Achievements in and Future Development of the Coordination Program "Very Large Information Integration and Application Platform"

Coordination Program of Science and Technology Projects
"Very Large Information Integration and Application Platform"
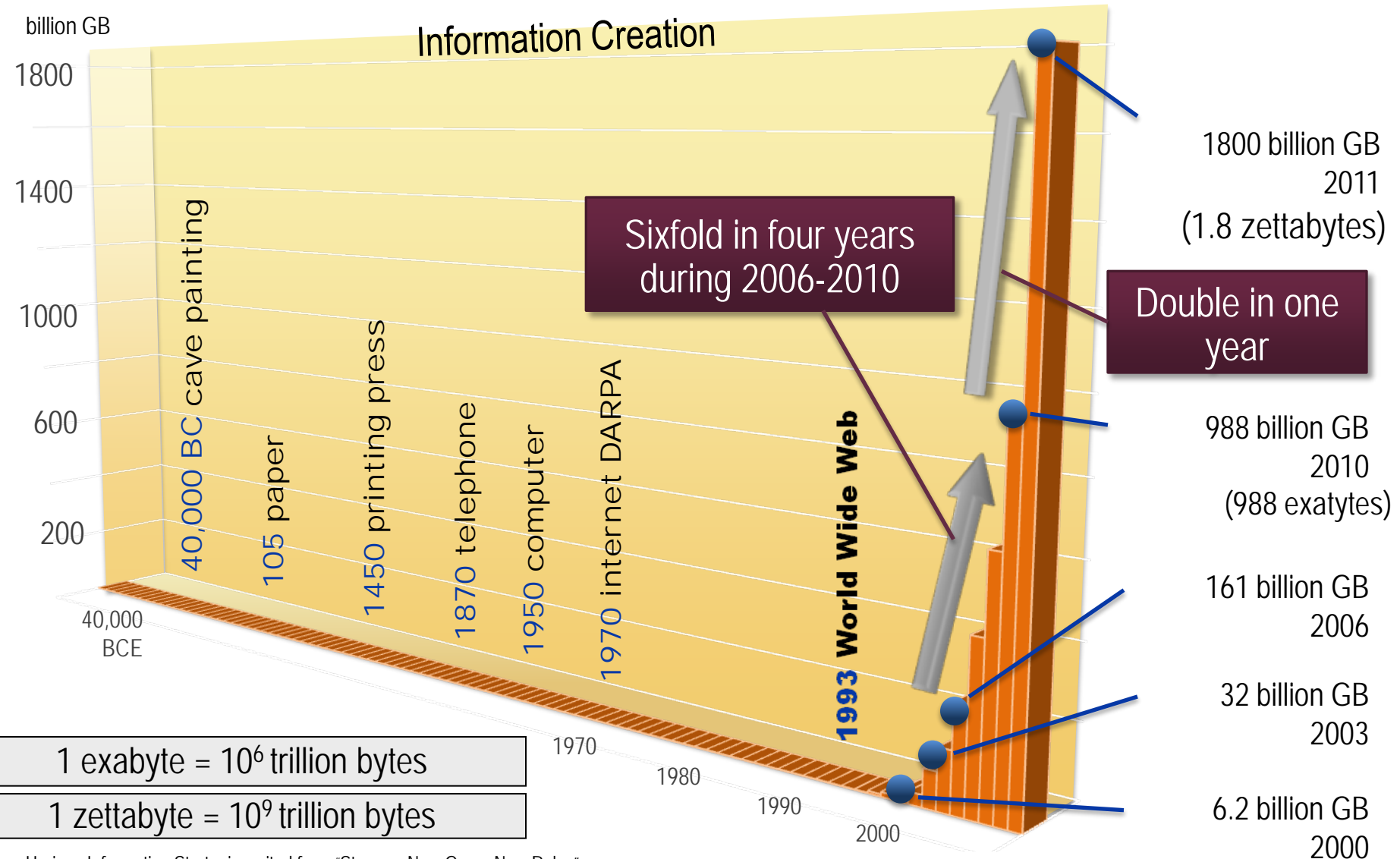Project Director: Shojiro Nishio, Ph. D.

December 1, 2009

Agenda

1. Background and Purpose

2. Coordination Program "Very Large Information Integration and Application Platform"

3. Major Achievements in the Projects of Each Ministry and the Complementary Subject

4. Summary of the Three-Year Activity of the Coordination Program

5. Future Development

1. Background and Purpose

Advent of the Age of Information Explosion

Information Creation

billion GB

1800

1400

1000

600

200

40,000 BC cave painting

105 paper

1450 printing press

1870 telephone

1950 computer

1970 internet DARPA

1993 World Wide Web

Sixfold in four years during 2006-2010

Double in one year

1800 billion GB
2011
(1.8 zettabytes)

988 billion GB
2010
(988 exatytes)

161 billion GB
2006

32 billion GB
2003

6.2 billion GB
2000

40,000
BCE

1970

1980

1990

2000

1 exabyte = 10^6 trillion bytes

1 zettabyte = 10^9 trillion bytes

Source: Horison Information Strategies, cited from "Storage: New Game New Rules"

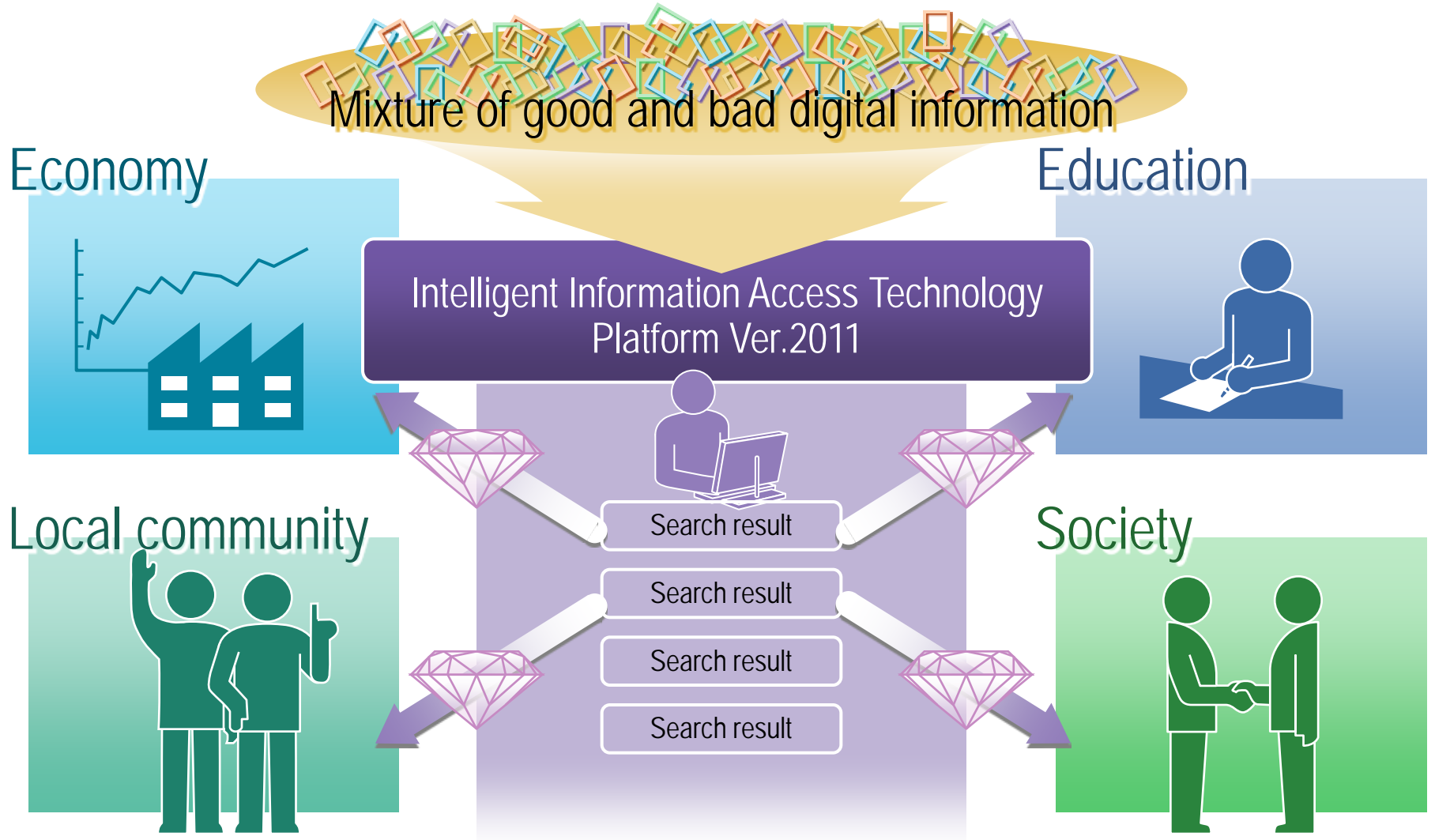Source: Information Data Corporation, "The Diverse and Exploding Digital Universe", 2008

Examples of Societal Problems Faced by Existing Search Services

- Current search engines does not analyze and present information from the user's perspective.
- As a result, the possibility of garbage and harmful information being presented at the top of search results exists.
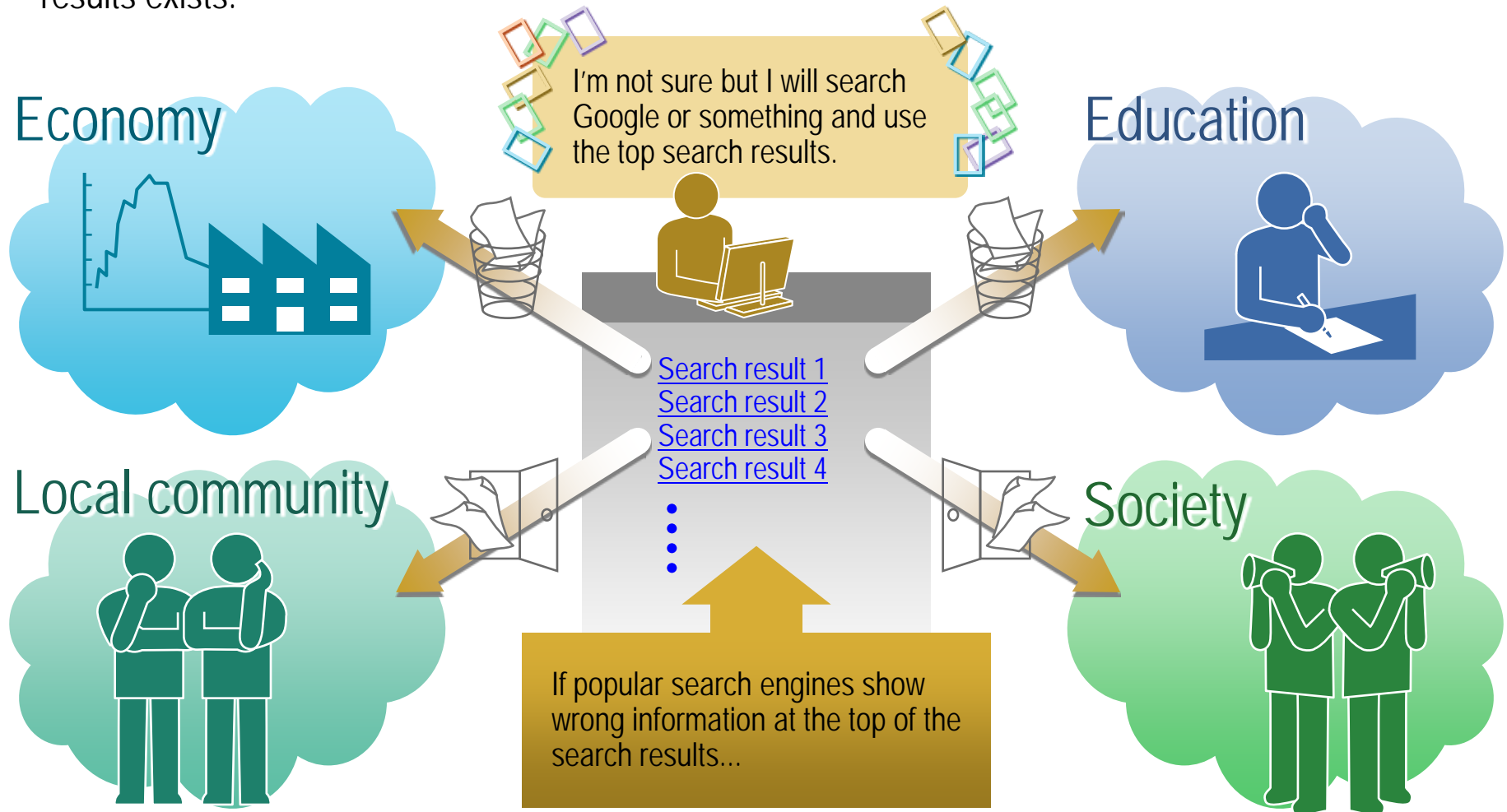
Visage of a Society Actively Using an Intelligent Information Access Platform

Target: Building society through the use of safe and intelligent services

Mixture of good and bad digital information

Economy

Education

Intelligent Information Access Technology Platform Ver.2011

Search result

Search result

Search result

Search result

Local community

Society

2. Coordination Program "Very Large Information Integration and Application Platform"

Promotion Scheme for Coordination Program of Science and Technology Projects

Council for Science and Technology Policy (plenary assembly)

Investigation Committee for Promoting Basic Policy

Cabinet Office

Minister in charge of science and technology policy

General PT (Project Team) for Promoting Sectoral Strategy

Coordination Program of Science and Technology Projects (6 programs)

Life Science PT

- Food and Biological Production Research
- Clinical Research and its Translational Research

Information and Communication PT

- Very Large Information Integration and Application Platform

Director-General for Policy Planning

Environment

- Research and Development on the Safe Use of Chemical Substances Based on Comprehensive Risk Evaluation

Nanotechnology and Materials PT

- Platform Development for the Promotion of Nanotechnology Research and Development And its Social Acceptance

Societal platform PT

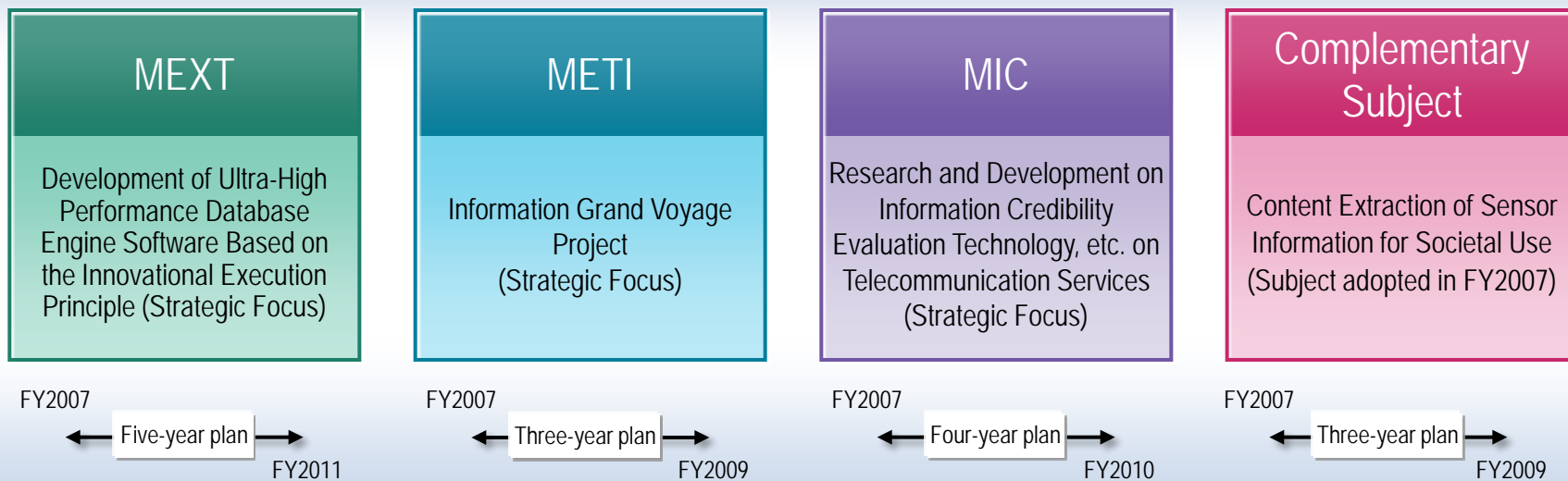- Research and Development for Counterterrorism

Secretariat

Coordination Program "Very Large Information Integration and Application Platform": Four projects were launched in FY2007
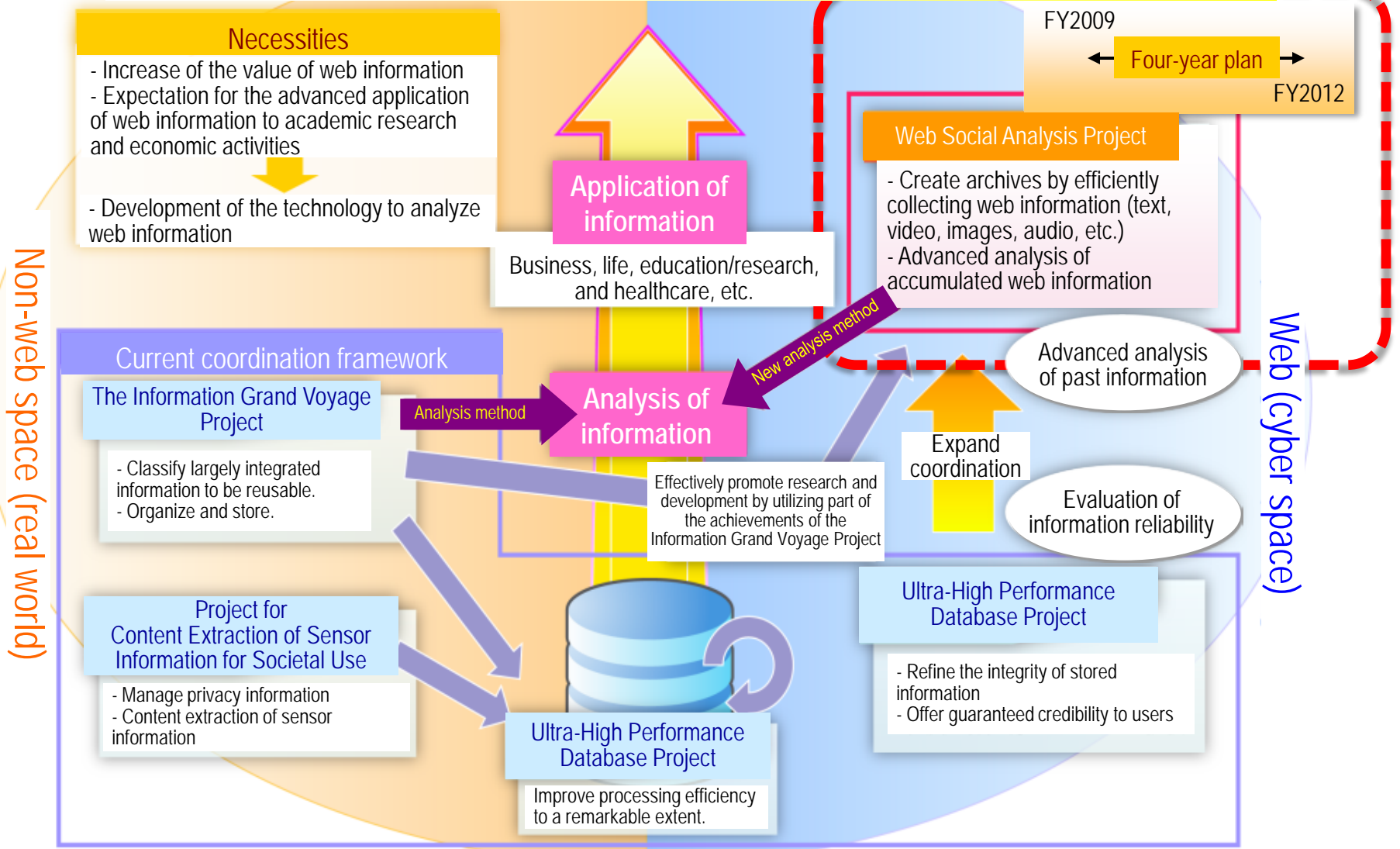
Objectives

To develop the platform technology for intelligent information application of next generation, for safe, appropriate, and secure information collection, analysis, and management of various information (content) to provide unique information service

Coordination of Each Ministry's Strategic Project

MEXT	METI	MIC	Complementary Subject
Development of Ultra-High Performance Database Engine Software Based on the Innovational Execution Principle (Strategic Focus)	Information Grand Voyage Project (Strategic Focus)	Research and Development on Information Credibility Evaluation Technology, etc. on Telecommunication Services (Strategic Focus)	Content Extraction of Sensor Information for Societal Use (Subject adopted in FY2007)

FY2007 ← Five-year plan → FY2011

FY2007 ← Three-year plan → FY2009

FY2007 ← Four-year plan → FY2010

FY2007 ← Three-year plan → FY2009

Expanding the Coordination Program "Very Large Information Integration and Application Platform" (Five projects since FY2009)

Goal
To build the next-generation information search and analysis technology platform that would enable a person to collect and analyze necessary information from diverse information inside and outside the web easily, accurately, and safely, and to enable the strategic application of super large-volume data during this age of information explosion by the year 2012

Non-web space (real world)

Web (cyber space)

Necessities
- Increase of the value of web information
- Expectation for the advanced application of web information to academic research and economic activities

- Development of the technology to analyze web information

Application of information

Business, life, education/research, and healthcare, etc.

FY2009
Four-year plan
FY2012

Web Social Analysis Project
- Create archives by efficiently collecting web information (text, video, images, audio, etc.)
- Advanced analysis of accumulated web information

New analysis method

Current coordination framework

The Information Grand Voyage Project

Analysis method

Analysis of information

- Classify largely integrated information to be reusable.
- Organize and store.

Effectively promote research and development by utilizing part of the achievements of the Information Grand Voyage Project

Advanced analysis of past information

Expand coordination

Evaluation of information reliability

Project for Content Extraction of Sensor Information for Societal Use

- Manage privacy information
- Content extraction of sensor information

Ultra-High Performance Database Project
- Refine the integrity of stored information
- Offer guaranteed credibility to users

Ultra-High Performance Database Project

Improve processing efficiency to a remarkable extent.

3. Major Achievements in the Projects of Each Ministry and the Complementary Subject

Information Grand Voyage Project

Ministry of Economy, Trade and Industry

http://www.igvpj.jp/index/

http://www.igvpj.jp/index_en/

Information Grand Voyage Project

Purpose

To promote further maturity of the "institution and environment" and further growth of "technology" with the success of the next-generation technology-related business as a trigger, and to establish an innovation creation mechanism leading to the creation of business, including more advanced public activity, in order to activate industry and strengthen competitiveness

Field testing next-generation technology

Evaluate the effectiveness and feasibility of next-generation information search and analysis technology developed by the field test business.

Establish a mechanism in which innovation is created spontaneously through application of information.

Activate industry and strengthen competitiveness

Upgrading the system and environment

- n Upgrade systems to protect privacy and copyrights.
- n Create a mechanism for smooth distribution of intellectual properties.
- n Upgrade the environment for development and field testing

Technical development

- n Develop the next-generation search and analysis technology.
- n Generalize and share the next-generation search and analysis technology (upgrading common technologies).

Major Achievements: Field Testing of the Next Generation Technology

The field testing of next-generation technology

Twenty-two field tests were performed from the following three directions, and it was found out that the services are highly feasible.

A: Future-oriented personal service considering privacy

By safely collecting and analyzing personal information (profiles, action histories, and preferences), user convenience and information-collecting capabilities can be improved and new services can be created.

B: Next-generation web services that generate new value

By solving problems caused by ambiguous and abstract user intention during information searches and by utilizing rich content, the improvement and expansion of market value can be expected.
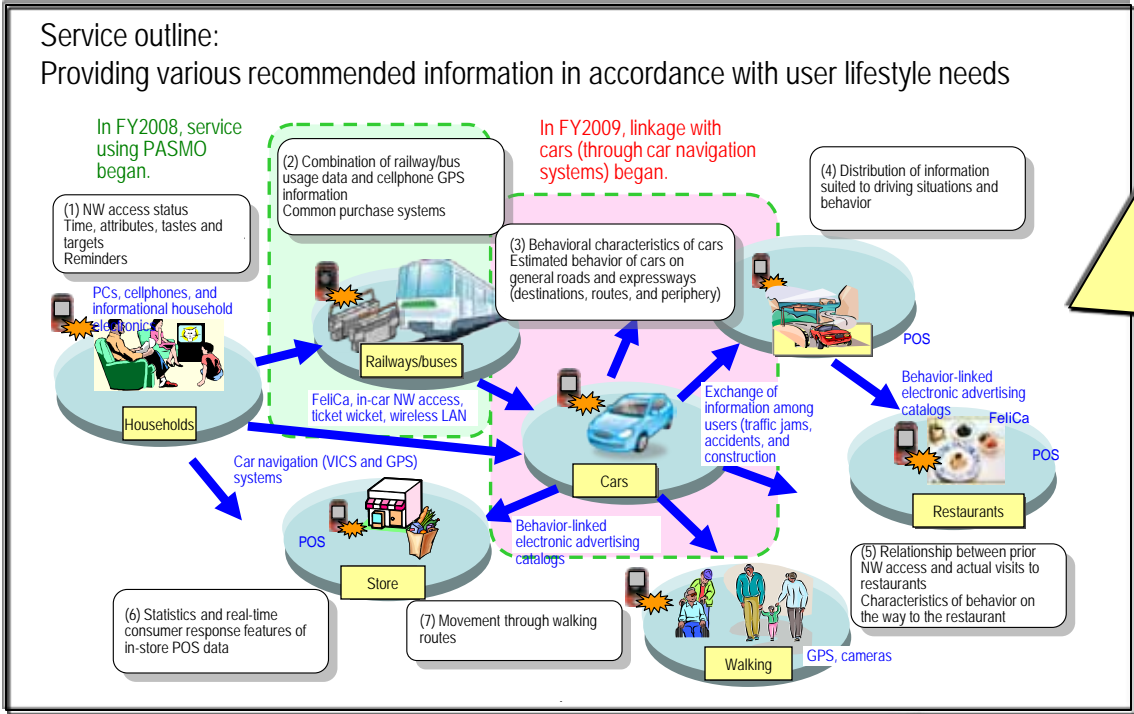
C: IT services of the new social infrastructure

By constructing a mechanism of detecting the possibility of accidents and diseases in advance and by taking appropriate countermeasures against them through the utilization of next-generation search and analysis technology, various risks can be avoided or reduced.

Concrete Example of Field Testing of the Next Generation Technology

The My Life Assist Service: NTT Docomo

- By implementing the "My Life Assist Service", which transmits information leading to new findings and discovery, based on the information of ordinary behavior collected from cellphones, the benefit of the service was evaluated.
- The know-how obtained from this service was expanded to "i-Concierge."

Service outline:
Providing various recommended information in accordance with user lifestyle needs

In FY2008, service using PASMO began.

In FY2009, linkage with cars (through car navigation systems) began.

(1) NW access status
Time, attributes, tastes and targets
Reminders

PCs, cellphones, and informational household electronics

Households

(2) Combination of railway/bus usage data and cellphone GPS information
Common purchase systems

Railways/buses

FeliCa, in-car NW access, ticket wicket, wireless LAN

Car navigation (VICS and GPS) systems

(3) Behavioral characteristics of cars
Estimated behavior of cars on general roads and expressways (destinations, routes, and periphery)

Exchange of information among users (traffic jams, accidents, and construction)

Cars

(4) Distribution of information suited to driving situations and behavior

POS

Behavior-linked electronic advertising catalogs

FeliCa

POS

Restaurants

(5) Relationship between prior NW access and actual visits to restaurants
Characteristics of behavior on the way to the restaurant

Behavior-linked electronic advertising catalogs

POS

Store

(6) Statistics and real-time consumer response features of in-store POS data

(7) Movement through walking routes

Walking   GPS, cameras

Service detail:
When opening a folding cellphone, contents pertinent to your everyday needs can be shown.

Propose a new finding by predicting a user's next behavior pattern from past data.

I worked really hard this week. I think I will give myself a reward.

Give a reward to yourself on a day off:
(1) Here is a list of good restaurants!
(2) Do you want to invest in yourself?

i-Concierge

Over 24 million subscribers
(as of October 14, 2009)

Unit: 10,000 subscribers

100

i-Concierge

08/11   12   09/1   2   3
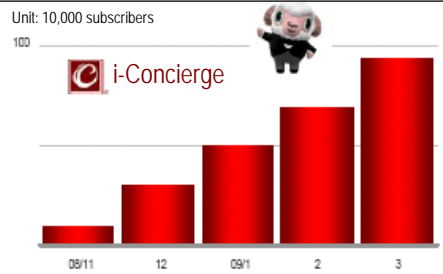
Target for FY2009:
38 million subscribers

Major Achievements: Common Technologies

Common technologies

In collaboration with taskforces consisting of those in industry, government, and academia, focused items and development details were formulated. As next-generation search and analysis technologies with high versatility and commonality, practical technologies that are directly related to services and cutting-edge basic technologies that complement them (a total of more than 50 technologies) were developed.

Service common technology	Practical technology that is directly related to the services devised by development and field-testing companies and that should be extracted and shared as a next-generation search and analysis technology, due to high versatility and commonality (focused items are determined by taskforces consisting of those in industry and academia and through benchmark investigations)

Platform common technology	Technology that is not devised through development and field-testing companies, but is required for social life platforms in the future; a next-generation search and analysis technology based on cutting-edge research that complements the service's common technology (development details are determined by taskforces consisting of those in industry, government, and academia, and through the council)

My Life Assist Service · Collaborative services utilizing traffic non-contact IC cards · Profile passport · Laddering search service · Multi-language video application platform · View search Hokkaido · SAGOOL TV · Time-space data mining services · Health control using mobile information terminals based on the next generation analysis technology · Sukoyaka Life Support Service · New comprehensive safe driving support systems · Mega research · "Kokonara" Communication Service · Distribution service through a cross-industrial common ID system · e-space service supporting the activation of local communities "Plat"

Extracted common technologies

Recommendations

Personal information protection

Video and text analysis

Service collaboration platform

User interfaces

Research and Development on Information Credibility Evaluation Technology, etc. on Telecommunication Services
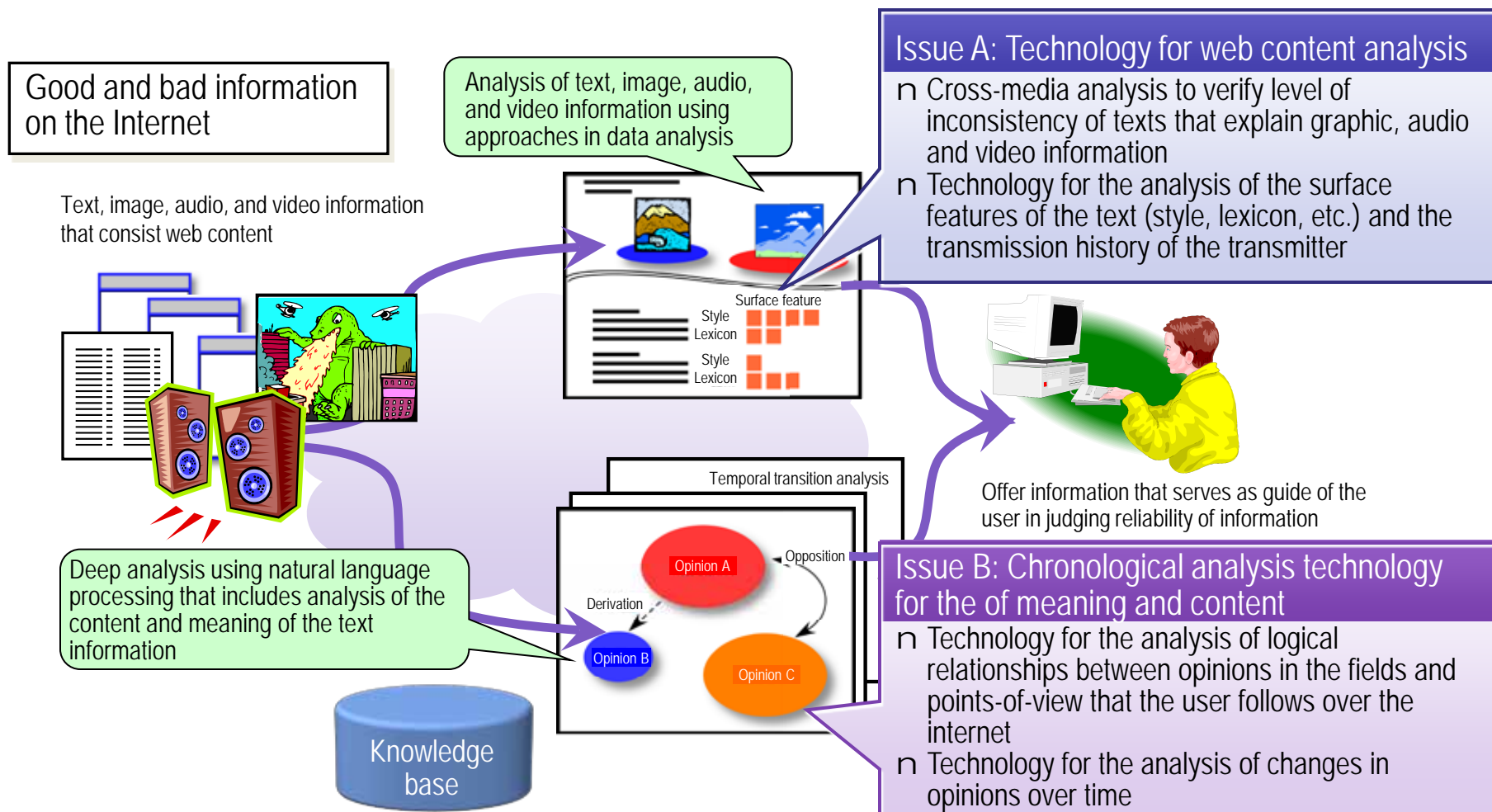
Ministry of Internal Affairs and Communications

http://kc.nict.go.jp/project1/

Information Credibility Evaluation Technology

information credibility evaluation technology

Develop technologies that enable users to use of web content to their benefit easily by analyzing and presenting good and bad information that flows through the Internet based on the user's judgment criteria.

Good and bad information on the Internet

Text, image, audio, and video information that consist web content

Analysis of text, image, audio, and video information using approaches in data analysis

Surface feature
Style
Lexicon
Style
Lexicon

Issue A: Technology for web content analysis

- ■ Cross-media analysis to verify level of inconsistency of texts that explain graphic, audio and video information
- ■ Technology for the analysis of the surface features of the text (style, lexicon, etc.) and the transmission history of the transmitter

Offer information that serves as guide of the user in judging reliability of information

Deep analysis using natural language processing that includes analysis of the content and meaning of the text information

Temporal transition analysis

Opposition

Opinion A

Derivation

Opinion B

Opinion C

Issue B: Chronological analysis technology for the of meaning and content

- ■ Technology for the analysis of logical relationships between opinions in the fields and points-of-view that the user follows over the internet
- ■ Technology for the analysis of changes in opinions over time

Knowledge base

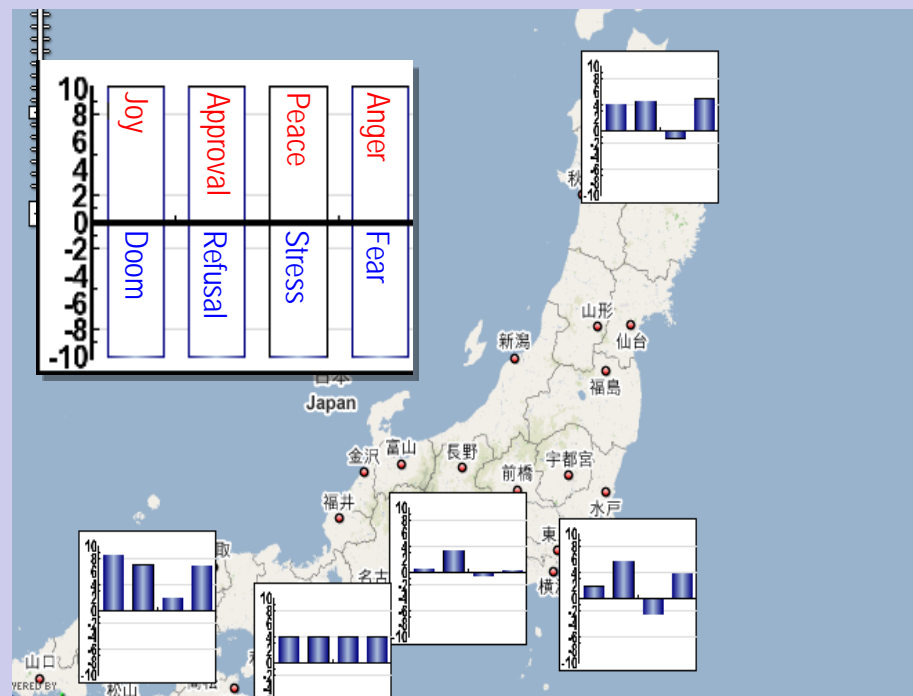# Major Achievements: "Technology for Web Content Analysis"

Develop an information transmitter analysis engine based on the transmitter's familiarity and sentiment analysis through use of call history information.

▮ What is the sentiment of the information transmitter who transmits information?
▮ To what degree is the transmitter familiar with the topic transmitted?
Technology to measure the reliability of transmitters, including newspaper sites, countries, and bloggers for each topic was developed by analyzing the above items

1. A sentiment expression dictionary was constructed and improved.
2. Sentiments of each article were extracted.
3. Sentiments by topic, transmitter, and organization were detected.
4. Differences in sentiments among topics, transmitters (individuals, organizations, and groups) and articles were analyzed.
5. Reliability judgment support was implemented by presenting differences.

Sentiments for certain topics were analyzed by region.
Example topic: the "Yomiuri Giants"

Major Achievements: "Technology for the Chronological Analysis of Meaning and Content"

Information reliability analysis

- **n** Purpose: Support unbiased judgment from a broader viewpoint.
- **n** Measure: Analyze the position in the language space for the information of interest and report its panoramic summary to users

You can lose weight easily by eating bananas in the morning.
It is very easy, and the results can be surprising.

You can lose weight by eating bananas in the morning!

Can we trust this information?

Position the information in
the logic space
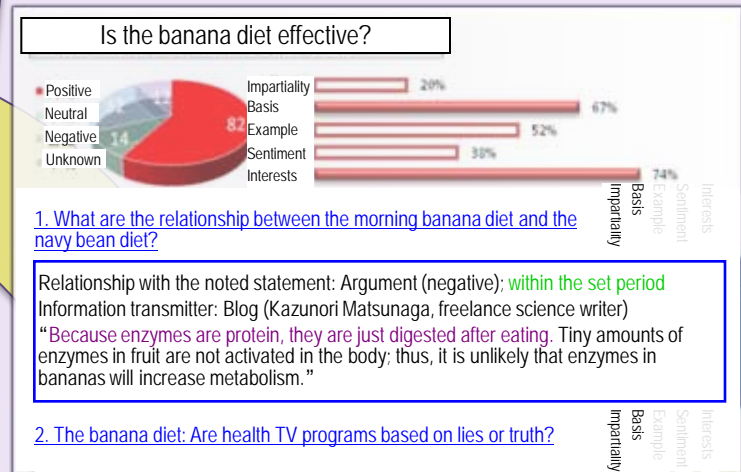What is written in other pages?

Are there any opposing views?

What is the basis for such information?

Position the information on
the time axis
What is the background of the information?

For what purpose was it written?

Is it effective even now?

Output the analysis result comprehensively

Is the banana diet effective?

- Positive
- Neutral
- Negative
- Unknown

82   14

Impartiality   20%
Basis   67%
Example   52%
Sentiment   30%
Interests   74%

Interests / Sentiment / Example / Basis / Impartiality

1. What are the relationship between the morning banana diet and the navy bean diet?

Relationship with the noted statement: Argument (negative); within the set period
Information transmitter: Blog (Kazunori Matsunaga, freelance science writer)
" Because enzymes are protein, they are just digested after eating. Tiny amounts of enzymes in fruit are not activated in the body; thus, it is unlikely that enzymes in bananas will increase metabolism."

2. The banana diet: Are health TV programs based on lies or truth?

Interests / Sentiment / Example / Basis / Impartiality

Reconstruct the language space
Explain the information in an easy-to-understand manner

What should be noted?

Development of Ultra-High Performance Database Engine Software Based on the Innovational Execution Principle

Ministry of Education, Culture, Sports, Science and Technology

http://cif.iis.u-tokyo.ac.jp/OoODE/#

Ultra-High Performance Database Engine Based on the Out of Order Execution Principle

Purpose

Develop "Ultra-High Performance Database Engine software" that enables "strategic application of super large-volume data", in the age of the information explosion, such as sensor network information analysis and distribution traceability systems.

Issues

1. The "strategic application of super large-volume data" is a key technology as a source of national power that ensures social platforms supporting the safety and security of the nation and enables the creation of various new industries.
2. The "strategic application of super large-volume data" in the age of the information explosion cannot be implemented just by improving existing database platforms.

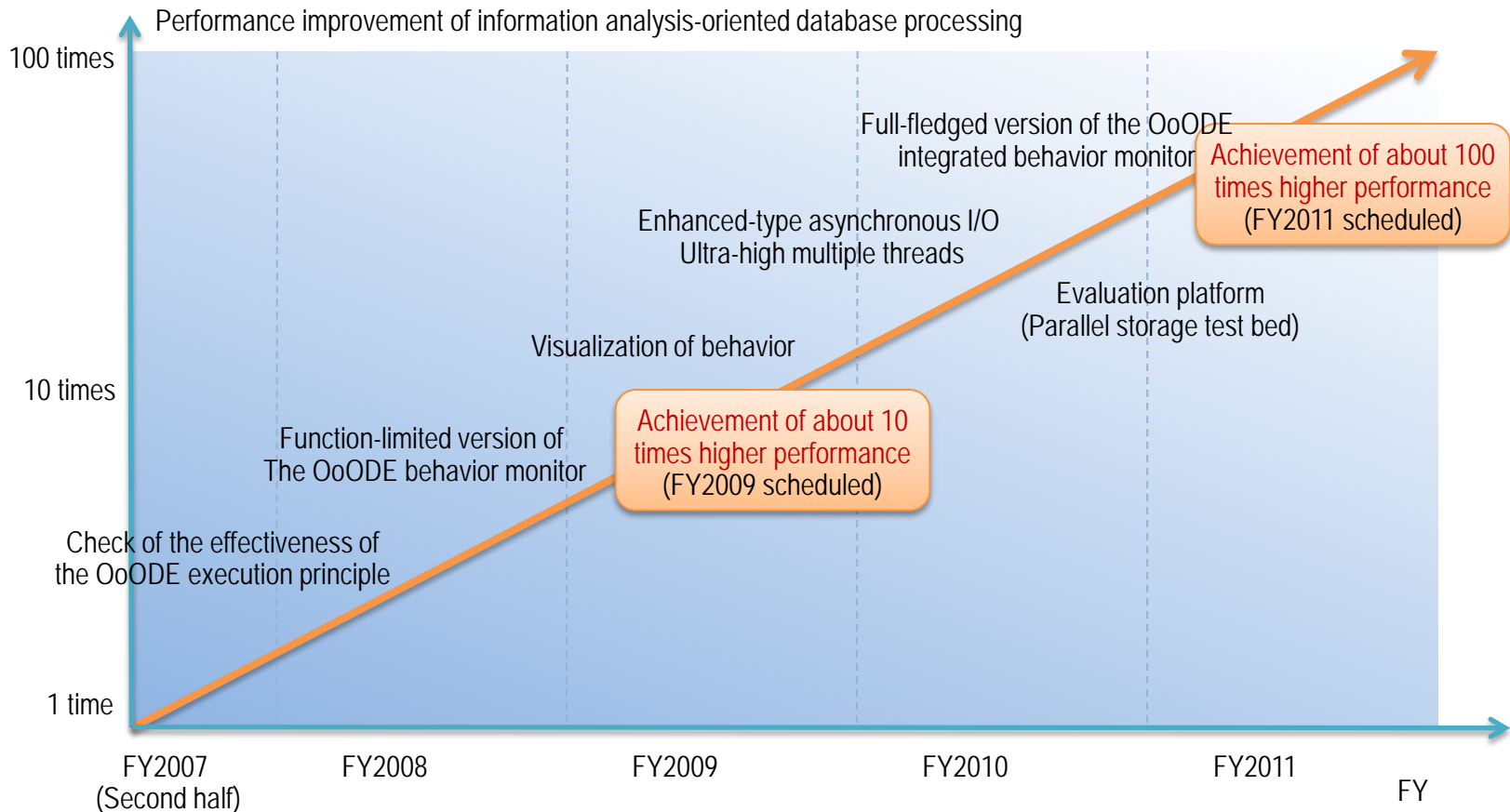**Performance breakthroughs with the innovation of the execution principle is required.**

Content of the research and development

n In order to develop an ultra-high database engine based on the out-of-order execution principle, conduct the following research and development.
   (1) Establish an innovative execution principle, the "out-of-order database execution principle".
   (2) Design and implement database platform software based on (1) above.
   (3) Verify the effectiveness through actual application use.

Achievement Goals

n By the end of FY2009, develop a database engine that applies the out-of-order execution principle to some database operations to achieve a tenfold performance improvement of information analysis-oriented database processing, compared to existing technology.

n By the end of FY2011, develop a database engine that applies the out-of-order execution principle in earnest to achieve a hundredfold performance improvement of information analysis-oriented database processing, compared to existing technology.

Performance improvement of information analysis-oriented database processing

100 times

Full-fledged version of the OoODE
integrated behavior monitor

Achievement of about 100
times higher performance
(FY2011 scheduled)

Enhanced-type asynchronous I/O
Ultra-high multiple threads

Evaluation platform
(Parallel storage test bed)

Visualization of behavior

10 times

Function-limited version of
The OoODE behavior monitor

Achievement of about 10
times higher performance
(FY2009 scheduled)

Check of the effectiveness of
the OoODE execution principle

1 time

FY2007
(Second half)   FY2008   FY2009   FY2010   FY2011
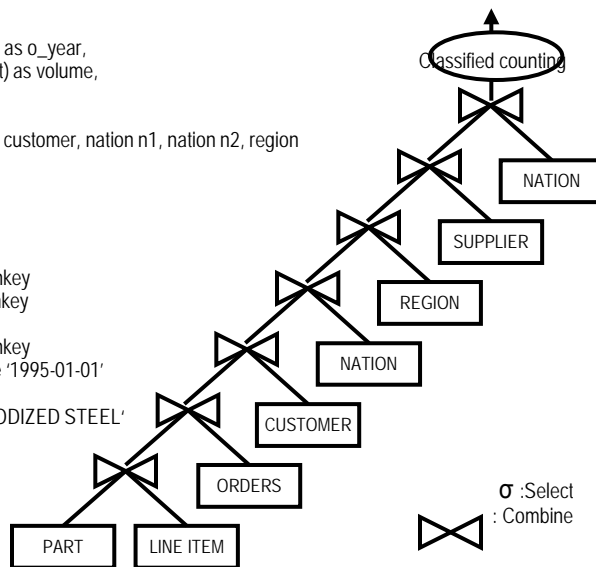
FY

Major Achievements: Experiment Results

■ Based on open source DBMS (MySQL) and commercial DBMS (HiRDB), we have developed an out-of-order database engine. In commercial DBMS-based implementation, approximately a twenty-fold performance improvement was achieved in query processing, which is very complicated processing comparing to industry standard benchmarks, while our goal was to achieve approximately a ten-fold performance improvement.

Industry standard benchmark TPC-H Query 8

By specified commodity in the specified region, changes in market share for two years was analyzed.

```
SELECT  o_year,
        sum(case when nation = 'BRAZIL' then volume else 0 end)  / sum(volume) as
mkt_share
FROM (
   SELECT
      extract(year from o_orderdate) as o_year,
      l_extendedprice * (1-l_discount) as volume,
      n2.n_name as nation
   FROM
      part, supplier, lineitem, orders, customer, nation n1, nation n2, region
   WHERE
      p_partkey = l_partkey
      and s_suppkey = l_suppkey
      and l_orderkey = o_orderkey
      and o_custkey = c_custkey
      and c_nationkey = n1.n_nationkey
      and n1.n_regionkey = r_regionkey
      and r_name = 'AMERICA'
      and s_nationkey = n2.n_nationkey
      and o_orderdate between date '1995-01-01'
          and date '1996-12-31'
      and p_type = 'ECONOMY ANODIZED STEEL'
      and p_size < 3
   ) as all_nations
GROUP BY o_year
ORDER BY o_year;
```

Classified counting

NATION

SUPPLIER

REGION

NATION

CUSTOMER

ORDERS

PART   LINE ITEM

σ :Select
: Combine

Performance improvement rate of out-of-order engines compared to conventional (sequential) engine

Out-of-order type

Conventional (sequential) type

Performance improvement rate (times)

Execution multiplicity
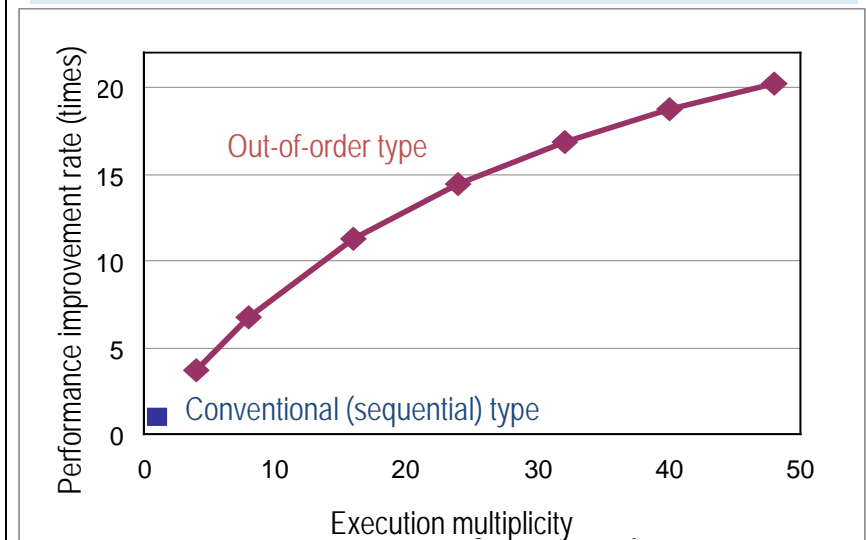
Processor	Intel Xeon 2.66GHz
Memory	32GB
Storage	20 HDDs (100Krpm)
Data set	TPC-H SF=1000 (equivalent to 1TB)

Construction of Multimedia Web Analysis Platform and Development of Social Analysis Software
(to be started in FY2009)

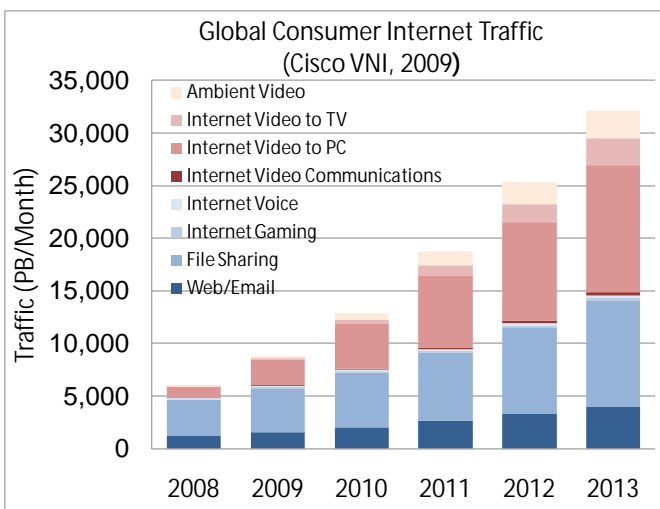Ministry of Education, Culture, Sports, Science and Technology

Construction of Multimedia Web Analysis Platform and Development of Social Analysis Software
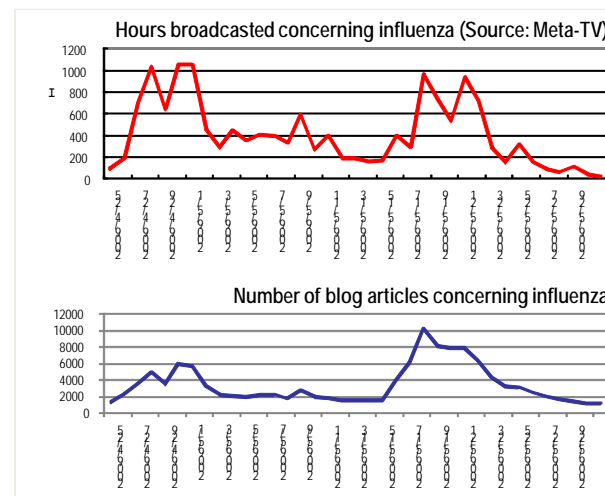
Purpose

To meet the needs of various types of social analyses including sociology, linguistics, risk management, and marketing, construct a multimedia web information analysis platform that enables the collection, accumulation, and various types of analyses of a vast amount of multimedia web information, as well as research & development and field testing of social analysis software.

Issues

- Web information is an essential information source to create new values in observation, investigation, and analysis of human society.
  - Ø With the rapid penetration of multimedia, influences mutually exerted between multimedia and real world information are increased.
- Collection and accumulation of multimedia web information, content analysis of multimedia, large-capacity and high throughput analysis platforms on a highly parallel computing environment, and realization of effective social analysis software are essential.

Global Consumer Internet Traffic (Cisco VNI, 2009)

Legend:
- Ambient Video
- Internet Video to TV
- Internet Video to PC
- Internet Video Communications
- Internet Voice
- Internet Gaming
- File Sharing
- Web/Email

Y-axis: Traffic (PB/Month) — 0; 5,000; 10,000; 15,000; 20,000; 25,000; 30,000; 35,000
X-axis: 2008, 2009, 2010, 2011, 2012, 2013

Increase of video traffic

Hours broadcasted concerning influenza (Source: Meta-TV)

Y-axis: 0, 200, 400, 600, 800, 1000, 1200

Number of blog articles concerning influenza

Y-axis: 0, 2000, 4000, 6000, 8000, 10000, 12000

Close relationships with the real world

Construction of Multimedia Web Analysis Platform and Development of Social Analysis Software

Content of Research and Development

■ Construct a multimedia web analysis platform and develop social analysis software based on four sub-themes and with the multimedia web-reasoning analysis as an axis.

Web information

Broadcasting image information

Multimedia web platform technology

Increase the speed of index-updating processing

Petabyte scale

Scheduling technology

Multimedia web analysis element technology

■ Keyword extraction technology
■ Linkage technology
■ Multidimensional analysis high-speed technology
■ Visualization technology

Aircraft

Sport

Building

Multimedia integrated processing

Image and video analysis

Reasoning analysis

Technology for the chronological analysis of meaning and content Natural language processing

Chronological analysis of link structure

Field-test evaluation of multimedia web analysis

Activities and books, etc.

Expansion of social entrepreneurs

Introduced in TV programs

Number of blog articles

The Sensing Web: Content Extraction of Sensor Information for Societal Use

Complementary Subject
(Representative Institution: Kyoto University)

http://www.mm.media.kyoto-u.ac.jp/sweb/index.html

Background and Purpose of Research

Background

- n The Internet is developing on a global scale and is becoming an enormous knowledge base.
- n Sensor networks are closed on purpose and are not widely utilized.
- n Research to develop sensor information into contents that can be used in a society is essential.

Purpose

Research and develop the following functions to realize sensor information applications (sensing web).

- n Mechanism for sharing sensor information (information sharing)
- n Responding to the problems unique to sensor information (access management)
- n Search and presentation of distributed sensor information (information application)

Major achievements: Sensing Web Technology Development

The basic functions of web-sensing technology were developed.

(1) Technology for removing the private data of observed persons in sensor information before releasing it publicly
(2) Technology for accepting various information demands regardless of the sensor type and installation status
(3) Technology for integrating data provided from sensors and presenting it in an easy-to-understand manner
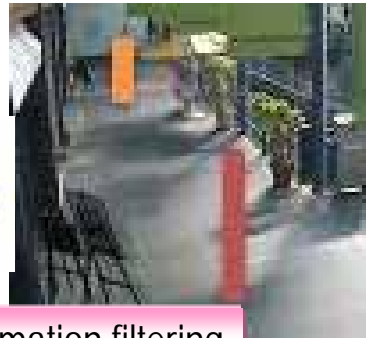
(1) Privacy information filtering


```
<data>
  <sensor id="130.54.15.135" />
  <time>
    <year>2009</year>
    <month>02</month>
    <day>27</day>
    <hour>15</hour>
    <minute>41</minute>
    <second>09</second>
    <millisecond>281</millisecond>
  </time>
```

(2) Description format for information sharing
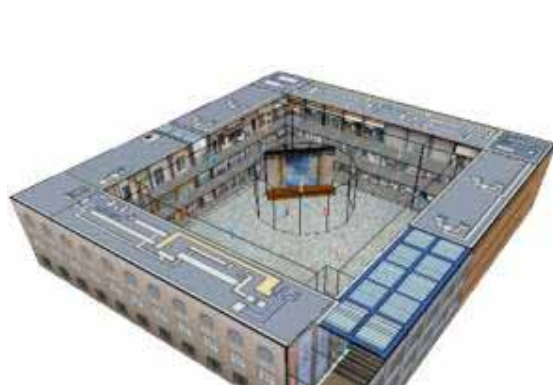
(3) Sensor information integration and content presentation

Major Achievements: Field-test Experiment and Acceptability Evaluations

Field-test experiments in an actual public environment

- **n** Experiment period: July through December 2009
- **n** Experiment outline: The following items were evaluated through regular testing
    1. Performance evaluations in the actual environment of the developed technology
    2. Evaluation of the effectiveness of the sensing web in the real world

1. Do you feel that the privacy is protected regarding this technology?

Yes and no   I seldom feel so

| I feel very much so | I sometimes feel so | |

0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

2. Do you have any concern about the spread of this technology?

| I have some concern | Yes and no | I have little concern | I do not have any concern |

0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

3. Do you feel that privacy protection is considered more in this technology compared to security cameras?

Yes and no   I seldom feel so

| I feel very much so | I sometimes feel so | |

0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

Results of a questionnaire administered to general visitors (interim result)

Cooperation during Each Project

Coordination Diagram between Each Technological Element of "Very Large Information Integration and Application Platform" Program (Details)

Hierarchical function \ Program	Information Grand Voyage Project (IGVP)			Information Credibility Evaluation Technology	Ultra-High Performance Database Engine software	societal use of sensor information

Promotion

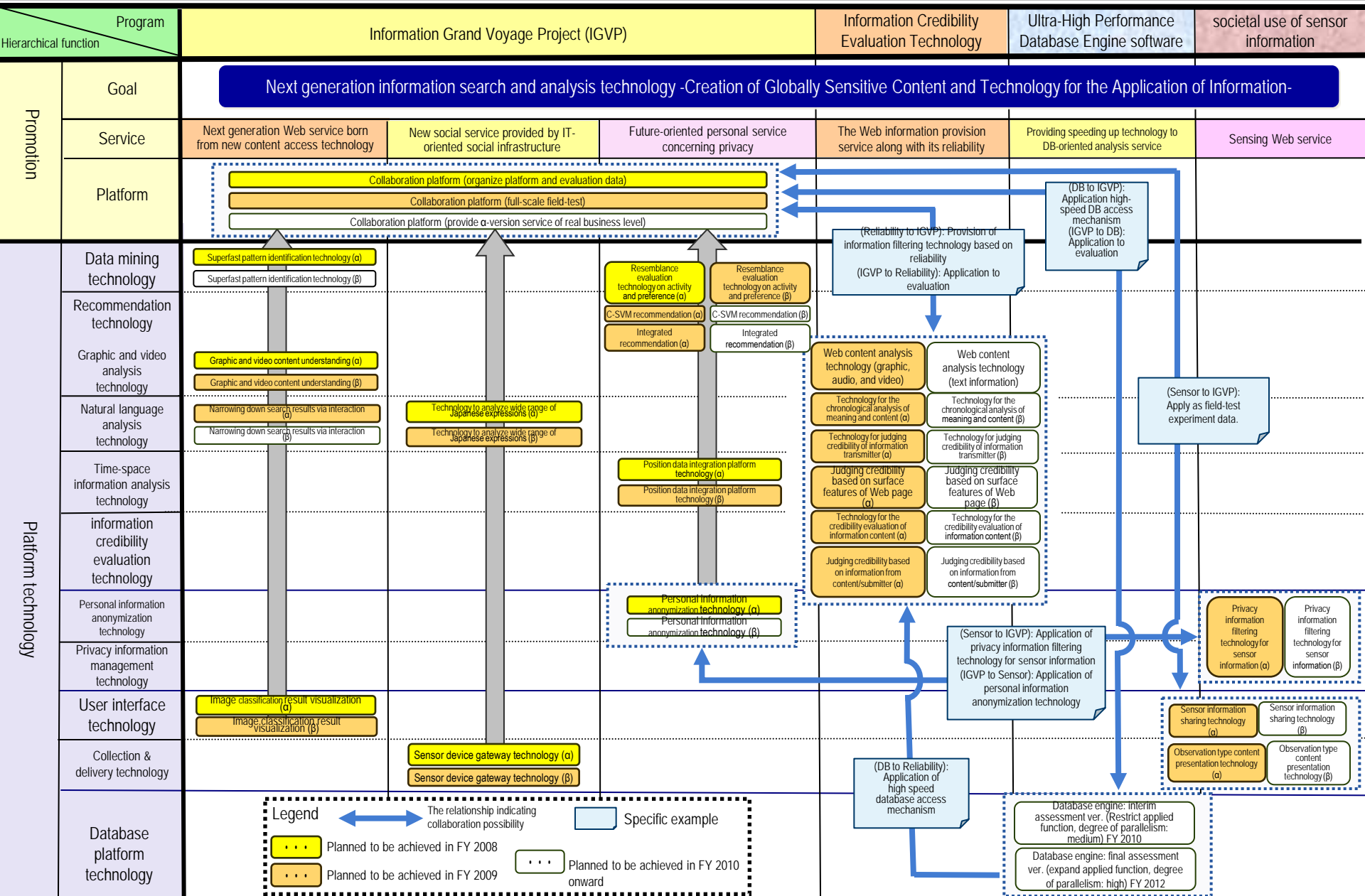|						
Goal	Next generation information search and analysis technology -Creation of Globally Sensitive Content and Technology for the Application of Information-					
Service	Next generation Web service born from new content access technology	New social service provided by IT-oriented social infrastructure	Future-oriented personal service concerning privacy	The Web information provision service along with its reliability	Providing speeding up technology to DB-oriented analysis service	Sensing Web service
Platform	Collaboration platform (organize platform and evaluation data)					
	Collaboration platform (full-scale field-test)					
	Collaboration platform (provide α-version service of real business level)					

Platform technology

Data mining technology
- Superfast pattern identification technology (α)
- Superfast pattern identification technology (β)
- Resemblance evaluation technology on activity and preference (α)
- Resemblance evaluation technology on activity and preference (β)

Recommendation technology
- C-SVM recommendation (α)
- C-SVM recommendation (β)
- Integrated recommendation (α)
- Integrated recommendation (β)

Graphic and video analysis technology
- Graphic and video content understanding (α)
- Graphic and video content understanding (β)
- Web content analysis technology (graphic, audio, and video)
- Web content analysis technology (text information)

Natural language analysis technology
- Narrowing down search results via interaction (α)
- Narrowing down search results via interaction (β)
- Technology to analyze wide range of Japanese expressions (α)
- Technology to analyze wide range of Japanese expressions (β)
- Technology for the chronological analysis of meaning and content (α)
- Technology for the chronological analysis of meaning and content (β)

Time-space information analysis technology
- Position data integration platform technology (α)
- Position data integration platform technology (β)
- Technology for judging credibility of information transmitter (α)
- Technology for judging credibility of information transmitter (β)
- Judging credibility based on surface features of Web page (α)
- Judging credibility based on surface features of Web page (β)

information credibility evaluation technology
- Technology for the credibility evaluation of information content (α)
- Technology for the credibility evaluation of information content (β)
- Judging credibility based on information from content/submitter (α)
- Judging credibility based on information from content/submitter (β)

Personal information anonymization technology
- Personal information anonymization technology (α)
- Personal information anonymization technology (β)
- Privacy information filtering technology for sensor information (α)
- Privacy information filtering technology for sensor information (β)

Privacy information management technology

User interface technology
- Image classification result visualization (α)
- Image classification result visualization (β)
- Sensor information sharing technology (α)
- Sensor information sharing technology (β)

Collection & delivery technology
- Sensor device gateway technology (α)
- Sensor device gateway technology (β)
- Observation type content presentation technology (α)
- Observation type content presentation technology (β)

Database platform technology
- Database engine: interim assessment ver. (Restrict applied function, degree of parallelism: medium) FY 2010
- Database engine: final assessment ver. (expand applied function, degree of parallelism: high) FY 2012

Notes:
- (DB to IGVP): Application high-speed DB access mechanism (IGVP to DB): Application to evaluation
- (Reliability to IGVP): Provision of information filtering technology based on reliability (IGVP to Reliability): Application to evaluation
- (Sensor to IGVP): Apply as field-test experiment data.
- (Sensor to IGVP): Application of privacy information filtering technology for sensor information (IGVP to Sensor): Application of personal information anonymization technology
- (DB to Reliability): Application of high speed database access mechanism

Legend
- ↔ The relationship indicating collaboration possibility
- (light blue box) Specific example
- (yellow box) Planned to be achieved in FY 2008
- (orange box) Planned to be achieved in FY 2009
- (white box) Planned to be achieved in FY 2010 onward

4. Summary of the Three-Year Activity of the Coordination Program

Summary of the Three-Year Activity of the Coordination Program "Very Large Information Integration and Application Platform"

(1) Projects

Initial plans

1. Information Grand Voyage Project
2. Information Credibility Evaluation Technology
3. Ultra-High Performance Database Engine Software
4. Complementary Subjects

Major achievements

- Existing projects: Progress was steadily made.
- Addition of a new project:
- "Construction of a Multimedia Web Analysis Platform and Development of Social Analysis Software" began as a four-year plan in FY2009.

The initial goal was completely achieved!

We achieved an increase of processing speed that was well above the initial plan.
® Toward a new proposal leading to the increase of speed by several orders of magnitude
(the Ultra-High Performance Database Engine)

(2) Legal system

Initial plans

Considering problems related to the legal system in information applications

Major achievements

- Our efforts to amend the copyright law resulted in the enactment of the Copyright Reform Act in June 2009.

Under the initiative of industry and academia, a consortium was founded to investigate the application of personal information to various services.

(3) Information transmission

Initial plans

- Holding a symposium
- Domestic and international conferences
- Utilizing the website

Major achievements

- Symposiums, and domestic/international conferences: Information was actively transmitted
- Transmission of achievements through use of the website (inside and outside Japan)

We invited lectures from QUAERO and Pharos, which are European programs that are promoting projects similar to this coordination program.

(1) Summary of Achievements in the Coordination Program (1/2)

The Information Grand Voyage Project

∎ Field testing of next-generation technology: Conducted for a total of 22 services.
∎ Development of common technologies: More than 50 basic technologies including video and text analysis technologies were developed.
∎ Upgrading of systems and environment:
- The copyright law was amended so that search service business can be performed.
- Under the initiative of industry and academia, an environment (consortium) for investigating the application of personal information to various services was founded.

∎ Many lectures were held in Japan.
∎ Presentations were made at international conferences.
∎ Media coverage: 236 articles.

Information Credibility Evaluation Technology

∎ Reliability of web information (images, video, audio content, and text) were developed; a field test was conducted.
∎ Analysis of the degree of the reliability of web transmitters
(Speech map creation systems, summarization and sorting of textual information, and development of chronological analysis)

∎ An international workshop was held.
∎ Presentations were made at international conferences: 31 presentations.
∎ Exhibitioned at a business show

(1) Summary of Achievements in the Coordination Program (2/2)

Ultra-High Performance Database Based on the Innovative Execution Principle

- An out-of-order DB engine was developed and approximately a twenty-fold performance improvement was verified (verification was performed in the real world, such as through an OSS, DBMS, and a commercially available DB).
  - The interim goal in FY2009 was a ten-fold performance improvement.

  - Presentations were made at symposiums in Japan.
  - Professor Masaru Kitsuregawa, a research promoter, received the 2009 SIGMOD Edgar F. Codd Innovations Award.

Sensing Web

- Sensor privacy information filtering was introduced.
- The sensor data description format was formulated, and data was opened to public.
- Multiple observation-type real world contents were introduced.
- Field-test experiments and acceptability evaluations were conducted at commercial facilities.

  - Special sessions were held at domestic and international conferences.
  - Events for general public were implemented.
  - Media coverage: Many newspaper and TV reports.

Achievements through collaboration

- Collaboration between projects was investigated and implemented.
  - Investigation for mutual applications of field-test experiment data (the Information Grand Voyage Project and sensing web)
- Collaboration with other coordination programs (a next-generation robot coordination program) was investigated and implemented.
  - In a sensing web, service specification descriptions and demand specification descriptions were prepared based on the interface specifications of the next-generation robot coordination program.

(2) Summary of Legal System

Upgrading of Systems and Environment

Copyrights

n We took continuous action towards the amendment of the copyright law by making proposals and submitting public comments, and our efforts resulted in the enactment of a bill for amending a part of the copyright law, which was passed by the Diet in June.
  Ø As a result, it became possible to provide search services.

n Arrangement of interpretation of the amended law
  Ø We supported business persons' interpretation of the amended copyright law as it helps the consideration of the development of the next-generation service.

n We extracted and investigated institutional problems and the challenges required to realize the development of the next-generation service (activation of industry) within the information industry and to realize cultural development in the future.

Personal Information Protection

n We organized the application methods of personal information including individual related information, such as purchase and action histories for which the handling method had not yet been clearly defined.

n For industry and academia, we promoted interest in the application of personal information by holding a personal information-related symposium.
  Ø Under the initiative of industry and academia, a consortium was founded to investigate the application of personal information to various services.

(3) Summary of Information Transmission (1/3): Website

■ As a part of information transmission, a website for the coordination program was set up to actively transmit information at home and abroad.

(1) Japanese version

(2) English version

http://www.renkei.jst.go.jp/infor/infor.html
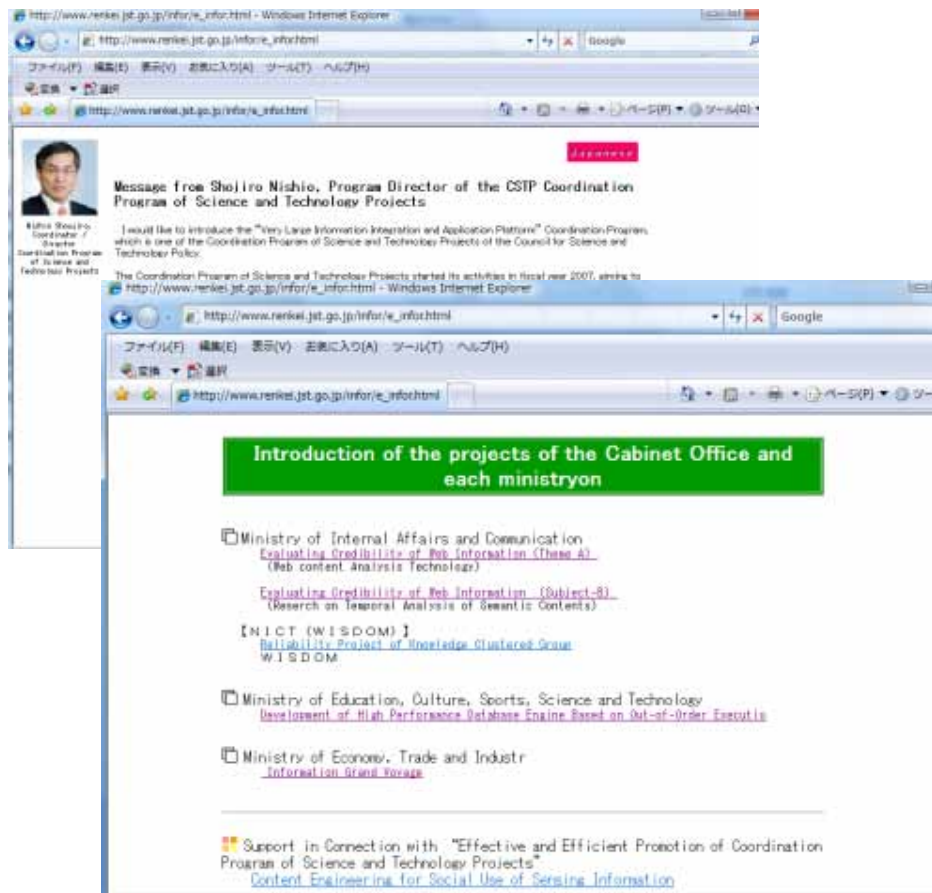
(3) Summary of Information Transmission (2/3): Lectures in Japan and International Conferences

- **n** The project director (Shojiro Nishio) made a keynote lecture to introduce the coordination program at international conferences.
    - Ø Second International Symposium on Universal Communication, December 2008
    - Ø International Conference on Complex, Intelligent and Software Intensive Systems, March 2009

- **n** As a special event in "FY 2009 Informatization Month", lectures were held.
    - Ø The lectures drew a lot of interest, as seen from the fact that audiences exceeded the capacity of the room.
- **n** Lectures were given at international conferences held by Chorus and OECD, which are search engine-related projects in the EU.
    - Ø Technical collaboration was requested from search engine development projects in EU countries
    - Ø The OECD delegation expressed an opinion stating that Japan has the possibility to lead discussions on the application of personal and privacy information; we have a high reputation internationally.

FY 2009 Informatization Month special event

October 1, 2009 (Thu.)
ANA Intercontinental Hotel Tokyo

Chorus Final Conference

May 26 (Tue.) to 27 (Wed.), 2009
Brussels (Belgium)

OECD Experts Conference

June 8 (Mon.) to 9 (Wed.), 2009
Lisbon (Portugal)

(3) Summary of Information Transmission (3/3): International Conferences and Academic Activity in Japan

(1) We transmitted information to a wide range of areas through, for example, holding a special session at the national convention of the Japanese Society for Artificial Intelligence (June 2009) and the national convention of the Information Processing Society of Japan (scheduled in March 2010).

(2) We showed our leadership in sensing web technology by holding a special session in SPC2009 (the International Conference on Security Camera Networks, Privacy Protection and Community Safety; held in Gunma in 2009).

(3) We held events for the general public several times in commercial facilities where field-test experiments were conducted. In addition, these events were picked up by newspapers and TV many times (newspapers: Asahi, Sankei, Mainichi, Yomiuri, Nikkei, and Nikkankogyo; TV: NHK, TV Tokyo, and KBS Kyoto, etc.)

(1) National convention of the Japanese Society for Artificial Intelligence

(2) SPC2009

(3) Holding of events and information transmission to the media

5. Future Development

Future Implementation of the Coordination Program

To explore new fields in the age of the information explosion, the application of information is required.

- n Promote the construction of the next-generation intellectual information access platform that would enable a person to collect and analyze necessary information from diverse information sources inside and outside the web easily, accurately, and safely.

- n Investigate and consider various issues for the application of information together with the consideration of the legal system to promote activity that enables effective utilization of the achievements.

It is necessary to establish a collaboration system between the projects that will be terminated this fiscal year and those that are to be continued.

- n Offer a place for discussion on the projects that will be terminated this fiscal year (Information Grand Voyage Project and the sensing web) and those to be continued (information credibility evaluation technology, the out-of-order DB, and web social analysis).

- n Further strengthen the collaboration system of related institutions for promoting sustainable research and development in the future and further improve international competitiveness.