

ニュアンス コミュニケーションズの音声ビジネス

平沢 純一 村上 久幸 田中 幸
ニュアンス・コミュニケーションズ・ジャパン株式会社

アブストラクト

ニュアンス コミュニケーションズ社 (旧名 スキャンソフト社) は 2001 年より音声ビジネスに参入し、いくつもの音声ベンダの吸収合併を重ねて、音声認識・音声合成に関するビジネスを成長させてきた。本稿では、ニュアンス社の世界市場における事業の概要を紹介し、併せて“音声”をビジネスとして成立させるために採用している方法論について論じる。さらにニュアンス社 日本法人が日本市場におけるビジネス展開で直面している課題にも言及する。

Introducing Nuance Communications And Its Speech Business.

Jun-ichi Hirasawa Hisayuki Murakami Miyuki Tanaka
Nuance Communications Japan K.K.

Abstract

Nuance Communications Inc., formerly known as ScanSoft, is one of the biggest companies in speech technology field. Nuance started its speech business in 2001, and has been contributing to global growth in speech market for five years. In this paper, we introduce the overview of our company and present how we achieve success in speech business. Furthermore, we discuss how we can grow speech market in Japan.

1. はじめに

ニュアンス コミュニケーションズ社 (旧名 スキャンソフト社。2005 年 10 月に社名変更) は、2001 年の L&H 社の吸収合併を機に音声ビジネスに参入した。以来約 5 年、いくつもの音声ベンダの吸収合併を重ねて、音声認識・音声合成に関するビジネスを成長させてきた。

本稿では、ニュアンス社の世界市場における事業の概要を紹介し、併せて“音声”をビジネスとして成立させるためにニュアンス社が採用している方法論について論じる。さらに、ニュアンス社 日本法人が日本市場においてビジネスを展開する現場で直面している課題に言及することにより、今後の日本の音声ビジネス市場を発展させるための課題を指摘する。

以下、次節 (2 節) ではニュアンス社の概要を紹介し、3 節ではニュアンス社の音声ビジネスに対するアプローチ (取組みと方法論)、4 節では日本市場での課題、を述べていく。

2. 会社紹介

ニュアンス社 (旧スキャンソフト社) について把握するには、数々の吸収合併の歴史を知ることが役立つだろう。図 1 にニュアンス社の変遷を示す。音声ビジネスに関する、主だった吸収合併は以下の通りである。

2001 年 12 月 : L&H 社 (ベルギー) 買収

2003 年 1 月 : Philips 社音声事業部 (ドイツ) 買収

2003 年 8 月 : SpeechWorks 社 (US) 買収

2005 年 1 月 : ART 社 (イスラエル) 買収

2005 年 9 月 : Nuance Communications 社 (US) 買収

2005 年 10 月 : 社名を ScanSoft から変更

2006 年 3 月 : Dictaphone 社 (US) 買収

日本市場においては、以下のように日本法人が設立されている。

1995 年 : エルアンドエイチジャパン設立

2001 年 2 月 : スピーチワークスジャパン設立

2001 年 4 月 : ニュアンスコミュニケーションズ (日本法人)

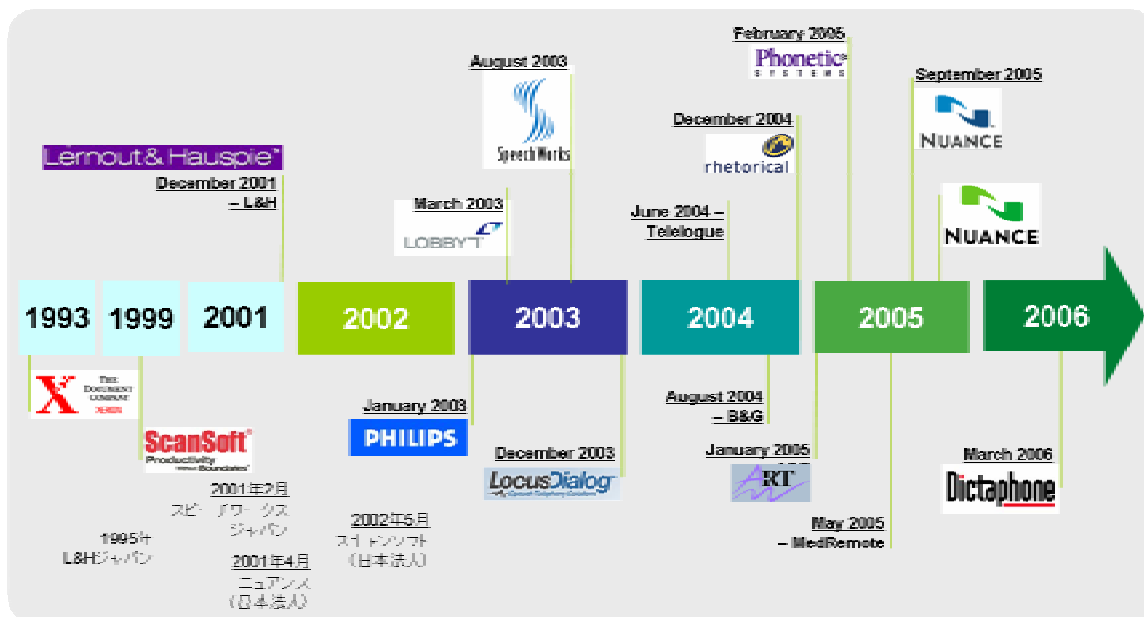


図 1：ニュアンス社の変遷

2002年5月：スキャンソフト（日本法人）設立
 （エルアンドエイチジャパンから社名変更）
 2005年10月：社名をニュアンスコミュニケーションズに変更

このような経緯の元、2006年10月現在、ニュアンス社は下記の4つを主な事業領域としている。

- **イメージング (ScanSoft) 事業**
PDF 編集 OCR ドキュメント管理
- **ディクテーション (DragonNaturallySpeaking) 事業**
PC 向けパッケージ IC レコーダ
- **ヘルスケア (Dictaphone) 事業**
医療機関向けディクテーション
- **音声認識・音声合成 (Speech) 事業**
組込み向け（車載 携帯電話）
ネットワーク向け（コールセンタ 通信）

このうち、イメージング事業を除く三事業が音声技術を扱っている。

以上述べた経緯と事業領域により、2006年10月現在、ニュアンス社は全社で社員数 1700 人ほどの企業規模となっている。内訳は、イメージング (ScanSoft) 事業に約 200 人、ヘルスケア (Dictaphone) 事業に 400 人程度、間接部門 (法務/財務/人事など) が 100 人以下で、それ以外の 1000 人程度がディクテーション事業 (DragonNaturallySpeaking) も含めた音声関連に従事

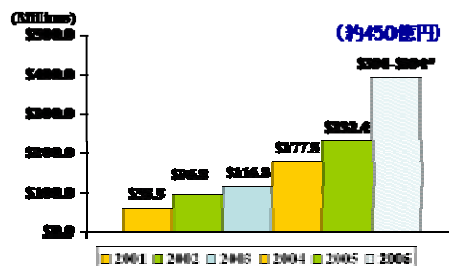


図 2：2001年～2006年（予測）の収益

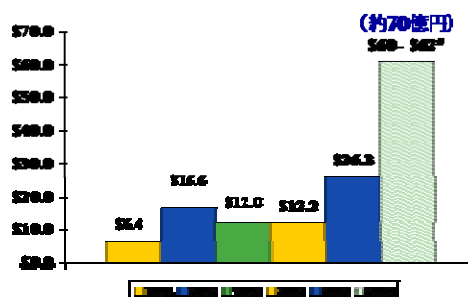


図 3：2001年～2006年（予測）の純利益

している。

音声関連 1000 人の職種による内訳は、営業・SalesEngineer・マーケティング部門などで 300～400 人程度、技術サポートなどのプロフェッショナルサービス部門で 300 人程度、研究開発部門で 300 人程度となっている。音声事業を専業としている企業

の中では、おそらく最も大規模な企業のうちの1つであると言えるだろう。

地域的には、事業拠点として15ヶ国、営業活動地域は78カ国に上る。従業員数で見ると、北米地域に1200人、テルアビブやブダペストの開発拠点も含めた欧州地域に400人、日本を含むアジアパシフィック地域には40人程度が配置されている。

財務面からの状況について、全事業部合算での収益を(図2)に、純利益を(図3)に示す。収益(図2)では、2005年度で232.4百万ドル、2006年度予測は約394百万ドル(約450億円)、純利益(図3)では、2005年度で26.3百万ドル、2006年度予測は約62百万ドル(約70億円)を見込んでおり、2001年の音声ビジネス参入以降、赤字を出すことなく毎年利益を伸ばし続けている。

ニュアンス社の音声関連製品の一覧を(図4)に示す。吸収合併を繰り返す過程で、各製品がどのように扱われていくのかに興味を持つ人は多いかもしれない。各製品を統合する作業は吸収合併が完了してから進められる。例えば、(図4)にあるように、ネットワーク向けの音声認識エンジンでは、旧ニュアンス(blue Nuance)の認識エンジンNuance8.5と旧スキャンソフトの認識エンジンOSR(OpenSpeechRecognizer; SpeechWorks由来)がまだ併存しているが、現在、両者を統合した新しいエンジンのリリース準備を進めている(2007年1月リリース予定)。

また、対応言語数(カバレッジ)の大きさもニュアンスの強みのひとつと言える。2006年10月現在、ネットワーク(telephony)向け製品では、音声認識48言語、音声合成24言語、

車載や携帯電話向けの組込み製品でも、音声認識20言語、音声合成24言語、に対応している。

3. 音声ビジネスへのニュアンスのアプローチ

本節では、音声技術の専門ベンダとしてのニュアンス社が、音声ビジネスで利益を上げていくために採用している取組みや方法論を紹介していく。

パートナーとの協業スキーム: ニュアンス社にとって、ネットワーク向けであれ、組込み向けであれ、音声でのビジネスを展開するのに、ニュアンス社だけでビジネスが成立することはありえない。言い換えれば、多くのパートナーとの協業スキームを成立させることが最重要課題となっている。音声サービス(システム)は音声ベンダだけでは決して成立しない。

(図5)には、ネットワーク向けプロジェクト、組込み向けプロジェクトのそれぞれについて典型的なプロジェクトのスキームを示した。例えば、ネットワーク向けプロジェクトでは、IVRベンダとして日本ローカルのパートナー会社が約8社、日本にも拠点を置くグローバルパートナー会社が約6社存在し、彼らとの協業の元でプロジェクトを遂行する。車載向け組込みプロジェクトでは、日本国内で6社のナビベンダと、海外でも10数社との間でビジネスを進めている。

プロフェッショナルサービスの提供: ニュアンス社では音声認識・音声合成エンジンをライセンス提供するだけでなく、各種の技術サポート(プロフェッショナルサービス)を提供している。例えば、音声サービス向けのアプリケーション仕様策定の

コンサルテーション、アプリケーション開発における技術サポート、音声データ(評価用・チューニング用)収録の請負、認識エンジンの各種パラメータチューニング、出力用の音声プロンプト(ガイドランス)収録の請負、音声合成(TTS)プロンプトのチューニングなど、開発の全フェーズにわたってカスタマへのサポートを提供している。

プロフェッショナルサービスの提供において着目したい

	音声認識	音声合成	ディクテーション	声紋認証
ネットワーク	Nuance 8.5 OSR 3.0 ・コールセンター ・ボイスポータル ・番号案内 ・プレミアムサービス	RealSpeak Telecom ・コールセンター ・ボイスポータル ・番号案内 ・Eメール読み上げ	Dragon Naturally Speaking SDK Dragon Audio Mining ・医療 ・オフィス	Nuance Verifier Speech Secure ・本人認証 ・アカウント管理 ・各種セキュリティチェック ・バージョンナリゼーション
組込みデスクトップ	VoCon 3200 VoCon Mobile VoCon Games ・カーナビゲーション ・ハンズフリーキット ・携帯電話 ・ゲーム	RealSpeak Solo RealSpeak Mobile ・カーナビゲーション ・ハンズフリーキット ・携帯電話 ・電子辞書 ・PDA	Mobile Dictation Dragon Naturally Speaking ・携帯電話 ・テレマティクス ・オフィスディクテーション ・医療 ・ICレコーダ	Speech Secure ・本人認証 ・各種セキュリティチェック ・バージョンナリゼーション

図4: ニュアンス社 音声関連 製品一覧

のは、単に個々のサポート作業を提供するだけでなく、音声サービス(システム)を実現していくための“開発プロセス”自体を提案しようとしていることである。音声サービス(システム)の開発においては、音声技術の独自性から通常のソフトウェア開発とは異なるノウハウや価値観が必要となることがある。そのため「音声認識・合成がそもそもどのような特性を持った技術なのか」という知識なしにプロジェクトを進めていくのが難しいことがある。一例としては、音声技術に馴染みがないと、単なる結果としての性能(認識率など)の“数値”だけに注目して振り回されてしまう危険がある。一方、音声プロジェクトの特性を踏まえた“開発プロセス”を実践することができれば、プロジェクトを混乱させることなく、結果として最善な性能を確保しやすい。

具体的には、ニュアンス社では、Speech Application Development Lifecycle (SADL; サドル) という枠組みを提唱している。SADL とは、

1. Discovery (音声インタフェース利用の戦略策定、ビジネスケース・要求仕様の分析・定義)
2. Design (ユーザインタフェースのデザイン・仕様設計)
3. Realization
(アプリ開発、グラマ開発、レコーディング、テスト)
4. Optimization
(エンジンチューニング、アプリチューニング、最適化)

という開発プロセスについて、各プロセスの順序を逆転させることなくプロジェクトを進めていこうとする枠組みの提案である。SADL の実践により、「蓋を開けてみないとわからない」「開発を終えてみたものの十分な性能が得られなかった」といった、プロジェクト失敗のリスクをできるだけ最小化することが可能となる。

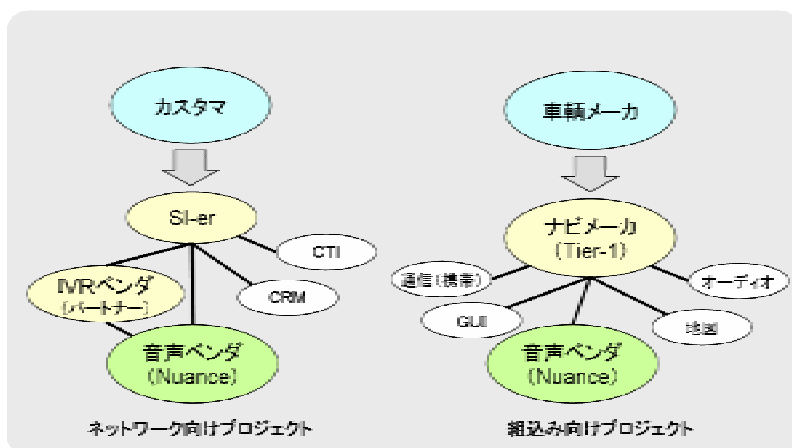


図 5: 音声プロジェクトにおける協業スキーム

Launch Program: 新たなサービス(システム)を導入するカスタマにとって“音声インタフェース”とはあくまで新たな機能を実現するための手段のひとつに過ぎず、必ずしも音声インタフェースの導入自体が目的化しているとは限らない。まして、音声サービス(システム)の導入を繰り返し経験しているカスタマなど稀であり、多くのカスタマにとって音声サービス(システム)の導入は初めての経験となる。

そこで、前述した SADL のような“開発プロセスへのサポート”だけにとどまることなく、サービス(システム)を導入する際の各種のプロモーションのようなプロセスでも、カスタマをサポートする体制(Launch Program)を提供している。Launch Program では、例えば、社内・社外(エンドユーザ)へのサービスの周知・告知はどの時点でどのように行なえばよいのか、社内でのチームの組織の仕方など、音声サービス導入にまつわるおおよそすべての側面に関して、カスタマをサポートする。

組込み向けポーティング開発: 車載機器や携帯電話といった組込み向けプロジェクトでは、それぞれのカスタマごとに特有のプラットフォームに合わせたポーティング開発作業を請負う。カスタマにとっては、“エンジンのポーティング”自体と“グラマ開発やチューニングなども含むアプリケーション開発”の両方を合わせて、全体として“システム開発”として捉えがちであるため、カスタマによっては、両者(ポーティング・アプリケーション開発)の区別を理解しにくいことがある。

ニュアンスでは、ポーティング作業とは「PC 向けに提供されているエンジン(SDK)と同じ機能・性能を当該プラットフォーム上に再現すること」と定義しており、それぞれのカスタマごとのアプリケーションに依存したグラマ開発やチューニング

(性能の最適化)などのプロフェッショナルサービスとは混同しないように整理している。ニュアンス社が製品として提供するエンジン自体の性能を向上させる取組みと、それぞれのプロジェクトごとに定義・設計されるアプリケーションにおいてエンジンの性能を最大限まで向上させる取組みを混同することは、音声プロジェクトに不必要な混乱を招く要因となりかねない。

研究開発戦略: エンジンを始めとした製品開発などの研究開発はすべて TTM(Time To

Market) プログラムという手続きに従って社内決定される。新たな言語カバレッジの開発方針も同様に TTM に従って行なわれる。TTM では、新技術・新機能・新言語の開発によって拡大される収益の予測と、開発に要するコストが比較され、開発方針が決定される。従って、R&D 部門や経営陣の一部の嗜好や意向で開発方針が策定されることはなく、すべては市場へのインパクトに応じて方向付けられていく。

4. 日本市場でのビジネス展開における課題

ネットワーク向け、組込み(車載・携帯電話)向けを中心としてニュアンス社日本法人は日本の音声市場でのビジネス展開を進めている。本節では、ビジネス展開の過程で課題となっているいくつかのポイントを挙げ、日本市場での音声ビジネスの発展の可能性を探りたい。以下に4つのポイントを挙げる。

- (a) プレイヤーとミッションは明確かつ適切か
- (b) 仕様の策定はいかになされているか
- (c) 製品の性能をどのように把握するか
- (d) アフターケアはどこまで行なわれるか

(a) プレイヤーとミッション: 前節(図 5)でも述べたように、実際のサービス(システム)が音声インタフェースだけで成立することはありえない。サービス(システム)の利用者にとって重要なのは、自分の口座を管理できるのか、目的地を設定できるのか、といった具体的なタスクであって、全体のサービス(システム)内には音声インタフェースの制約を受けないモジュールも存在することになる。従って、音声技術の関係者だけでサービス(システム)をイメージしていてもプロジェクト全体の成功は覚束ない。

言い換えれば、全体のプロジェクトには音声インタフェース技術に目利きも馴染みもないプレイヤーが加わるのが現実である。音声技術について彼らに何をどこまで理解してもらう必要があるのか。非音声のモジュールを担当するプレイヤーから、音声技術特有の事情を考慮していない要求や価値観がプロジェクトに持ち込まれた際に、どこでどのようにプロジェクト全体の整合性を調整することができるのか。その役割を担うのは誰なのか。音声エンジニアに求められている役割は、認識率向上だけではない。

音声技術の関係者だけでプロジェクトを企画してしまっていないか。また、非音声の関係者とのミッションの分担は明確かつ適切になされているか、がプロジェクトの成否を握る。

(b) 仕様の策定: サービス(システム)全体の要求仕様、それに基づいて策定される各モジュールの詳細仕様の品質が、最終的なサービス(システム)自体の品質を左右する。そこで、これらの仕様策定は誰が担っているのか、その仕様策定プロセスには音声技術に関する目利きやノウハウが必要とされるだけ投入されているか、がサービス(システム)の品質を決めることになる。

例えば、実際の現場では、ある時は音声のエキスパートを擁さない IVR ベンダが詳細仕様を策定するケースもありうる。ある時は最上流の車輛メーカーがすべての仕様を掌握して策定し、音声ベンダはおろかナビベンダすら関われないケースもありうる。そのような現実のビジネスの制約の中でも、音声認識・合成の現状の技術的な特性を踏まえた適切な仕様の策定を確保しないと、結局、プロジェクト全体を迷走させるリスクを高めてしまうことになる。

卑近な例を挙げれば、仕様策定のミッションを担っているプレイヤーは、対話フロー仕様を記述するための適切なフォーマットを持っているだろうか。excel の作図機能を駆使して涙ぐましいまでに巨大なフローを作成するのが精一杯だとしたら、細かいフローの更新管理が適切になされるだろうか。

エンジンの技術的な特性とコンテンツ(アプリケーション)の関係を整然と把握した上で仕様を策定することが、音声プロジェクトを成功させるために必要なのである。エンジンとコンテンツ仕様(グラム・フロー)の関係は、CDプレイヤーとCD自体の関係に例えられることがある。素晴らしいCDコンテンツ(仕様)を実現しようとするあまり、いつの間にかCDプレイヤー(エンジン)自体の改善に取り掛かってしまっていたのではプロジェクトは混乱する。逆に、折角のCDプレイヤー(エンジン)の長所を活かすことのないCDコンテンツ(仕様)しか策定できないのではCDプレイヤーも報われない。

(c) 性能の把握: あるカスタマが、せっかく音声サービス(システム)の導入を思い立っても、認識性能や読み上げ性能が「やってみないとわからないものだ」と聞かされたら不安に陥るだろう。認識対象の音声(発話内容・発話者・発話環境)や読み上げ対象のテキストの全パターン・全数チェックを行なうことが現実的でない以上、「やってみないとわからない」側面があるのは事実だが、購入・導入する立場からすればこれほど恐ろしい商品もない。しかしこれまで音声技術は“性能保証”に積極的だったとはいえない。

カスタマが音声システム(エンジン)の性能を把握したいと

思うタイミングは、エンジン選定時、仕様策定時、導入（量産）前テスト時など、プロジェクト中に複数存在する。例えば、ネットワーク向けプロジェクトでは IVR ベンダの選定がそのまま認識エンジンを定めてしまうケースもあるため、「認識エンジンの選定」というモチベーションがさほど高くないこともある。しかし、組込み向けプロジェクトでは「エンジン選定」は一大イベントであり、エンジンに必要な性能の保証が求められることは必至である。

性能を把握するには「評価データ」を定める必要がある。音声技術に馴染みが少ない人の中には「エンジン」自体が何か固有の性能を持っているかのように錯覚する人がいるが、評価データを定めずに性能は定義できない。では何を評価データにすればよいか。

資金やノウハウのある車輛メーカーやナビメーカーであれば、自前でデータ収集ができるかもしれない。しかし、実は「どんな音声データをどの程度収集すれば、エンジン選定において後悔のない十分な評価データと言えるのか」は自明でない。評価データを無限に調達することはできない。用意するデータが性能保証の観点から十分と言えるのか、不十分だとしたら何がどのくらい足りないのか、実は手探りなのが現実である。

自前で評価データを収集・設定できないカスタマには「性能を把握する」ことはさらに深刻な課題となる。第三者機関が公平な「ベンチマーク向け評価データ」を提供できればよいのだろうか。誰がどんな資金を使って、このような中立なデータを作成できるのかも大きな課題であるが、より難しい課題は「カスタマが自らのサービス（システム）で扱いたい対象の範囲（空間）と、ベンチマーク向け評価データがカバーしている対象の範囲（空間）との関係を把握できない」ことである。カスタマが自らのサービス（システム）との関係（距離）を把握できなければ、ベンチマークが示す性能は、個々のカスタマにとって単なる「参考記録」でしかなくなる。

本質的な課題は、それぞれのカスタマが実現したい対象をカバーする範囲を「評価データ」として表現する明確な方法を、カスタマが手にできていないことにある。

(d) アフターケア： エンジン選定、仕様策定、適切なチューニングも含めて、すべてのアプリケーション開発までを正しいプロセスで進めることができたと仮定しよう。しかし「やってみないとわからない」音声技術では、この段階に及んで不幸にもカスタマの求める性能を満たせず、「どうも特定のコマンドがうまく認識されないのだけれど」という個別のクレームを受ける

ことがある。そんな状況で、現状の音声技術はどんなアフターケアを提供できるだろうか。

まず、第一は「音声ファイル（ログ）を取得・保存してもらう」ことである。音声データがなければ現象を再現・確認することすらできない。しかし組込み製品では音声ログの保存が許されないことも多い。カスタマからは「クレーム内容の報告だけ」しか得られない時にどんな対処（アフターケア）が可能だろうか。

運良く音声データを取得でき、現象の再現ができたとして、当該のクレーム事項だけを改修することができるだろうか。プロジェクトの最終フェーズに近づくほど、何らかの改修は往々にして予期せぬ副作用も併発してしまうことがある。クレーム事項を修復させる代わりに、新たに別のクレーム項目を作ってしまうようなイタチごっこを避けるには、クレームに対してどう対応するのがよいだろうか。たとえ性能を改善させることができないとしても、「認識できない」理由を明確に説明できればカスタマに納得してもらえる可能性もある。しかし「転んだから骨折した」のように「〇〇のせいで認識できなかった」という単純な因果律で回答をできるとは限らない。

そもそも、「認識しない」というお馴染みのクレームを受けずに済むことを目指したい。それには実際の開発に入る前に「対象とする入力音声の範囲」についてカスタマと合意した上で、提供可能な性能の限界についてカスタマからの了解を得ておくべきである。しかし前項でも述べたように「性能の事前把握」が十分に保証されるとは限らない。個別のクレームに対する誠実なアフターケアとは何だろうか。

5. おわりに

本稿では、ニュアンス社の事業概要から始めて、音声をビジネスとして成立させるための数々の取組みを紹介した。さらに実際の現場で切実な課題となっているポイントについて議論してきた。日本市場はカスタマの要求水準が高い。議論してきたポイントは、単にニュアンス社の課題であるだけでなく、音声技術を普及させたいという情熱を持つすべての人にとっての課題と思えてならない。

謝辞： 本発表の機会を与えていただいた 旭化成株式会社 庄境誠さんに感謝いたします。

連絡先： email: jun-ichi.hirasawa@nuance.com