

TOSHIBA



オープンソースカンファレンス 2020 Tokyo/Spring

多様性時代のDB選択

東芝デジタルソリューションズ株式会社

栗田 雅芳

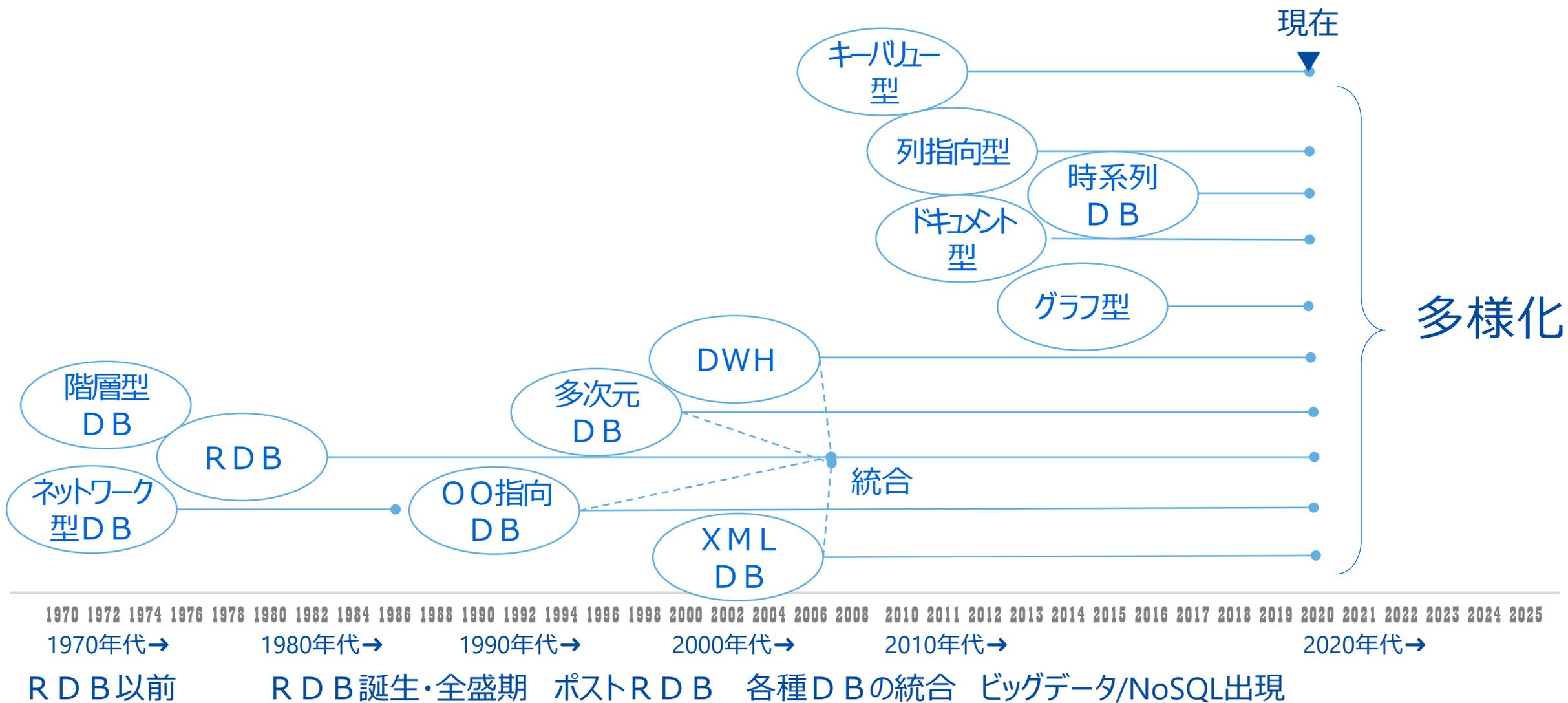
2020/2/21

多様性とは？

幅広く、性質の異なる群が存在すること。

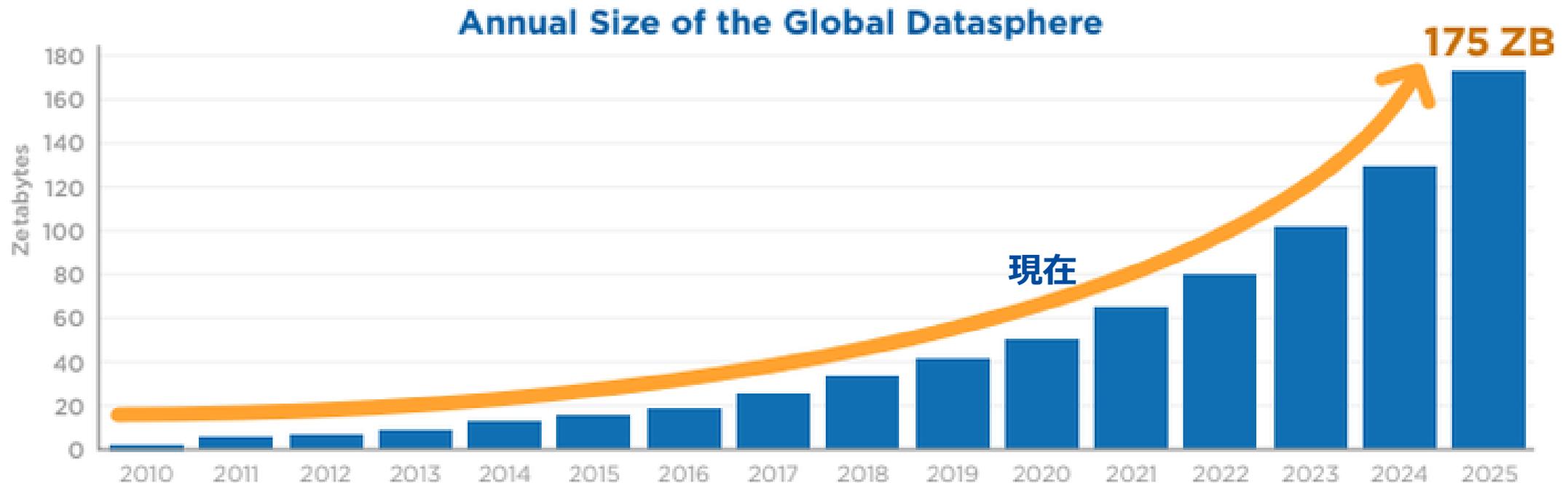
性質に類似性のある群が形成される点が特徴で、単純に「いろいろある」こととは異なる。

DBの50年の歴史



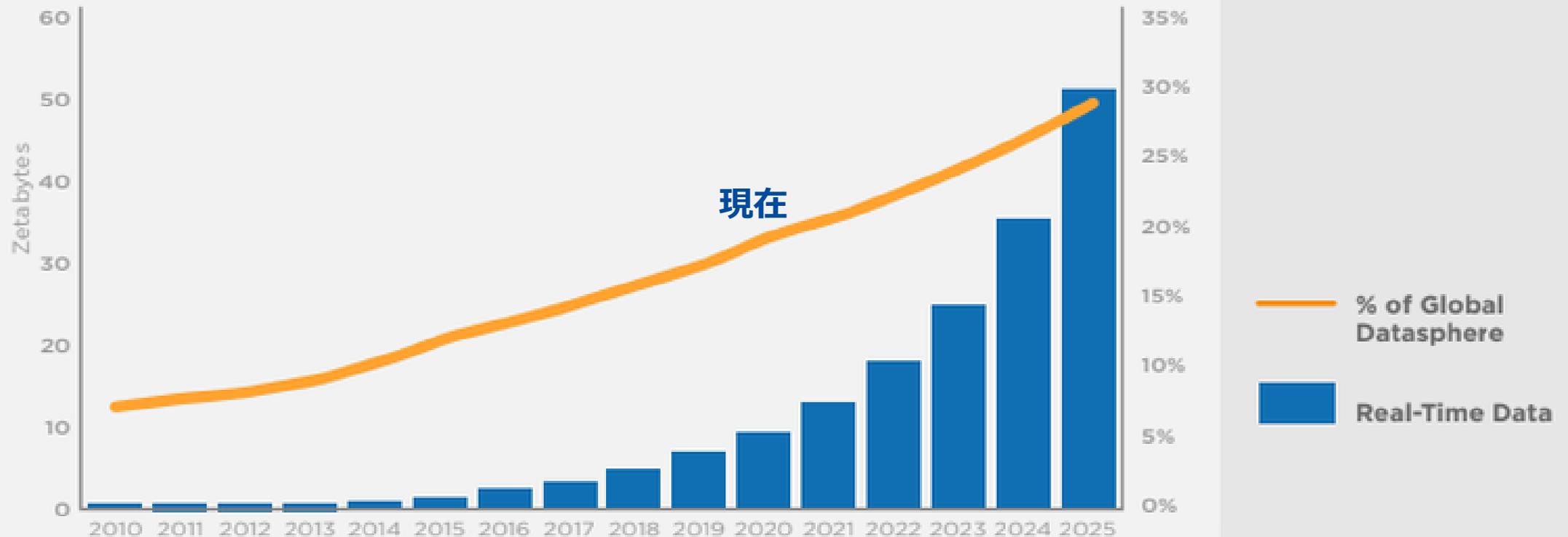
ターゲットを広めるため多機能化の傾向があり、
表面的に機能・非機能ともに、類似性が高い
場合が多いので、DB 選択が困難（？）

2025年には世界のデータ量は、
2017年の23ゼタバイトから175ゼタバイトへ

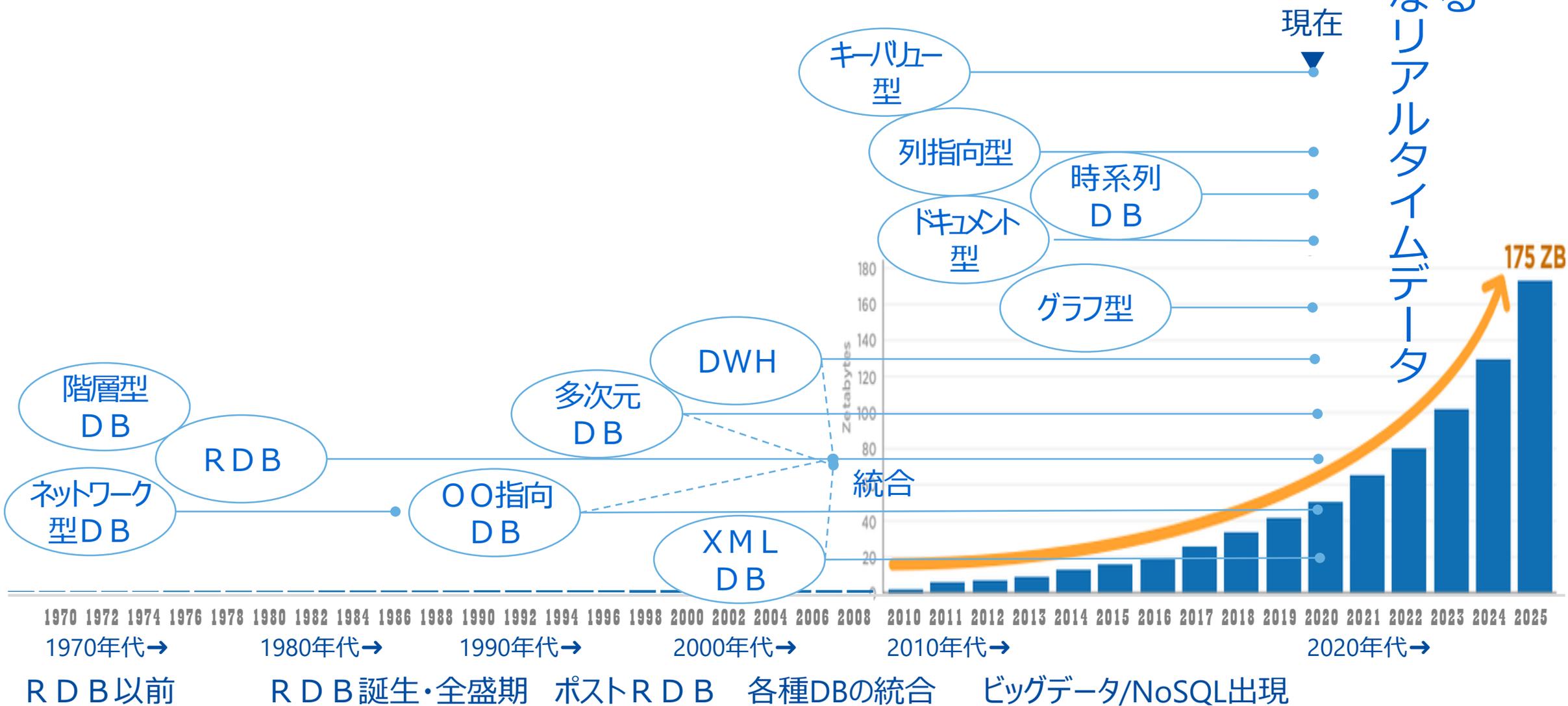


2025年には世界で生成されるデータの約30%がリアルタイムデータに

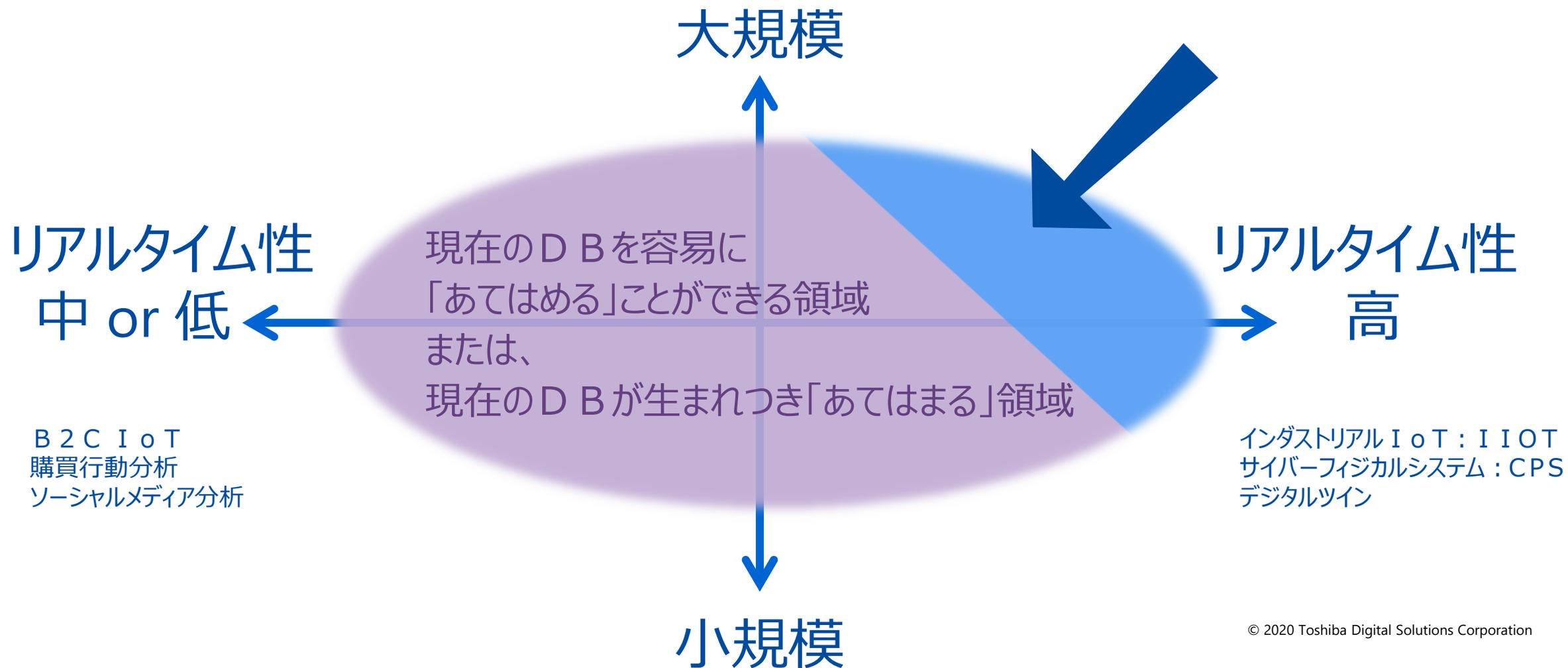
How Much of Global Datasphere is Real-Time?



DBの50年の歴史



更なる膨大なリアルタイムデータ



	B2C IoT 購買行動分析 ソーシャルメディア分析	インダストリアルIoT: IIoT サイバーフィジカルシステム: CPS デジタルツイン
データ	サンプル (一部欠損は許容範囲)	すべてのデータが必須(欠損不可)
分析手法	統計手法	小さな異常も見逃さない
分析正確性	低精度でもよい	高精度がとめられる
リアルタイム性	中～低 バッチでもよい	高
システム信頼性	中程度のもも十分	高信頼が必須

更なる膨大なリアルタイムデータへの領域

実世界のデータをサイバー空間にリアルタイムに再現し、アナリティクス。その結果を現実世界へフィードバックする

これを実装するために「あてはめる」のではなく、**生まれつき「あてはまる」DBも必要**



GridDB

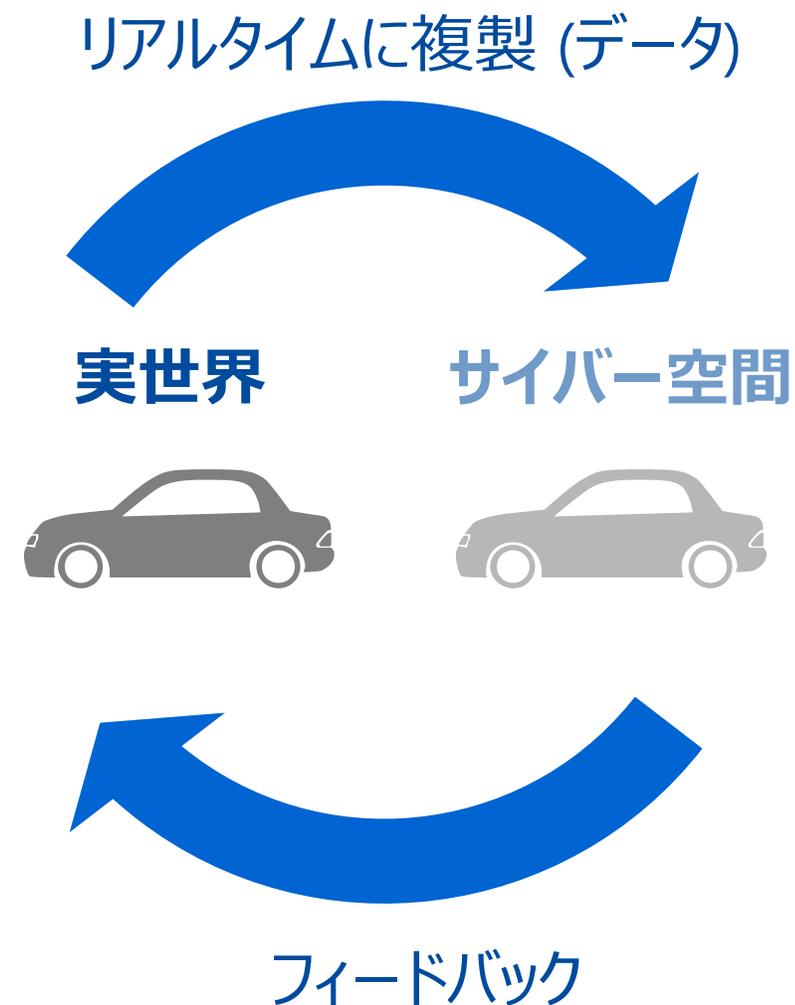
GridDB とは

- 実世界のデータをサイバー空間でリアルタイムに再現することをターゲットしてデザインした DB
- スケールアウトとスケールアップのベストミックスで、ペタバイト級のデータ管理を実現
- NoSQL / SQL (キー・バリュー) のデュアル A P I

次世代車両管理システム

運転情報をリアルタイムに集めて、
保険会社に提供。

保険会社は保険料に反映する
とともに、リアルタイムにドライバー
に運転技術について警告



HDD製造会社 品質管理システムの再構築

製造レコードを全件保管し、ありのままの状態をリアルタイムに把握を目指す。

- データ蓄積量：1.9PB / 5年
- 登録データ量：267 GB / 日

DBアプライアンス以上の性能を標準的なIAサーバで実現

多様性の時代

幅広く性質が異なるが類似性のある D B
が存在する

お互いの生まれつきの特性を認め、
適所適材に「あてはまる」D Bを採用



