

インターネット学術情報インデックス (IRI)の構築・運用とネットワーク情 報資源

継続教育コース2
Open Access Informationの徹底活用
基礎編

東京大学情報基盤センター
小山憲司

目次

- 東京大学情報基盤センターの概要
- 図書館電子化部門の概要
- デジタル・ライブラリ系の概要

- インターネット学術情報インデックス(IRI)の概要
- IRIにおける情報管理プロセス
- IRIとネットワーク情報資源

情報基盤センターの概要

- 1999年4月、大学の情報基盤整備を目的として発足した、研究・業務部局
 - 情報メディア教育部門
 - 教育用計算機システム(ECCS)など
 - 図書館電子化部門
 - キャンパスネットワークング部門
 - UTnetの運用・管理
 - スーパーコンピューティング部門
 - スパコンの運用・管理(全国共同利用)

図書館電子化部門の概要

- 図書館情報係
 - 図書館業務システムの運営・管理
- デジタル・ライブラリ係
 - デジタル・コンテンツに係るさまざまなサービスの提供
- 学術情報リテラシー係
 - デジタル化されたさまざまなサービスの利用に係る教育支援

<http://www.dl.itc.u-tokyo.ac.jp/>

デジタル・ライブラリ系の概要

- 学内研究成果の収集・発信
 - 東京大学学術機関リポジトリ UT Repository
 - 学位論文データベース
 - Academic Navi U-Tokyo
- デジタル・コンテンツの利用環境の整備
 - 電子ジャーナルリンク集
 - E-Journal Portal
 - インターネット学術情報インデックス(IRI)
- デジタル・コンテンツの作成と発信
 - 電子版貴重書コレクション
 - ブック・コンテンツDBのデータ作成

デジタル・ライブラリ系の概要2

- 広い意味でオープンアクセス情報を収集、構築・整理、発信する係
 - 学内で生産された学術情報の収集・整理・発信
 - 機関リポジトリ
 - 学位論文データベース
 - 学内(主に総合図書館)で所蔵する資料の電子化
 - 電子版貴重書コレクション
 - 学内外のネットワーク情報源の収集・整理、および提供
 - インターネット学術情報インデックス
 - Academic Navi U-Tokyo

参考: 東京大学学術機関リポジトリ



- UT Repository
- 2006年4月1日公開
- 登録コンテンツ数 1,753件
(2006年7月15日現在)
 - 学術雑誌論文 38件
 - 学位論文 233件
 - 紀要 1,473件
 - 図書、図書の一部 1件
 - 教育資料 3件
 - その他 5件

<http://repository.dl.itc.u-tokyo.ac.jp/>

本日の内容

- インターネット学術情報インデックス(IRI)の概要
- IRIにおける情報管理プロセス
- IRIとネットワーク情報資源

インターネット学術情報インデックス(IRI) の概要

- IRIとは
- IRI作成の動機
- IRIの現状



IRIとは

- Index to Resources on Internet
- 1999年3月公開
- インターネット上の学術情報を収集し、
- それらについて記述したメタデータをデータベースとして組織化し、
- キーワードや分類で検索可能とした、サブジェクト・ゲートウェイ

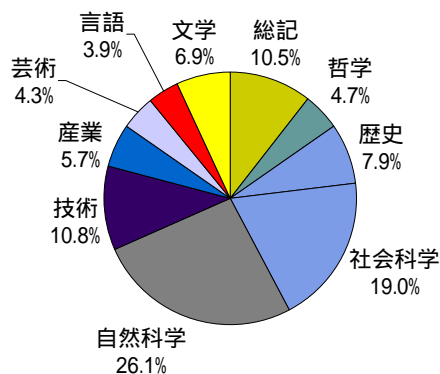
http://resource.lib.u-tokyo.ac.jp/iri/url_search.cgi

IRI作成の動機

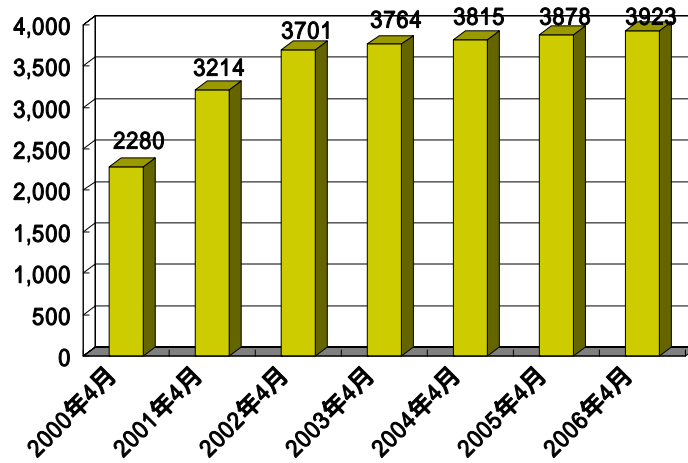
- 膨大な数のネットワーク情報資源の中から学術情報を検索することの困難さ。
 - 検索エンジンによる検索にはノイズが混ざる、適合情報を見逃してしまうなどの限界がある。
 - リンク集では、分野の偏りやその量、検索の方法において限界がある。
- 学術情報を対象とした、より適切で、有意義なネットワーク情報源の提供手段はないか。

IRIの現状

- データ件数: 4,002件
(2006年6月26日現在)
- 分野別(重複あり)

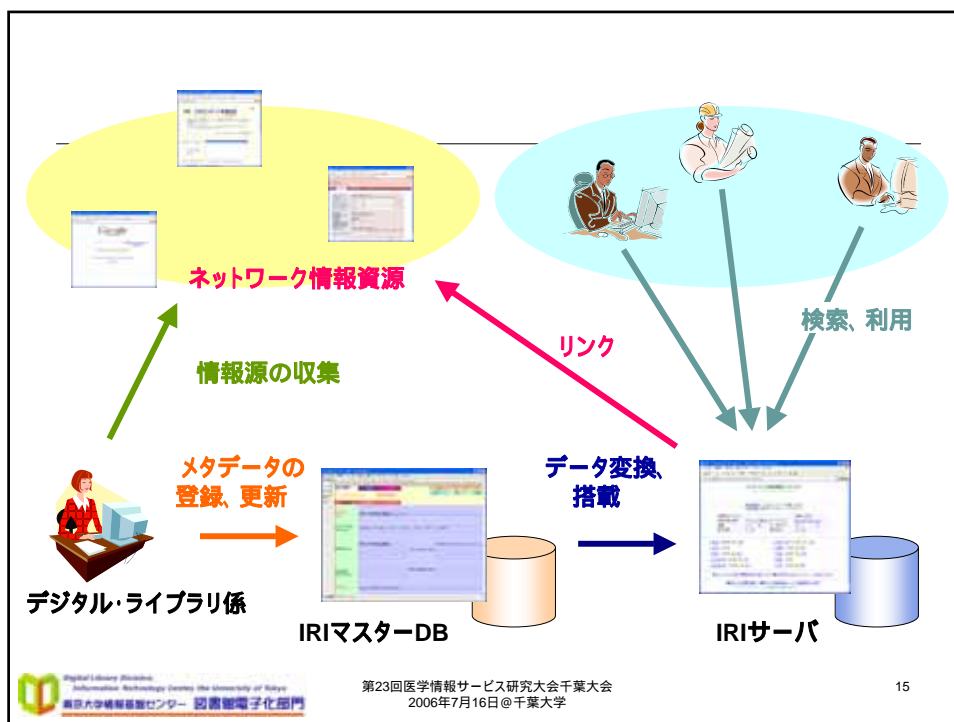


参考:IRIのデータ登録件数の推移



IRIにおける情報管理プロセス

- IRIにおける情報管理プロセスの概要
- 情報源の収集と選定
- メタデータの記述と登録



情報源の収集と選定

- 収集の対象
 - 蔵書、書誌、出版情報(ただし、特定の主題に関するもの)
 - 各分野に関する総合的な情報リソース(各分野のコア・ページ)
 - 特定主題に関するリンク集など
 - 辞書、用語集
 - データベースなど

情報源の収集と選定

- 収集の方法
 - 人手による
- 収録情報の情報源
 - 各専門分野の単行書・学術雑誌
 - 大学・学術機関・学会・学術出版社等のWebページ
 - 学術情報源を中心としたリンク集
 - 上記3点から得られるリンク先

情報源の収集と選定

- 選択の基準
 - 内容が学術的で、研究・教育に有用である。
 - 信頼性の高い学術情報源である。(典拠などが明示されている等)
 - 継続的に、安定した運営がおこなわれている。
 - 量的にある程度充実している。
 - 複数の学術情報機関からリンクされている。
 - 大学、研究機関、学会、学術出版者が運営している。
 - 企業体のサイトの場合、営利性が低く、内容が研究・教育を中心としている。
 - 個人のページの場合、更新頻度が高く、内容が学術的である。
 - 個人および特定の団体を宣伝するサイトでない。
 - 作者あるいは情報ソースが明示されている(匿名サイトはなるべく採用しない)。

メタデータの記述と登録

- DCを基本に、独自のエレメントを追加
 - 追加項目の例
 - ヨミ(タイトル、作成者、責任団体)
 - NDCを利用した分類の付与
 - LCSHを利用した検索キーワードの付与
 - フリーワードの付与
 - データタイプNIIの付与
- 入力基準は、NACSIS-CATに準拠

IRIとネットワーク情報資源

- 収集における課題
- 組織化における課題
- メンテナンスにおける課題

収集における課題

- 何を掘り所とするか。
 - 情報源の問題
 - 収集方針の問題
- 収集の方法は今のままでよいか。
 - IRI担当の体制の問題
 - システムの問題
- 何を収録の対象とするか。

何を掘り所とするか

- 広範なインターネット情報資源の中から、いかに情報を収集するのか。
 - そのためのリソース(情報源)
 - 学術情報としての評価 = 収録のための基準

組織化における課題

- 何を組織化の対象とするか。
 - 情報単位の問題
- 何を記述の情報源とするか。
 - 情報源そのものの問題
 - Web情報源の性質の問題
- どのように記述するか。
 - 記述規則の問題

何を組織化の対象とするか

- 組織化の対象 (= 収録の対象) となる情報の単位 (ユニット) が明確ではない。
 - 例: 土木学会のデジタルアーカイブス
 - <http://www.jsce.or.jp/library/page/report.html>

何を記述の情報源とするか

- IRIで用意しているメタデータ・エレメントの記述の根拠があいまいである。
 - タイトルは？
 - 記述言語であるHTMLによる表現のあいまいさ
 - 作者は？
 - 責任性(著者性)のあいまいさ
 - 公開日は？
 - 出版事項のあいまいさ
 - 同じもの？
 - 同定識別の困難さと情報の可変性

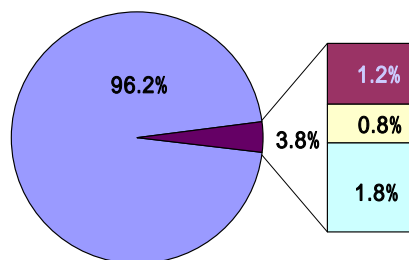
どのように記述するか

- 記述規則の明確化
 - 記述項目および記述方法について、他の動向を踏まえた改訂
- 相互利用可能性を視野に入れたシステムづくり
 - 他のシステムとのメタデータの相互利用

メンテナンスにおける課題

- リンクチェックソフトによるチェック
 - URLの変更
 - 情報そのものの消滅

2006年5月実施のチェック結果



■ OK ■ URL変更 □ サイトの消滅など □ 一時的にアクセス不可

ご清聴ありがとうございました