

## SeC ( Semantic Computing ) 計画と

### CDL ( Concept Description Language )

セマンティックコンピューティング ( SeC ) 計画と CDL の概要説明です。説明は次の 4 点からなります。

- ( 1 ) セマンティックコンピューティング ( SeC ) とは何なのか、SeC 計画の目的とは
- ( 2 ) SeC を構成する技術とは、SeC の技術アーキテクチャとは
- ( 3 ) SeC が目標とする実現機能とは、SeC 技術の応用とは
- ( 4 ) SeC 計画の推進、どのように進めるのか

#### 1. SeC 計画の目的

SeC の目的は、コンピュータと人間・社会が意味を共有すること、共通の意味世界を持つことによって、人間・社会における情報処理とコンピュータにおける情報処理をシームレスにつなぐ新しい IT 技術を実現することである。

シームレスにつなぐ技術とは、人間・社会における情報の形態をコンテンツ ( ドキュメントを含め ) であるとし、コンテンツに関してコンピュータと人間が共通理解できるようにする技術である。そして、この技術によって、人間・社会の知的生産性を飛躍させる支援機能をコンピュータ上に実現することが SeC 計画の目標である。

SeC は、これからの IT システムを構成する次のような基本技術の中に位置付けられる。機能面から、

- セマンティックコンピューティング
- 知的で創造的な基盤実現のための技術
- ユビキタスコンピューティング
- 何時でも、何処でも、何でも状態の基盤実現のための技術
- セキュアコンピューティング
- 安全で、安定し、安心できる基盤実現のための技術

構成面から、

- エージェントコンピューティング
- グリッドコンピューティング

人間・社会とコンピュータとの共通となる意味世界を構成する技術として、オントロジー技術がある。しかし、オントロジーは、語彙のレベルだけで意味世界を対象にしようとい

うものである。重要なのは、その語彙を用いてまとまった情報を表現するための言語を含めて意味世界共通化の仕組みを作ることである。SeC の目標はそこにある。

言語としては、かつてのように知識表現言語のような人工的な形式言語を工夫する方法もある。もうひとつの方法は、人間・社会が用いている言語やメディアをコンピュータも用いることができるようにする方法である。SeC はこの方法に関する初めての本格的な取り組みとなる。SeC はこの方法を実現する手立てのために共通のインタフェースとなる概念記述言語 (CDL : Concept Description Language) を中核に置く。

概念記述言語は次のような仲介役を果たすように設計される。

メディア間の仲介言語、メディア間には、多言語間、自然言語と形式言語 (数学言語、プログラム言語など) 間、イメージメディアやビジュアルメディアと言語メディア間  
概念レベル意味処理からより深い意味処理や知識処理への仲介言語  
構文構造処理 (XML) と意味構造の仲介言語

表現メディアを用いて、まとまった情報 (知識や感情) を表現したものがコンテンツ (ドキュメントなどを含む広義のコンテンツ) である。コンテンツは人間・社会での情報の外在化の形態である。一方、コンピュータにおける情報の形態はソフトウェアやデータである。コンピュータが人間・社会と意味世界を共通にすると、コンピュータにとってコンテンツもソフトウェアとなるということである。そのために、概念記述言語は表現メディアからコンテンツの概念構造までを連続的に対象とするように設計される。

すべてのサービスは、コンテンツを仲介役とすることによって成り立つ。コンテンツを直接の対象とするサービス産業がいわゆるメディア産業 (放送、出版、映画など) である。それには、コンテンツの制作サービスと配信サービスが含まれる。

## 2. SeC の技術アーキテクチャ

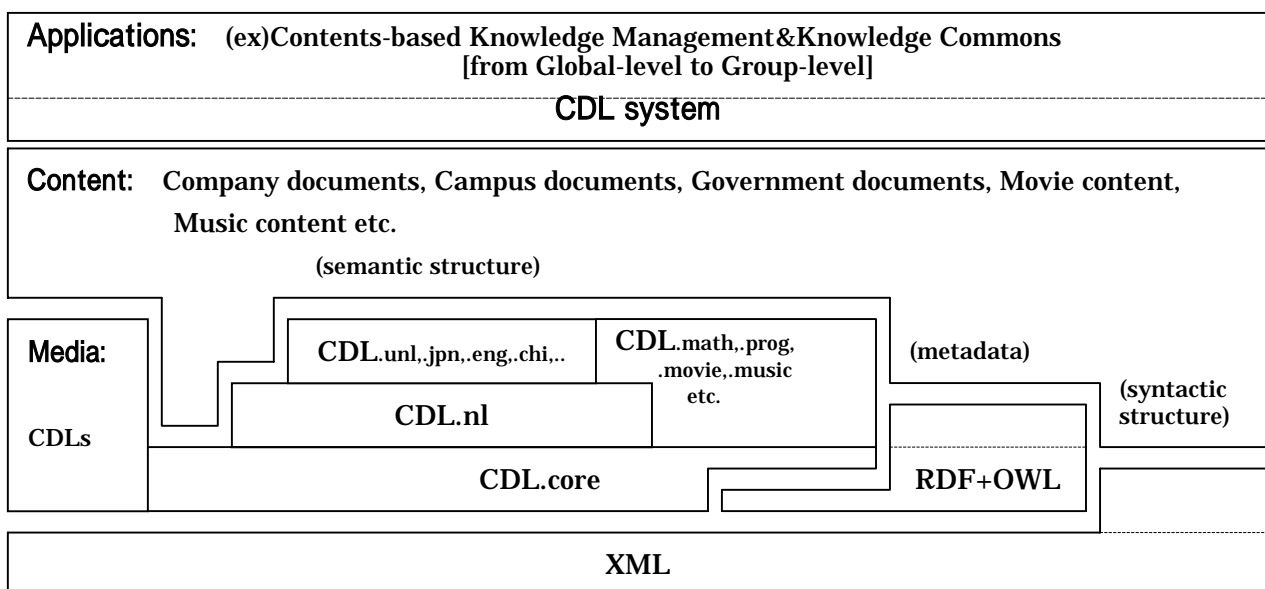
CDL と CDL システムが大枠としての SeC の基幹技術を構成する。CDL は、CDL.core を核として多数の言語群 (CDLs) から構成され、CDL システムは、多数のソフトウェアやデータベース (知識ベースや辞書) から構成される。CDL と CDL システムを詳細化することによって、SeC のミドルウェアとしての技術アーキテクチャが明らかとなる。

### 2.1 CDLs

CDL 全体、および、XML やセマンティック Web との関係を図に示す。CDL 全体を CDLs と名付ける。この CDLs の内で、中心となるのは、コア部分となる CDL.core と自然言語共通部分となる CDL.nl である。この CDL.core と CDL.nl の上に各種表現メディア対応の CDL や各種コンテンツ対応の CDL が開発される。これは、ちょうど XML (eXtended Markup Language) の上に各種表現メディア対応や各種コンテンツ対応のマークアップ言語 (タグセット) が開発されるのと同様のことである。

XML がコンテンツの構文構造（文書構造）のマークアップを行うのに対して、CDL はコンテンツの意味構造（概念構造）のマークアップと記述を行う。XML では、DTD（Document Type Definition）によって目的ごとのタグセットを定義するのに対して、CDL では、CTD（Concept Type Dictionary）によって目的ごとの概念セットの定義を行う。

XML では、XSL（XML Stylesheet Language）とブラウザによって人間に見える形に文書構造を可視化するのにに対して、CDL では、変換辞書と文生成プログラムによって人間に読めるように概念構造を可読化する。XML には、エディタやブラウザやチェッカなどさまざまなソフトウェアが開発されているが、CDL においても、これら変換辞書や文生成プログラムを含め CDL システムと呼ばれるソフトウェア群が開発される。



CDL.core を共通コアとして、概念タイプ辞書（CTD:Concept Type Dictionary）の内容によって様々な表現メディア対応のファミリー言語が生まれる。さらに、それらファミリー言語を用いて、それぞれのコンテンツ形式対応の CDL が作られることになる。それらは、各々のコンテンツ形式の概念構造を記述するための概念タイプを定義した CTD を用意することによって実現される。

さまざまな表現メディアを用い、互いに連携させることにより、コンテンツが表現される。したがって、コンテンツの意味処理（概念処理）には、それらの表現メディアを一貫して概念を扱う仕組みが必要になる。概念の分節化と記号化によって、すべての表現メディアを概念レベルで一貫する役割をになうのが自然言語である。そこで、CDLs においても、自然言語が表現メディアの中心的な役割を担うことになる。

自然言語も世界では数千といわれる言語からなる。言語にまたがるコンテンツという視点が必要になる。そこで、概念を扱う仕組みに関して、言語に共通となる部分と各言語固有の部分に適切に分けるという観点が重要になる。その観点から設定されたのが、自然言語共通部分に対応する CDL.nl である。

**CDL.core :**

すべての基盤となる言語部分である。データモデル、概念タイプ定義（クラス定義）、概念タイプ階層の定義が定められる。それらにかかわる実体概念と関係概念が定義された CTD からなる。

**CDL.nl :**

すべての自然言語（natural language）に共通の概念が定義された CTD によって定められる。概念タイプ階層の最上位レベルに位置する語概念、それらによって構成される事象概念や状況概念の定義、それらの定義に用いられる関係概念や属性 - 属性値のセットが CTD に定義される。

**CDL.unl :**

国連大学・UNDL 財団（<http://www.undl.org/>）で開発され普及が進められている国際共通言語、UNL（Universal Networking Language）の CDL 版である。CTD に世界共通となる語概念や事象概念が定義される。

**CDL.jpn, CDL.eng, CDL.chi etc. :**

日本語、英語、中国語など、各国言語対応の CDL である。言語ごとに、その語概念や事象概念を定義する CTD が用意される。日常言語対応の CTD、専門分野やコミュニティごとの CTD、制限言語対応の CTD などを用意することによって、多様な言語使用に対応することができる。

**CDL.math, CDL.prog, CDL.movie, CDL.music etc. :**

数式、プログラム言語、動画像、音楽音などに対する CDL である。それぞれの表現メディア固有の概念と自然言語から援用した概念によって CTD が構成される。

CDLs は XML を用いて実装される。CDL.core は SW（Semantic Web）における RDF+OWL 部分におおよそ対応する。ただし、設計理念は対極的である。RDF+OWL の設計理念は、リソースの特性記述である。一方、CDL の設計理念は、コンテンツの構造記述である。

コンテンツもリソースである。したがって、RDF+OWL は、コンテンツの特性記述、すなわち、コンテンツのメタデータをマシン理解処理可能にするのが目的である。何らかのメディアで表現されたものとするリソースもコンテンツとなる。CDL はリソース本体の概念構造をマシン理解処理可能にするのが目的である。

構造記述と特性記述は、対極的ではあるが、相補的である。したがって、CDL と RDF+OWL とは対極的であり、相補的である。そこで、CDL の仕様策定にあたっては、

類似部分は極力 RDF+OWL の仕様を活用し、また、RDF+OWL 上に実装する方式も提案する。

コンテンツの構造に関しては、XML が構文的構造のマークアップ機能を提供する。コンテンツの構文構造は、入れ子ツリー構造によって近似される。一方、コンテンツの概念構造は、入れ子ネットワーク構造によって近似されるとする。構文構造は意味構造をガイドするものであることから、XML によるコンテンツの要素タグに意味構造を付加するという機能も CDL に求められる。そして、コンテンツ各部の意味を概念タイプ定義辞書の概念で近似することによって、コンテンツから分離された概念構造として処理可能な記述データが得られることになる。

## 2.2 CDL システム

CDL に関わるソフトウェア全体をまとめて CDL システムと呼ぶ。CDL システムは人、および、コミュニティとコンピュータが CDL シンボルの意味（概念）を共有し、創生していくためのシステムである。柔軟なシステム構成が可能となるよう、コミュニティシステムはエージェント技術を活用して実装される。

CDL システムの基本機能には、以下のようなものが含まれる。

表現メディアと CDL 間の相互変換機能（機械変換とアノテーション）、いわゆる、  
グラウンディング機能

辞書ベース、用語ベース、知識ベースの協調構築・利用機能、グループやコミュニティにおける知識共有、知識共通化機能

CDL と各種推論・知識処理との連携機能

CDL システムの基本構成には、以下のようなものが含まれる。

### (1) CDL.lang オーサリングシステム (lang = { unl, jpn, eng, chi, etc. } )

NtoC コンバータ：自然言語 (Natural Language) テキストから CDL.lang テキストへの変換を簡単なアノテーションだけで行うソフトウェア

CtoN コンバータ：CDL.lang テキスト、あるいは、CDL.lang でマークアップされたテキストを自然な自然言語テキストに変換するソフトウェア

CDL.lang エディタ：自然言語テキストへの CDL.lang によるマークアップを知的に支援するソフトウェア

### (2) CDL.nl 概念タイプ定義辞書システム

語概念、事象概念、関係概念、さらに、テキスト概念に渡る定義ベースとその推論・検索エンジン、CDL.nl と自然言語間の変換辞書を含む

### (3) CDL ドキュメント知識ベースシステム

CDL ドキュメント、あるいは、CDL でマークアップされたドキュメントと推論規則の知識ベースとその分類・検索エンジンや変換・推論エンジン

### (4) CDL イントラコミュニティシステム

コミュニティ内でネットワークを介して辞書や知識ベースの共同開発・維持管理・共有利用を行うためのエージェントで構成されたシステム

(5) CDL インターコミュニティシステム

コミュニティ間で辞書や知識ベースを交換し合うためのエージェントで構成されたシステム

### 3. SeC 技術の応用

SeC ミドルウェアの上に、セマンティック Service、あるいは、セマンティック SPM (Service Process Management) と呼ばれる各種の応用ソフトウェアが開発される。

セマンティック SPM の分野としては、以下のようなものが考えられる。

#### ビジネスプロセス分野

##### 技術ビジネスプロセスマネージメント

特許情報ドキュメント、技術新聞記事、知的財産権法令・判例集、技術情報ウェブページ、技術論文等を対象に技術ビジネスにおけるダイナミックで高度な知識マネージメント機能の実現を目指す。企業が主体となって推進する課題である。

##### コンテンツビジネスプロセスマネージメント

コンテンツビジネスにおける企画、制作、製作、配信という全工程を統合的に知的にマネージメントする機能の実現を目指す。

#### 研究開発プロセス分野

##### 研究開発プロセスマネージメント

論文・報告書ドキュメント、研究業務管理情報、学会誌・会議録、協調・協創作業環境等を対象に研究開発プロセスの柔軟で高度な知識マネージメント機能の実現を目指す。研究開発機関が主体となって推進する課題である。

##### 設計・製造・試験・運用保守プロセスマネージメント

システム仕様書、設計書、試験運用保守マニュアル等の記述を統一し、機能レベルでの相互参照を容易にし、機械の判読を可能にすることで、ノウハウの明示化、知識の継承、設計品質・開発効率の向上を図る。IT 業界と各製造業界が協力して構築する課題である。

#### 教育学習プロセス分野

##### 教育学習プロセスマネージメント

大学における教育・学習環境を統合的に高度化する新システムを研究・開発・構築する。いわゆる教務システムなどは除き、教員と学生による教育・学習活動に焦点を絞る。複数の大学が共同で行うプロジェクト、それぞれの大学ごとのプロジェクト、外国語教育など分野を絞ったプロジェクトなど、様々な形態のプロジェクトを想定する。大学が主体となって推進する課題である。

分野ごとのセマンティック SPM に対して、共通のサービス機能として以下のようなものを設定する。

セマンティック特許情報処理

セマンティック Documentation を含むセマンティック Authoring (コンテンツの制作支援)

セマンティック Specification を含むセマンティック Programming (ソフトウェアの制作支援)

#### 4. SeC 計画の推進

- SeC 技術説明会：3月10日(木) 13:00～17:00 於産総研
- CDL (.core&.nl) の国際標準化に関する SCOPE-I (3年間) への応募：4月末採否決定
- ワークショップ “ Semantic Computing Initiative –From Semantic Web to Semantic World ”：5月10日(火) 於幕張メッセ  
W3C 主催 WWW2005 の併設ワークショップとして開催する。CFP が <http://www.instsec.org/2005ws/>にある。

#### [ 連携するプロジェクト ]

- AIST-SOA ( Service Oriented Architecture ) プロジェクト  
産総研が目指す次世代情報サービス基盤、サービス主導アーキテクチャ研究開発プロジェクト(平成17年度～21年度)。SeC計画と密に連携する。具体的には、CDLシステムの研究開発はAIST-SOAで行い、ISeCがCDLの仕様策定、国際標準化、共通辞書の開発・維持管理を行うなどの協力体制を作る。
- Semantic Web プロジェクト
- UNL プロジェクト

## セマンティック・コンピューティング技術説明会 - Concept Description Language (CDL) とその応用 -

ISeC が検討を進めてきました CDL ( Concept Description Language ) の基本仕様第 1 版とその役割・応用について説明会を開きます。CDL はセマンティック・コンピューティング ( SeC ) 技術における様々な研究開発活動を連携させ効果的に推進させるための共通言語です。また、日本においても立ち上がってまいりましたセマンティック Web ( SW ) 技術開発の活動と適切に連携するための技術手段としての役割もあります。説明会は、SeC と SW を総合的に議論し、意見交換する場でもあります。奮ってご参加ください。また、説明会における議論を 5 月幕張にて開催されます WWW2005 の併設ワークショップ “ The Semantic Computing Initiative (SeC 2005)--- From Semantic Web to Semantic World--- ” ( <http://www.instsec.org/2005ws/> ) へと発展させたいと思っております。

---

### 共 催

特定非営利活動法人セマンティック・コンピューティング研究開発機構 (ISeC)  
独立行政法人産業技術総合研究所 (AIST) 情報技術研究部門

### 日 時

平成 17 年 3 月 1 0 日 ( 木 ) 1 3 : 0 0 ~ 1 7 : 0 0

### プログラム

13:00~13:20 セマンティック・コンピューティングと CDL  
13:20~13:50 産業技術総合研究所におけるセマンティック・コンピューティング  
13:50~14:50 CDL.core について  
14:50~15:10 休憩  
15:10~16:10 CDL.nl について  
16:10~16:50 総合討論  
16:50~17:00 SeC2005 と ISeC

### 場 所

独立行政法人産業技術総合研究所 臨海副都心センター 第 1 会議室  
住所 東京都江東区青海 2-41-6  
地図 [http://www.aist.go.jp/aist\\_j/guidemap/guidemap\\_main.html](http://www.aist.go.jp/aist_j/guidemap/guidemap_main.html)

### 申込方法

準備の都合上、3 月 4 日 ( 金 ) までに、氏名、所属機関名、役職、メールアドレス  
または電話番号を、下記の申込・問合せ先へメールにてお送りください。  
(定員 42 名、無料)

### 申込・問合せ先

[info@instsec.org](mailto:info@instsec.org)