Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

A Data Management Infrastructure for Intelligent Systems

Antonello Ceravola, Frank Joublin, Heiko Wersing, Stephan Hasler, Behzad Dariush, Yi-Ting Chen

2018

Preprint:

This is an accepted article published in Japanese Society of Artificial Intelligence (JSAI). The final authenticated version is available online at: https://doi.org/[DOI not available]

特　集「AI 計算資源」

インテリジェントシステム用データ管理インフラストラクチャ

A Data Management Infrastructure for Intelligent Systems

Antonello Ceravola　Honda Research Institute Europe GmbH
Antonello.Ceravola@honda-ri.de, http://www.honda-ri.de/

Frank Joublin　（同　上）
Frank.Joublin@honda-ri.de, http://www.honda-ri.de/

Heiko Wersing　（同　上）
Heiko.Wersing@honda-ri.de, http://www.honda-ri.de/

Stephan Hasler　（同　上）
Stephan.Hasler@honda-ri.de, http://www.honda-ri.de/

Behzad Dariush　Honda Research Institute USA, Inc.
dariush@hra.com, http://usa.honda-ri.com/

Yi-Ting Chen　（同　上）
ychen@hra.com, http://usa.honda-ri.com/

Keywords: big-data, data recordings, multi-sensor systems, data management, intelligent systems, AI, automotive, robotics.

1. は　じ　め　に

自動車分野では，ここ数十年，エレクトロニクスやコンピュータに基づく機能への注目がしだいに高まってきている．この傾向は，新しい機能を提供するカスタム電子制御ユニット（ECU）を車内で使用したことから始まり，しだいに，より汎用性の高いコンピュータシステムの採用へと移行していった．これにより車内で実現できる機能の範囲が拡大し，インテリジェントシステムやロボットシステムもその機能に含まれるようになった．ECU を用いていた時代は，実装できる機能は単純なものに限定され，少数の限られた（毎秒数キロバイト範囲の）センサしか使うことができなかった．インテリジェントシステムやロボットシステムを導入すると，より複雑な機能が必要であり，（毎秒数メガバイト以上の）高速広帯域の信号送信を伴う多数のセンサを扱う必要がある．さらに，最近のインテリジェントシステムやロボットシステムでは，大量の学習データにアクセスを要するディープラーニングなどの AI 技術が採用されている．このようなインテリジェントシステムの研究，開発，テストにおけるパラダイムシフトによって，ビッグデータといってよい量のセンサデータ処理が必要である．このため，センサデータ収録フェーズを車両機能開発と一元的に管理することは難しい．この二つのパッケージは，いずれも複雑なものとなるため，データ管理パッケージと車両機能パッケージは，それぞれ別のプロジェクトやコンセプトとして，必然的に分離することになった．こうした観点から，インテリジェントシステムやロボットシステムを構築するプロジェクトに導入できる，収録センサデータ用のビッグデータ処理をサポートしたデータ管理インフラストラクチャの設計・構築プロジェクトを立ち上げることにした．本稿では，本データ管理インフラストラクチャを初めてリリースするにあたって，そのプロセスと直面した課題について述べる．

2. データの取扱い

複数の大容量センサデータストリームを使用するインテリジェントシステムを開発する際に直面する重要な課題の一つは，さまざまなストリームデータを，データサイズ，データ構造およびパッケージングといった観点で，アクセス可能なフォーマットにすることである．複数のセンサをリアルタイムで管理するには，十分な帯域幅，適当なデータ収録ソフトウェアシステム，高性能のコンピュータハードウェアおよび記憶装置をもつコンピュータシステムが必要である．ここでは，頻度，データ構造および（情報の伝送バイト数換算での）サイズが異なる複数のデータソースを使用するシステムを念頭に置いている．インテリジェントシステムを開発する際には，そのようなデータが以下のように異なる方法で供給され得る．

- オンラインデータ：このデータはセンサから実行中の開発・テストシステムへリアルタイムに供給される．
- オフラインデータ：このデータはいったんリアルタイムで記憶装置に保存され，保存されたデータは，次のフェーズで，再生機能を有する開発・テストシステムで使用される．
- 合成センサデータ：このデータは仮想的な環境で生成され，開発・テストシステムに供給される．

システムの開発・テストでリアルタイム処理やインタラクションが重要となる場合には，オンラインデータが必須となる．一方，オンラインデータでは，特定のシーケンスのセンサデータを再生する，システムを異なるサイクルで開始・停止する，テスト実施後に停止して不具合を修正する，1ステップごとに実行する，といったオペレーションを実現することは難しい．オフラインデータを使用することにより，こうした制約のほとんどが解消できる．この場合，データを一度収録しておいて，あとからシステムで再生することになる．オフラインデータには，収録に大容量の記憶装置が必要となることがある，リアルタイムでインタラクションができない，センサデータのフォーマットが変わるとこれまでに行った収録が無効になることがある，といった問題がある．最適なソリューションは，シミュレーション環境を使用してセンサをシミュレーションし，合成データストリームを生成することであろう．シミュレータが所定のセンサ品質を提供できれば，オンラインおよびオフラインの利点を併せもつことができる．しかし，合成データには，シミュレーションの際に十分な品質を確保する必要があること，静的，動的，およびこれらが相互に作用するシナリオのシミュレーションを扱おうとすると処理が複雑になる，などの課題がある．

本稿では，オフラインデータを使用する場合を分析するとともに，大量の収録データを処理するために我々が構築したデータ管理システムについて述べる．

説明のための事例として，先進運転支援システム（ADAS）や自律走行（AD）などのインテリジェントシステムを実現するプロジェクトを取り上げる．この車には以下のセンサが備わっている．

- カメラ10台（白黒およびカラーカメラ）
- レーダセンサ6個
- レーザセンサ6個
- GPSセンサ1個
- IMU 1ユニット
- 車からのCANネットワークストリーム数本

この場合，約30本の独立したデータストリームを毎秒約200 Mbで収録することになる．この場合，以下の課題がある．

- 計算用インフラストラクチャ：（USB，イーサネット，CANなどを介して）全ストリームをリアルタイムで受信できる計算用インフラストラクチャの構築
- センサの同期：センサから収録システムに送られる情報の各パケット・フレームにラベリングするための同期プロトコルの実装
- 収録システム：複数のストリームを並列で受信し，保存できる収録ソフトウェアシステムの構築
- データ転送：収録プラットフォームからオフィスネットワークにデータを転送するためのプロセスの編成
- 再生システム：収録されたデータを収録時と同様に再生できるシステムの構築

3. 技 術 の 現 状

センサ収録管理をサポートする現行のデータ管理システムを分析した結果，標準化共同体ASAM（自動化システムと測定システムの国際標準化団体）[ASAM]のプラットフォームが自動車分野で最も使用されているプラットフォームの一つであることがわかった．この共同体は約30の企業（OEM会社，サプライヤ，ソフトウェアベンダ）で構成されている．そのほとんどが欧州および米国の会社であり，車内測定データのための標準を定義することを目指している．その主な標準がODS（オープンデータサービス）である．これは，標準化されたサーバ記憶装置用に測定データ（車両データストリーム）を統一的に表現するようモデル化を行うものである（図1参照）．その目的の一つが，会社内あるいは会社間でのデータ交換を可能にすることである[ODS]．

この共同体は，提案した標準が採用されることを促進するために，仕様の一部といくつかのツールをオープンソースとして公開している．我々が調査した時点で，一つのサーバソリューションと一つの構成可能なグラフィックユーザインタフェース，複数の収録

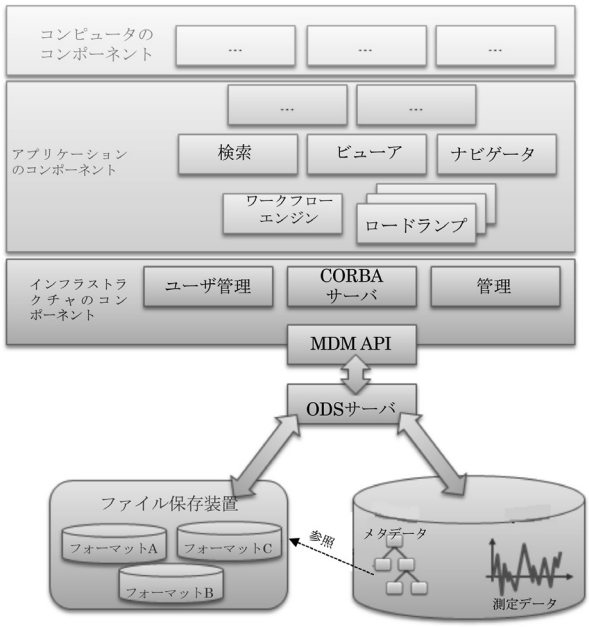

図 1　ASAM ODS のアーキテクチャ

処理モジュールが備わった ODS が入手可能であった [OPENMDM]．しかし，これは低帯域幅のデータ系列用に実装されたもので，我々のビッグデータの要件を満足するものではなかった．後年，ビッグデータの使用が自動車会社の間に広まったので，共同体はこの要件も考慮するようになった．もう一つのデータ管理システムとして，米国の複数の大学と運輸機関により実施された InSight SHRP2 NDS（Strategic Highway Research Program on Natural Driving Study；自然な運転行動に関する戦略的道路研究計画）[SHRP] という安全運転研究 [Campbell 12] に関連して構築されたシステムがあげられる．この研究の目的は，運転状況（運転手，車両，道路および環境因子）が運転手の挙動にどのような影響を及ぼすかを分析することである．このシステムでは，運転手に収録システムとして一連のセンサ（カメラ，GPS）を車両に取り付けてもらい，データを収集する．収集したデータは中央データベースに集約し，Web インタフェースを介して入手できるようになっている [INSIGHT]．大学はこのインタフェースを介してデータにアクセスし，データ解析を行う特定のセグメントの可視化，アノテーション，検索などを行うことができる．

また，別のデータ管理システムに関連した類似のイニシアティブが，道路の安全性を主目的とした非営利団体である CEESAR により実施されている．この非営利団体は，InSight と同様のワークフローを取り扱う，UDRIVE 共同体（European Naturalistic Driving Study；欧州自然運転研究）[ROB 14, UDRIVE] と関連しており，SALSA と名付けられたシステムを構築した．この二つのシステムはデータ管理システム自体を提供することを目的としたものではなく，第三者に収録されたデータへのアクセスを提供するものである．これらのシステムは各プロジェクトのデータ用にカスタマイズされており，第三者がホストしたり，データを投入したりすることを意図したものではない．それでもなお，これらのシステムは，我々が目指すシステムの概念を十分に裏付けてくれるものである．

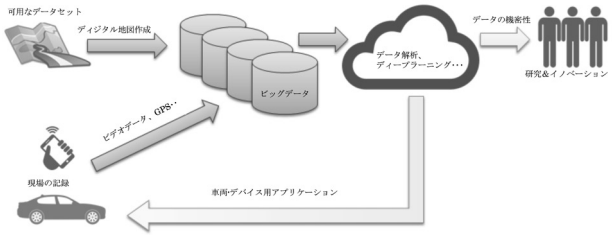

図 2　クラウド LSVA ターゲットの概要

アノテーションに焦点を当てたデータ管理システム（図 2 参照）の実現を目指したまた別の欧州イニシアティブが Cloud LSVA [CLSVA]，すなわちクラウドベースの大規模ビデオアノテーションプラットフォームである．この欧州プロジェクトはイノベーションプログラム Horizon 2020 の一環として立ち上げられたもので，リアルタイムでの車両センサデータ収録からクラウドサーバへのデータストリーム送信，アノテーションプロセス，インデキシングおよび収録データの検索までを考慮したものである．この欧州プロジェクトの目的は，センサデータとメタ情報を表現するための一連の標準を策定するとともに，さまざまな会社から送られてくる収録データを保存したり，可視化したり，アノテーションしたりすることが可能なクラウドサーバを構築することである．このプロジェクトは，車からクラウドへのデータ直接送信ができることを除けば，我々と多くの要件を共有している．この部分に関しては，我々は全収録セッションをクラウドではなく，社内の中央ファイルサーバにコピーすることを目指している．

より一般的な方法で収録データを処理するよう設計された商用ソリューションが XCube 社 [XCUBE] の DSSC（Distributed Storage and Simulation Cluster；分散型ストレージ & シミュレーションクラスタ）[XCUBEPD] である．DSSC は（地理的にも）分散型の記憶装置であるとともに，データを，それが作成されたノードから最も近くてアクセス可能なノードに保持するという原理に基づいた計算システムでもある．データ計算時には，（一連の仮想マシンをホストする）目的のデータを含む各ノードにプログラムが配布される．仮想マシンを使用したコードモビリティ原理 [Fuggetta 98] は，（我々の場合と同様に）ビッグデータを処理するデータ管理システムに特に適している．このアプローチは標準的なコンピュータ構成を使用する場合には適しているが，我々のプログラムは，さまざまな特定のライブラリやハードウェア構成に依存することがあるので，我々のアプリケーションに用いることは難しい．さらに DSSC のアーキテクチャではデータやプログラムの強力なカプセル化

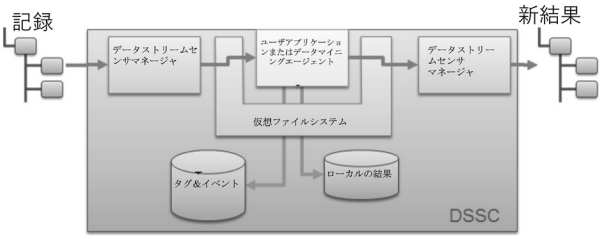

図3 DSSCノードのアーキテクチャ

が行われるので，ユーザアプリケーションの開発やテストが面倒になる（図3参照）．

このようなデータ管理システムの最も重要な機能の一つが，収録データ内にある関連するイベントやオブジェクトを検索すること，およびその検索結果を使用して，インテリジェントシステム開発用の検証あるいはテストデータを収録データセグメントとして抽出することである．そのような検索機能は，収録データの中の，特にアノテーションあるいはラベリングされたデータストリームから検索可能な情報を収集する処理に基づいている（この処理は，収録データをシステムにインポートする際に実行される）．

機械学習やコンピュータビジョン分野での最近のブレークスルーにより，画像認識，自己位置推定，動作計画，意思決定といった機能の性能が加速度的に向上した．特に，ディープニューラルネットワークは，セマンティックラベルを伴う大規模データセットが必要となるさまざまな画像研究分野で，一般的に有用な表現を学習できることが実証されている．自動運転では，特に，画像から周囲の交通参加者に対する自車の相対的位置，方向，速度などの意味のあるアフォーダンス指標へのマッピングを学習する際に，正解ラベルを有する非常に大規模で慎重に作成されたデータベースが重要となる．

セマンティックラベリングされた大規模なデータセットの可用性は，プロセスに多大な時間や費用を要するため，ある程度制限されてきた．こういった要因で新たな大規模データセットの開発は遅れているものの，研究コミュニティとしては，ImageNet [Deng 09]，NYU-DepthV2 [Silberman 12]，PASCAL-Context Dataset [Mottaghi 14] や Microsoft COCO [Lin 14] といったさまざまなデータセット開発のための投資が続けられている．これらのデータセットは確かに屋内シーンのセマンティックセグメンテーションを行ったり，共通の対象物を認識したりする機能の性能向上を加速するものである．しかし，自動運転に関係するタスクなどのより特有なタスクには適していない．

自動車分野で評判の良いセマンティックセグメンテーション用ベンチマークデータセットとしてCAMVID [Brostow 09] や，より新しいCityscapes [Cordts 16] があげられる．後者は，画素レベルやインスタンスレベルのセマンティックラベリングの学習用サンプルとして導入されたデータセットである．Cityscapesは，欧州のさまざまな都市で収集された画素ベースのラベルを含んだ5 000個の画像とおおざっぱなラベルを含んだ20 000個の画像で構成されている．このデータセットは，データセットのサイズ，アノテーションの豊富さ，シーンの多様性および複雑さの点でそれ以前に構築されたデータセットを超えるものではあるが，このデータセットで学習したアルゴリズムは，季節や照明，天候，交通条件，シーン構造，センサ特性などの変化を含むさまざまな交通シーンのドメインのテストデータに適用できるほど十分に一般化されたものではない．

この種のラベリングの規模を十分に拡張するには，さまざまなタスクや画像ドメインにおいて，要求される多様性を満たすよう十分な数の画像を収録する必要がある．このためには，莫大な経済的投資が必要である．この問題に取り組むため，さまざまな条件で実際の都市景観をシミュレーションし，適切なアノテーションを含んだ合成画像を使用する代替法が有望であろう [Ros 16]．特にSYNTHIAデータセットは，運転シーンの状況下でのセマンティックセグメンテーションや関連するシーン理解の問題を支援する目的で開発されたものである．SYNTHIAは仮想都市からレンダリングされた一群のフォトリアリスティックなフレームで構成され，13クラスの厳密な画素レベルのセマンティックアノテーションを含んでいる．学習段階でSYNTHIAを，公開されている現実の都市データセットと組み合わせることにより，セマンティックセグメンテーションタスクの性能が著しく向上することが実証されている．

研究者は，手動アノテーションの代わりに，セマンティックラベリングされた大規模なデータセットを自動的に生成する方法を探求している．そのアプローチを用いれば，大規模ビデオデータベース用二次元ラベルを街路景観から作成することができ，結果的にセマンティックアノテーションや瞬時画像のアノテーションを生成する自動あるいは半自動のプロセスが提供できる．

4. 動　　機

ADAS/AD分野のインテリジェントシステム研究用として，複数センサを備えたテスト車数台の新規開発に伴って，データ管理システムを導入した．これらの車には複数の高解像度カメラとLiDARセンサが搭載されており，以前のテスト車よりもはるかに多くのデータが生成される．さらに，従来の実験ではデータをネットワーク上の複数の場所にコピーして管理していたが，この手法は利用できない．グローバルなネットワーク全体にまたがって，効率的にデータを扱い，格納し，共有する必要がある．データ管理システムの初期設計における要求仕様には，以下の項目があげられる．

- （1収録当たり数百Gbの）ビッグデータを処理できるシステム

- 単一のリージョンおよび複数のリージョンをサポートする機能
- グラフィックインタフェースやコマンドラインを介してのアクセス，およびプログラムによるアクセスが可能なシステムおよびデータ
- ビデオデータ，点群データおよび時系列データを扱う能力
- ロボットや車などさまざまなプラットフォームから送られてくるさまざまな収録フォーマットのサポート
- 複数センサ情報の統合するアルゴリズムを利用できるようにするため，データストリームは時間的に同期していること
- マニュアルまたは自動でのイベントへのアノテーション付与と，階層化された分類によるオブジェクトクラスがサポートされていること
- アノテーションはすべてサイトごとにあるいはサイト間で共有できるように一括でインデックスが付けられて検索可能であること
- 収録したデータを，まとめて，あるいは一部を指定してエクスポート可能であること

我々のこれまでの経験では，収録データをどこに保存し，どのように構造化，可視化，アクセスするか，といったインテリジェントシステムに使用される収録データの管理に関する取決めは，各プロジェクトの裁量に任されていることが多かった．この仕組みは，収録データが少量である，他プロジェクトとの交流がほとんどない，時間が経っても再利用する収録データがそれほど多くないといった場合は，うまくいっていた．しかし，収録データサイズや，プロジェクト内あるいはプロジェクト間でのデータの再利用頻度が増大した途端，そのような自由は逆に足かせとなる．例えば，収録データをどこに保存するかについては特に方針がないため，一部の収録データのコピーをネットワーク上のさまざまな場所に保存するのがこれまでの慣行であることがわかった．これにより，一度しか使用されておらず，今後も使用される見込みのない収録データのコピーが大量に存在していた．さらに，もう一つ持ち上がった問題は，構造およびフォーマットの両面で大きく異なる収録データがネットワーク上に拡散してしまっていたということである．収録データの構造やフォーマットは以下の点から異なることがある．

- 収録データの構造：どのようなファイルやディレクトリを使用して収録データ情報を構造化あるいは配布するか．
- モジュール性：収録データの中には，（すべてのセンサストリームを含む）単一のメインファイルを用いる場合，センサストリームごとに一つのファイルを使用する場合，センサフレームやパケットごとに一つのファイルを割り当てる場合など，さまざまな場合がある．
- 時間同期：収録データの中には時間的に同期されていることを暗黙的に仮定するものもあれば，ファイル名，インデックスファイルあるいはセンサデータの中に同期情報をエンコードするものもある．
- ストリームのフォーマット：ストリームの中には（ASCII, json などの）人が読むことができるフォーマットで収録されるものもあれば，アクセスに特定のライブラリを要するバイナリエンコーディングが行われるものもある．

このように大きな違いがあるので，単一の収録フォーマットへの共通化を目指すことは難しい．場合によっては，収録フォーマットが収録プラットフォームの実際のストリーミングとリンクされていることがあるので，この問題はさらに難しくなる．例えば，ビデオカメラセンサ用の車両プラットフォーム環境（センサネットワーク–コンピュータ）で，カメラから取得したデータに対して，1 フレームに 1 枚ずつ（圧縮または生の）画像ファイルを作成・保存することは困難を伴う．このように高い頻度でファイルを作成すると，ファイルシステムのオーバヘッドによってコマ落ちが大量に発生してしまうためである．一方，すべてのフレームを単一のファイルに保存すると raw フォーマットでもコマ落ちは発生しない．我々はこのような制約から，オリジナルの収録データをさまざまなフォーマットで保存することができるデータ管理インフラストラクチャを検討しなければならなかった．

データ管理インフラストラクチャのアーキテクチャを決める際には，以下に例示する項目を含めて，さまざまな要件を考慮しなければならなかった．

- さまざまな場所や部門のニーズに合ったインフラストラクチャであること．
- 各部門が収録データにすばやくアクセスできること．
- 各部門は主にローカルに保存した収録データで作業するが，収録データの共有は可能であること．
- 各部門，場合によっては各プロジェクトで異なる収録フォーマットが使用されることがあること．
- データ管理の全体的なワークフローはすべての部門およびプロジェクトでほぼ同じであること．

これまでデータ管理に自由度を許容していたため，それぞれのチームが，それぞれの用途に合った独自のデータ管理ツールを開発していた．これらのツールは，各プロジェクトに高度にカスタマイズされたものであり，いずれも収録データの再生機能とともに，閲覧やアノテーションを行う機能を実現したものだった．こうした構造によって，以下の問題が生じていた．

- 収録データのコピーの管理：ネットワーク上の記憶装置に配布された収録データをユーザが管理する（もう使われていない収録データや差分，関連性な

どを追跡する）ことは困難である．

- 収録データの使用：通常，目的のシステムの開発やテストに際して，実際には，各収録データの一部しか使用されない．
- 収録データのアクセス：（再生時に）収録データへの迅速なアクセスが頻繁に必要となる．

5. アプローチ

我々の構築したデータ管理システムは，以下の構成要素からなる．

(1) 収録データの構造

我々の経験によると，収録データ構造は，そのデータのアクセス，処理，共有する際の性能に重大な影響を及ぼす．プロジェクトのメンバは，往々にして何セットもの収録データをローカルコンピュータ内に保存していて，それらを目的のシステムの実行時やテスト時に用いている．我々は，情報の種類に応じて，収録データをさまざまなディレクトリに入れて，以下のように構造化することにした．

- **.../Recording Name/**
  - **On-line streams/**：ここには，収録プラットフォームから収録される全ストリームを保存する．「on-line」なる用語は，ストリームは，車内でライブ実行中のアプリケーションに対して使用可能な唯一のデータであることを意味している．一度作成されると，このディレクトリは書込みに対しロックがかかる．
  - **Off-line streams/**：ここにはon-line streamsあるいは他のoff-line streamsから事後生成された全ストリームが保存される．これらのストリームはプログラム（後処理，自動アノテーション，認識アルゴリズム）で生成したり，人手で生成（マニュアルアノテーション）したりすることができる．
  - **Information streams/**：ここには，on-lineおよびoff-lineディレクトリの配下にあるストリームに対して，これを記述するメタデータ，校正情報やプラットフォーム情報など，オリジナルの収録データからの情報抽出に役立つ全ストリームが保存される．
  - **Preview streams/**：簡易可視化可能なフォーマットでサブサンプリング表現された全ストリームが保存される．この機能は収録データをすばやくプレビューできるようにするために使用され，収録データの内容をすばやく可視化するためにグラフィックインタフェースから使用することができる．
  - **Searchable information/**：ここにはタグのストリームが保存される．タグは，データベースに保存してクエリを介して検索することができる，属性が付加された記述実体である．車両プラットフォームに関連するタグの例として，運転速度，特定の時刻におけるGPS位置情報，ある期間における映像中の歩行者の有無などが考えられる．

この構造では，各収録データは一つのディレクトリに完全に保存され，リスト化されたセクションやサブディレクトリを使用して構造化される．この構造は，収録フェーズ（on-line streams）で得られるオリジナル情報と（収録データを説明したり拡張したりするために）後で付加されるデータとを分離することを目的として決められた．この収録データ構造はインポート時にオリジナル収録データに適用され，収録データ用の中央記憶装置に保存される．

(2) 中央記憶装置のアーキテクチャ

我々は，部門に対応して地理的に分散配置した中央記憶装置を考案した．これによりローカルユーザ（ローカル部門）は高速アクセスが可能になり，必要に応じて収録データのコピーを他部門に送って収録データを共有することができる．このソリューションは，それぞれの場所における特有のニーズに応じて，部門ごとに異なるストレージソリューションが使用できるようにするものである．部門によっては，基本的なNFSによる収録データへのアクセスとデータ完全性の管理が必要となる．この場合，収録データ用の中央記憶装置には，データリプリケーション機能とCephFSによるアクセス機能を有するCEPHアーキテクチャ[CEPH]を採用する．例えば，別の部門では，大規模な収録データを処理するいくつもの計算プロセスが必要となる．この場合の中央記憶装置は，クラスタコンピュータを用い，Map-Reduceプロセスを通じて同一の収録データに並列アクセスが可能なHadoopアーキテクチャ[HADOOP]に基づくものとなる．

(3) 収録データのインポートプロセス

我々は，新しい収録データを保存するプロセスをシステム化するために収録データのインポート手順を定義した．新しい収録データは前述の収録データ構造にカプセル化され，この構造からさまざまなオフラインストリームに基づく計算を行うことができる．このインポートプロセスにはオリジナルの収録データからデータが供給されるので，収録データフォーマットやプラットフォームに応じて異なるデータアクセス機能が使用される．そのためインポートプロセスは，おのおのが異なる収録データフォーマットやストリームフォーマットを使用することができる一連のプラグインに基づいたものになっている．

(4) 収録データの可視化

ユーザに関する調査の結果，収録データを目視で検査して収録データの品質を評価したり，収録された特定のイベントや条件を見つけたりすることがユーザに求めら

れることがしばしばある．情報の中には，収録データの説明や収録日，場所，ビデオや他のストリームの再生などの収録データと一緒に可視化する必要のあるものがある．こうした機能のために，我々は，収録データのメタ情報を保存する一連のメダデータファイルを設計した．また，インポート時に，ほとんどの収録データストリーム用のプレビューファイルを作成する．この情報はファイルシステムから直接あるいはWebベースのグラフィックインタフェースを介して閲覧できるので，記憶装置内に存在する利用可能な全収録データに容易にアクセスすることができる．

(5) アノテーションツールとラベリングツール

インテリジェントシステムの開発プロセスでは，データのアノテーションは重要なステップである．このステップで，正解情報すなわち収録されたデータのより高レベルの表現を作成することができる（オフラインストリーム）．そのため我々は，一連のマニュアルおよび自動アノテーションツールを設計するとともに，このフェーズで第三者のツールが使用できるようにした．我々のWebグラフィックインタフェースには手動アノテーションツールが含まれていて，ユーザが収録データ内に直接アノテーションを生成できるようになっている．しかし手動アノテーションは費用のかかるプロセスでエラーも起こりやすいので，検証と品質管理を要する．そのため我々は，さまざまな種類の自動アノテーションサブシステムについても取り組んだ．車のプロジェクトでは，場所の三次元表現をダイナミックに獲得するアノテーションアルゴリズムを使って，交通参加者の三次元位置や方向，速度といったより高次の意味記述子用の正確なラベルを生成できるようにすることが非常に重要である．これらのより高次の記述子は，画像のマニュアルアノテーションでは容易に得られない．我々はここで，意味のあるタグを画像に自動的に割り当ててADAS/ADに関連するベンチマーク用データセットの構築を促進した．このため，LIDAR-カメラ融合に基づく新しい機械学習フレームワークに取り組んだ．我々のアプローチでは，同期した画像ストリームと点群データストリームを入力として，意味的にセグメント化された画像および街路景観，交通参加者とその速度プロファイルの三次元ダイナミック再構築を含む意味的アノテーションが画像および点群内に生成される（図4参照）．



図4　シーンアノテーションの例

(6) タグ抽出とインデキシングのプロセス

何時間もかけて収録されたデータと膨大な数のイベントやアノテーションが収録データに含まれている場合には，ファイルストリームを直接アクセスして情報を直接検索すると膨大なコンピュータ演算が必要となる．我々は（メタデータ，アノテーション，ラベルといった）検索可能な全情報をバッチで後処理して，それらをタグストリームに変換することにした．このようなストリームはタイムスタンプが付いたテキストフォーマットの情報を表現するので，データベースへの保存に適している．タグはあらかじめ定義された構造をもたないので，我々はタグをNoSQLデータベースに保存し，タグがもつさまざまな関係をフレキシブルにモデル化して，タグにすばやくアクセスできるようにした．我々はこのデータベースにさらに，タイムスタンプ（場合によっては継続時間）に従ってタグをグループ分けする構造を構築した．この構造により，特定のイベントやアノテーションが有効な収録セグメントを見つけるための検索をより効率良く行うことができる．

(7) 収録データの検索

タグをデータベースに保存することにより，プレフォーマット検索を行って特定の条件に適合する収録データのセグメントを見つけることができる．我々は，データベースの上部に簡単なクエリ構文を作成し，実行時にこの構文が特定のクエリ言語に翻訳されるようにした．さらに，使用されたタグの値や種類の統計を保存する特定の構造をデータベース内に構築した．これによりユーザは，収録データ内で何が検索されるかを事前に知ることができる．

検索要件の分析で，主に3種類の検索があることがわかった．（例えば，車が交差点を通過している，あるいは雨天で走行しているなどの収録データを見つけるといった）存在の検索，（例えば，車の速度がある値より高い，車が特定のGPSエリア内にいるといった）値の検索，および（例えば，車が左側車線のトラックに追いつかれた，より高速の車が通過しているので車が減速しているといった）複雑な条件の検索である．最初の2種類の検索は我々のクエリ言語で構築できるが，3番目の検索には，収録データストリームにアクセスしたり，イベントや収録データのコンテキストをさらに翻訳したりする必要がある．最初の2種類の検索はすでに実現したが，3番目は未着手である．

(8) 検索結果

ある検索クエリを実行して得られる検索結果は，そのクエリに適合する（開始時間と終了時間の付いた）収録

データセグメントのリストになっている．検索結果はさまざまな方法で使用できる．検索結果は，見つかったセグメントをマニュアル検査用に可視化するために使用することもできるし，ファイルに保存して目的のシステムのプログラムで使用し（収録用中央記憶装置を介して）収録データセグメントに直接アクセスすることもできるし，エクスポート機能を使用して見つかったセグメントのコピーを作成することもできる．後者の場合は，ユーザはローカルマシンに収録結果のコピーを置いて，高性能マシンで構成される目的のシステムで再生することができる．

6. EMI システム

我々のデータ管理システム（実験管理インフラストラクチャ EMI）の設計開発は欧州および米国の部門共同プロジェクトとして実施された．両部門は，多数のセンサから得られる車の収録データの処理要件を共有し，収録データを集中的な方法で保存した．

我々の収録データ管理ワークフローでサポートされているステップとツールのリストを図 5 に示す．

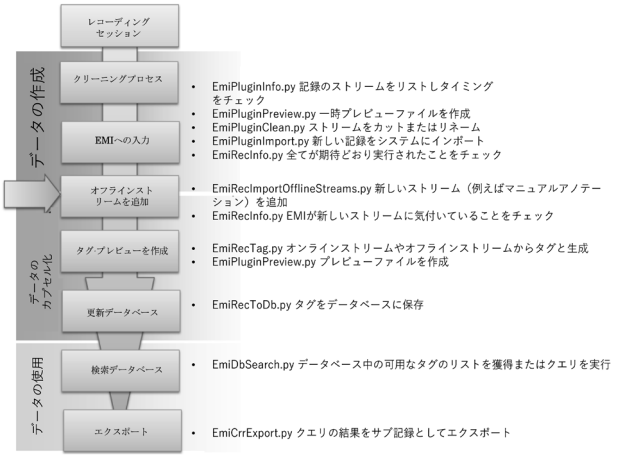

図 5　EMI ワークフローの各ステップ

ワークフローは収録セッション（車，ロボットなどの収録プラットフォームから得られる生データ）から始まる．このデータはプラットフォームから社内ネットワーク上のローカル記憶装置にコピーされる．最初に行われるのは，収録データの前処理で，データのクリーニングと不要部の切取りが行われる．このフェーズで，収録の目的を記述した一般情報がメタデータとして付加される．次のステップにはインポート，収録データの統一構造への再成形，オリジナルストリームに基づいて計算されたストリームの付加，およびプレビューの作成が含まれる．この段階で，手動アノテーションまたは自動アノテーションを生成することができる．その後，全ストリームからタグが抽出されて，検索可能なデータベースにインポートされる．

この段階で，ユーザは収録データ全体に直接アクセスしたり，検索クエリを使用して特定の条件に適合するセグメントを見つけたりすることができる．検索結果（インタバルリストファイル）を使用して収録データセグメントに直接アクセスしたり，収録データセグメントをサブ収録データとしてエクスポートしたり，ユーザのローカルコンピュータにコピーしたりすることができる．検索クエリの例を図 6 に示す．

検索クエリ	説明
EmiDbSearch.py -r "project=TFP OR project=PAZ"	プロジェクトTFPまたはプロジェクトPAZから記録を獲得
EmiDbSearch.py -r "country=eu" -s "ego_vehicle.mean_speed > 15"	自車が15m/sより高速で走っている国euから間隔を獲得
EmiDbSearch.py -s "traffic_objects_extended.objects.most_typename=car AND traffic_objects_extended.objects.mean_velocity_abs_x>20"	Carであるトラフィックオブジェクトとの距離および速度20m/s超のトラフィックオブジェクトとの距離を獲得
EmiDbSearch.py -s "(traffic_objects_extended.objects| most_typename=car AND mean_velocity_abs_x>20)"	速度20m/s超のcarであるトラフィックオブジェクトとの距離を獲得

図 6　EMI 検索クエリの例

我々は，アーキテクチャの観点から，各部門からデータをすばやくアクセスできるデータ管理システムを目指した．この要件は，分散アーキテクチャを有する，部門ごとの中央データ管理システム（各部門がサーバインスタンスを一つもつ）をサポートするデザインアーキテクチャにつながる．このアーキテクチャにより，Web インタフェース（図 7 の（EMI フロントエンドの）EMI UI 参照）を介して中央の収録データアクセスビューを使用してローカルの収録データ（図 7 の EMI バックエンド参照）への高性能アクセスが可能になる．ここで各サーバは，特定の収録データフォーマット，収録データへのさまざまなアクセス（CEEPH，Hadoop など），ローカルインデックスを使用したより高速な検索，あるいは特定のプロジェクトアノテーションサービスを処理するために一連のローカルな選択肢をもつことができる．



図 7　EMI アーキテクチャの全容

我々は，システムの性能を調べるために，継続時間の異なる収録データを使用してテストを実施した．平均すると，1 時間の収録データはセンサストリーム 20 本分に相当する 680 ギガバイトに達する．実行時間結果の例を図 8 に示す．

機能	実行時間
記録のインポート（1時間の記録）	30分
タグの計算（1時間の記録）	1時間
データベースの更新（1時間の記録）	2分
検索クエリの実行（DBでの6時間の記録）	2秒
検索クエリの実行（DBでの133時間の記録）	7-12秒
記録のエクスポート（1時間の記録）	30分

図 8　EMI 機能の実行時間の例

7. ま　と　め

本稿では，学習アルゴリズム用の大規模な学習データを生成するシステムやマルチセンサプラットフォームを扱う際に AI 研究者が直面するニーズや課題を紹介した．収録データフォーマットに事前に制約を設定できないマルチサイトの組織構造でこの種のビッグデータのセットを保存，共有，使用するために設計，構築したソリューションについて解説した．同様な課題の解決を目的にした国際プロジェクトが複数存在することは，こうした経験が AI コミュニティにとって価値があることを示している．次のステップでは，Web ベースのユーザインタフェースのパーツとアノテーションやタグ生成用ストリームで動作している低次機能の向上・完成に焦点を当てる．システムは，最初のリリースを行うフェーズにあり，ユーザからのフィードバックを今後の開発計画に生かしていく予定である．

◇ 参 考 文 献 ◇

[ASAM] ASAM WIKI, https://wiki.asam.net, access 2017

[Brostow 09] Brostow, G. J., Fauqueur, J. and Cipolla, R.: Semantic object classes in video: A high-definition ground truth database, *Pattern Recognition Lett.,* Vol. 30, No. 2, pp. 88-97（2009）

[Campbell 12] Campbell, K. L.: The SHRP 2 Naturalistic Driving Study, TR News 282, September-October 2012, https://insight.shrp2nds.us/documents/shrp2_background.pdf（2012）

[CEPH] CEPH, http://ceph.com, access 2017

[CLSVA] Cloud LSVA - Large Scale Video Annotation, http://cloud-lsva.eu, access 2017

[Cordts 16] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S. and Schiele, B.: The cityscapes dataset for semantic urban scene understanding, *Proc. IEEE Computer Vision and Pattern Recognition*（*CVPR*）（2016）

[Deng 09] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K. and Fei-Fei, L.: A large-scale hierarchical image database, *Proc. IEEE Computer Vision and Pattern Recognition*（*CVPR*）（2009）

[Fuggetta 98] Fuggetta, A. Picco, G. P. and Vigna, G.: Understanding code mobility, *IEEE Trans. on Software Engineering,* Vol. 24, No. 5, pp. 342-361（1998）

[HADOOP] Apache Hadoop, http://hadoop.apache.org, access 2017

[INSIGHT] InSight Data Access Website - SHRP2 Naturalistic Driving Study, https://insight.shrp2nds.us, access 2017

[Lin 14] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P. and Zitnick, C. L.: Microsoft COCO: Common objects in context, *Proc. European Conf. on Computer Vision*（*ECCV*）（2014）

[Mottaghi 14] Mottaghi, R., Chen, X., Liu, X., Cho, N. G., Lee, S. W., Fidler, S., Urtasun, R. and Yuille, A.: The role of context for object detection and semantic segmentation in the wild, *Proc. Computer Vision and Pattern Recognition*（*CVPR*）（2014）

[ODS] ASAM ODS Base Model, http://www.highqsoft.de/download/ao_base.htm, access 2017

[OPENMDM] OpenMDM toolkit, https://www.openmdm.org/about-openmdm/asam-ods（2017）

[Rob 14] Rob, E., et al.: UDRIVE: The European naturalistic driving study, *Proc. Transport Research Arena,* IFSTTAR（2014）

[Ros 16] Ros, G., Sellart, L., Vazquez, D. and Lopez, A. M.: The SYNTHIA dataset: A large collection of synthetic images for semantic segmentation of urban scenes, *Proc. IEEE Computer Vision and Pattern Recognition*（*CVPR*）（2016）

[SHRP] Strategic Highway Research Program（SHRP）2 - Revised Safety Research Plan, http://onlinepubs.trb.org/onlinepubs/shrp2/RevisedSafetyResearchPlanMarch2012.pdf, access 2017

[Silberman 12] Silberman, N., Hoiem, D., Kohli, P. and Fergus R.: Indoor segmentation and support inference from rgbd images, *Proc. European Conf. on Computer Vision*（*ECCV*）（2012）

[UDRIVE] UDRIVE - European Naturalistic Driving Study, http://www.udrive.eu, access 2017

[XCUBE] XCube, http://www.x3-c.com, access 2017

[XCUBEPD] XCube Product Description, http://www.x3-c.com/wp-content/uploads/2017/02/XCube-Products-Description_Jan2017.pdf（2017），access 2017

2017 年 11 月 6 日　受理

著　者　紹　介



Antonello Ceravola

イタリアのピサ大学でコンピュータサイエンスを専攻．IT ソフトウェア分野で 5 年間勤務しマルチメディアシステム，通信システム用大規模ソフトウェアインフラストラクチャ，マルチティアアプリケーションおよびプロセス管理システム用ワークフローエンジンに取り組んだ．2001 年にドイツの Honda Research Institute Europe に移り，現在は主任研究員．研究課題はソフトウェアコンポーネント，ミドルウェア，大規模システム，統合環境など．



Frank Joublin

1993 年フランスのルーアン大学で神経科学のヨーロッパ博士号を取得．1994 ～ 98 年ドイツのボーフム大学神経情報科学研究所の博士研究員．1998 ～ 2001 年 Philips Speech Processing Aachen 社の顧客プロジェクトマネージャ．2001 年からドイツの Honda Research Institute Europe の主任研究員．研究課題は発達ロボティクス，意味獲得，データマイニングなど．



Heiko Wersing

ドイツのビーレフェルト大学で 1996 年に物理学の学位を，2000 年に科学博士号をそれぞれ取得．2000 年にオッフェンバッハ市の Honda R&D Europe GmbH に入社し，現在ドイツ オッフェンバッハ市の Honda Research Institute Europe GmbH の主任研究員として活躍中．2017 年にドイツのビーレフェルト大学から名誉教授を授与．研究課題はインクリメンタルなオンライン学習，パーソナライゼーション，適応 HMI，コンピュータビジョンなど．



Stephan Hasler

ドイツのビーレフェルト大学で 2010 年に工学博士号を取得．現在ドイツの Honda Research Institute Europe の上席研究員．研究課題は視覚認知用モデル，生涯学習，人間と機械のインタラクションなど．



Behzad Dariush

オハイオ州立大学でロボット工学と制御システムを専攻して電気工学博士号を取得．現在 Honda Research Institute USA の主任研究員として Knowledge Discovery グループを管理し，グローバルホンダの自動運転プロジェクトやロボティクスプロジェクトをサポートしている．これまでの研究課題はヒューマノイドロボティクス，ウェアラブル技術，人体動作解析など．



Yi-Ting Chen

パデュー大学で結像システムと画質解析を専攻して電気情報工学博士号取得．卒業後マーセッドのカリフォルニア大学に移り，グーグル社でコンピュータビジョンの研究やプロジェクトに従事．現在 Honda Research Institute USA の研究員としてコンピュータビジョンと機械学習に取り組んでいる．特に，自動運転用のさまざまなアプリケーションのための三次元ダイナミックシーンの理解に取り組んでいる．