

# Food Informatics

Application and development of state of the art information and communication technology to support a powerful and cost effective R&D environment

D.J.P. Out  
Unilever Research  
Vlaardingen

30 November 2004



## Background

*The prime task of Food Research and Development organisations is to deliver and implement, in time, all the know-how in science & technology, whilst aware of the customer needs, that will drive the successes of their respective companies and institutes.*

## Background

*The prime task of Food Research and Development organisations is to deliver and implement, in time, all the know-how in science & technology and awareness of the customer needs that will drive the successes of their respective companies and institutes.*

Proper execution of this task requires access to and use of information available from numerous sources, inside and outside the organisation.

## Background

*The prime task of Food Research and Development organisations is to deliver and implement, in time, all the know-how in science & technology and awareness of the customer needs that will drive the successes of their respective companies and institutes.*

Proper execution of this task requires access to and use of information available from numerous sources, inside and outside the organisation.

This information is in general not for free. The retrieval and discovery processes can require substantial effort and costs. Execution of new (experimental) work can be an alternative.

## Goal Food Informatics:

*To increase quality, speed and cost effectiveness of the contribution of R&D to Food Science & Technology and innovation processes by sharpening our ICT tools for information and knowledge deployment.*

## Why “FOOD” Informatics?

Why “FOOD” Informatics?

What are the specific opportunities?

## To safeguard and support the strong position of the Dutch Agro-Food Industry and Research Institutes

- Presence of large number of national and multinational companies
- Export Value > MEuro 40,000.-
- Supported by a well developed scientific infrastructure of universities and institutes (Wageningen 'Food Valley')
- Generating high quality employment
- Knowledge export; significant impact on international food society



Why “FOOD” Informatics?

What are the specific problems?

## Changes in approach to industrial research

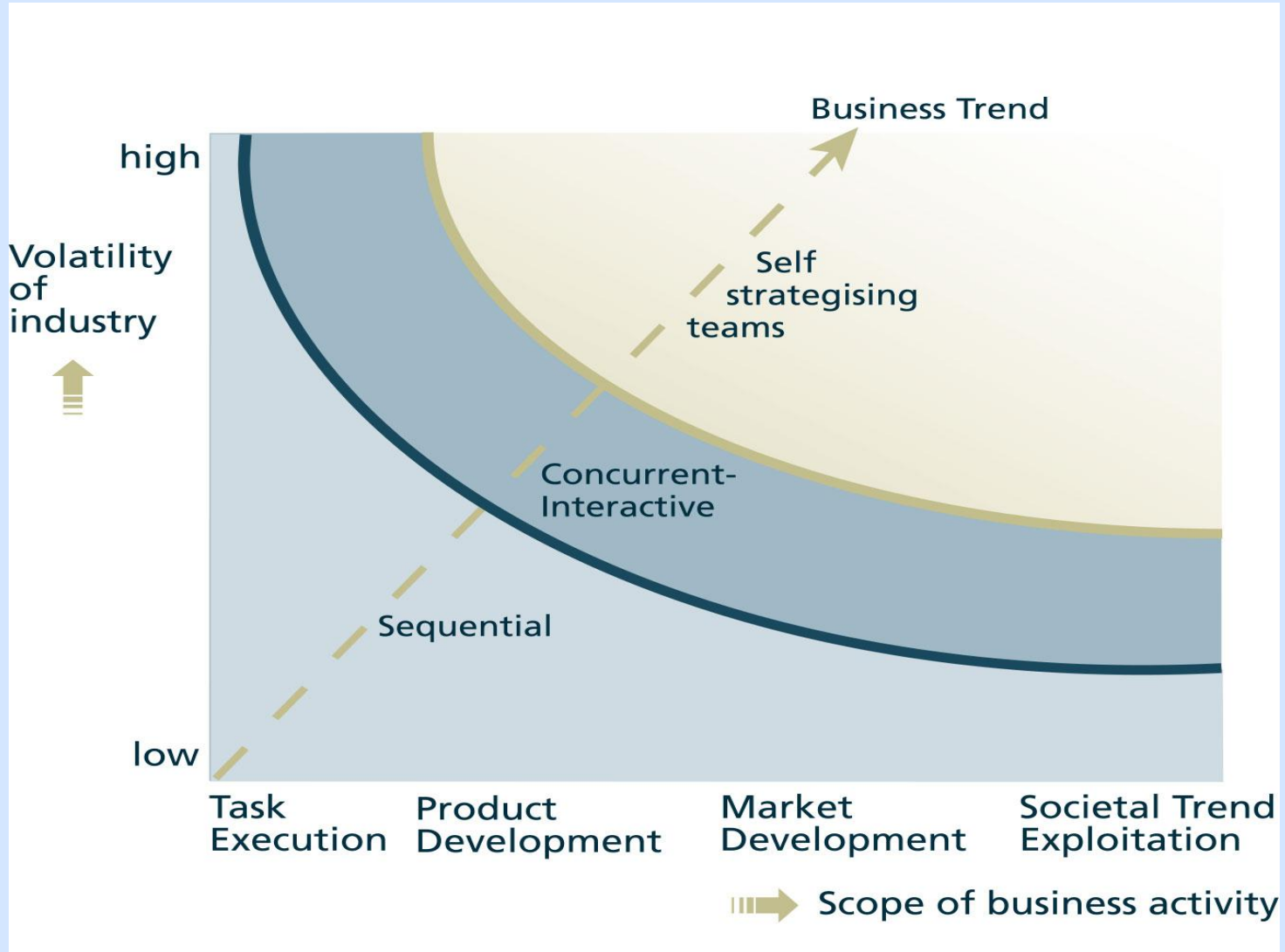
### From

- Functional
- Sequential
- Learning while developing
  
- Do it yourself
- Specialised information

### To

- Integrated
- Concurrent
- Strategic investment in capabilities
- Network of partnerships
- Systems networks

# Business environment > strategy integration



How to react properly on a business request?

How to react properly on a business request?

> Example

## How to react on a business request?



How to react on a business request?

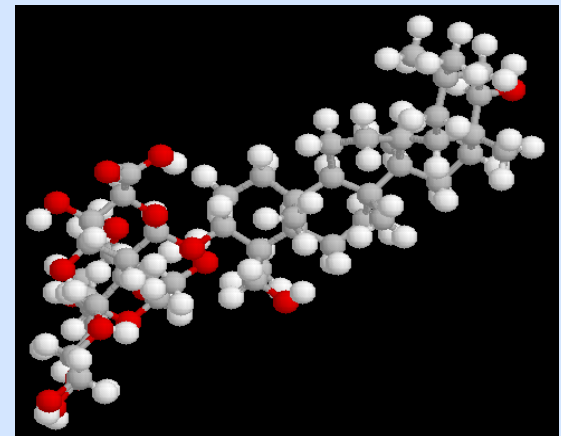
**“Develop  
Packed  
Sandwich”**





## Storage tests: Lettuce develops a bitter taste

- Type of lettuce
- Contamination (Micro-organisms, enzymes, ...)
- Preservatives
- Acid level (pH)
- Water activity ( $A_w$ )
- Headspace composition ( $O_2$ )
- Inactivation by heat treatment
- Complex formation
- Bitterness perception (Masking)
  - Taste buds (Receptor system)
  - Signal processing
  - .....



Supply chain

Bakery

Packaging

Spreads

Barriers

Culinary

Microbiology

Olive oil

Hygienic Processing

Meat

Quality Assurance

Lettuce

Consumer understanding

Marketing

Cheese/dairy



Successful development requires information from  
a multitude of domains

## The data universe:

- Product/technical information
  - scientific and technical publications
  - models (mathematical)
  - ingredient databases (specifications)
  - formulation databases (recipes)
  - process descriptions
- Consumer & market information
  - consumer insights
  - habits, trends
  - sales (market) data
  - health and wellness
  - risks
  - legislation
- Business and management information
  - project management
  - supply chain

## The data universe: format and access

- Different formats like:
  - Office applications (MS, ..)
  - Data bases (Oracle, ..)
  - Web info (HTML, XML, ..)
  - Programming languages (Fortran, C++, ..)
- Different locations
  - Network
  - Language
- Different ownership
  - Public
  - Company
  - Department
  - Publisher
  - Scientist

## Problems in Information Retrieval:

- Information is missed because sources are not known
- Access to individual sources can be difficult due to the various procedures (passwords) and differences in look & feel
- Searches might miss valuable information due to spelling differences, minor typing errors or the use of synonyms (poor recall)
- Queries that return “hits” outside the domain of interest require additional screening work (lack of precision)
- Searches may produce a data overload that requires much time and effort for proper assessment
- The format of the retrieved information may impede easy reuse

## Consequences:

- Valuable knowledge remains unused
- Retrieval is slow
- Retrieval causes high costs

*As R&D workers spend a substantial part of their time on information retrieval the impact of these consequences is high!*

Improvement of this situation requires the following steps:

- (1) **Inventory** of content and tools
- (2) Clear **presentation** of content and tools
- (3) Implementation of advanced & adaptive information **disclosure** techniques
- (4) Implementation of tools to improve data and information **assessment and processing (Knowledge discovery)**



Improvement of this situation requires the following steps:

- (1) **Inventory** of content and tools
- (2) Clear **presentation** of content and tools
- (3) Implementation of advanced & adaptive information **disclosure** techniques
- (4) Implementation of tools to improve data and information **assessment and processing (Knowledge discovery)**

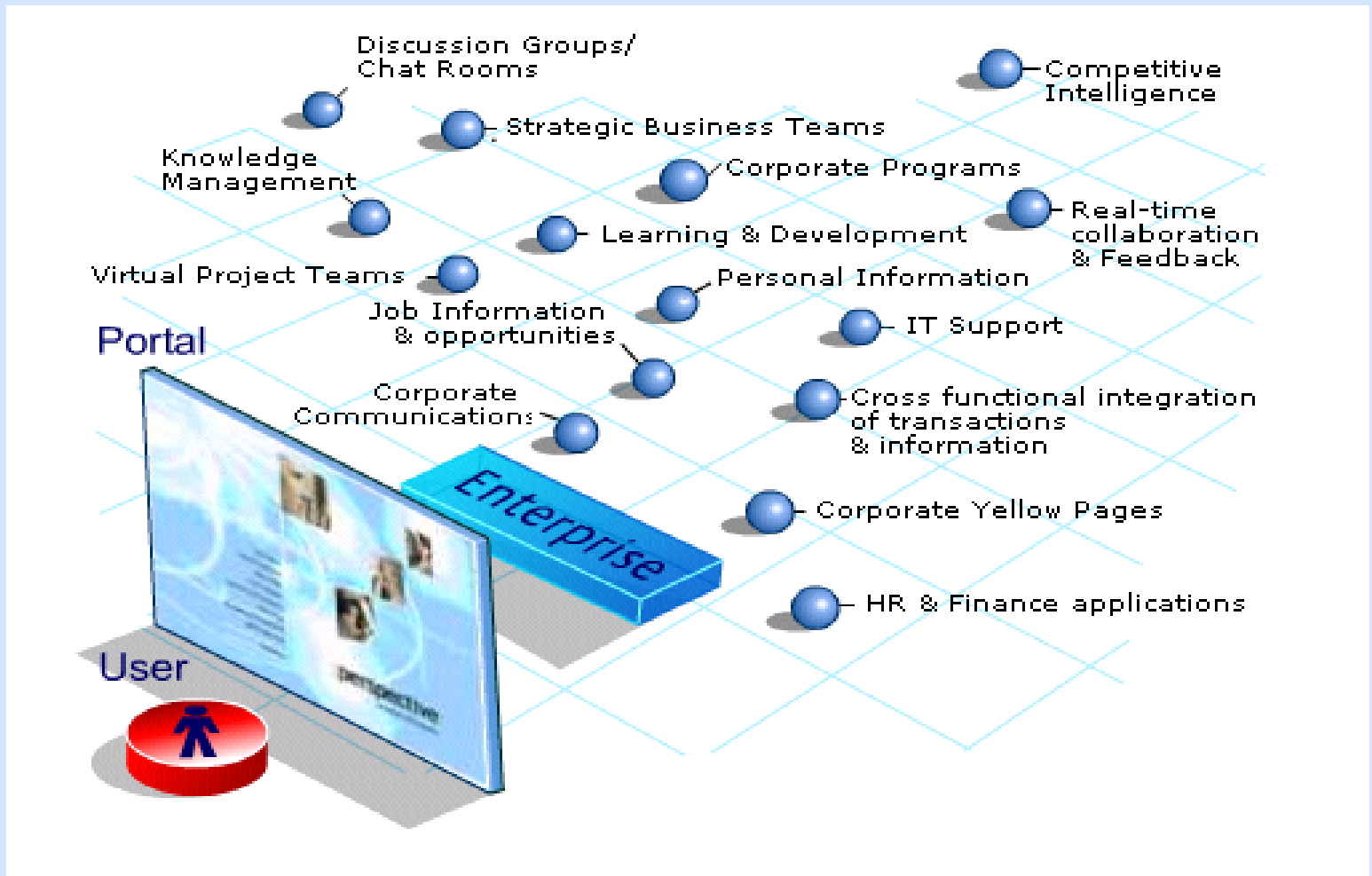
*And: Always realise that food scientists are **NO** ICT Experts!*

## (1) Inventory of content and tools

- Proper overview of the information landscape
- Proper overview of processing and modelling capabilities
- Connect to the network (Information owner!)
  - > Gap analysis
    - > Requirements

## (2) Clear Presentation of content & tools:

## Clear Presentation of content & tools



## (2) Clear Presentation of content and tools by building and customising a portal.

- Simplified access management; Single “Sign-on”
- Easy browsing through uniform “Look & Feel” and architecture
- Tailored environments for individuals and groups
- Possibility to implement advanced search techniques
- Reduced maintenance costs

## (3) Advanced & Adaptive Information Disclosure Techniques

- Selection and implementation of search tools:
  - Parallel searches in dispersed & disparate sources
- Application of thesauri, taxonomies, ontologies and annotation tools (domain specific)
  - Improvement of completeness of searches
  - Improvement of precision of searches by context sensitive searching and browsing

>> *VL-e/Food Informatics Project*

## (4) Improve data and information assessment and processing

- Select, adapt and implement summarisers
  - Benefit: Effort reduction of assessment of voluminous query results
- Implementation of tools for re-use of know how captured in models
  - Easy access and use of know how represented in models
  - Easy reuse of numerical data
  - Integration of models allowing in-silico evaluation of complex systems
  - Preservation of specific know how in a rapidly changing software and hardware environment
- Data mining by federation of data/information: knowledge discovery
  - Make implicit knowledge explicit

>> *VL-e/Food Informatics Project*

## Virtual Laboratory e-Science/FI project

(Funded by ICES/KIS-III)



## VL-e Project Mission

---

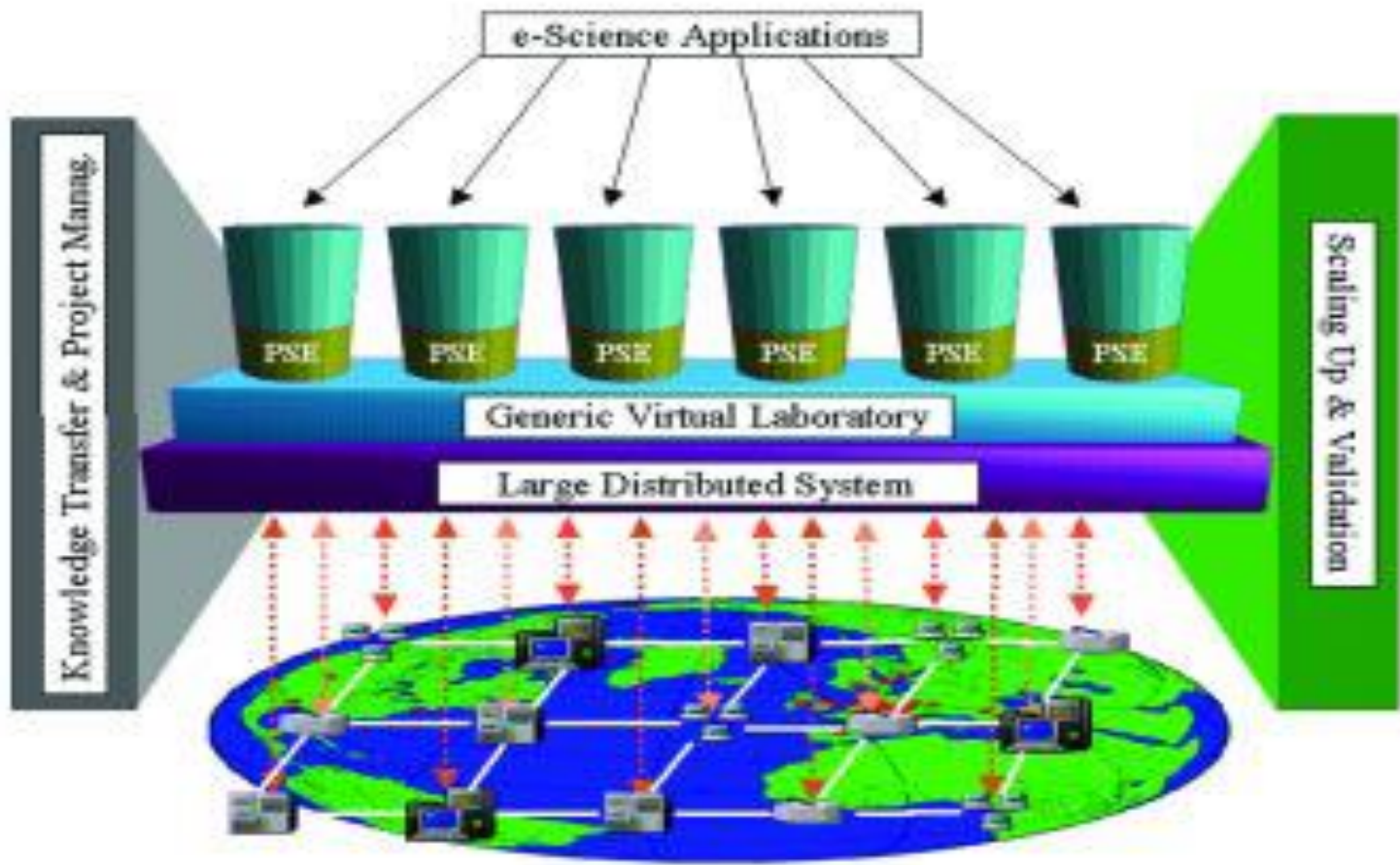
To boost e-Science by the creation of an e-Science environment and doing research on methodologies.

## Project Aims

---

- To make a significant contribution to the realisation of an e-Science stronghold in the Netherlands.
- To be a portal to major international developments in this area.
- To help building up the required e-Science infrastructure.
- To spread the knowledge to other parties in both public and private domains.
- To enhance the quality of our knowledge intensive society.

# VL-e Project



**FIGURE 1.1:** THE VL-E PROGRAM DEFINES THREE LAYERS OF RESEARCH ACTIVITIES. IN ADDITION IT STUDIES SCALING UP & VALIDATION AND IT PROVIDES KNOWLEDGE TRANSFER. (PSE: PROBLEM SOLVING ENVIRONMENT)

## Partners

---

### Technology providers

- WTCW
- UVA (ILLC, ....)
- VU
- CWI
- SARA
- TNO-TPD
- TU Delft
- IBM
- CMG
- .....

## Partners

---

### Technology providers

- WTCW
- UVA (ILLC, ....)
- VU
- CWI
- SARA
- TNO-TPD
- TU Delft
- IBM
- CMG
- .....

### Technologies

- Network (Grid)
- AAA
- Text mining
- Ontology building, annotation
- Summarisers
- Image retrieval
- Modelling/Integration
- Q&A systems
- GUI/Visualisation
- Problem solving environments

## Partners - Applications

---

### Food:

- TNO-Voeding
- Wageningen/ATO
- WCFS
- FCDF
- Unilever

## Partners - Applications

---

### Food:

- TNO-Voeding
- Wageningen/ATO
- WCFS
- FCDF
- Unilever

### Content (Food cases)

- Nutrition & Health risk awareness and analysis
- From ingredient to perception (crispiness, creaminess) with modelling
- Food choice in consumer populations
- Bitter taste control



Workflow analyses



Project Goals

## *VL-e/Food Informatics Goals*



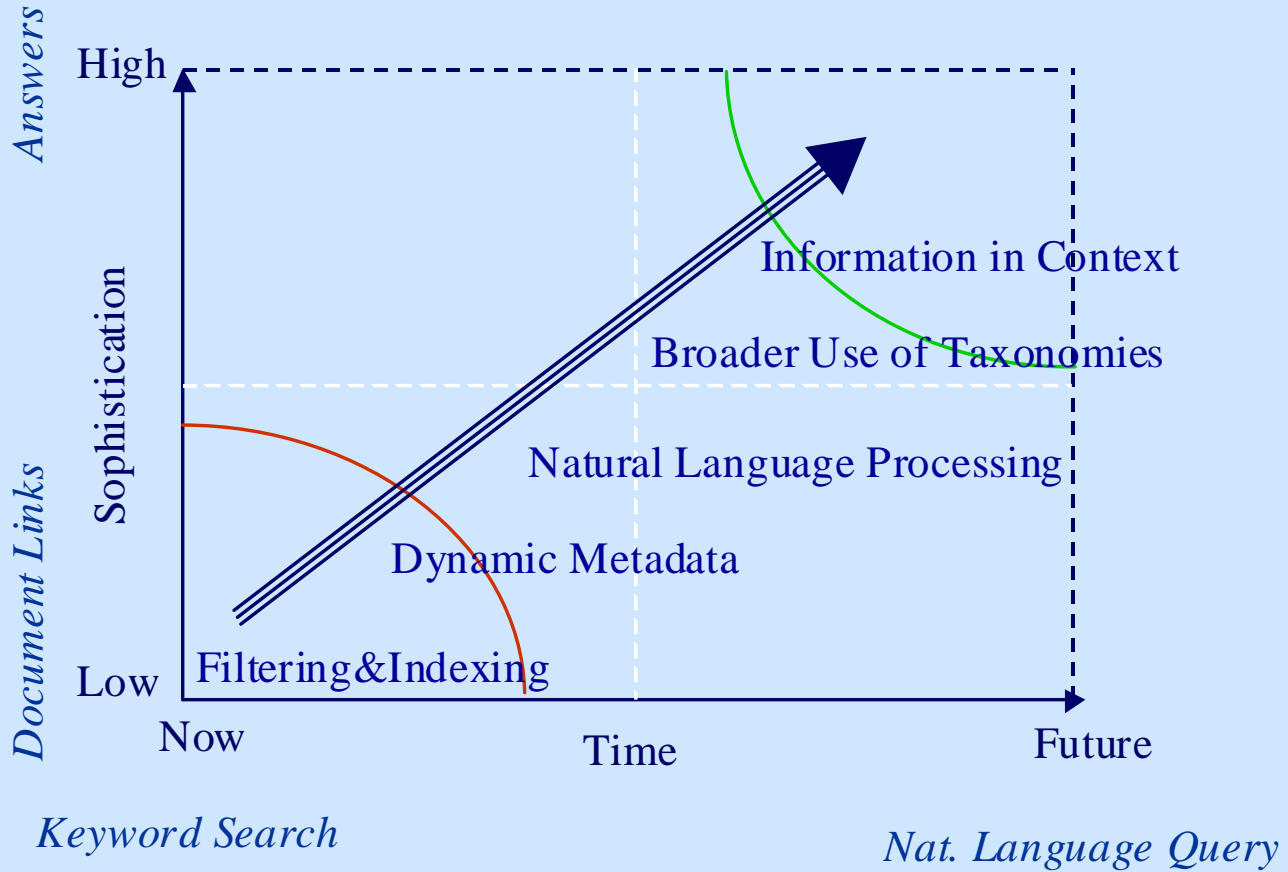
1. Increase recall and precision of searches in large distributed and disparate data sets by exploiting (new) high speed network facilities and application of the latest data structuring techniques (annotation)



## *VL-e/Food Informatics Goals*

1. Increase recall and precision of searches in large distributed and disparate data sets by exploiting (new) high speed network facilities and application of the latest data structuring techniques (annotation)
2. For the generation and maintenance of information (meta-data generation, ontologies, taxonomies...) required to improve the quality of searches new software has to be developed which reduce time consuming and expensive processes.

## Trend in Information Retrieval



## *VL-e/Food Informatics Goals*

1. Increase recall and precision of searches in large distributed and disparate data sets by exploiting (new) high speed network facilities and application of the latest data structuring techniques (annotation)
2. For the generation and maintenance of information (meta-data generation, ontologies, taxonomies...) required to improve the quality of searches new software has to be developed which reduce time consuming and expensive processes.
3. To improve retrieval and reuse of information captured in mathematical models. This includes the construction of an environment that allows running and linking mathematical models from remote systems and with different formats.

## Model Integration - Example

---

Optimisation Thermal Preservation Process: Product Taste vs Microbiological Safety and Stability

M1 Process line configuration (geometry, lay-out, flow)

> temperature ( $T$ ) and residence time distribution ( $t$ ) profiles

M2: Maillard kinetics ( $T$ ,  $t$ , pH,  $A_w$ , composition)

> concentration M-reaction products > off-taste

M3: Micro flora inactivation kinetics ( $T$ ,  $t$ , pH,  $A_w$ , composition)

> reduction rate > storage stability/risk

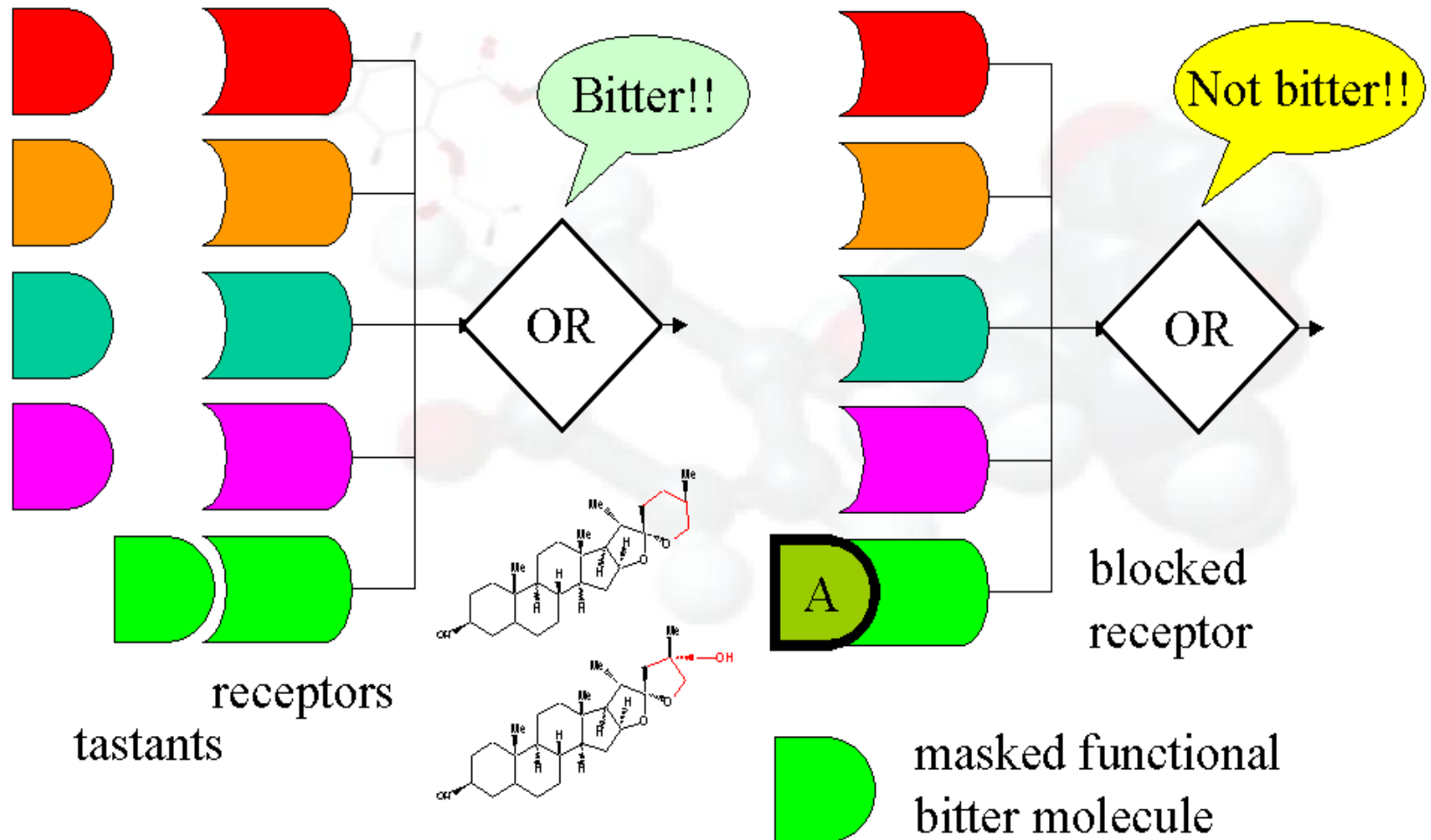
>>In Silico Product Design

## *VL-e/Food Informatics Goals*

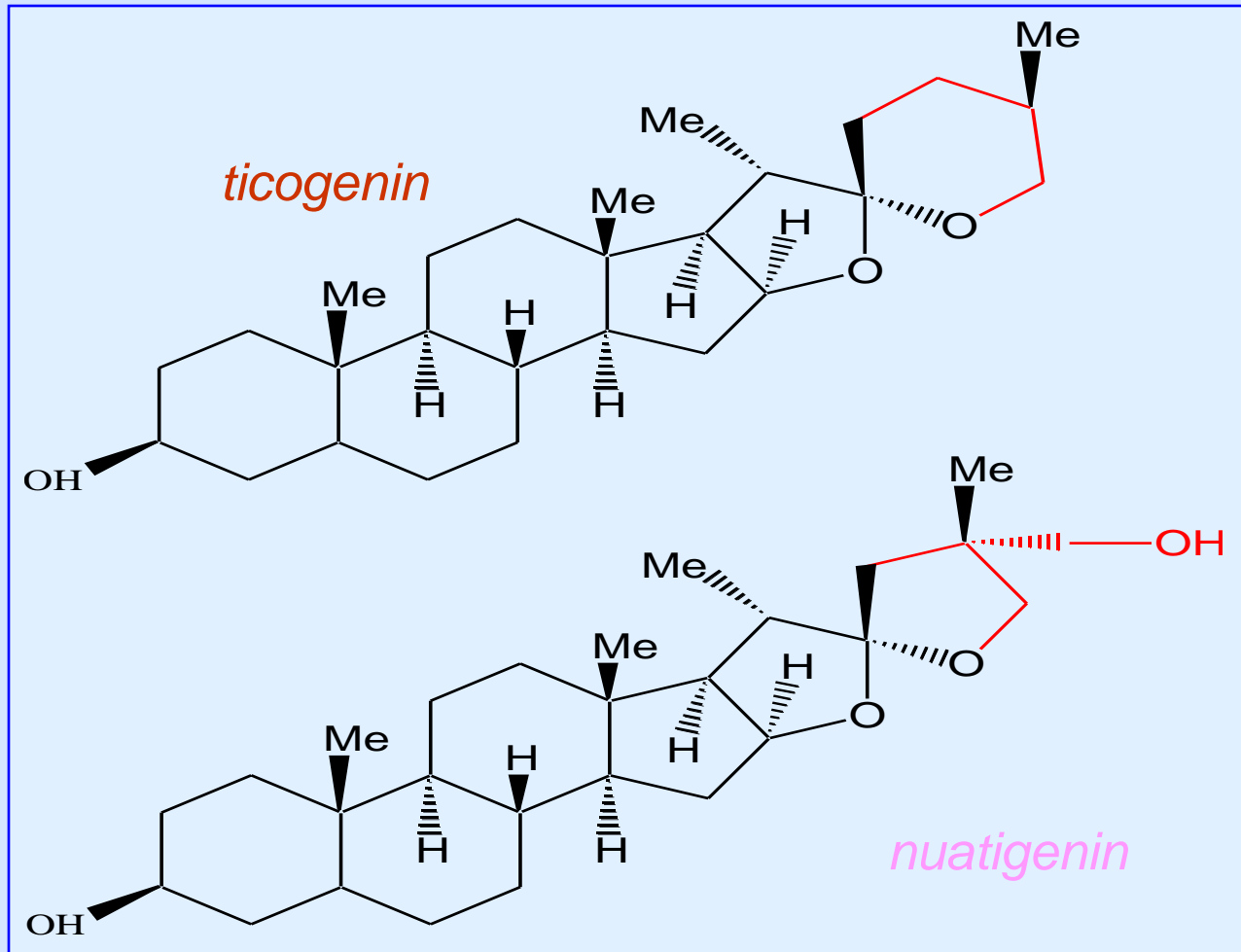


1. Increase recall and precision of searches in large distributed and disparate data sets by exploiting (new) high speed network facilities and application of the latest data structuring techniques (annotation)
2. For the generation and maintenance of information (meta-data generation, ontologies, taxonomies...) required to improve the quality of searches new software has to be developed which reduce time consuming and expensive processes.
3. To improve retrieval and reuse of information captured in mathematical models. This includes the construction of an environment that allows running and linking mathematical models from remote systems and with different formats.
4. To select and develop tools to search and assess 2D/3D molecular structure information.
5. Integration of the different tools and solutions mentioned in the working environments from the various users.

# Bitter taste



## Bitter taste



## *VL-e/Food Informatics Goals*



1. Increase recall and precision of searches in large distributed and disparate data sets by exploiting (new) high speed network facilities and application of the latest data structuring techniques (annotation)
2. For the generation and maintenance of information (meta-data generation, ontologies, taxonomies...) required to improve the quality of searches new software has to be developed which reduce time consuming and expensive processes.
3. To improve retrieval and reuse of information captured in mathematical models. This includes the construction of an environment that allows running and linking mathematical models from remote systems and with different formats.
4. To select and develop tools to search and assess 2D/3D molecular structure information.
5. Integration of the different tools and solutions mentioned in the working environments from the various users.

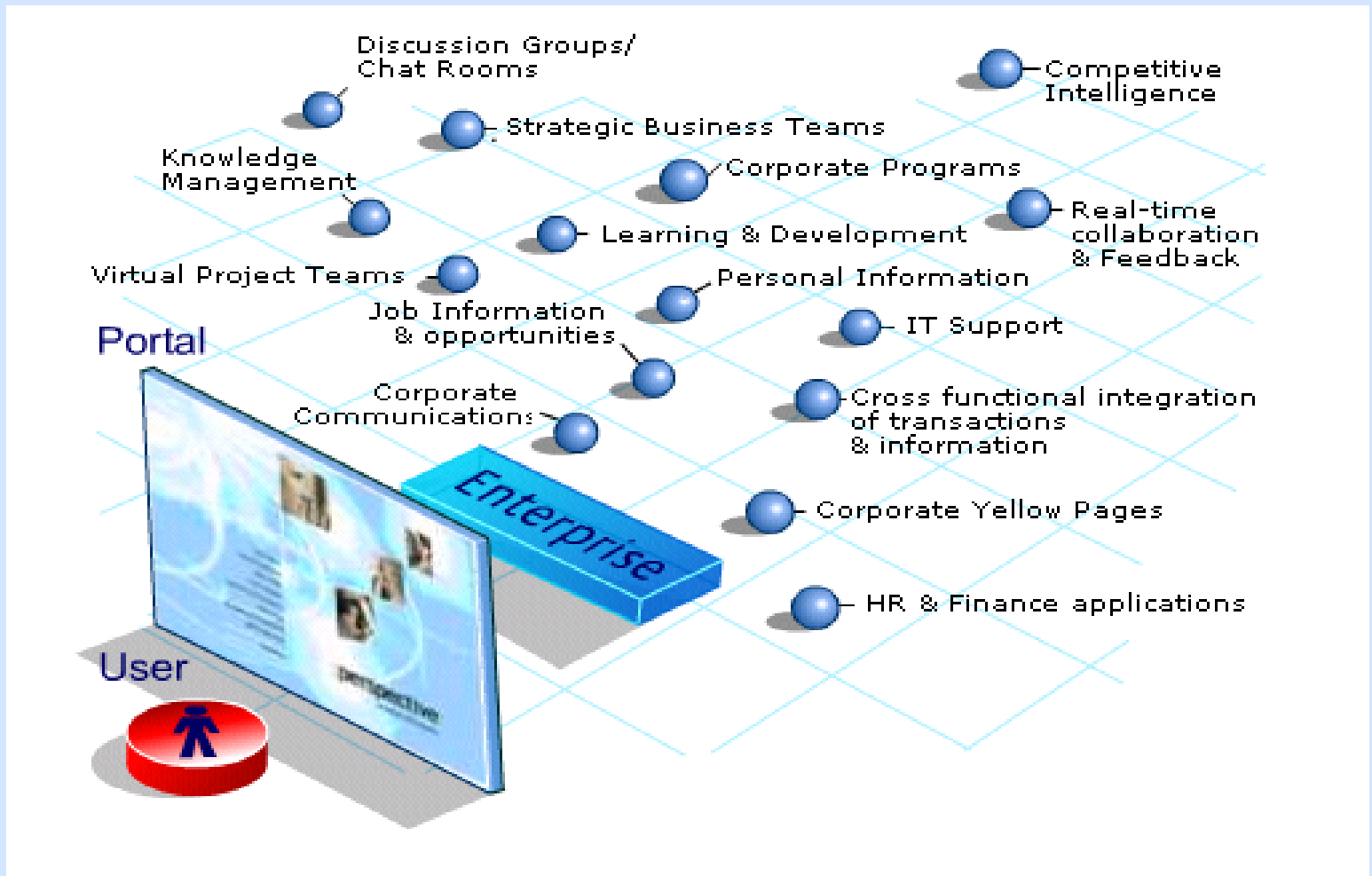


## Grid expectations

---

- Tackling problems involving large amounts of data and/or computing power
- Better sharing dispersed data and software
- Knowledge discovery from semantically indexed databases
- Building 'virtual work spaces'

## Personalised Portal > Virtual Workspace



## Key Figures

---

Starting date:	01-01-2004
Duration:	5 years
Total VL-e budget:	~40,000 kE
> Total Food Informatics budget:	3,600 kE
Total subsidy:	20,000 kE

# Food Informatics

Application and development of state of the art information and communication technology to support a powerful and cost effective R&D environment

D.J.P. Out  
Unilever Research  
Vaardingen

30 November 2004

## Stellingen

---

1. Food bedrijven weten dat bioinformatica belangrijk is, maar hebben niet de know-how om ermee te starten.
2. Binnen de bedrijven is weinig bekend over bioinformatica toepassingen voor food industrie. Hoe komt dit en hoe kunnen we dit verbeteren?
3. Kennisinstellingen moeten toegankelijkheid van databases en open source software faciliteren voor bedrijven, al dan niet via BioASP.
4. Food bedrijven moeten bioinformatica in huis hebben, omdat genomics data vertrouwelijk zijn.
5. Food bedrijven moeten gaan investeren in nieuwe medewerkers met bioinformatica expertise
6. Het dilemma van de kleine bedrijven: wat uitbesteden en wat zelf doen?
7. De door geavanceerde ict geboden mogelijkheden om binnen ondernemingen kennis te delen wordt door velen als een bedreiging gezien.
8. De toekomstige voedingswijzer zal een adviseur zijn die aanbevelingen doet op grond van persoonlijke behoefteprofielen en zal gebruik maken van gedetailleerde informatie over de productsamenstelling.

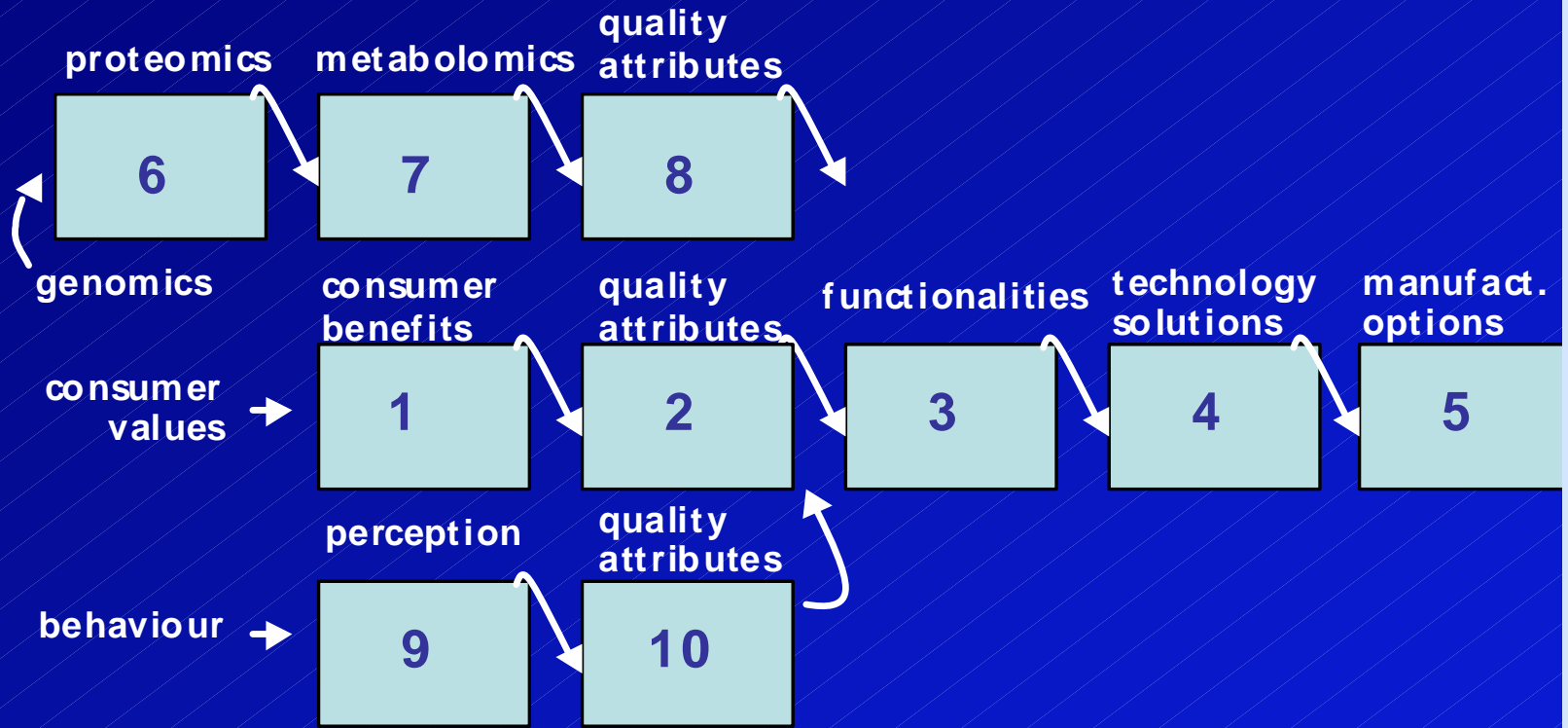
# Food Informatics

Application and development of state of the art information and communication technology to support a powerful and cost effective R&D environment

D.J.P. Out  
Unilever Research  
Vaardingen

30 November 2004

# How to structure Information?















## Networks

- + Who's Who
- + Organisation Foods Division
- 10 Golden Rules of Sharing
- + Academies
- + Events
- + Company Address Book

## 10 Golden Rules of Sharing

Owner: [Anthony Simon](#)

We are driven by the **10 Golden Rules of Sharing**:

-  Promote learning and sharing as an obligation
-  Accelerate documentation and sharing of best practices and learnings
-  Support and reward sharing, particularly "borrowing with pride"
-  Establish and maintain "yellow pages" directories of key personnel by functional expertise
-  Include sharing/borrowing as a measurement component of the Balanced Scorecard
-  Encourage the use of information systems/ processes (Intranet, Internet, virtual networks, video conferencing, visits, benchmarking, knowledge search facilities, etc.)
-  Aim for common understanding and a common "shared context"
-  Support experimentation as key to learning
-  Require that action follows knowledge
-  Encourage willingness to give and take, mutual trust, and curiosity as the drivers of sharing

## Links

### Web sites

[HRweb](#)

<http://hrweb.unilever.com>

[Path to Growth](#)

<http://pathtogrowth.com>

