

Dr. Richard Isida, W3C



W3C's goal of universal access

جعل شبكة الويب العالمية عالمية حقاً !

缔造真正全球通行的万维网

締造真正全球通行的萬維網

የዓለም አቀፉን ድር በእውነት አለም አቀፍ ማድረግ!

Κάνοντας τον Παγκόσμιο Ιστό πραγματικά Παγκόσμιο

ליצור מהרשת רשת כלל עולמית באמת

वर्ल्ड वाईड वेब को सचमुच विश्वव्यापी बना रहें हैं !

$\vec{c} \cdot \vec{L} = \Delta P \sin^2 \theta$

Making the World Wide Web world wide!

ワールド・ワイド・ウェブを世界中に広げましょう

Hogy a Világháló valóban az egész világé lehessen!

"Дүниежүзілік торды" нағыз дүниежүзілік етеміз!

전세계의 월드 와이드 웹으로 만들기!

ਵਰਡ ਵਾਈਡ ਵੈਬ ਨੂੰ ਵਾਕਈ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਬਨਾਉਣਾ !

Сделаем "Всемирную паутину" действительно всемирной!

การทำให้ World Wide Web แพร่หลายไปทั่วโลกอย่างแท้จริง

U ita uri Webu Nyangaredzi ya Dzhangho i vhe nyangaredzi ngangohol

Characters, the foundation

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | 0 | @ | P | . | p | | | | ° | À | Ä | à | ä | |
| 1 | | ! | 1 | A | Q | a | q | | | | ± | Á | Ñ | á | ñ | |
| 2 | | " | 2 | B | R | b | r | | | | ² | Â | Ò | â | ò | |
| 3 | | # | 3 | C | S | c | s | | | | ³ | Ã | Ó | ã | ó | |
| 4 | | \$ | 4 | D | T | d | t | | | | ¼ | Ä | Ô | ä | ô | |
| 5 | | % | 5 | E | U | e | u | | | | ½ | Å | Ö | å | ö | |
| 6 | | & | 6 | F | V | f | v | | | | ¾ | Æ | Ø | æ | ø | |
| 7 | | ' | 7 | G | W | g | w | | | | § | Ç | × | ç | × | |
| 8 | | (| 8 | H | X | h | x | | | | ¨ | È | Ø | é | ø | |
| 9 | |) | 9 | I | Y | i | y | | | | © | É | Ù | ê | ù | |
| A | | * | : | J | Z | j | z | | | | ® | Ê | Ú | ë | ú | |
| B | | + | ; | K | | k | { | | | | « | Ë | Û | ü | û | |
| C | | , | < | L | \ | l | | | | | ¬ | ¼ | Ì | Ü | í | ü |
| D | | . | = | M |] | m | } | | | | · | ½ | Í | Ý | î | ý |
| E | | . | > | N | ^ | n | ~ | | | | ® | ¾ | Î | Þ | ï | þ |
| F | | / | ? | O | _ | o | | | | | ™ | ¿ | Ï | ß | ì | ÿ |

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 5

Characters, the foundation

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | 0 | @ | P | . | p | | | | ° | ı | Π | ϕ | π | |
| 1 | | ! | 1 | A | Q | a | q | | | | ± | Α | Ρ | α | ρ | |
| 2 | | " | 2 | B | R | b | r | | | | ² | Β | □ | β | ς | |
| 3 | | # | 3 | C | S | c | s | | | | ³ | Γ | Σ | γ | σ | |
| 4 | | \$ | 4 | D | T | d | t | | | | ¼ | Δ | Τ | δ | τ | |
| 5 | | % | 5 | E | U | e | u | | | | ½ | Υ | μ | Ε | Υ | ε |
| 6 | | & | 6 | F | V | f | v | | | | ¾ | Ζ | Φ | ζ | φ | |
| 7 | | ' | 7 | G | W | g | w | | | | § | Η | Χ | η | χ | |
| 8 | | (| 8 | H | X | h | x | | | | ¨ | Θ | Ψ | θ | ψ | |
| 9 | |) | 9 | I | Y | i | y | | | | © | Η | Ι | Ω | ι | ω |
| A | | * | : | J | Z | j | z | | | | ® | Ι | Κ | Ι | κ | ι |
| B | | + | ; | K | | k | { | | | | « | Λ | Υ | λ | υ | |
| C | | , | < | L | \ | l | | | | | ¬ | Ο | Μ | ό | μ | ο |
| D | | . | = | M |] | m | } | | | | · | ½ | Ν | έ | ν | ή |
| E | | . | > | N | ^ | n | ~ | | | | ® | Υ | Ξ | ή | ξ | ώ |
| F | | / | ? | O | _ | o | | | | | ™ | Ω | Ο | ι | ο | □ |

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 6

Characters, the foundation

| Standard | Country | No. of Characters |
|-----------------|---------|-------------------|
| JIS X 0208:1997 | Japan | 6,879 |
| JIS X 0212:1990 | Japan | 6,067 |
| KS X 1001:1992 | Korea | 8,224 |
| GB 2312-80 | China | 7,445 |
| Big Five | Taiwan | 13,494 |
| CNS 11643-1992 | Taiwan | 48,711 |

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 7

Characters, the foundation

European alphabetic scripts

Latin
Greek
Cyrillic
Armenian
Georgian
Runic
Ogham
Modifier letters
Combining characters

East Asian scripts

Han
Hiragana
Katakana
Hangul
Bopomofo
Yi

Middle East scripts

Hebrew
Arabic
Syriac
Thaana

Symbols

Currency symbols
Letter like symbols
Mathematic operators
Numeric forms
Technical symbols
Geometrical symbols
Miscellaneous symbols & dingbats
Enclosed & square
Braille

South & South East Asian scripts

Devanagari
Bengali
Gurmukhi
Gujarati
Panjabi
Oriya
Tamil
Telugu
Kannada
Malayalam
Sinhala
Thai
Lao
Tibetan
Myanmar
Khmer

Additional scripts

Ethiopic
Cherokee
Canadian Aboriginal
Syllabics
Mongolian

Etc....



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 8

Characters, the foundation

| Character | Bytes |
|-----------|-------------|
| A | 41 |
| á | C3 A1 |
| あ | E3 81 82 |
| 不 | F0 A3 8E B4 |

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 9

Characters, the foundation

axあaxあ

61 D7 90 E3 81 82 61 D7 90 E3 81 82



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 10

Characters, the foundation

ax|あ ax あ

61 D7 90 E3 81 82 61 D7 90 E3 81 82



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 11

Characters, the foundation

ax あ|ax あ

61 D7 90 E3 81 82 61 D7 90 E3 81 82



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 12

Internationalized Resource Identifiers

<http://raksmorgas.josefsson.org/mal/franzen.html>

<http://räksmörgås.josefsson.org/mål/franzén.html>

Easier to create

- ... memorize
- ... transcribe
- ... interpret
- ... guess / find things
- ... relate to (branding)

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 15

Internationalized Resource Identifiers

<http://raksmorgas.josefsson.org/mal/franzen.html>

<http://räksmörgås.josefsson.org/mål/franzén.html>

domain name

path


<http://rksmrqs-5wao|o.josefsson.org/m%C3%A5l/franz%C3%A9n.html>


- Phishing (www.paypal.com)

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 16

Internationalized document formats

 The title says "W3C, פעילות הבינאום" in Hebrew.

 Using the bidi algorithm only

The title says "W3C, פעילות הבינאום" in Hebrew.

Markup (characters ordered as in memory!)

"פ'עילות הבינאום, W3C"

Copyright © 2005 W3C (MIT, ERCIM, Keio)

side 18

Presentation Matters

国際サッカー連盟（FIFA）本部での理事会を前に記念撮影におさまる理事ら。前列右から2人目がヨハンソン（Johansson）欧州サッカー連盟会長。

Copyright © 2005 W3C (MIT, ERCIM, Keio)

side 22

Internationalized document formats

Volcanic eruptions have literally devastated large inhabited areas. During the 1914 eruption of Sakurajima in Kyushu, 687 houses in Kurokami were buried in hot ash. What remained of this shrine gate, previously five meters tall, was left as a reminder.



Kurokami maibutsu gate
(腹五社神社黒神埋没島居),
Sakurajima Island.

```
<image src="kk-torii.jpg" height="180" width="240"
caption="Kurokami maibutsu gate (腹五社神社黒神埋没島居), Sakurajima Island."/>
```

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 19

Internationalized document formats

Volcanic eruptions have literally devastated large inhabited areas. During the 1914 eruption of Sakurajima in Kyushu, 687 houses in Kurokami were buried in hot ash. What remained of this shrine gate, previously five meters tall, was left as a reminder.



Kurokami maibutsu gate
(腹五社神社黒神埋没島居),
Sakurajima Island.

```
<image src="kk-torii.jpg" height="180" width="240">
<caption>Kurokami maibutsu gate
(<span xml:lang="ja">腹五社神社黒神埋没島居</span>),
Sakurajima Island.</caption>
</image>
```

Other things: bidi markup, abbr, styling, etc.

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 20

Presentation Matters

国際サッカー連盟（FIFA）本部での理事会を前に記念撮影におさまる理事ら。前列右から2人目がヨハンソン（Johansson）欧州サッカー連盟会長。

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 23

Presentation Matters



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 24

Presentation Matters

| |
|---|
| Richard Ishida Consulting, Design of International User Interfaces |
| ريشارد ايشيدا - مستشار (123-456) تصميم عالمي واجهات (global design) للتفاعل |
| ריצ'רד אישידה - יועץ, חכנון מישקי למשתמש 123-456 בין לאומי |
| 理查德·伊希达、国际化软件界面设计顾问，国际化软件界面、设计顾问「国际化软件界面设计顾问」 |
| 石田リチャード ユーザーインターフェース「国際化対応設計」コンサルタント。 |
| 리차드 이시다 각국어 인터페이스 기술컨설팅도 |
| ਅਰਿਚਾਡ ਇਸ਼ਿਦਾ ਅੰਤਰਰਾਸ਼ਟਰੀ ਆਰਗਨਾਈਜ਼ੇਸ਼ਨ ਆਫ ਇੰਟਰਨੈਸ਼ਨਲ ਯੂਜ਼ਰ ਇੰਟਰਫੇਜ਼ (UI) |
| Ριτσαρντ Ισιδα Εὐμβουλός, Εξειδικευμένος διαβήτων διασυνδέσεων χρήστη |
| Ричард Ишида Консультант по интернационализации интерфейса пользователя |
| रिचर्ड ईशिदा परामशदाता, डिजाइन ऑफ इंटरनेशनल यूजर इंटरफेज |

<http://people.w3.org/rishida/scripts/samples/wrapping.html>

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 25

Presentation Matters

والدهان والمخرج والتاجر
والضابط والحي والميت، ويومئ
الينا فو از حديد عوسم ٤٩ اذ
يهيئ حسن فكرت عشر
مسرحيات، ثم يومئ اذ تراود

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 26

Presentation Matters

والدهان والمخرج والتاجر
والضابط والحي والميت، ويومئ
الينا فو از حديد عوسم ٤٩ اذ
يهيئ حسن فكرت عشر
مسرقيات، ثم يومئ اذ تراود

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 27

Supporting cultural differences

"... one Latin American teacher recently complained to me that the US-manufactured and well-translated educational software currently being used in his country's primary schools presupposed 'solitary problem solvers', whereas his culture stressed collective problem-solving."

Kenneth Keniston,
Language International, May 1996

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 34

Enabling localization

Tokyo hotel:

- ◆ Cooles and heates. If you want the condition of cool in your room, please control yourself.

Paris hotel:

- ◆ Guests are expected to complain at the office between the hours of 9am and 10am each day.

Tokyo hotel:

- ◆ You are invited to take advantage of the chambermaid.

Vienna hotel:

- ◆ In case of fire, do your utmost to alarm the porter.

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 29

Enabling localization

They are speaking to her from my new house.

Están hablándole desde mi casa nueva.

私の新しい家から彼女と話しています。

تكلّمونها من بيتي الجديد

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 30

Enabling localization

There were <X> spelling mistakes in file: <Y>

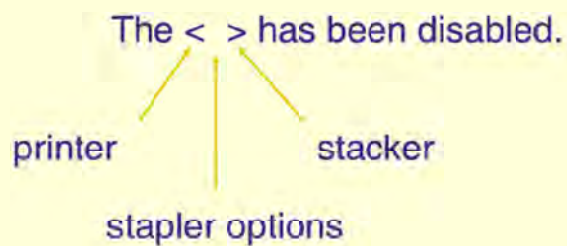
Datei <Y> enthält <X> Rechtschreibfehler.

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 31

Enabling localization

The < > has been disabled.



printer stacker

 stapler options

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 32

Supporting cultural differences

Website Directory Sites organised by subject.

- Arts & Humanities**
Literature, History, Photography...
- Business & Economy**
B2B, Shopping, Investments, Property...
- Computers & Internet**
Internet, Reviews, Software, Games...
- Education**
UK, Ireland, Universities...
- Entertainment**
Humour, Movies, Music, Actors...
- Government**
UK, Ireland, Politics, Law...
- Health**
Medicine, Drugs, Diseases, Fitness...
- News & Media**
Newspapers, Weather, TV...
- Recreation & Sport**
Sport, Hobbies, Travel, Motoring...
- Reference**
Maps, Dictionaries, Phone Numbers...
- Regional**
UK, Ireland, Countries...
- Science**
Animals, Geography, Engineering...
- Social Science**
Economics, Languages, Psychology...
- Society & Culture**
People, Food & Drink, Environment, Sexuality...

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 35

Supporting cultural differences

Guide Web - Classement thématique de sites

- Actualités et médias**
Journaux, Télévision, Météo...
- Commerce et économie**
B2B, Shopping, Emploi, Immobilier...
- Informatique et Internet**
Internet, Logiciels, Fonds d'écran...
- Santé**
Diététique, Médecine, Thalasso...
- Enseignement et formation**
Primaires, Secondaire, Supérieur...
- Institutions et politique**
Ministères, Droit, Politique...
- Sciences et technologies**
Animaux, Astronomie, Physique...
- Sports et loisirs**
Foot, Tourisme, Auto/Moto, Jeux...
- Art et culture**
Littérature, Cinéma, Musique, BD...
- Divertissement**
Tests/Quiz, Loteries, Humour, Soma...
- Classement géographique**
Pays, Europe, France, Paris...
- Références et annuaires**
Dictionnaires, Annuaires, Cartes/Atlas...
- Société**
Enfants, Gastronomie, Rencontres...
- Sciences humaines**
Archéologie, Histoire, Psychologie...

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 36

Supporting cultural differences

Yahoo!カテゴリ

エンターテインメント
映画,音楽,芸能人,コミック,占い

メディアとニュース
テレビ,ラジオ,新聞,雑誌...

趣味とスポーツ
アウトドア,ゲーム,車,スポーツ,競馬...

ビジネスと経済
ショッピング,B2B,雇用,金融...

芸術と文化
写真,建築,美術館,歴史,文学

各種資料と情報源
図書館,辞書,郵便,電話番号...

生活と文化
子ども,環境,グルメ,健康者...

コンピュータとインターネット
ハードウェア,ソフトウェア,WWW

教育
大学,専門学校,小中高,資格...

政治
政治,行政,国会,法...

健康と医学
病院,病気,ダイエット...

自然科学と技術
動物,エコロジー,地球,天文,工学...

社会科学
経済学,社会学,言語,政治学...

地域情報
日本の地方,世界の国...

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 37

Supporting cultural differences

Directory siti Web - I siti ordinati per argomento.

Economia e aziende
Lavoro, Finanza, Shopping, B2B

Società e culture
Sesso, Amore, A tavola, Cronaca

Aree geografiche
Paesi, Città italiane, Meteo, Mappe

Divertimento e spettacolo
Musica, Cinema, Celebrità, Fumetti

Attualità e media
Giornali, TV, Google

Informatica e Internet
Giochi, Canale, Chat, Software

Arte e cultura
Letteratura, Racconti, progetti, Moda

Sport e tempo libero
Viaggi, Sport, Calcio, Auto

Scienza e tecnologia
Cellulari, Animali, Ecologia

Medicina e salute
Gravidanza, Benessere, Malattie

Politica e istituzioni
Governo, Diritto, Fisco, Partiti

Scienze umane e sociali
Sessuologia, Storia, Geografia, Blog

Risorse
Calendari, Elencchi, telefonici

Istruzione e formazione
Università, Scuole, Dizionari, Libri

Copyright © 2005 W

slide 38

Current Areas of Activity

How do I ...

- Ensure that XHTML forms return data in the right encoding?
- Make my Hebrew or Arabic text display correctly?
- Declare language and encoding for XML documents?
- Order XSL output according to French rules?
- Approach the creation of multilingual documents in HTML?
- Help users navigate to the right localized page?
- Ensure the table I'm about to write has all the right i18n features?
- etc

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 41

Current Areas of Activity

General

- International & multilingual Web sites
- Monolingual vs. multilingual Web sites
- Serving XHTML 1.0

Characters & encoding

- Character encodings
- Document character set
- Who uses Unicode?
- Upgrading to Unicode
- An Introduction to Multilingual Web Addresses
- Language tags in HTML and XML
- The HTTP Charset parameter
- Checking HTTP headers
- Checking the character encoding using the validator
- CSS character encoding declarations
- Multilingual form encoding
- HTML, XHTML, XML and control codes
- Unexpected characters or blank lines
- Setting encoding in web authoring applications
- Setting 'charset' information in .htaccess
- Missing characters and glyphs
- Non-English tags

Language

- Why use the language attribute?
- Using HTTP and meta for language information
- Language tags in HTML and XML
- 2-letter or 3-letter language codes
- Styling using the lang attribute
- Setting language preferences in a browser

- When to use language negotiation
- Apache MultiViews language negotiation set up
- Accept-Language used for locale setting

Bidirectional text

- What you need to know about the bidi algorithm and inline markup
- Bidi space loss
- XHTML/HTML & bidi formatting codes vs. markup
- CSS vs. markup for bidi support

Resource identifiers

- An Introduction to Multilingual Web Addresses

Local formats

- Displaying formatted dates
- Select box sorting
- Dates and time

Miscellaneous

- Display capabilities
- Script direction & languages
- Introduction to Indic Scripts
- Ruby

Navigation

- Global gateway pull-downs

Other

- XSLT for multilingual output
- Using the link element

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 42

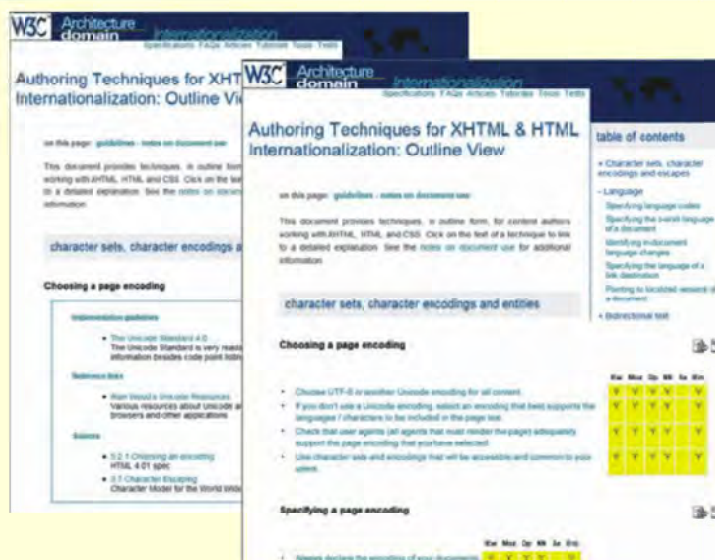
Current Areas of Activity



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 43

Current Areas of Activity



Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 44

Current Areas of Activity

W3C Internationalization (i18n) Activity
Making the World Wide Web truly World Wide!

The W3C Internationalization Activity has the goal of progressing and coordinating any techniques, conventions, guidelines and activities within the W3C and together with other organizations that allow and make it easy to use W3C technology worldwide, with different languages, scripts, and cultures.

Announcements

- 15 Aug 05 FAQ: Migrating from language-specific pages encoding to Unicode encoding **available for review**
- 12 Aug 05 New translations see bottom left of this page
- 17 Aug 05 FAQ: encoding in XML document schemas **available for review**
- 13 Aug 05 Tutorial: creating an XML schema, nmtoken, nmtoken, nmtoken **available for review**

Web Services Internationalization (WS-I18N)

New First Working Draft
 The Internationalization Core Working Group has released the First Public Working Draft of Web Services Internationalization (WS-I18N). The draft enhances SOAP messaging for local and international preference negotiation and defines a locale policy, without using Accept-Language and user identity, implementations can handle the requester's locale, locale policy and language preference.

Changing XHTML page encoding to UTF-8
 New FAQ based article
 How do I change the encoding of my XHTML pages to UTF-8?

W3C Participates in 28th Int'l. News Item
 The 28th Internationalization Conference will be held 7-9 September in Berlin, Germany. The conference will be a technical conference worldwide for software and Web internationalization. Read about Unicode and the W3C Internationalization Activity.

More news:
 More announcements [RSS] - Resources [RSS] - Reviews [RSS] - Changes [RSS] - Translations [RSS] - Tests [RSS] - General [RSS]

<http://www.w3.org/International/>

Copyright: © 2005 W3C (MIT, ERCIM, Keio)

slide 45

Getting Involved

- ensure that the needs of your scripts and languages are brought to the attention of the W3C developers
- always use standards and validate your code
- push for adoption of international technology by user agents
- editing environments are also key to success
- importance of the mobile web

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 47

Getting Involved

- visit the I18n Activity Home Page
- join a Task Force, the Interest Group, or the public mailing list
- offer to help with reviews, or provide local knowledge for other WGs
- provide translations of W3C specifications or articles
- take advantage of the i18n-readiness of W3C technology

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 48

Current Areas of Activity

Internationalization Activity

Core Working Group

ITS (Internationalization Tag Set) Working Group

GEO (Guidelines, Education & Outreach) Working Group

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 40

Getting Involved

- this is your Web – not the W3C's – if something isn't right, get involved to fix it

Thank you
<http://www.w3.org/International/>

Copyright © 2005 W3C (MIT, ERCIM, Keio)

slide 49