

Knowledge Sharing System

0. Introduction

1. Information Cycle

武田英明 Hideaki Takeda

takeda@nii.ac.jp

国立情報学研究所

National Institute of Informatics

Hideaki Takeda / National Institute of Informatics



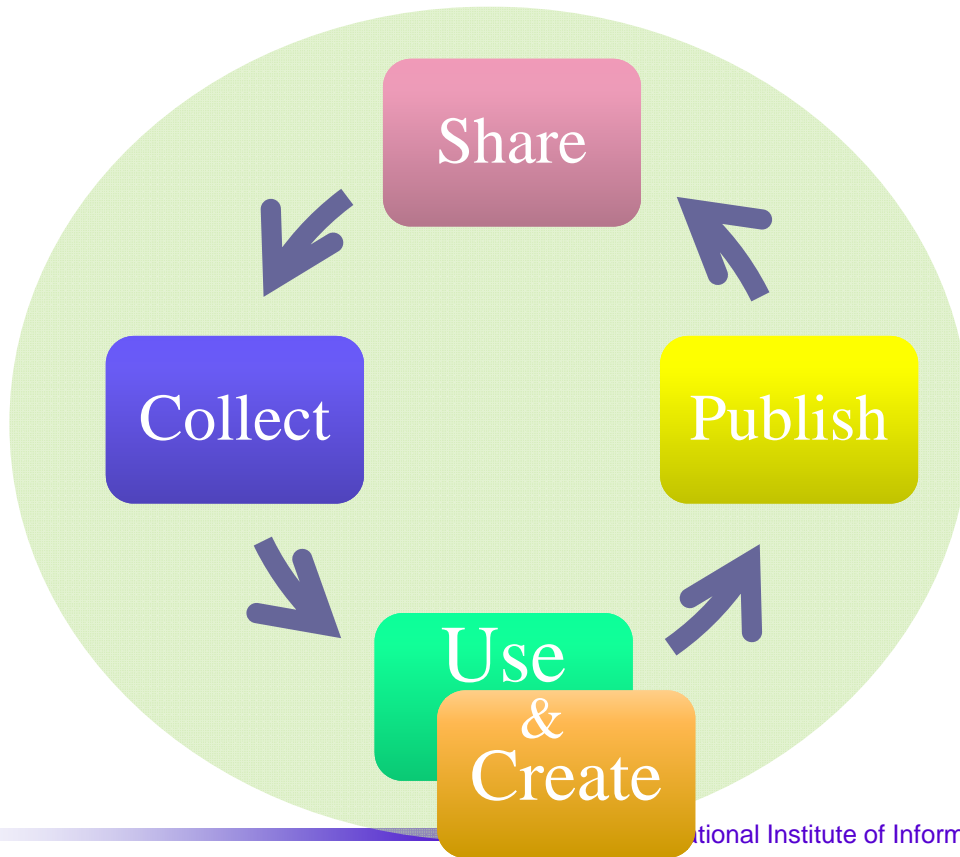
Outline of this lecture

- Table of Contents
 - Information Cycle
 - Information Integration
 - Semantic Web
 - Linked Data
 - Social Network
- Style of the lecture
 - 2-3 talks
 - Assignments
 - Exercise and Presentation

Hideaki Takeda / National Institute of Informatics



Information Cycle

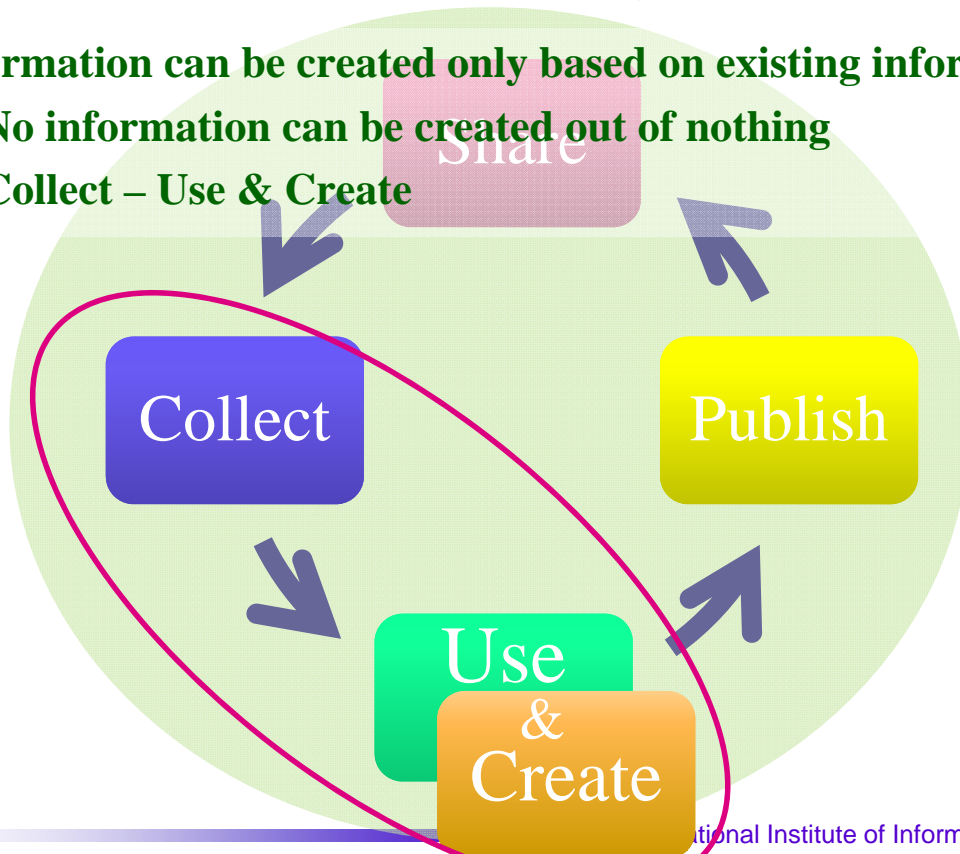


ational Institute of Informatics

NII

Information Cycle

- **Information can be created only based on existing information**
 - **No information can be created out of nothing**
 - **Collect – Use & Create**

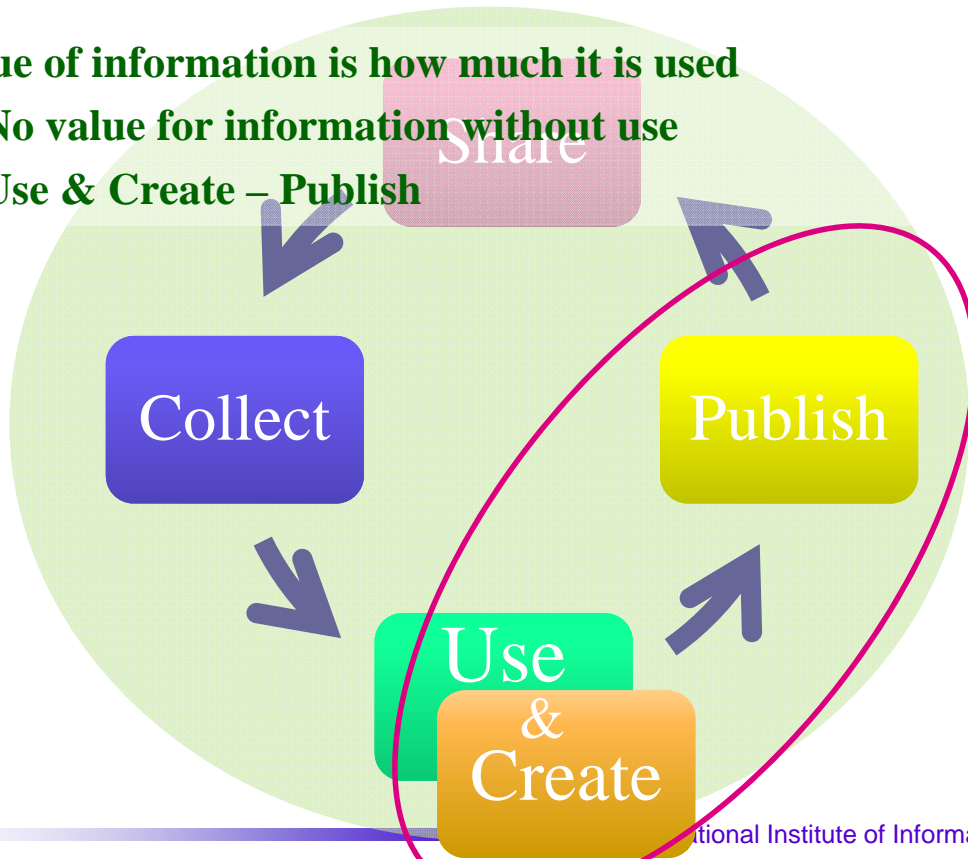


ational Institute of Informatics

NII

Information Cycle

- Value of information is how much it is used
 - No value for information without use
 - Use & Create – Publish

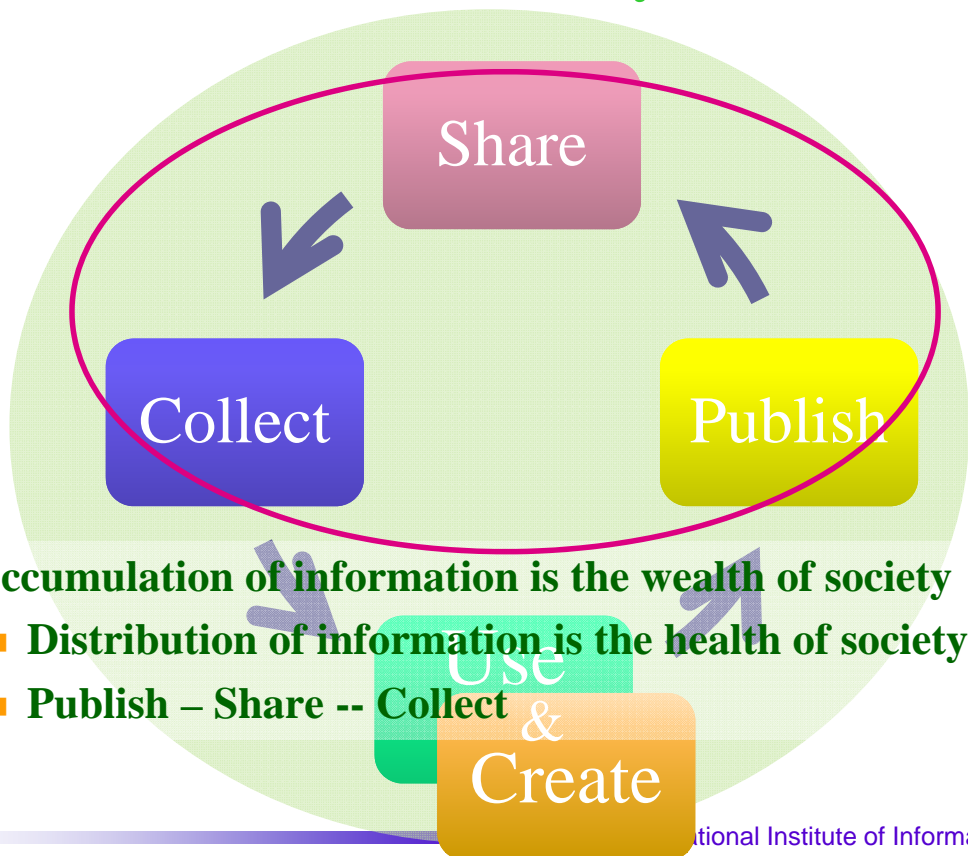


ational Institute of Informatics

NII ■

Information Cycle

- Accumulation of information is the wealth of society
 - Distribution of information is the health of society
 - Publish – Share -- Collect

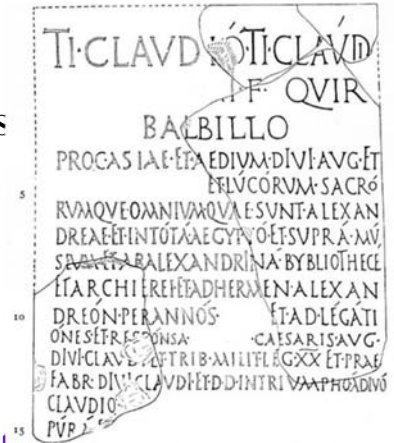


ational Institute of Informatics

NII ■

Ancient times

- Oral communication
 - Coverage: Everyone
 - Speed of diffusion: very slow
 - Style of diffusion: one-to-one
 - Quantity: small
- Hand writing
 - Coverage: very limited
 - ◆ Aristocracy, monks, bureaucrats, scholars
 - c.f. Library of Alexandria (3c BC)
 - Speed of diffusion: very slow
 - Style of diffusion: one-to-some
 - Quantity: small



Hideaki Takeda / National I

Gutenberg Revolution

- Letterpress Printing (活版印刷) 1440
 - Coverage: split-off between authors and readers
 - ◆ Author: professional writers, scholars
 - ◆ Reader: many people
 - Speed of diffusion: faster
 - Style of diffusion: one-to-many
 - Quantity: larger



Hideaki Takeda / National I

Communication Revolution

- Rapid Growth of communication technology since mid. 19c
 - Telegraph (Samuel Morse ca.1836)
 - Phonograph (Record Player, Thomas Edison ca. 1877)
 - Motion Picture (Cinema, Auguste and Louis Lumière, 1895)
 - Radio Broadcasting (Reginald Fessenden,1906)
 - TV Broadcasting (1928)



Birth of Mass media



Hideaki Takeda



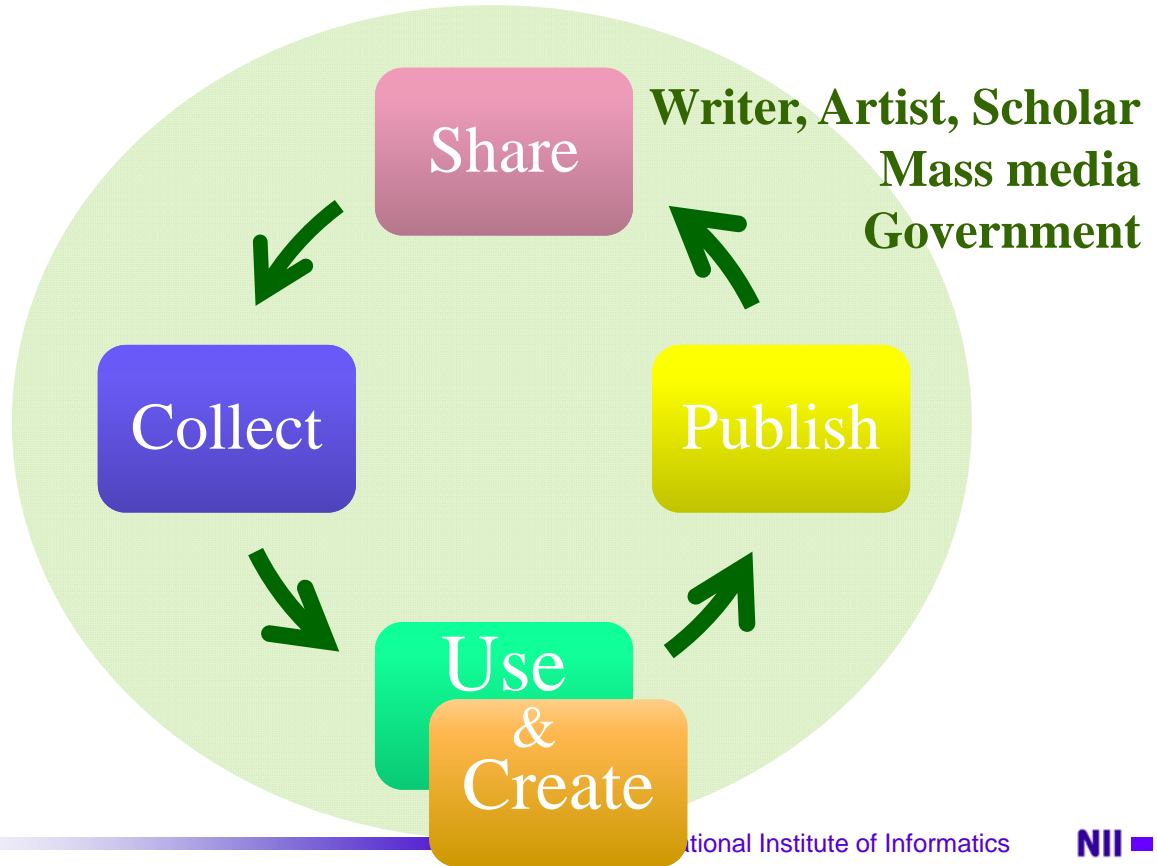
informatics

NII

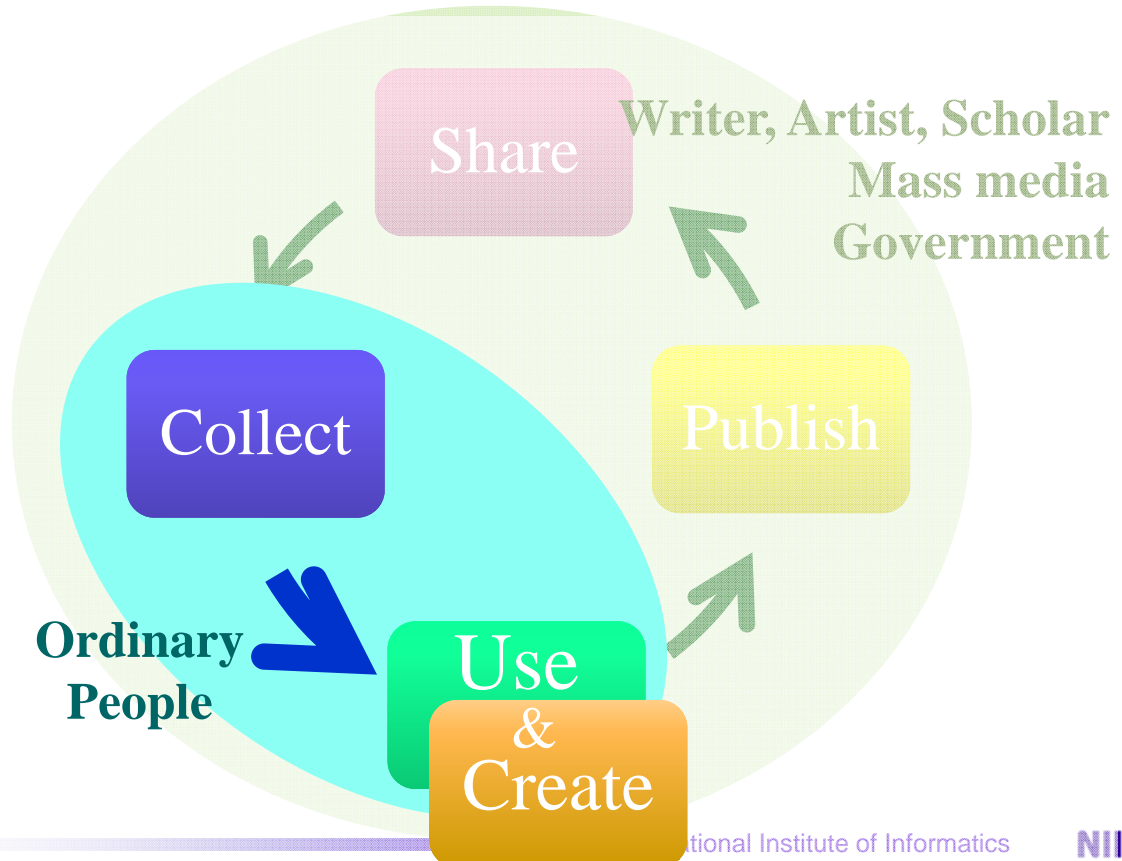
Mass media

- Newspapers, Magazine, Radio broadcasting, TV broadcasting
 - Coverage:
 - ◆ Author: **press**, scholars, politician, government, professional writers
 - ◆ Reader: **Almost everyone**
 - Speed of diffusion: from **simultaneous** to a couple of days
 - Style of diffusion: one-to-so-many
 - Quantity: much larger

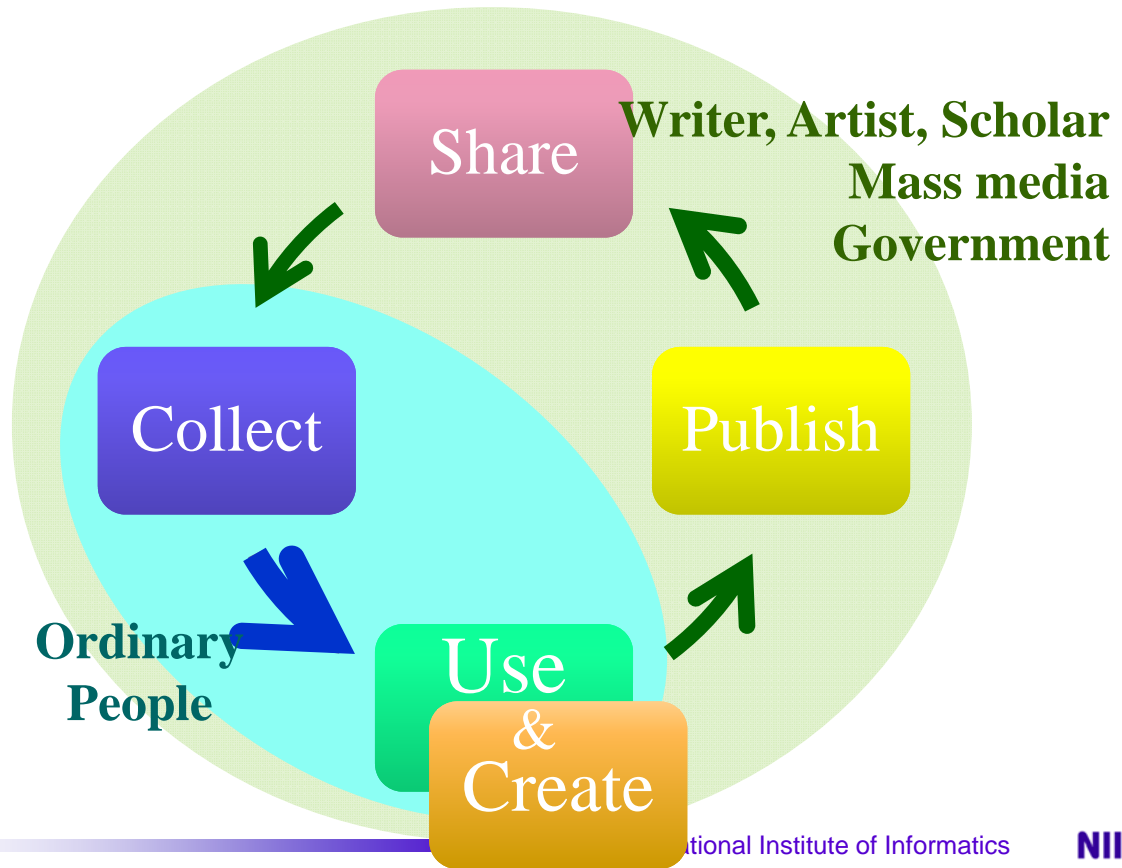
Two social layers on information cycle with Mass Media



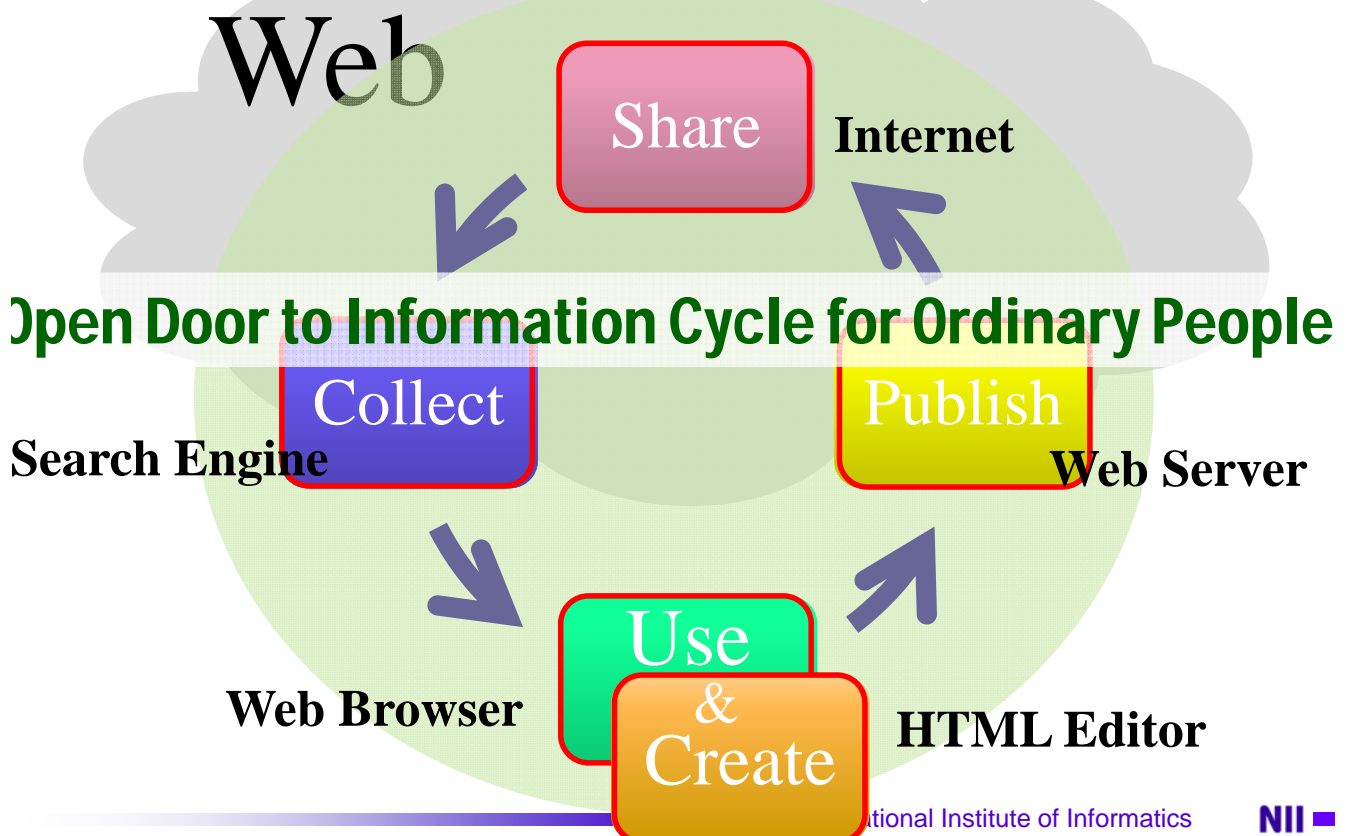
Two social layers on information cycle with Mass media



Two social layers on information cycle with Mass Media



Information Cycle with Web



Internet - history

- 1969: First interconnection between UCLA and, SRI
- 1974: Specification of TCP/IP (IETF RFC675)
 - *"Internet" as a single global TCP/IP network*
- 1983: APRANET switched to TCP/IP
- 1984: Domain Name Services started
- 1984: Over 1,000 hosts on Internet
- 1980s: Many interconnection among universities
- 1988: Commercial Internet Providers started in US.
- 1989: Over 10,000 hosts, Countries connected to NSFNET: au, de, il, it, jp, mx, nl, nz, pr, uk
- 1992: Over 100,000 hosts
- 1974: Unix code released
- 1970s: Migration of Unix onto different machines
- 1976: Apple I released
- 1981: IBM PC released
- 1984: Macintosh released
- 1980mid: Unix workstations widely used in universities
- 1991: Windows 3.1

Hideaki Takeda / National Institute of Informatics

NII ■

Internet

- The early period: Infrastructure of personal communication
 - Protocol: TCP/IP, uucp
 - Usage: Telnet, ftp, email, netnews
 - Functions: very slow, narrow, and unstable TCP/IP connection, no TCP/IP connection (uucp)
- The middle period: Infrastructure of public communication
 - Protocol: TCP/IP
 - Usage: email, ftp, netnews, WWW
 - Functions: reliable TCP/IP connection
- Now
 - Protocol: TCP/IP (http)
 - Usage: WWW, email
 - Functions: real-time connection

networking

storage

Hideaki Takeda / National Institute of Informatics

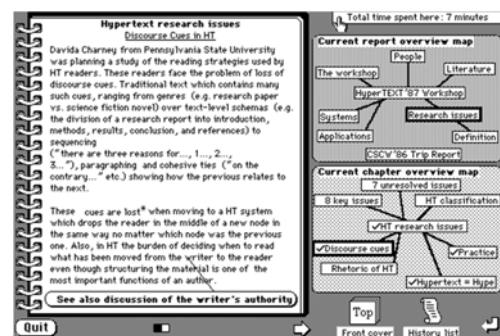
NII ■

Web (World Wide Web, WWW)

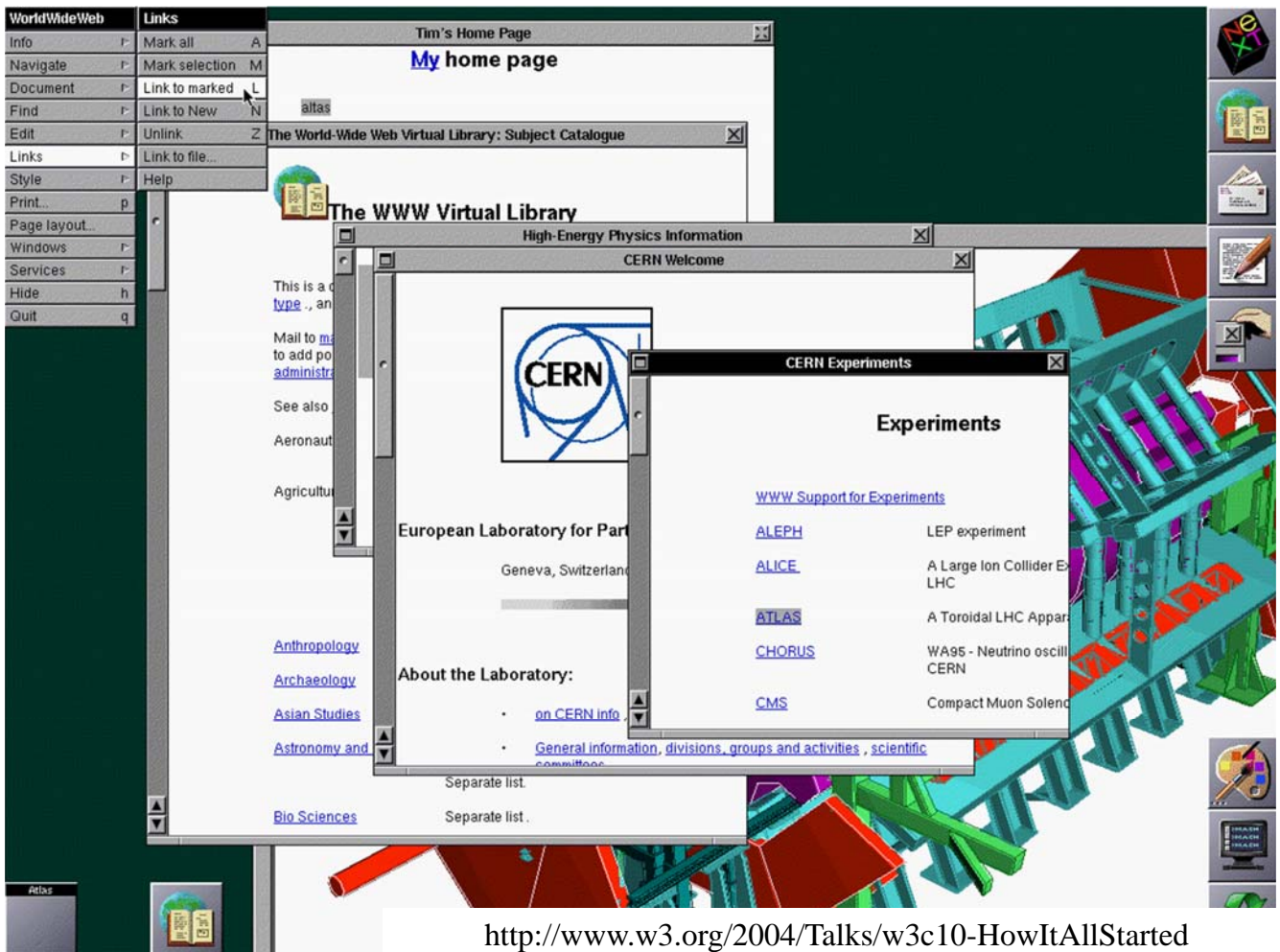
- 1980: Tim Berners-Lee (TBL) developed ENQUIRE
- 1989: TBL wrote a proposal “Information Management: A Proposal”
- 1990: TBL open the first WWW page and HTML browser (editor, called “WorldWideWeb”)
- 1991: TBL posted summary of WWW to alt.hypertext newsgroup
- 1992: National Center for Supercomputing Applications at the Univ. of Illinois at Urbana-Champaign (NCSA-UIUC), led by Marc Andreessen, developed NCSA Mosaic browser (images can be used)
- 1994: TBL founded the World Wide Web Consortium (W3C). *The W3C decided that their standards must be based on royalty-free technology*
- 1996-1998: Commercialization of Web
- 1999-2001: “Dot-com” boom and burst

Before Web

- Ancestor Hypertext Systems
 - ◆ Memex (V. Bush, 1945)
 - ◆ Hypertext Editing System (A. van Dam, T. Nelson, 1969)
 - ◆ HyperCard (Apple Computer, 1987)



- ◆ SGML (Standard Generalized Markup Language)



<http://www.w3.org/2004/Talks/w3c10-HowItAllStarted>

World Wide Web

The WorldWideWeb (W3) is a wide-area [hypermedia](#) information retrieval initiative aiming to give everyone access to everything there is online about W3 is linked directly or indirectly to this document, including an [example](#)

[What's out there?](#)

Pointers to the world's online information, [subjects](#) , [W3 servers](#), etc.

[Help](#)

Information on the browser you are using

[Software Products](#)

A list of W3 project components and their current state. (e.g. [Line Mode](#) ,[X11](#) [Viola](#) , [NeXT](#))

[Technical](#)

Details of protocols, formats, program internals etc

[Bibliography](#)

Paper documentation on W3 and references.

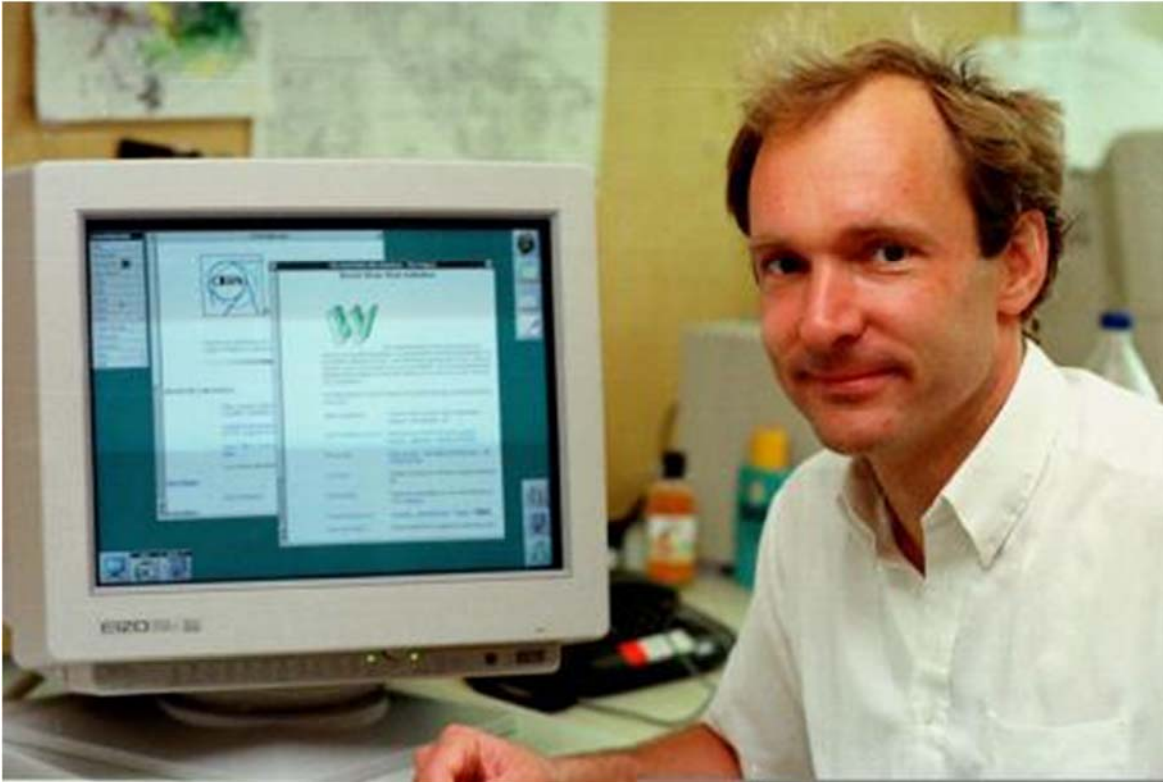
[People](#)

A list of some people involved in the project.

[History](#)

A summary of the history of the project.

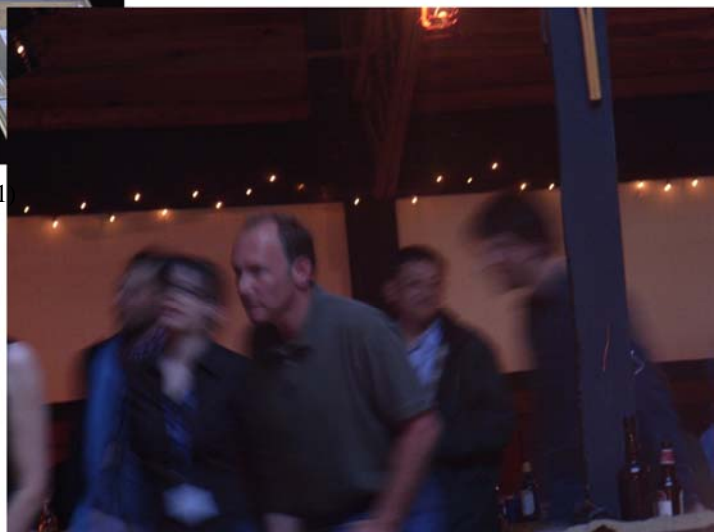
[How can I help ?](#)



Hideaki Takeda / National Institute of Informatics



話すTim Berners-Lee (Hyderabad, India, 2011)



踊るTim Berners-Lee (Banff, Canada, 2007)

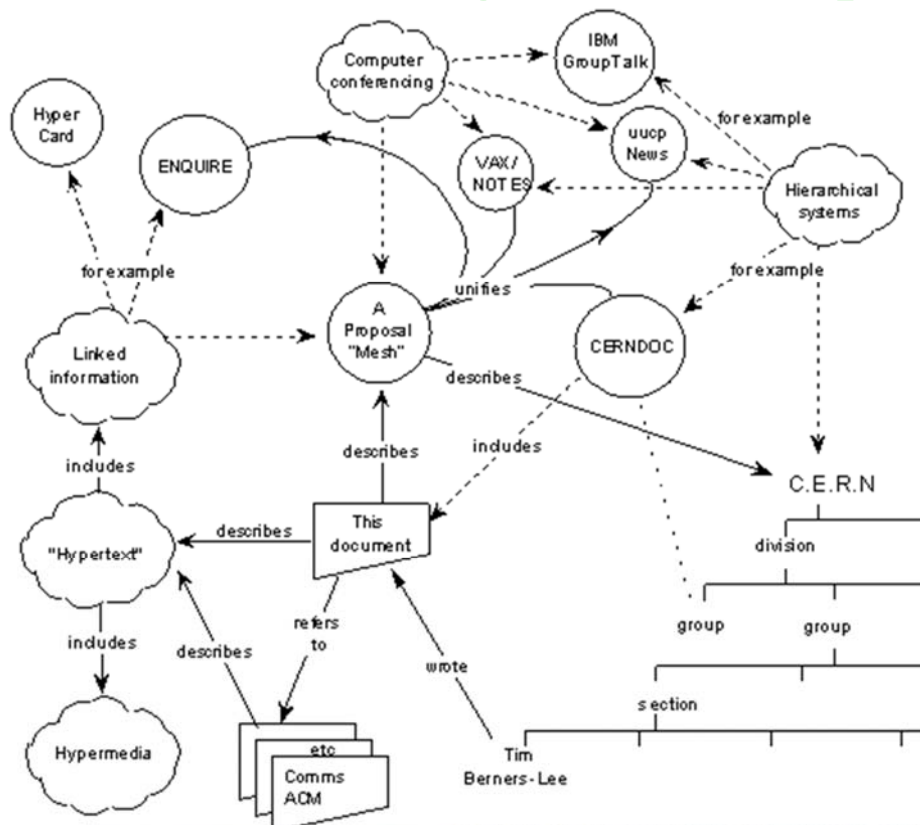
Hideaki Takeda / National Institute of Informatics



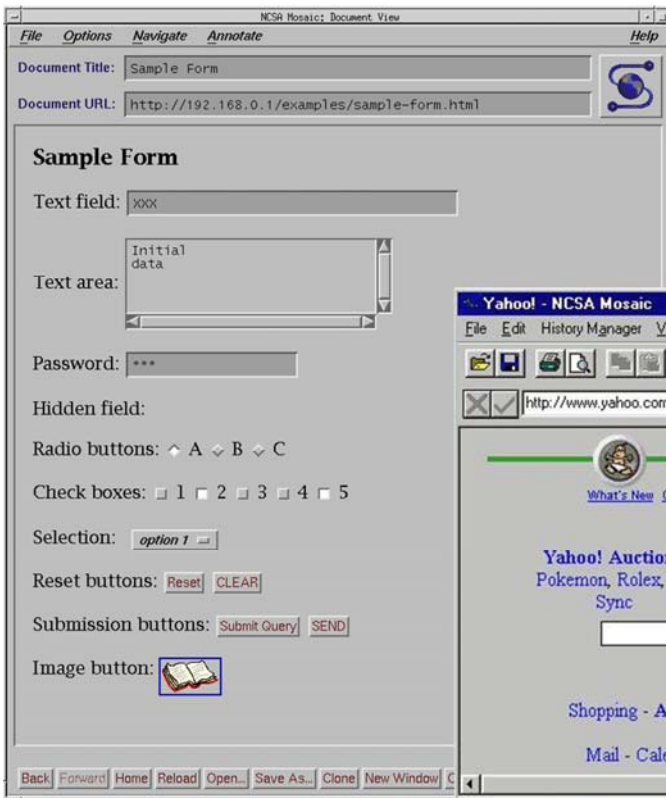


Hypertext '91

Information Management: A Proposal



NCSA Mosaic



Web History - Web Architecture and Information Management

Spring 2010 — INFO 190-02 (CCN 42509)

Wilde and Ryan Shaw, UC Berkeley School of Information, 2010-01-25

<http://dret.net/lectures/web-spring10/history>

Hideaki Takeda / National Institute of Informatics



図2 1996年ごろの研究者のWebページ例

Hideaki Takeda / National Institute of Informatics

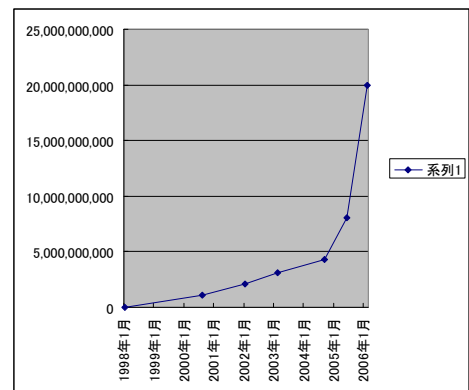


Web Architecture

- Essential technologies
 - Uniform Resource Identifier (URI): a system of globally unique identifiers for resources on the Web and elsewhere
 - HyperText Markup Language (HTML): the publishing language;
 - Hypertext Transfer Protocol (HTTP)
- Server-Client Architecture base on URI, HTML, HTTP
 - Fully distributed systems
 - ◆ No center needed
 - ◆ How to access?
- Royalty-free policy
 - Rapid software development (everyone can develop software)
 - Rapid software deployment (everyone can use software)

Some statistics (but old!)

- Over 20 billion pages (google, Yahoo!)
 - 2006 Feb: Yahoo!



- Over 80 million servers (80,655,992)
 - 2006 April
 - Web Server Survey

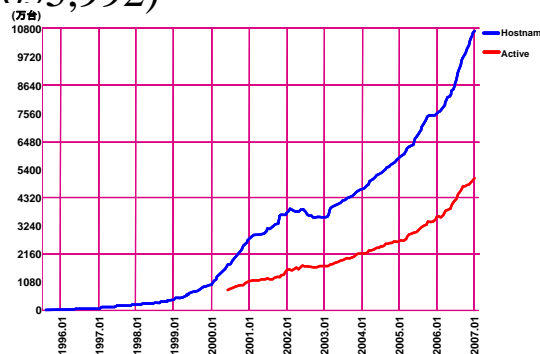


図1: Webサーバの台数の変化
(NetCraft, <http://www.netcraft.com>)

- 0.69 billion users (14% of population)

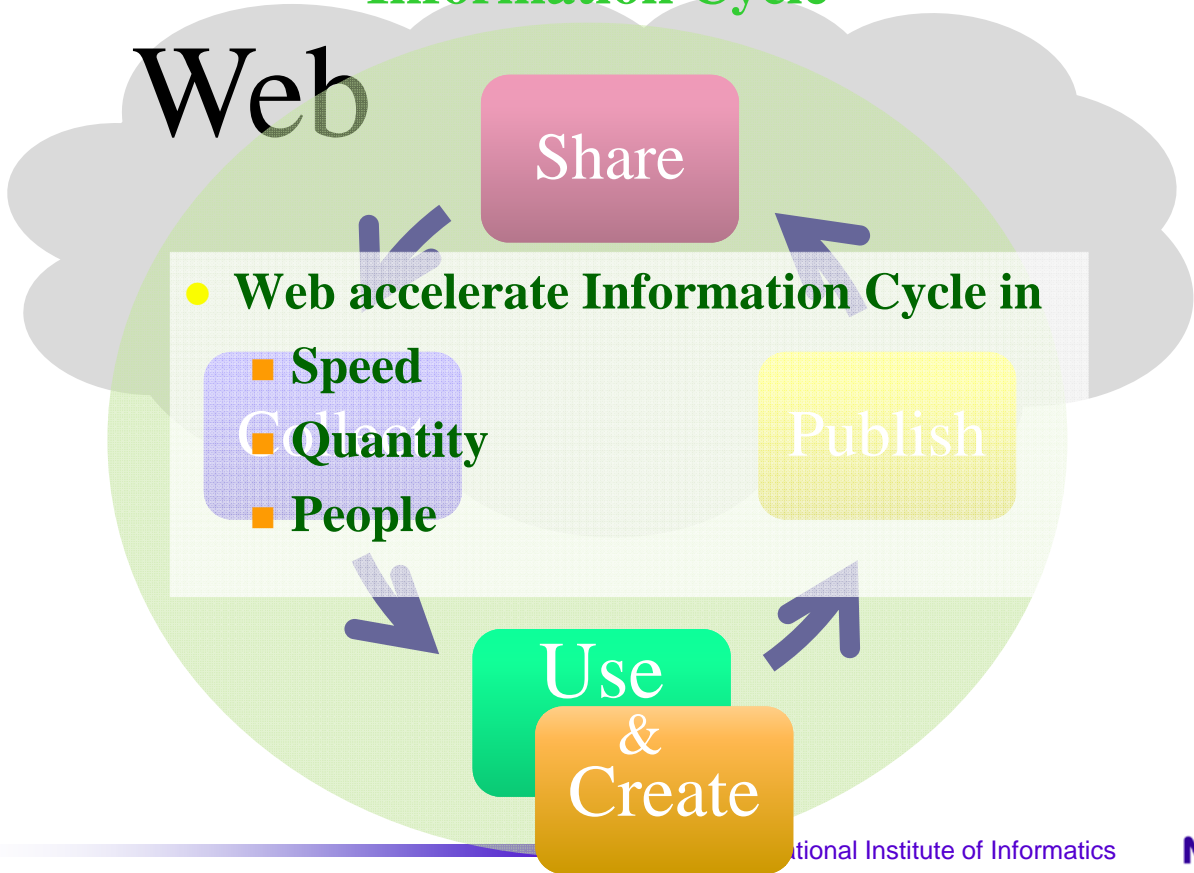
- 2006 March
- comScore <http://www.comscore.com/press/release.asp?press=849>

Top 15 Online Populations by Country, Among Visitors Age 15+*
March 2006
Source: comScore World Metrix

	Unique Visitors (000)
Worldwide Total	694,260
United States	152,046
China	74,727
Japan	52,100
Germany	31,813
United Kingdom	30,190
South Korea	24,645
France	23,884
Canada	18,996
Italy	16,834
India	16,713
Brazil	13,186
Spain	12,452
Netherlands	10,969
Russia	10,833
Australia	9,735

Information Cycle

Web



National Institute of Informatics

NII

Why is Web diffused so much?

- Openness
- Economics
- Easiness

Why is Web diffused so much? - Openness

- Openness:
 - Open to participate
 - Open to relate
- Internet
 - Open to participate
 - ◆ Open to join the computers to the network
 - Open to relate
 - ◆ Open to communicate to any computers
- Web
 - Open to participate
 - ◆ Open to make web pages
 - Open to relate
 - ◆ Open to make links to others

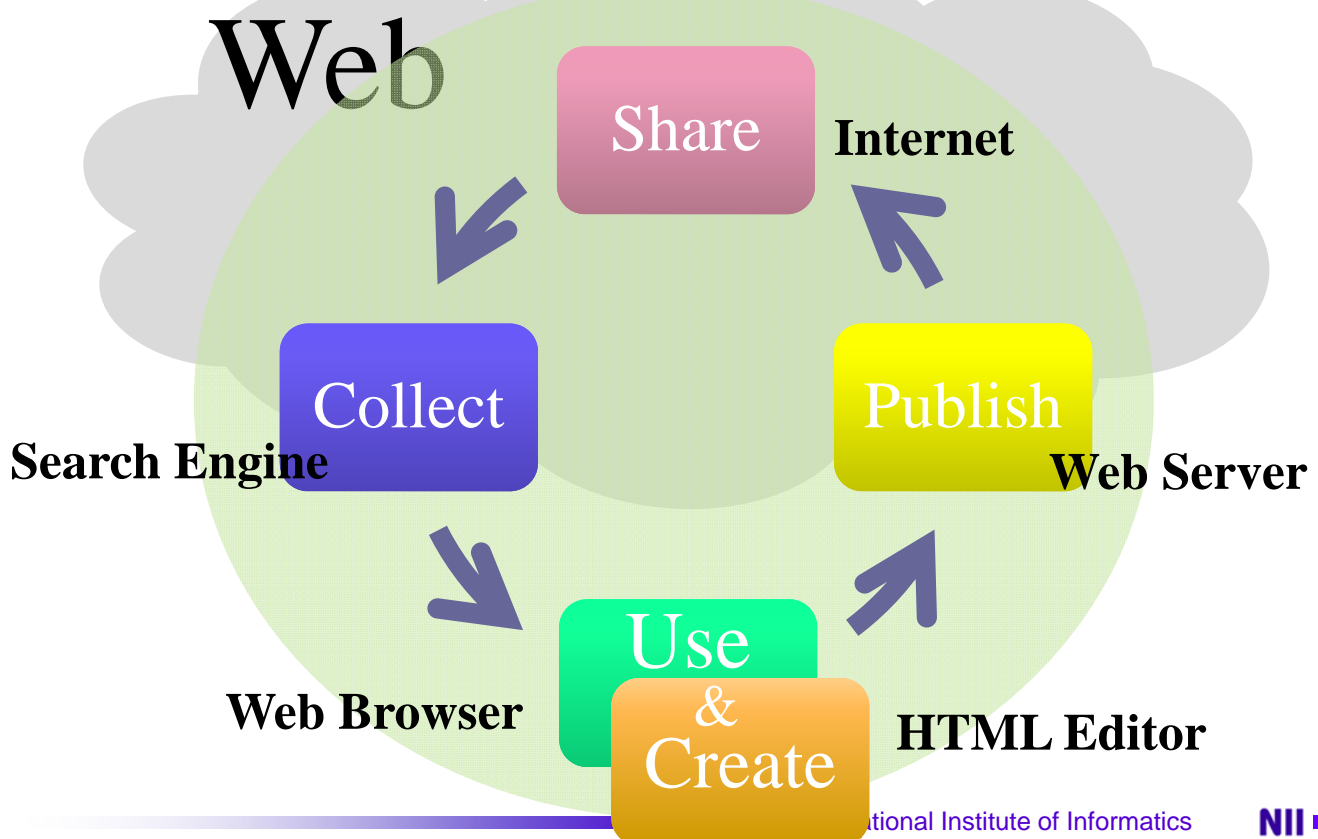
Why is Web diffused so much? - Economics

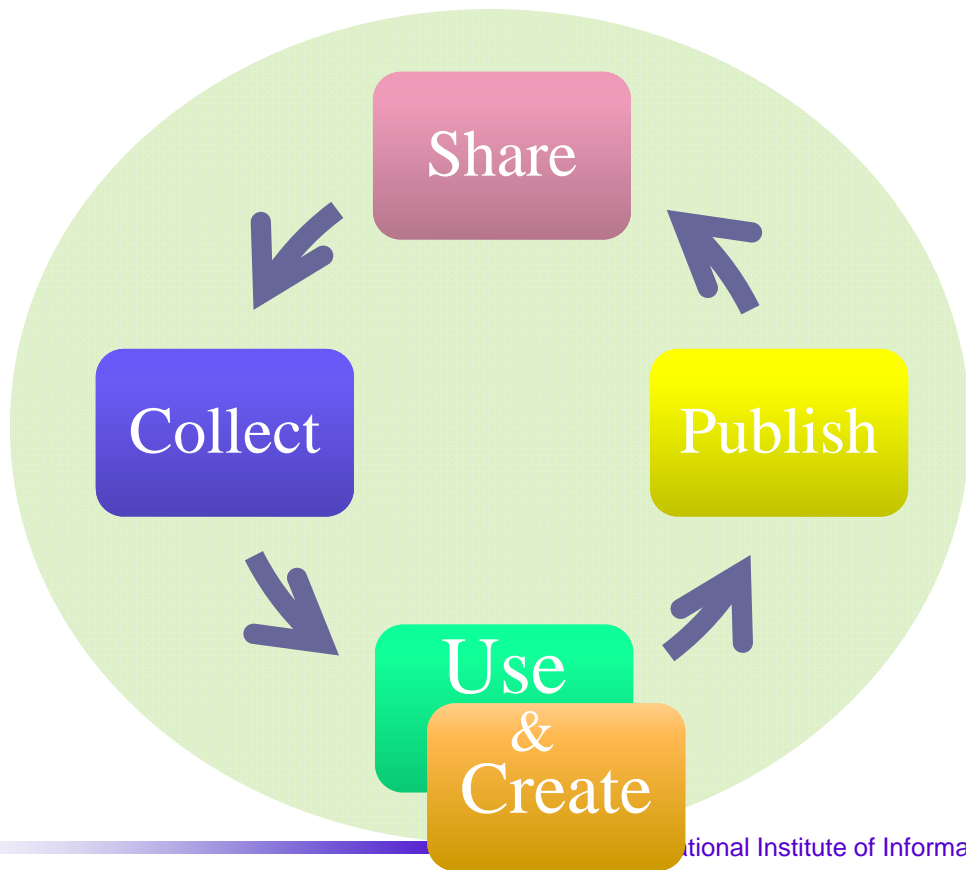
- Cost is not proportional to quantity of information
 - Packet communication (TCP/IP)
- Cost is proportional to quantity of information in legacy media
 - For people: A really new media that people can use

Why is Web diffused so much? - Easiness

- The first computer systems for ordinary people
- Internet is a system for computer professionals (or *geeks*)
 - Read manuals, solve problems by yourself
- TBL designed Web not for computer professionals but for researchers (physics)
 - Simplicity
 - Robustness

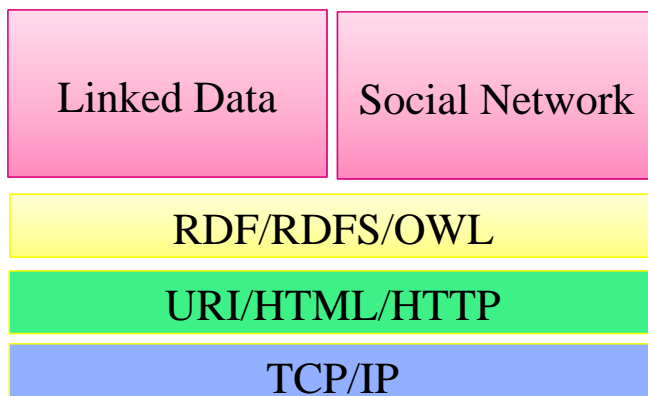
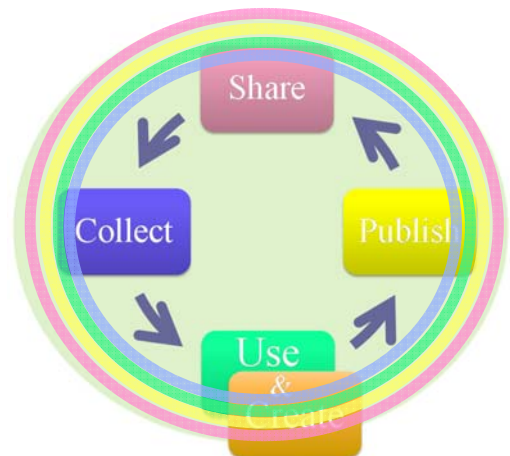
Information Cycle with Web





Layers for Information Cycle

- Internet
- Web
- Semantic Web
- Linked Data
- Social Network



Semantic Web

Web

Internet