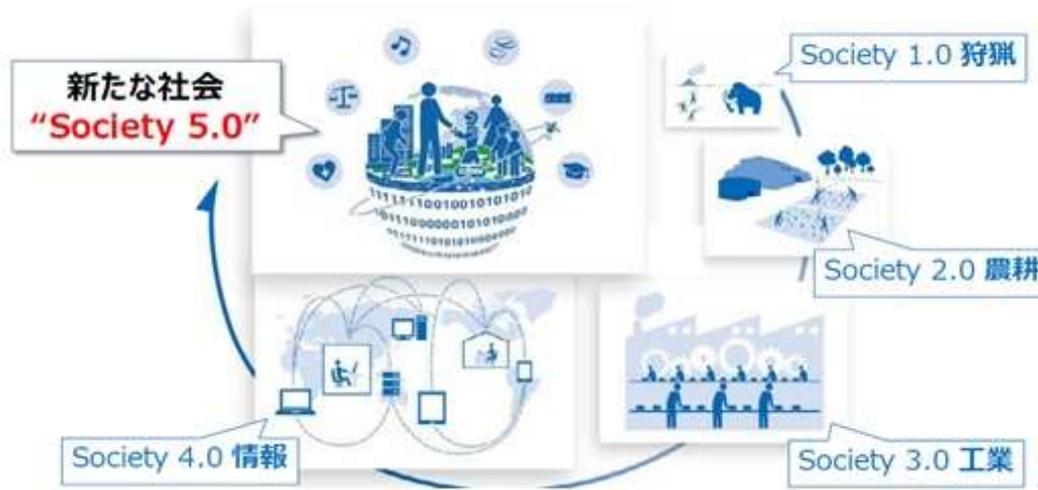
■ 具体的な調査内容

- ① データ種類と利活用視点でのデータ特性
- ② データ利活用におけるILMの考え方
- ③ 各種メディアの特長

- 本活動の目的
- **背景**
- データ特性の調査・検討
- ILM調査・検討
- メディア特性調査
- まとめ
- 活動メンバ

■ デジタルデータの利活用がますます重要になる

- ✓ 内閣府未来投資戦略2018 Society5.0や総務省/情報通信白書等、IoT/AI/BigData時代におけるデータ利活用の重要性が記載されている。
- ✓ 本部会のメンバーもデータの利活用意識が高まっていると感じている。

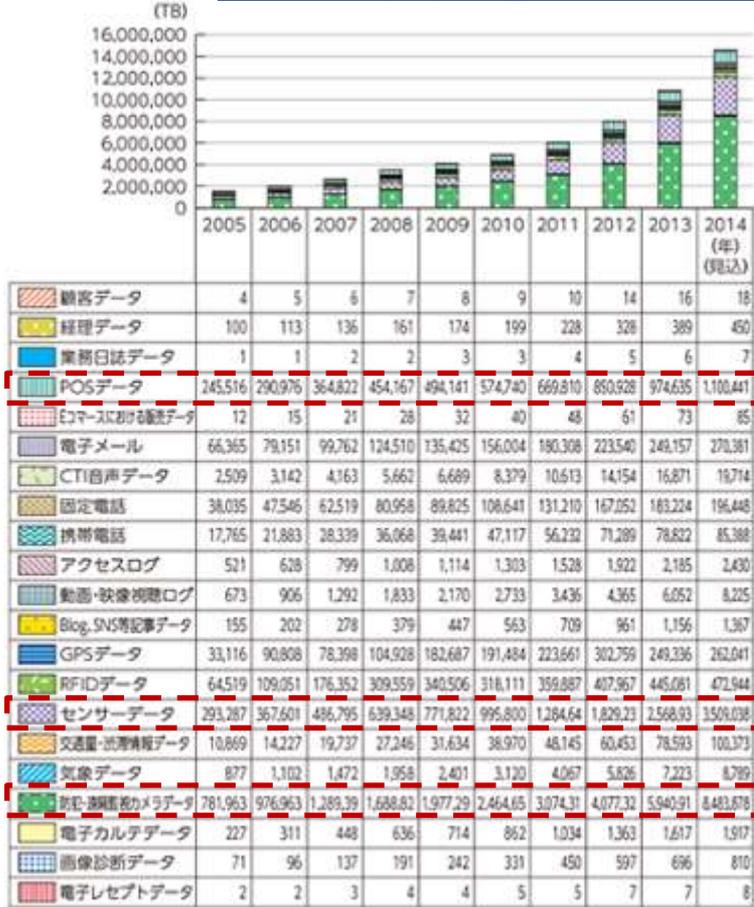


出展：内閣府未来投資戦略2018 ~Society5.0/データ駆動型社会の絵

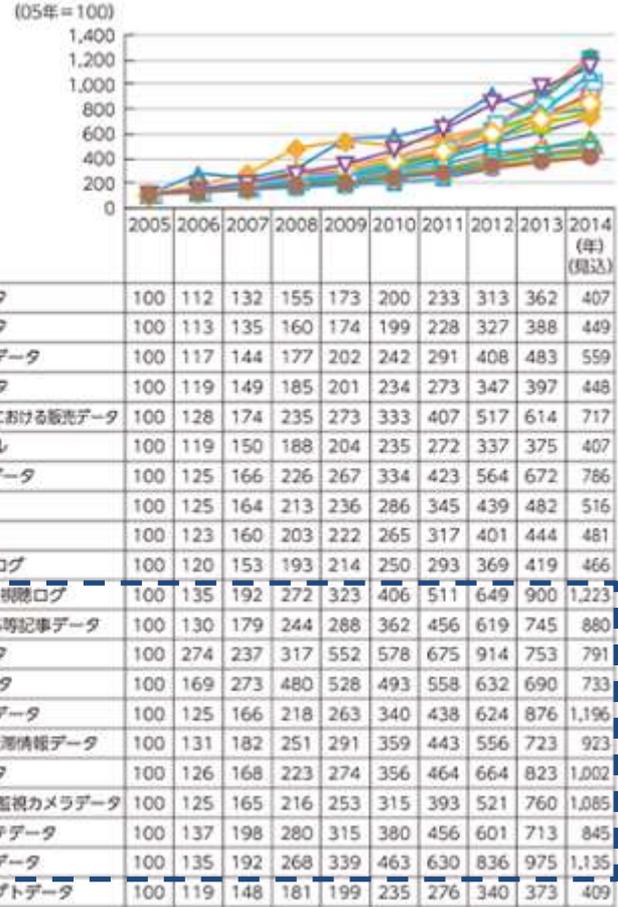


間違いなく、データ量は増加しています

出展：H27年の情報通信白書のデータ流通量の絵



データ量：大

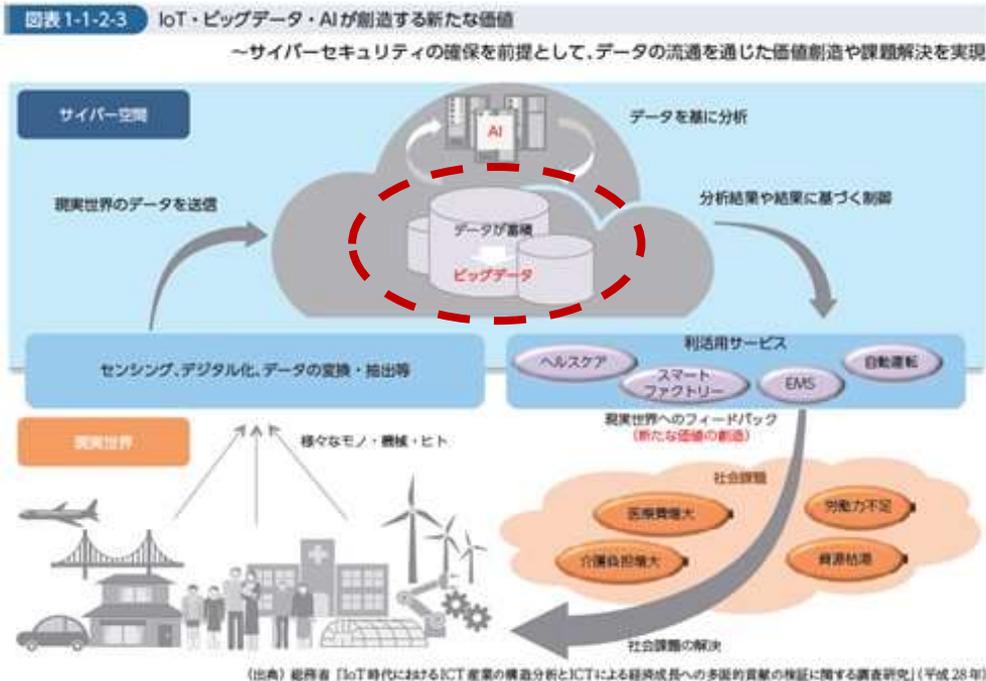


増加率：大

今後
更に
増加

■ データの重要性を支えるデータアーカイブ

Society4.0情報社会においても資産と位置づけられるデータ・将来における再利用・コンプライアンス観点から長期保存が必要となるデータが存在する。Society5.0データ駆動型社会においては、価値あるデータを如何に多く保有しているかが企業価値を左右する。現在利活用できないデータであっても、将来AI/データ分析技術の向上により新たな価値を生みだし、そのデータ自体も価値あるものとなる。しかしながら、肝心のデータ蓄積方法に関しては触れられていない。本部会ではデータを蓄積（アーカイブ）部分に関して活動を行っている。



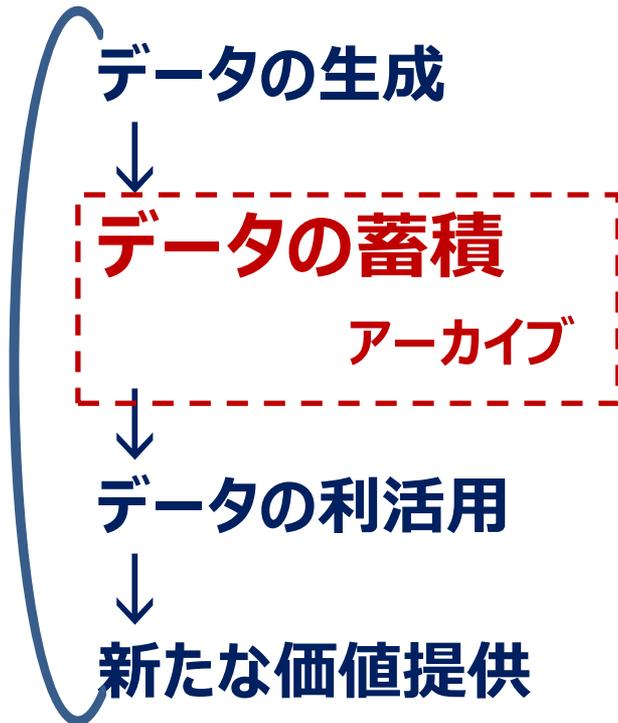
データの生成

データの蓄積

アーカイブ

データの利活用

新たな価値提供



- Society5.0の実現に向けて、データ利活用のベースとなるデータの蓄積、すなわちデータアーカイブが重要になるという共通認識を得た。
- データ蓄積（アーカイブ）をどのように実現するかの方法に関して活動する事に意義はある。

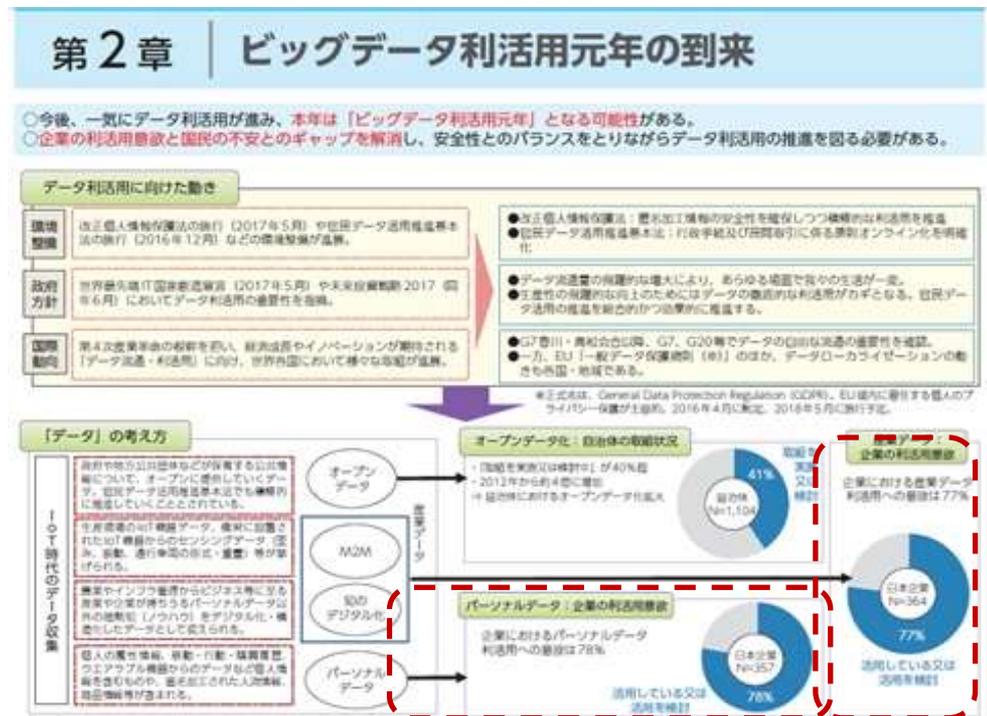
- 本活動の目的
- 背景
- データ特性の調査・検討
- ILM調査・検討
- メディア特性調査
- まとめ
- 活動メンバ

データ特性の調査・検討

■ データ利活用に向けて記載されている 産業データ・パーソナルデータに関して利用意欲が高い

- オープンデータ
- 産業データ
- パーソナルデータ

出展：総務省白書(H29)から引用



■ 映像（放送/コンテンツ）に関しては、アーカイブが運用と一体化している

■ SET部会で行ったデータ調査結果

=> データには種類によって幾つかの**アーカイブの動機**に分類できる

今回調査したデータ

CG画像(映画)
動画/映像 (放送)

設計データ(CAD/CAE)
製造データ
医療データ

動画/音声/画像 (著作権)
設計/製造/監視カメラ (製造物責任)
医療 (医師法)

センサ (保管しない) 等

分類

アーカイブの動機

積極的な
利活用

活用の可能性があ
り取っておく

捨てられない

法規制

廃棄

総務省白書
の調査

パーソナル
データ
産業データ

データ特性の調査・検討：まとめ

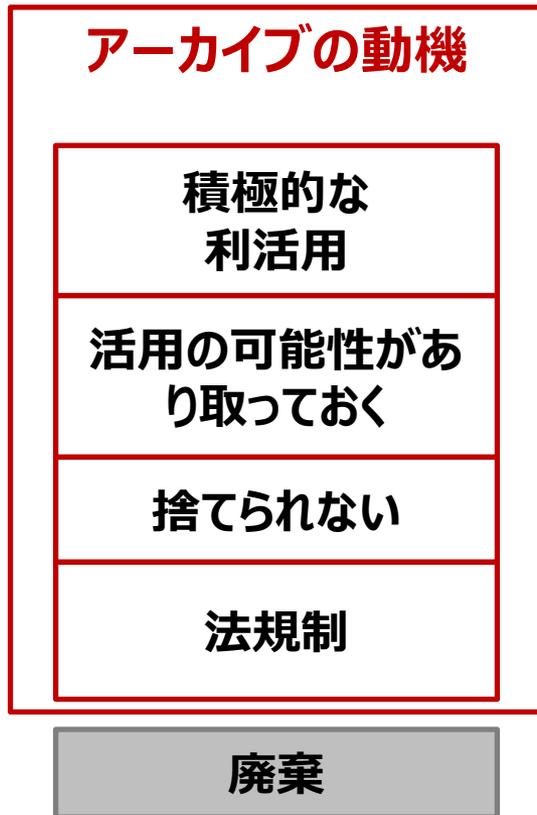
- 調査結果の詳細をエクセルでまとめています。別ファイルで提供。
- データ種類/業務/特性/アーカイブの動機として一覧化

アーカイブ関連調査の調査フォーマット(JDSF/SET部会)

©調査フォーマット 20180626
公開権限付 20180620

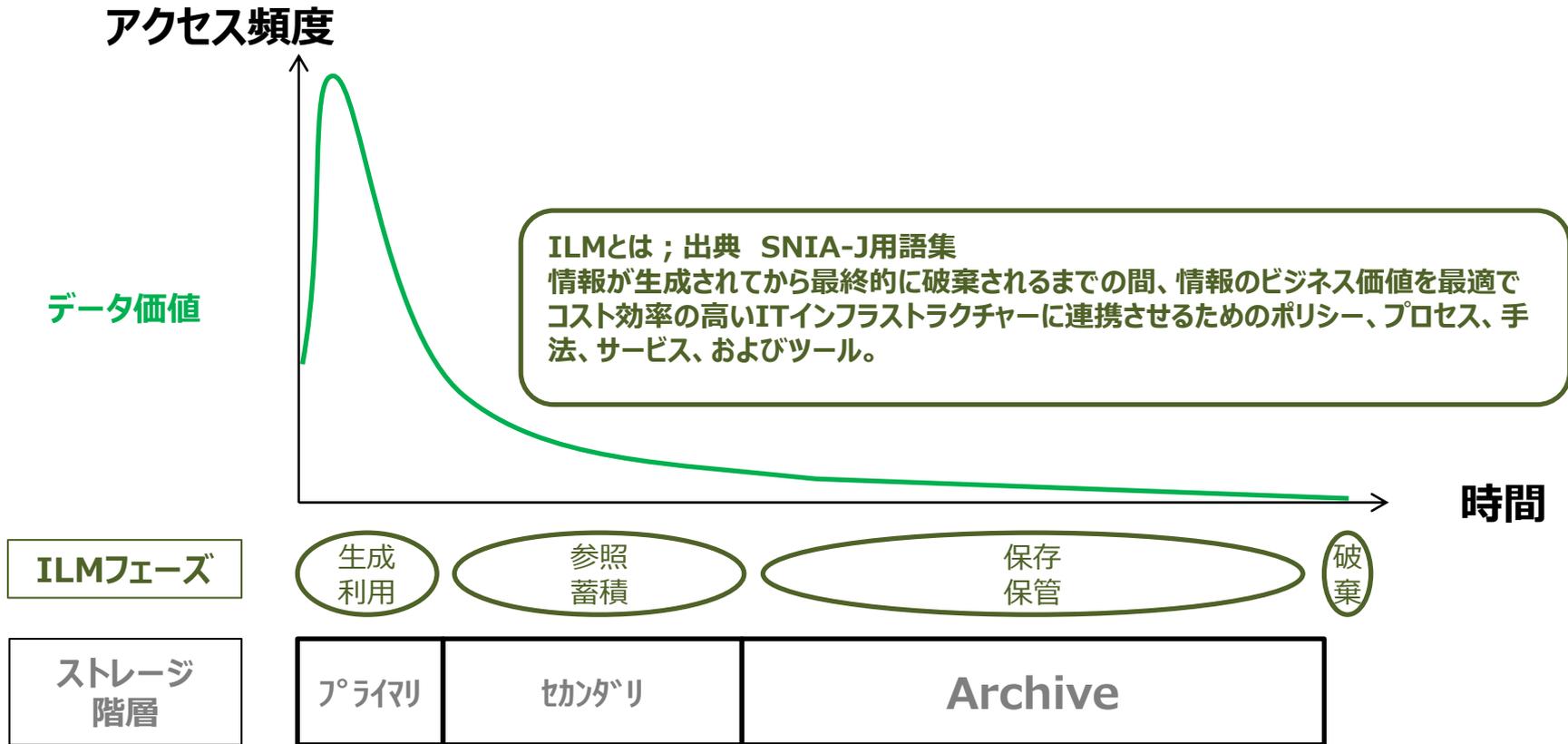
※必須に応じて「予」の追加をお願いします。

調査社	業種	用途	データの利用用途		データの種類		お持ちの状況			今後の必要性		コストの範囲		今後の条件		備考	アーカイブの動機	
			単体規模などの条件	データの総量	ファイル数	容量	アクセス頻度	法規制	保存期間	消去の必要性	コストの範囲	アクセス頻度						
IoT/センサーデータ																		
日立	技術実証実験データの運用管理	センサーデータ	西小牧地域の海域で約10ヶ所の観測地点	モニタリングデータ	モニタリングデータ	年毎最大20TB、最終的には100TB以上	不明	不明	不明	不明	不明	不明	不明	不明	不明	不明	不明	NEC事例1 クラウドにデータ移す クラウドにデータ移る
映像データ																		
NEC	映像	監視映像データ				100TB/年												HDD B. 法規制等による義務 C. 今後活用の可能性がある
SonyIPS	調査	映像制作データ	数十名	動画ファイル(数GB/file)	不明	不明	不明	不明	不明	不明	不明	不明	不明	不明	不明	不明	不明	B. 法規制等による義務
映像/放送データ																		
NEC	テレビ放送	動画/録音ファイル	(準キー局)	録音素材	約13万時間分													NEC事例5 A. 将来的に利益を生む
NEC	テレビ放送	動画/録音ファイル	(準キー局)	録音素材	年毎500TB以上生成													NEC事例5 A. 将来的に利益を生む
富士通	映像	放送コンテンツのバックアップ		映像	23PB/初年、15PB/年以降													A. 将来的に利益を生む
富士通	映像	映像アーカイブ		映像	300TB、今後数年													A. 将来的に利益を生む
富士通	映像	CD映像			20TB(2011)→150TB(2015)													HDD A. 将来的に利益を生む
SonyIPS	放送局	動画/音声/画像	百名程度	動画ファイル(1GB/file)	約1TB/day/1PB													A. 将来的に利益を生む
SonyIPS	放送局	動画/音声/画像	数十~数十名	動画ファイル(数GB/file)	約100GB/日													A. 将来的に利益を生む
SonyIPS	映像制作会社	動画/音声/画像	数十~数十名	動画/音声/画像(2GB/file)	数TB/day/日													A. 将来的に利益を生む
SonyIPS	映像制作会社	動画/音声/画像	数十	動画/音声/画像(数GB/file)	数GB/日													A. 将来的に利益を生む
SonyIPS	映像制作会社	動画/音声/画像	数十	動画ファイル(10MB/file)	約100GB/日													A. 将来的に利益を生む
SonyIPS	映像制作会社	動画/音声/画像	数十	動画ファイル(1.5GB/file)	数GB/日													A. 将来的に利益を生む
監視データ																		
日立	家電	生体・状態監視		産業カメラ画像	1MB~100MB/ファイル													D. 記録/記録して保管
日立	製造	監視カメラ		AVIファイル	100MB~5GB													D. 記録/記録して保管
日立	食品製造	監視カメラ(製造工程、入退室)	生鮮食品	VGA画像	~500MB/日													B. 法規制等による義務 C. 今後活用の可能性がある
日立	食品製造	監視カメラ(製造工程、入退室)	一時食品	VGA画像	~500MB/日													B. 法規制等による義務 C. 今後活用の可能性がある
日立	食品製造	監視カメラ(製造工程、入退室)	長期保管食品(加工等)	VGA画像	~500MB/日													HDD、テープは検討 B. 法規制等による義務 C. 今後活用の可能性がある
日立	食品製造	監視カメラ(生鮮、動物監視)		SKVGA画像														D. 記録/記録して保管
研究データ																		
SonyIPS	各種研究開発	研究/監視データ	不明	動画ファイル(数十MB/file)	数GB/日													B. 法規制等による義務
研究データ																		
日立	研究	ゲノム解析		ゲノム解析データ	200GB~1TB/人													C. 今後活用の可能性がある
富士通	研究所	HPCデータ			50~300PB													C. 今後活用の可能性がある
富士通	大学	HPC/AIデータ			50PB													C. 今後活用の可能性がある
富士通	科学技術	HPCデータ			50PB~300PB													C. 今後活用の可能性がある
富士通	x x IT	気象データ			数GB/日													C. 今後活用の可能性がある
SonyIPS	研究所	研究データ	数十~数十名	ゲノム解析情報データ	数十GB/日													C. 今後活用の可能性がある
SonyIPS	研究所	研究データ	数十名	回折データ(数十GB/file)	数GB/日													C. 今後活用の可能性がある
設計/製造データ																		
NEC	製造	設計/製造データ																A. 将来的に利益を生む



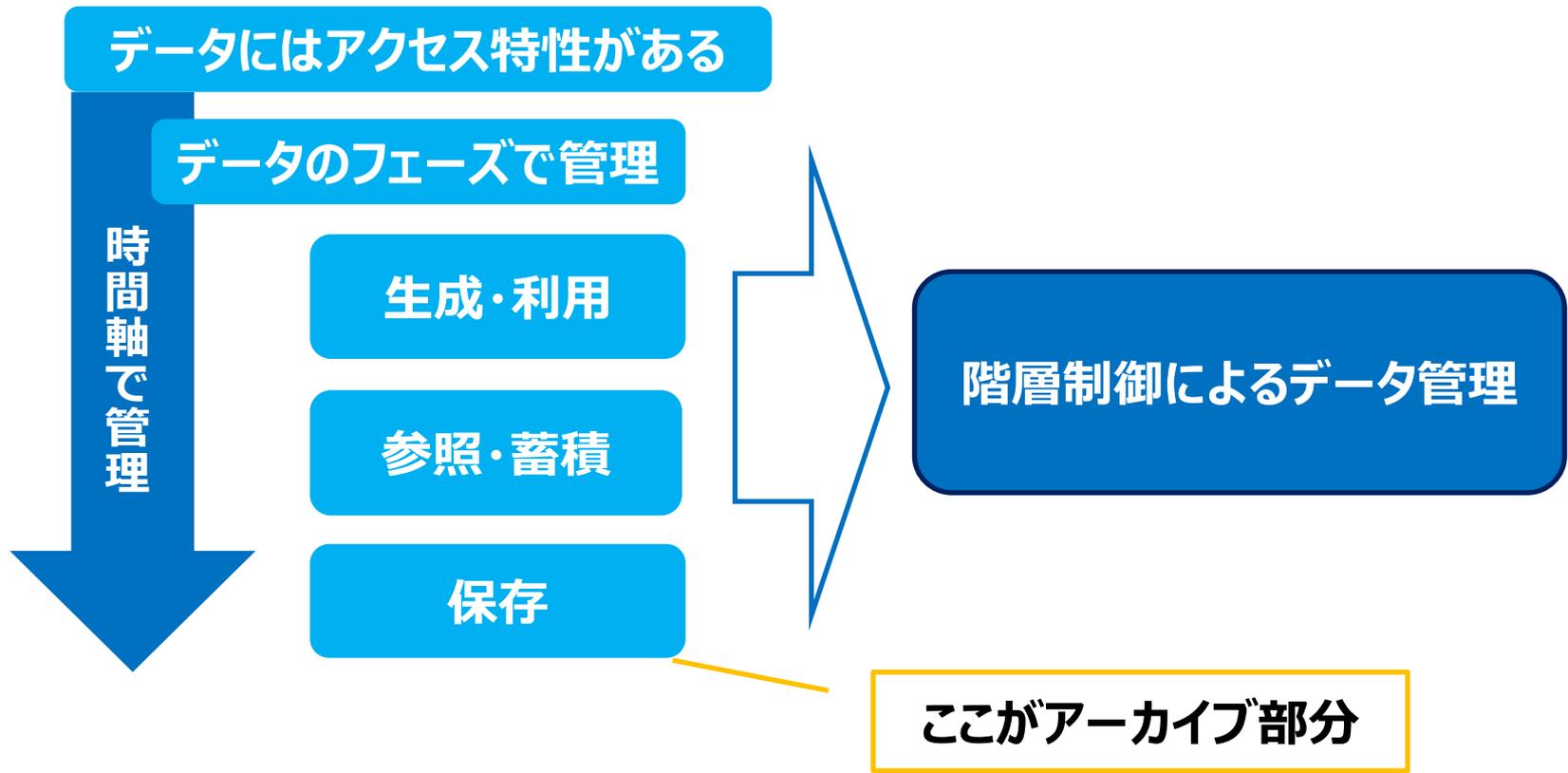
- データは、その種類によって「アーカイブに対する動機」に分類できる。
- アーカイブの動機に応じたアーカイブを検討する必要がある。

- 本活動の目的
- 背景
- データ特性の調査・検討
- **ILM調査・検討**
- メディア特性調査
- まとめ
- 活動メンバ

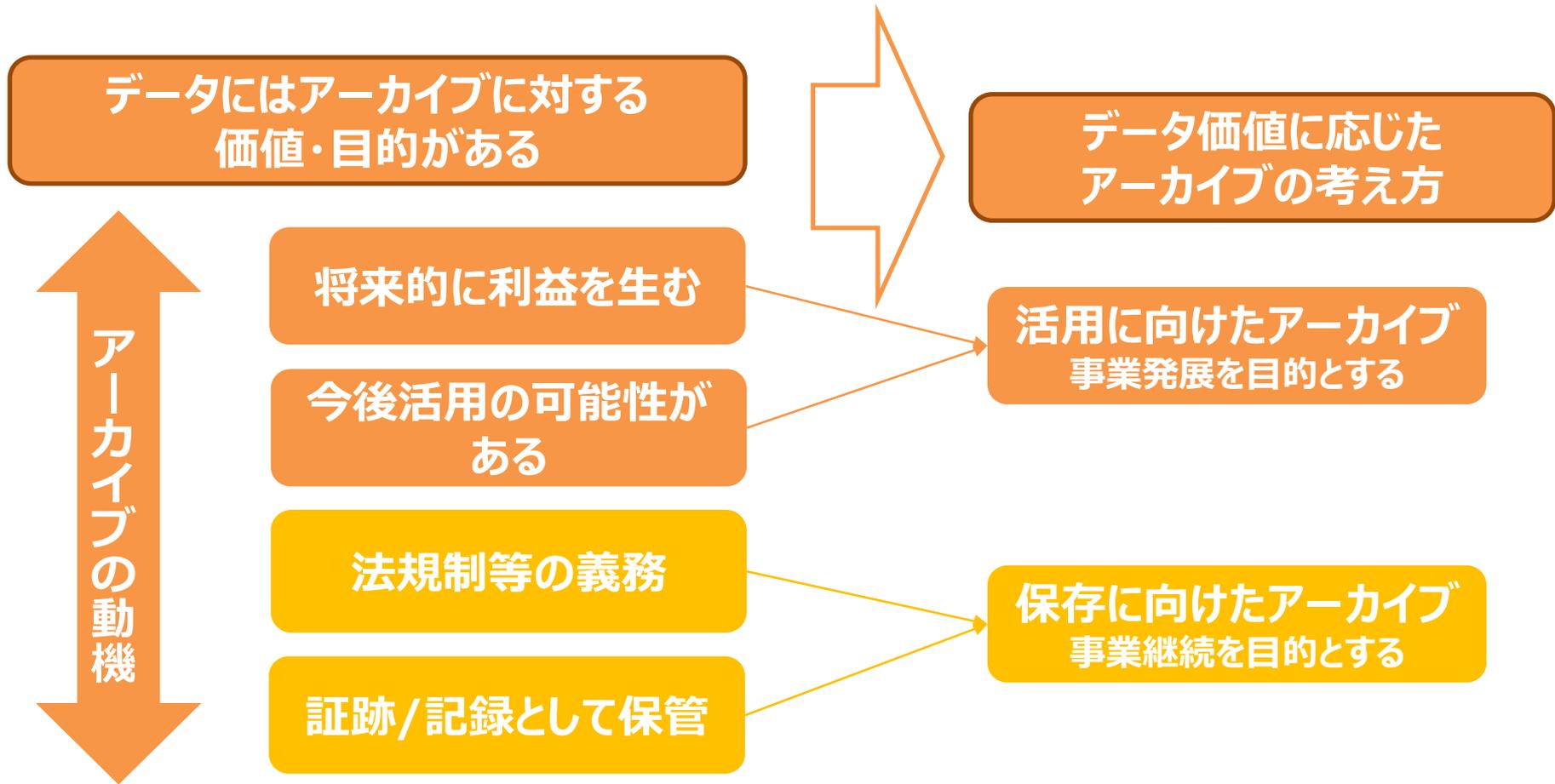


データのアクセス頻度をデータ価値と捉え、データ生成から時間と共にアクセス頻度・ビジネス価値が低下し最終的に破棄されるという生成から破棄までのモデル。この実現にはストレージ階層化の考えを当てはめるのが一般的。

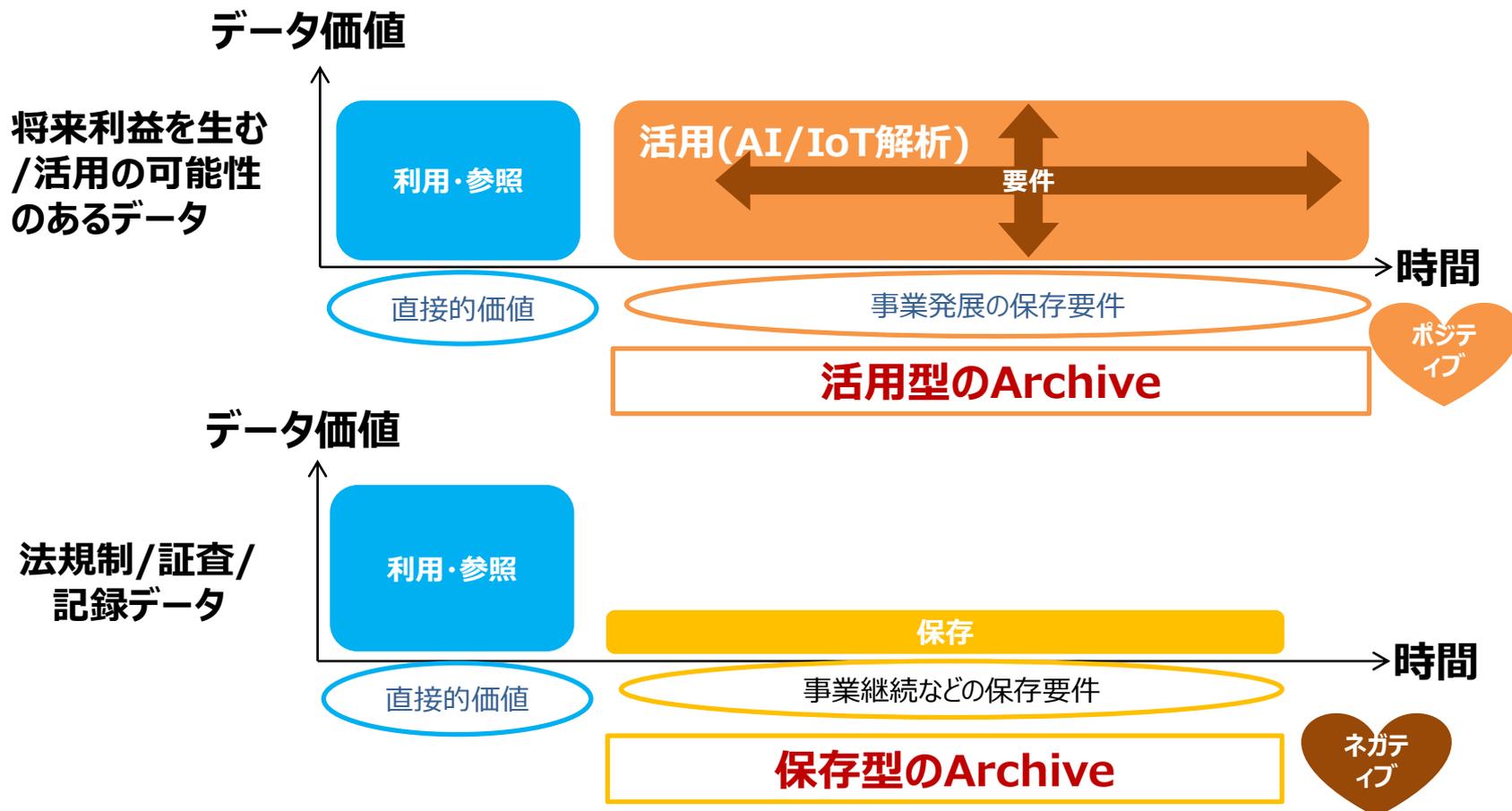
ILMの考え方 (information Lifecycle Management)



ILM調査・検討：データ価値とアーカイブ

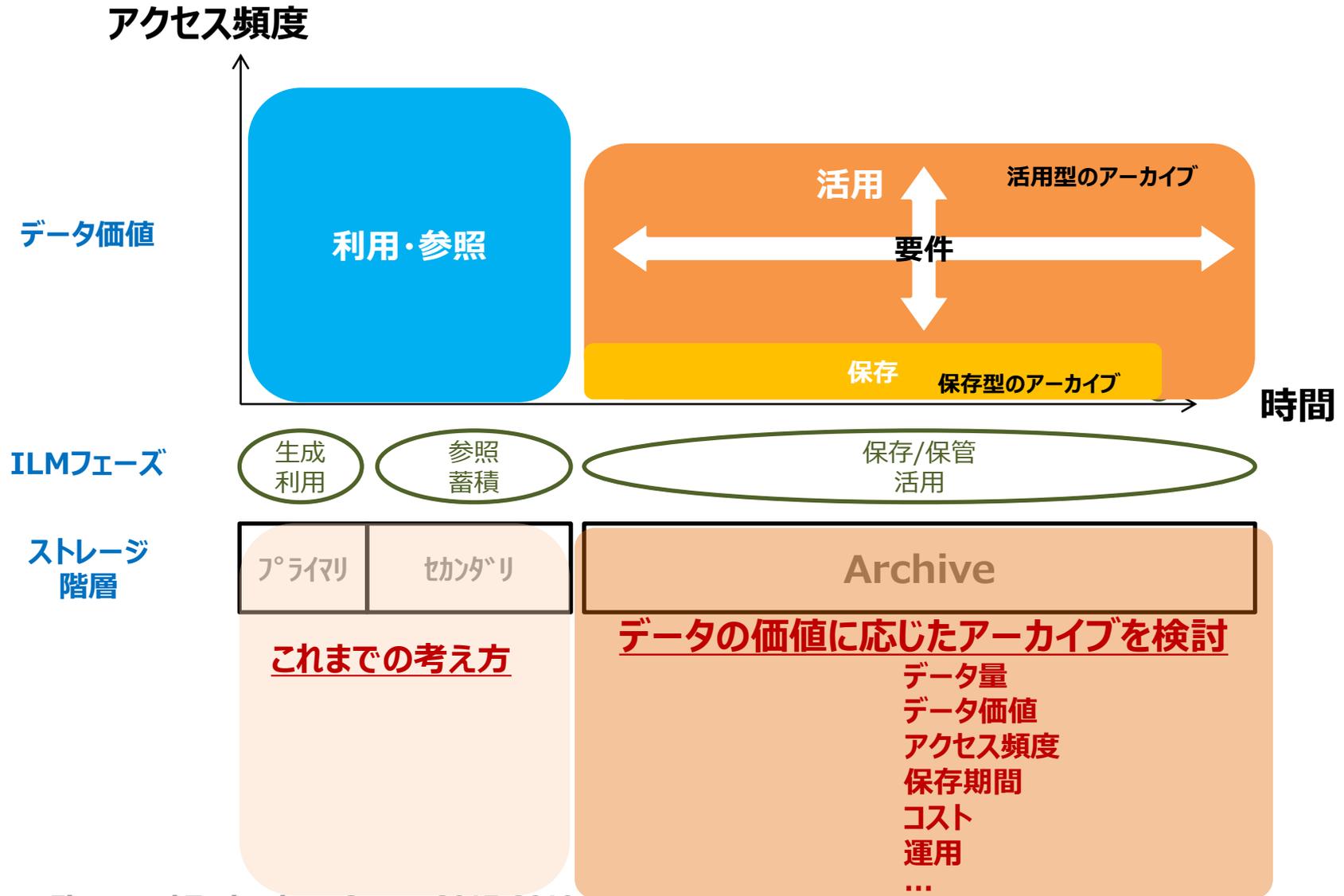


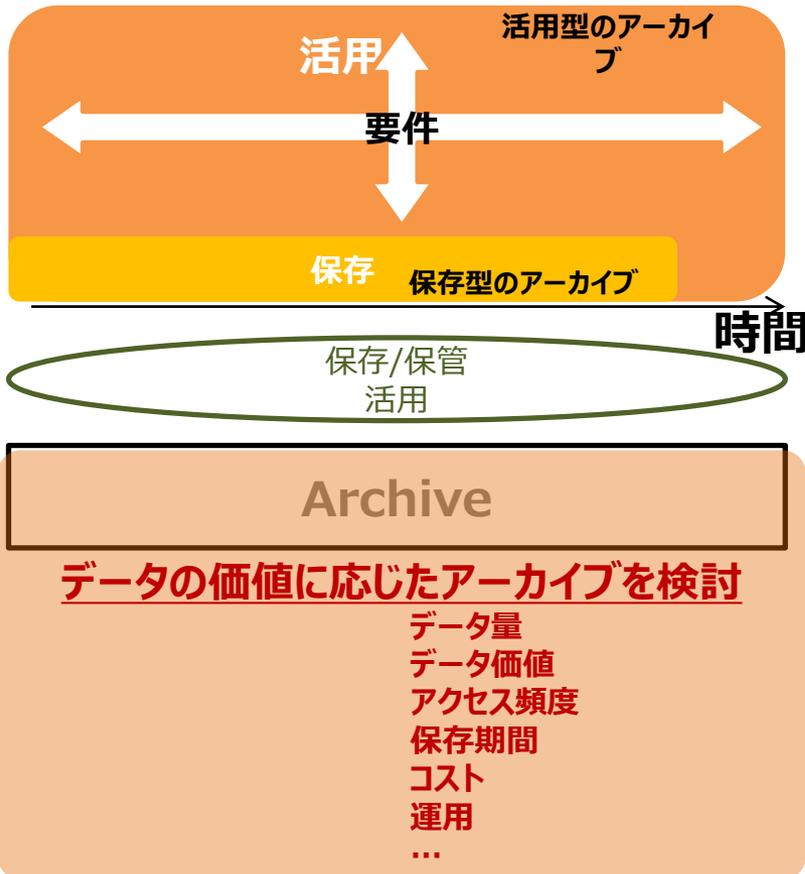
ILM調査・検討：データ価値とアーカイブ



データを保存する要件から、活用を目的とする「活用型のArchive」と保存を目的とする「保存型のArchive」に分けて考える。

ILM調査・検討：データ価値とアーカイブ





■ アーカイブの動機は、活用型のアーカイブと保存型のアーカイブに分類することが出来る。

■ 活用型/保存型に対して、データ価値に応じた要件を加味したアーカイブシステムを検討する必要がある。

- 本活動の目的
- 背景
- データ特性の調査・検討
- ILM調査・検討
- **メディア特性調査**
- まとめ
- 活動メンバ

アーカイブ要件の実現に適したストレージとは



- データ量
- データ価値
- アクセス
- 期間
- コスト
- 運用

...



メディアの特性：SSDの特長

■ 高速

サブミリ単位のレスポンス

■ 大容量

多層化/多値化/積層化等

■ 実効容量

重複排除/圧縮による容量拡大

■ NVMe

アクセス手法の変更による高速性



3D NANDフラッシュ技術の世代推移(予測)

開発発表年	ワード線の積層数	多値化方式(bit/セル)	シリコンダイ当たりの記憶容量	備考
2013年	24層	MLC(2bit/セル)	128Gbit	
2014年	32層	TLC(3bit/セル)	128Gbit	
2015年	48層	TLC(3bit/セル)	256Gbit	
2016年	64層	TLC(3bit/セル)	512Gbit	2ティア(32層×2)
2017年	96層	TLC(3bit/セル)	512Gbit	2ティア(48層×2)
	64層	QLC(4bit/セル)	1Tbit	2ティア(32層×2)
2018年	96層	QLC(4bit/セル)	1.33Tbit	2ティア(48層×2)
2019年?	128層/144層	QLC(4bit/セル)	2Tbit	2ティア(64層×2)/3ティア(48層×3)
2020年?	192層	QLC(4bit/セル)	2.66/3Tbit	3ティア/4ティア?
2021年?	256層	QLC(4bit/セル)	4Tbit	4ティア?
2022年?	384層	QLC(4bit/セル)	6Tbit	5ティア???
2023年?	512層	QLC(4bit/セル)	8Tbit	???

©2018 Copyright by Akira Fukuda. All rights reserved.

<https://pc.watch.impress.co.jp/docs/column/semicon/1139321.html>

メディアの特性：HDDの特長

■ 大容量

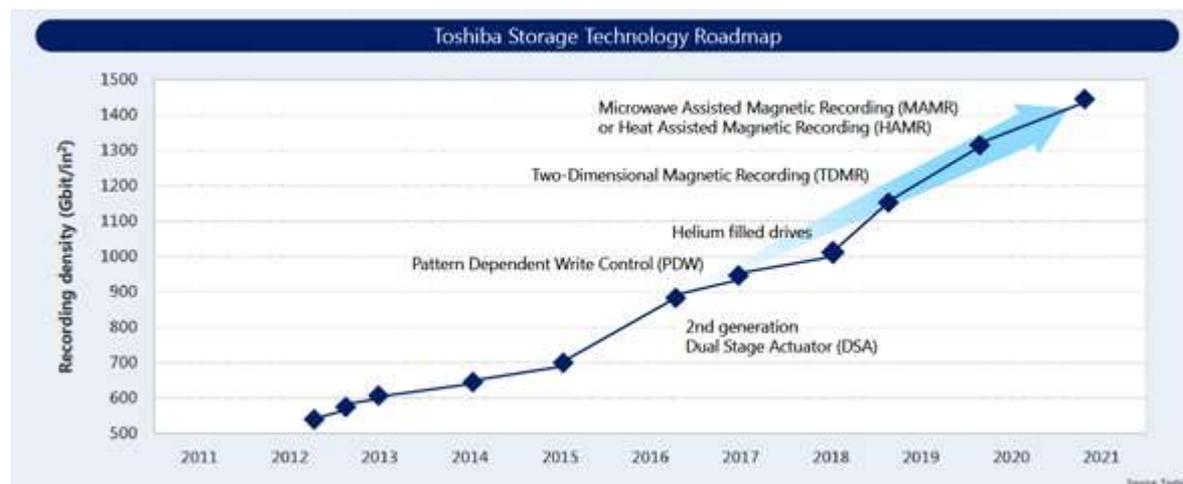
ヘリウム封止/ディスク枚数増加/CMR/TDMR

■ コスト

大容量化によるビットコスト優位(対SSD)

■ 高性能

ミリオーダのレスポンス



<http://eetimes.jp/ee/articles/1809/25/news012.html>

メディアの特性：TAPEの特長

低コスト

GB単価と低運用コスト

Gbit単価は各メディア比較で一番安価。
データの保管には電気代を必要とせず省エネ・低コスト



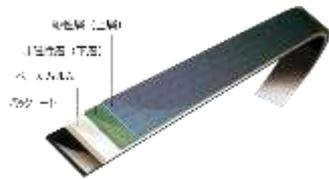
長期保管に最適

長期保管

JIS規格により標準化された長期保管手法を確立
磁気テープの寿命は50年以上(室内保管)
継続的に向上される素材の耐久性

長期供給性

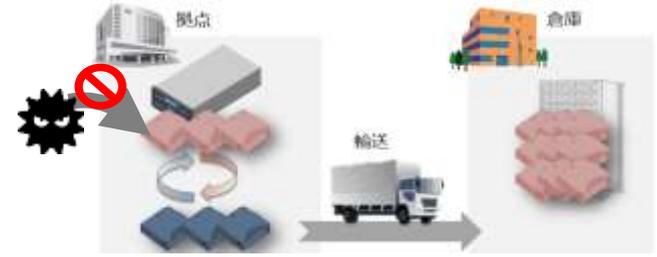
LTOはコンソーシアムにて標準化
ドライブ、メディアとも複数企業により開発、供給
⇒ ベンダロックインの心配なし



安心・安全

オフラインと可搬性

可搬性を活用し、災害対策も容易に実現
オフラインストレージのためランサムウェア等のウィルス対策も万全



大容量

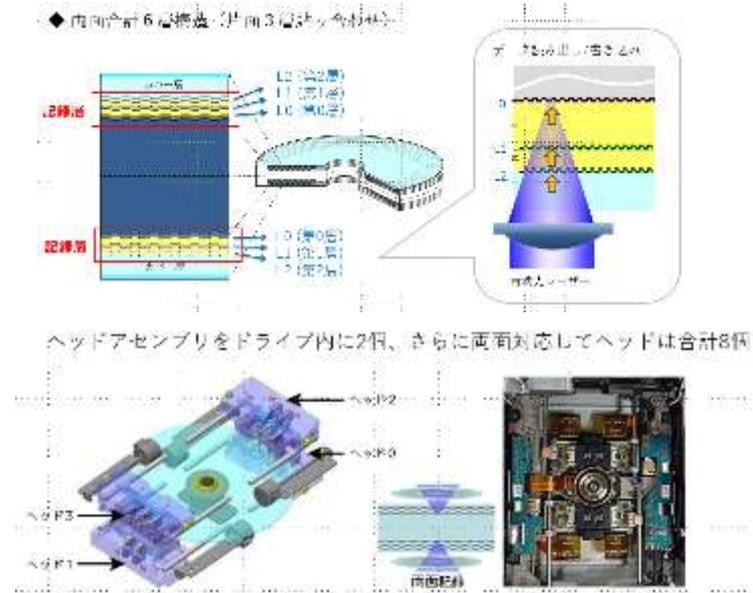
進化し続ける確かな技術

2017年10月に第8世代のLTO 8が登場。
LTO 12(192TB/巻)まで拡張されたロードマップ



300TB/巻以上を実現した実証実験が示す磐石の将来性

メディアの特性：光ディスクの特長



JDSF Open Tech Forum 2018 ソニービジネスソリューション株式会社
「データ管理に希望の光！100年の長期保存を実現する光ディスクと活用事例のご紹介」提供資料より抜粋引用

光ディスクの特長

- ◎データ長期保存性、環境耐性
(温度、湿度、磁気)
- ◎真正性 (Write Once)
- 可搬性
- ランダムアクセス性

最新アーカイブ用光ディスク技術動向

- ・大容量化
多層/両面記録、記録密度向上
- ・スループット向上
複数ピックアップ
- ・メディア耐久性/可搬性/保管性向上
記録膜材質改良、カートリッジ利用、
後方世代互換性

メディアの特性：クラウドの特長

クラウドの特長を一言で表すのは難しいが、概ね以下の特長がある

- **種類**：現在提供されているサービスとしては、オンライン/ニアライン/オフライン相当のものがあ、それぞれコストおよびアクセス性能が異なる
- **データ量**：SLAによるが原則として制限はない
- **アクセス**：オンライン・サービスはリアルタイム、ニアライン・サービスは数分、オフライン・サービスは数時間
- **期間**：SLAによるが明示的な制限はない。サービスの継続は保障されていない
- **コスト**：一般的にはオンプレより低いとされているが、オンライン・サービスの場合データのリトリブはかなり割高になるので注意を要する
- **運用**：運用負荷がほとんどないのは、クラウドの大きなメリット
- **セキュリティ**：データの機密性を担保するためにSLAで明示的に規定する必要あり
- **データ価値**：比較的アクセス頻度の高い活用型アーカイブよりは、寧ろ保存型アーカイブに向けたメディアと考えられる。特に運用負荷の観点から優位性がある。



- アーカイブシステムを構成するメディアには多くの種類があり、それぞれの特長がある。
- データ価値や要件を加味したアーカイブシステム、メディアを検討する必要がある。

- 本活動の目的
- 背景
- データ特性の調査・検討
- ILM調査・検討
- メディア特性調査
- **まとめ**
- 活動メンバ

アーカイブシステムを考える時のポイントは3つ

- **データ種類とデータ価値**
=> **アーカイブの動機**
- **ライフサイクルマネジメントの考え方**
=> **保存型と活用型のアーカイブ**
- **ストレージ・メディアの特長**
=> **要件に応じた階層制御システム**

今後の活動として、これらの情報と組み合わせに対する情報を提供したいと考えています

- 本活動の目的
- 背景
- データ特性の調査・検討
- ILM調査・検討
- メディア特性調査
- まとめ
- **活動メンバ**

■ ストレージ要素技術(SET)部会 2018年 参加社

(五十音順)

イーグローバレッジ株式会社

AR アドバンステクノロジー株式会社

株式会社MO

ソニーイメージングプロダクツ&ソリューションズ株式会社

東芝デジタルソリューションズ株式会社

日本電気株式会社

株式会社日立情報通信エンジニアリング

富士フイルム株式会社

富士通株式会社

