



Language Resources and Legal Issues: Problems and Solutions for Basic and Industrial Research

John Hendrik Weitzmann

EEAR gGmbH
Creative Commons Germany
legal@creativecommons.de

Prodromos Tsiavos

LSE
Creative Commons UK and Greece
prodromos.tsiavos@gmail.com

META-FORUM 2010: Challenges for Multilingual Europe
Brussels, Belgium, November 17/18, 2010



Co-funded by the 7th Framework Programme of the European Commission through the contract T4ME, grant agreement no.: 249119.

1. Cultivating an environment for effective research
2. Main legal issues
3. Features of a Charter
4. Standardised Licences
5. Flows of rights/ value/ content
6. The role of an MoU
7. Technology
8. Workflows and their importance
9. Solutions suggested for META-SHARE and the iterative process of their creation

1. Cultivating a good research environment

In an ideal world ...

- * LRs are freely **traceable, accessible, re-usable** and **derivatives** are allowed.
- * Sets of data where **personality rights, moral rights and Personally Identifiable Information** are easily infringed are **cleared** so that researchers do not have to worry about this.
- * Monetary matters are either not an issue (**no charges**), pre-cleared or subject to a **standardised remuneration system**.

2. Main legal issues [I]

-
- A. The existing legal system is not designed for the use of language resources
 - B. Different basic legal terms as well as different exceptions and limitations to protective rights apply across different national legal systems.
 - C. Issues beyond Copyright: obtaining consent from the Data Subject or data processing in accordance to Data Protection Principles

2. Main legal issues [II]

D. Access is often subject to non-standard legal terms (over time producing a lot of legal noise).

E. Licensing terms are not readily accessible or machine readable

F. Licensing does not always provide the level of security required by the users of a resource

2. Main legal issues [III]

G. Re-use is subject to even more non-standard legal terms

H. Re-use is often restricted to certain types of users, usage scenarios or usage objectives, which are not consistent across the same pool of resources.

I. In many cases the chain of rights is entirely invisible to the re-user (risk of unintentional infringement and liability – low possibilities of extensive institutional use or commercialisation)

J. Use of custom-made Share-Alike/ Copyleft clauses is likely to cause islands of commons that cannot be mixed with each other

3. A Charter [I]

A. An important step toward the setting up of getting the ground rules for specific communities right

B. An increasingly growing trend:

- * Europeana

- * COMMUNIA

- * Panton Principles (Science Commons)

C. A fundamental step in the direction of solving some of the key environment problems for META-SHARE

3. A Charter [II]

What should the META-SHARE charter cover:

1. **Access** to language resources must be as **open** as possible and should allow **transformative uses** and **super distribution**.
2. Content entering the network must be properly documented, carry sufficient metadata as well as a **rights clearance**, commercial rights to be included or easily obtainable.
3. **Attribution and copyleft** must be governed by **consistent rules** and the data itself should be free of any licensing or other conditions.
4. **Internal sharing** must be as **restriction-free as possible** and each LR must have a network member or repository as **curator**. For the purposes of **external dealings, standard commercial deals** have to be made available.

3. A Charter [III]

What should the META-SHARE charter cover:

5. Draw up the **framework for sharing**, f.e. dissemination via a separate MoU for each resource.
6. State the **bottom line of the MoU**: Granting a permanent license to use, re-use, disseminate and sublicense the LR.
7. No **third party rights** granted, licensee liable for obtaining such.
8. Restrictions on **commercial use** (if applicable) and requirement to keep intact **notices**.
9. **Additional clauses** concerning metadata, indemnification, governing law, completeness etc.

4. Standardised Licences [I]

Why are Standardised Licences important:

- * thorough check and sustainability
- * longevity
- * interoperability

Why Creative Commons:

- * large user base
- * network of legal experts
- * three layers of expression

▫ [1]

▫ Human-Readable: Commons
Deed

▫ [2]

▫ Lawyer-Readable:

▫ Legal Code

▫ [three expressions]



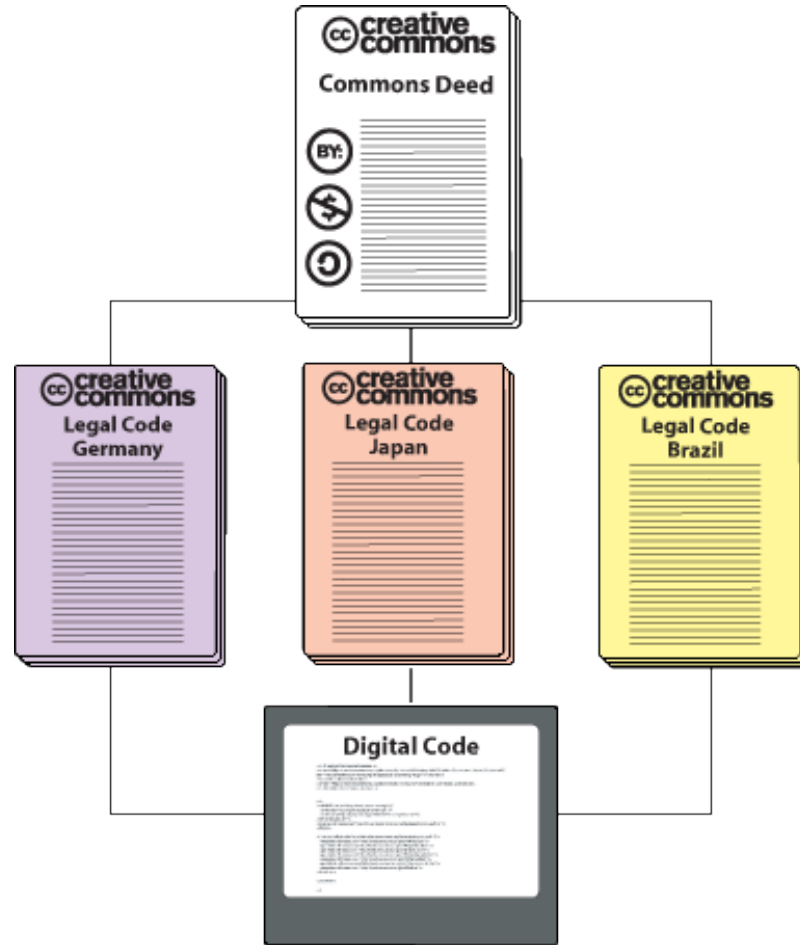
▫ Logo + Link

▫ [3]

▫ Machine-Readable:

▫ Meta Data

□ schematic



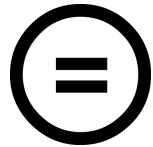
▫ CC logos



▫ Attribution



▫ Non Commercial



▫ No Derivatives



▫ ShareAlike

▫ CC Combinations



▫ Attribution



▫ Attribution - Share Alike



▫ Attribution - No Derivatives



▫ Attribution - Non Commercial



▫ Attribution - Non Commercial -
ShareAlike



▫ Attribution - Non Commercial -
▫ No Derivatives

□ Is Commercial Use Allowed?

□ Yes

□ No

□ Can the User Make Changes?

□ Yes

□ No

□ Yes

□ No

□ CC_BY_ND

□ CC_BY_NC_ND

□ Must the Derivative Work Be Distributed Under the Same Terms and Conditions

□ Yes

□ No

□ Yes

□ No

□ CC_BY_SA □ CC_BY

□ CC_BY_SA_NC □ CC_BY_NC

□ Open

□ Closed



5. Flows

* Rights/ Value/ Content do not always follow the same path

* Models:

(a) Clean Hands

(b) Star Shaped,

(c) Hybrid



▫ Flows of rights



▫ Flows of content

[flow models]

Star-shaped
model



[A]

Example:

Birmingham Museum and Art Galleries

Welcome to BMAGiC - Birmingham Museums & Art Gallery Information Centre - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.bmagic.org.uk/

Most Visited Creative Commons ((i)) Indymedia :: Ath... Comments on Educati... A Swarm of Angels :: Commons.org Articl... Blue-White GT RSS feed from the LS...

BMAGiC Birmingham Museums & Art Gallery

About the Museums · For Visitors · What's On · Learning · Contact · Help

BMAGiC Home Search Browse BMAGiC Help About BMAGiC Contact Us

Welcome to BMAGiC


BMAGiC is Birmingham Museums & Art Gallery's online database. It gives you information about the objects in our collection. Every month we add more images of objects for you to look at. For further information please see [About BMAGiC](#).


Designated as an Outstanding Collection

My Bookmarks

You currently have no objects bookmarked. You can bookmark objects to help you find them again easily when using this computer.






Featured Objects:


[19th Century Birmingham Halfpenny Token - School of Art](#)


[Covered Button Box](#)

Browse the Collections

You can use the links below to see all objects in a particular collection. You can also [browse the list of all collections](#).

-  [Antiquities](#)
[Ancient & Mediaeval Coins](#)
[Ancient Cyprus](#) more...
-  [Applied Art](#)
[Asian](#) [Costume](#) more...
-  [Coins & Medals](#)
[Hull-Grundy Collection](#)
-  [Ethnography](#)
[Africa](#) [Asia](#) more...
-  [Fine Art](#)

Search the Collections

Quick Search

Enter a keyword or words to search all the objects in the collections:

For more control over your search you can use the [Advanced Search](#)

Top 5 Searches

Click on any of the following to see objects that other people are searching for:

- [Pre-Raphaelite](#)
- [Chinese](#)
- [Burne-Jones Stained Glass](#)
- [Ford Madox Brown](#)
- [Japanese Print](#)

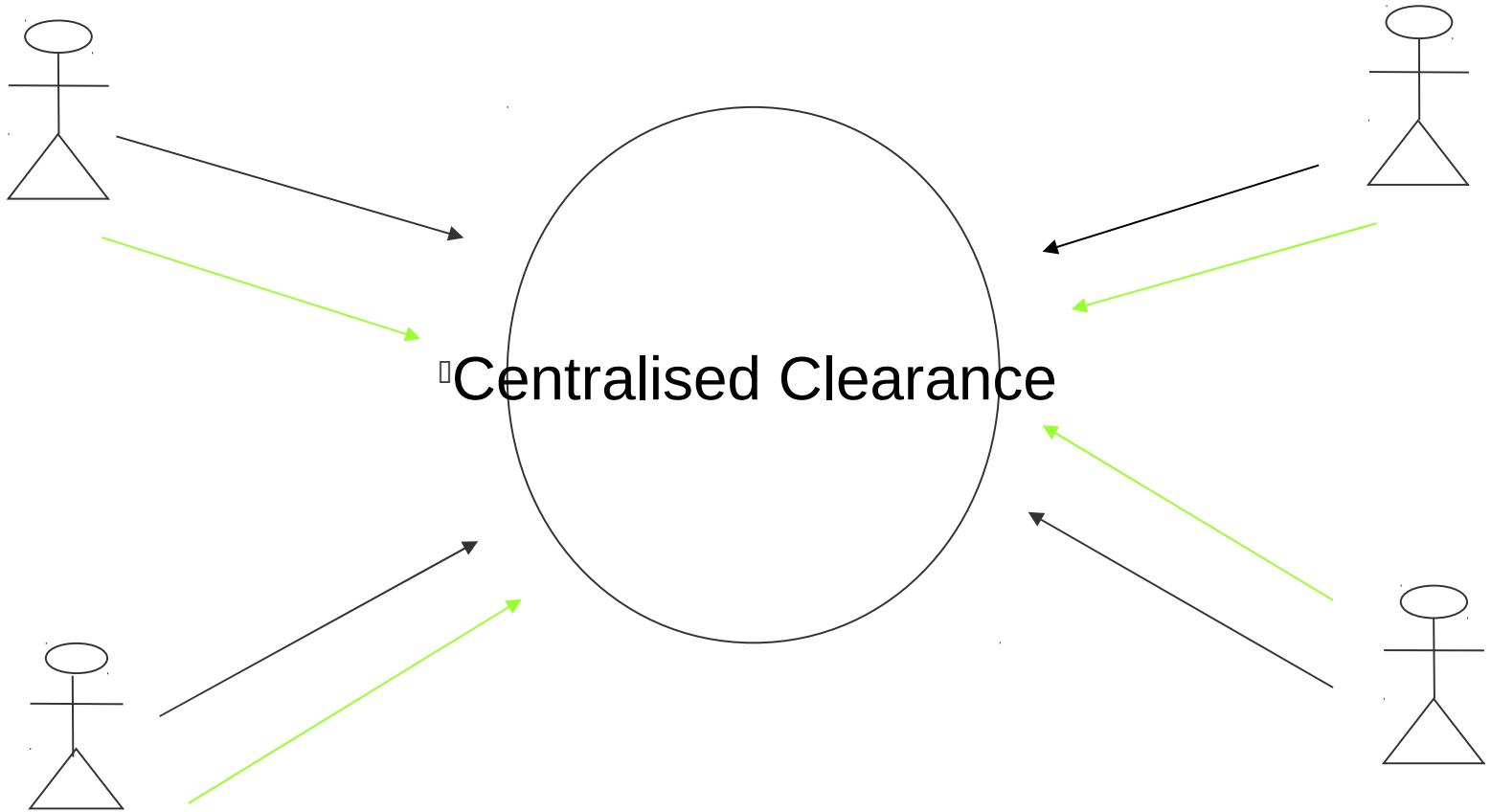
Done

start veech zotero Now: Partly Sunny, 1° C Mon: 5° C Tue: 6° C

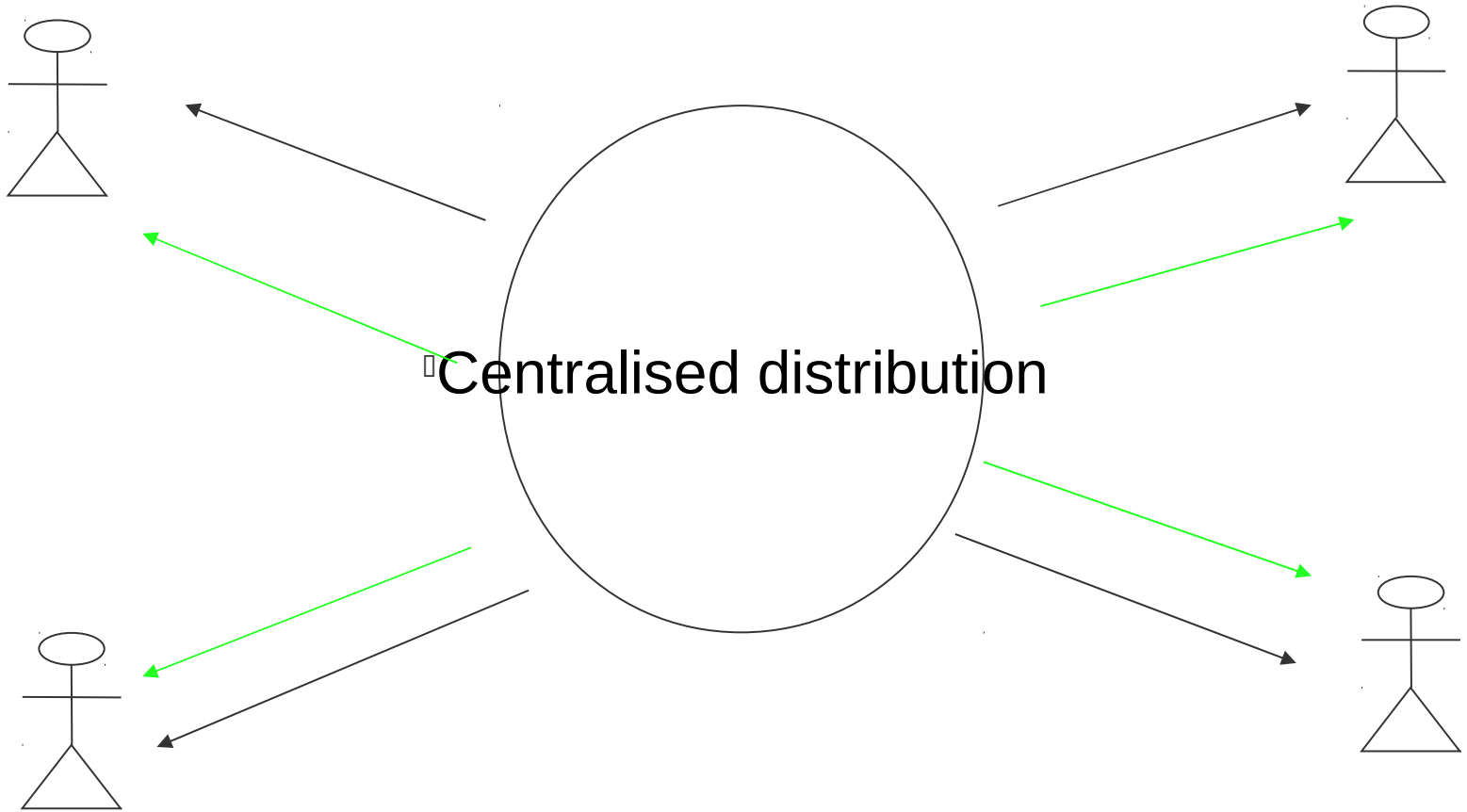
Microsoft Office P... Inbox - Outlook Expr... MELODIA 99,2 - Mozil... Welcome to BMAGiC ... peer_review_agend...

EN 11:08

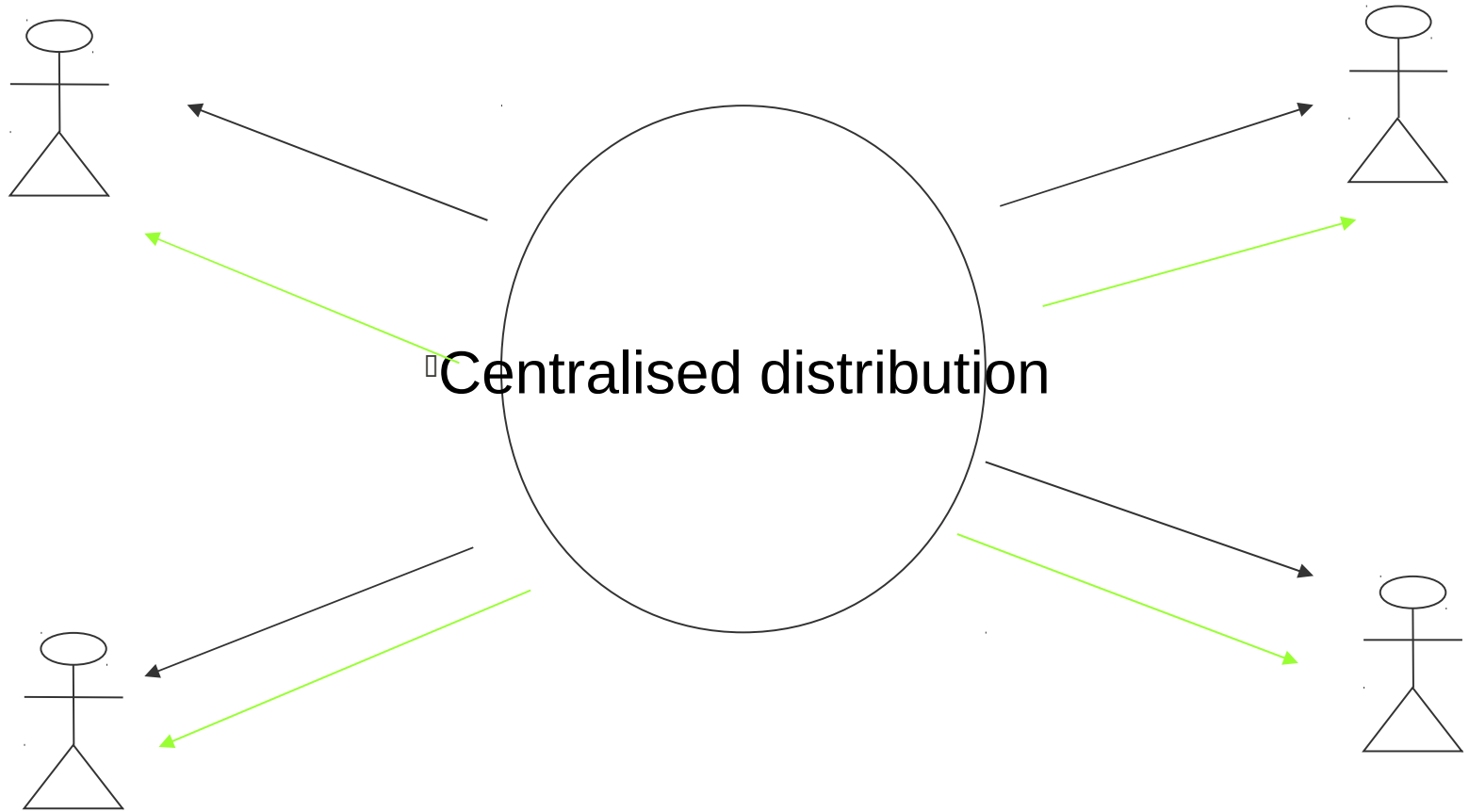
▫ Inputs



▫ Outputs



▣ Outputs



British Library Archival Sound Recordings (ASR I and II)

The screenshot shows the website <http://sounds.bl.uk/> in a Mozilla Firefox browser. The page features a navigation menu with links for Home, About this site, Project, Audio tools, Contact us, Blog, and Help. The main header includes the text "ARCHIVAL SOUND RECORDINGS" and a sub-header "Explore 12,000 selected recordings of music, spoken word, and human and natural environments". A call to action asks users to log in via their home organisation (UK higher and further education only). The page is associated with JISC. A search bar is present with a "Limit to: All collections" dropdown and a "GO" button. A "BROWSE all collections" sidebar lists various categories such as Accents and dialects, African Writers' Club, Amphibians, Art and design interviews, Beethoven String Quartets, British wildlife recordings, David Rycroft Africa recordings, Ethnographic wax cylinders, and Klaus Wachsmann Uganda recordings. The "HIGHLIGHTS" section features three featured recordings: Elisabeth Frink talks about her background and early life, A train passing along the Dartmoor railway, and Princess Magogo sings the folktale song "Uhamba namanzini", 1964. The browser's taskbar at the bottom shows several open applications and the system tray with the time 11:14.

Archival Sound Recordings - Mozilla Firefox

File Edit View History Bookmarks Tools Help

<http://sounds.bl.uk/>

Most Visited Creative Commons (()) Indymedia :: Ath... Comments on: Educati... A Swarm of Angels .: iCommons.org Arbid... Blue-White GT RSS feed from the LS...

Home | About this site | Project | Audio tools | Contact us | Blog | Help

BRITISH LIBRARY

ARCHIVAL SOUND RECORDINGS

Explore 12,000 selected recordings of music, spoken word, and human and natural environments

Please **log in** via your home organisation (UK higher and further education only)



In association with **JISC**

Anyone can search this site and play over 1,500 of the recordings. These include the wax cylinder, UK dialect and wildlife collections. For copyright reasons, the other recordings can only be played in licensed UK higher and further education institutions or in our reading rooms. Downloading is available in **licensed institutions**. You can explore our full audio holdings in the **Sound Archive catalogue**.

SEARCH Limit to: All collections

[Search tips](#)

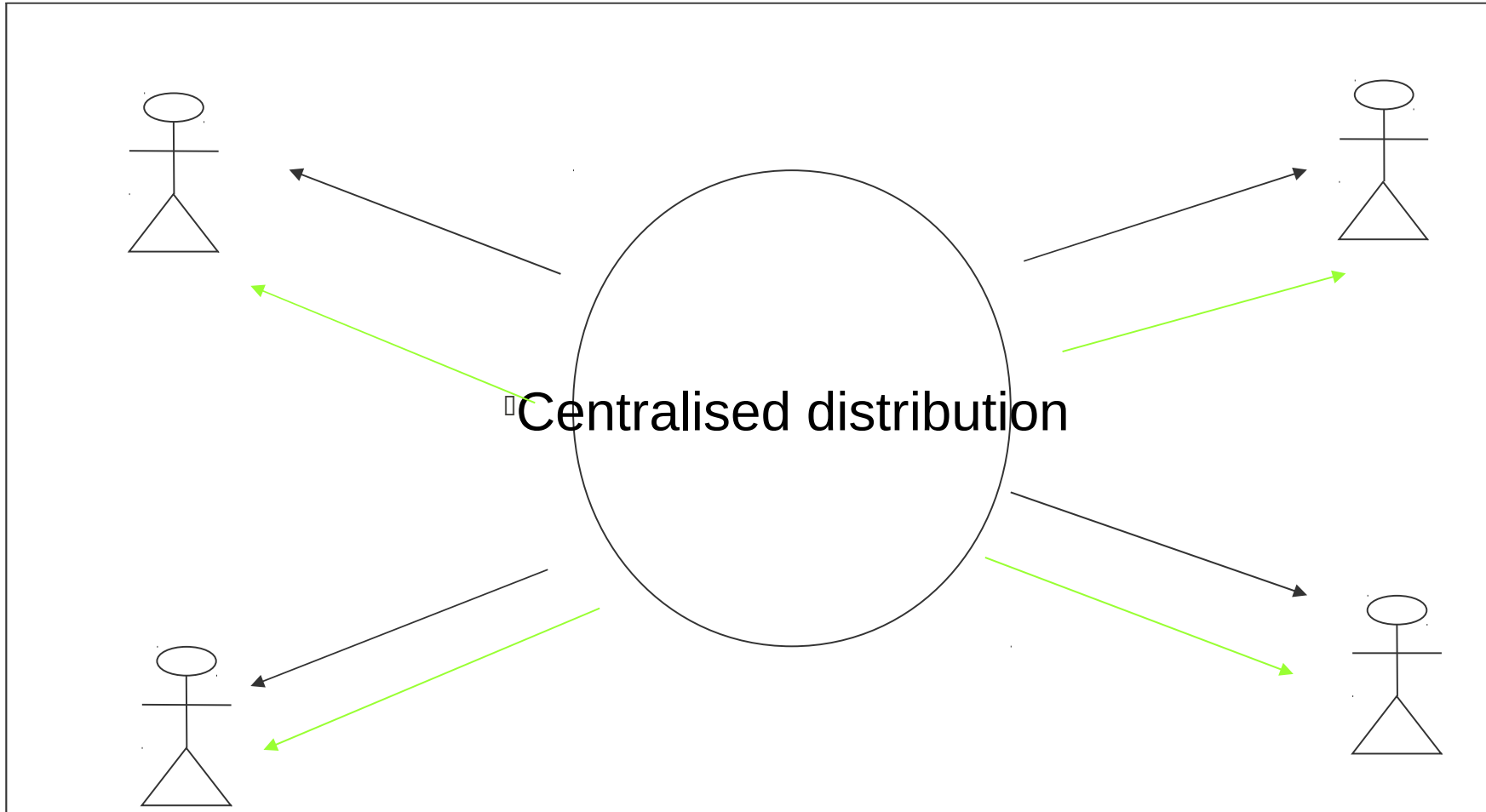
HIGHLIGHTS

-  Elisabeth Frink talks about her background and early life. From the "Art and design interviews"
-  A train passing along the Dartmoor railway. From the "Soundscapes" collection.
-  Princess Magogo sings the folktale song "Uhamba namanzini", 1964. From the "David Rycroft Africa"

Done

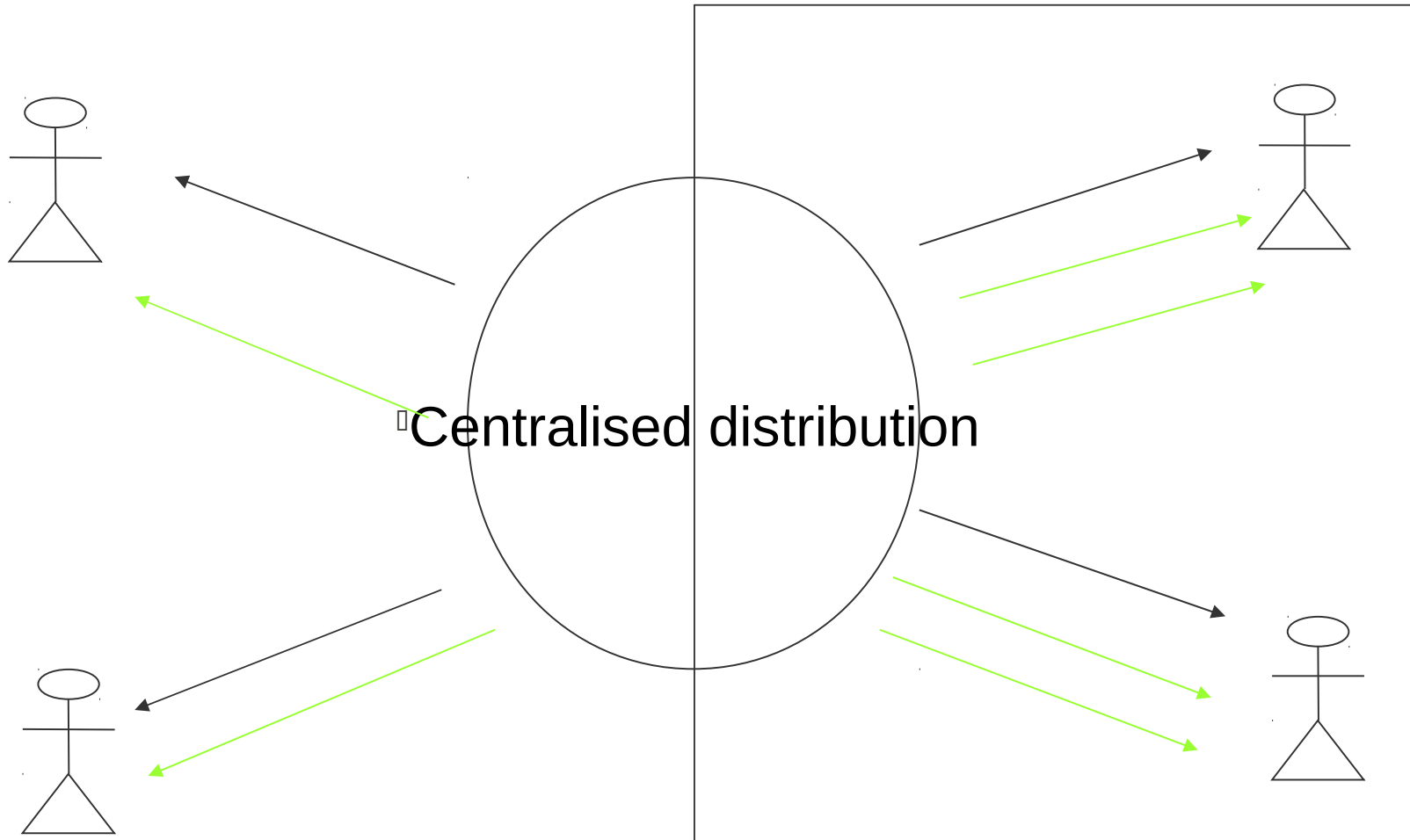
start 4 Microsoft Offi... Inbox - Outlook Ex... AGENDA: SCA Inte... MELODIA 99,2 - M... Archival Sound Re... peer_review_age... EN 11:14

▫ Outputs



▫ Controlled distribution environment

▫ Outputs



▫ Extra rights flowing within the controlled environment

□ Clean
Hands



□ [B]

National Centre for eResearch (MyExperiment/ OMII)

The screenshot shows the myExperiment website in a Mozilla Firefox browser. The address bar displays <http://www.myexperiment.org/>. The page features the myExperiment logo and a tagline: "myExperiment makes it really easy to find, use and share scientific workflows and other files, and to build communities." Below this is a search bar with a dropdown menu set to "All" and a "Search" button. On the left, a sidebar titled "Use myExperiment to..." lists various actions: Find Workflows, Find Files, Share Your Workflows and Files, Create and Find Packs of Items, Create and Join Groups, Find People and Make Friends, Send Messages, Get Feedback, Tag and Rate things, Write Reviews and Comments, and Build your Profile and Reputation. The main content area is divided into three sections: "Explore" with a workflow diagram and a "Find Workflows" button, "Register" with a login form (Username or Email, Password, Remember me checkbox), and "About myExperiment" with links to the mailing list, myGrid Project, Taverna Workflow Workbench, and myExperiment Publications.

The screenshot shows the OMII-UK website in a Mozilla Firefox browser. The address bar displays <http://www.omii.ac.uk/>. The page features the OMII-UK logo and tagline: "Software Solutions for e-Research". Below this is a search bar with a "Search" button and links for "Login" and "Registration". On the left, a navigation menu includes: Home, About Us, Software, Community, Resources, Support, New to OMII-UK, Software News, and Community Wiki. The main content area is titled "Welcome to OMII-UK" and features a large image of a lightbulb with the text "Software Solutions make e-Research easy." Below this, there are sections for "New to OMII-UK" (providing information for new users), "Software News" (with a link to "CALL FOR PAPERS - 5th International Conference on..."), "Events" (with a link to "2008-11-04: National Grid Service Innovation Forum 2008: 4-5 November 2008 - MoSI, Manchester"), and "News Feed" (with a link to "Digital Economy Research Hubs").

National Library for Health eLearning Object Repository (LOR)

The screenshot shows the National Library for Health website in a Mozilla Firefox browser window. The browser's address bar displays the URL <http://www.library.nhs.uk/>. The website's header includes navigation links for [About NLH](#), [For Librarians](#), [Tools](#), and [Help](#). The main navigation bar features [Home](#), [News & RSS](#), [My Library](#), and [Clinical Knowledge Summaries](#).

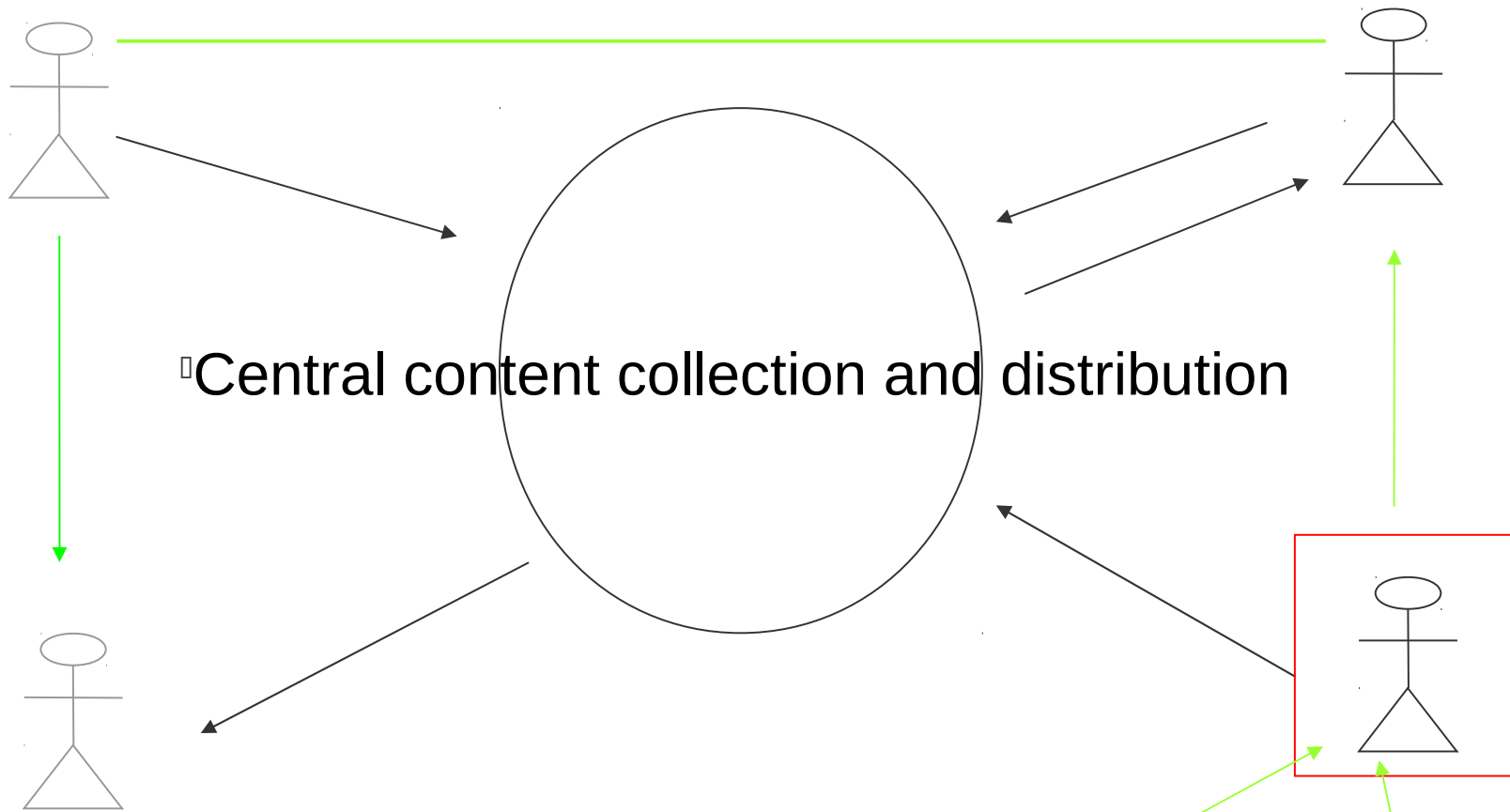
The page is divided into several sections:

- LOG IN:** A section for user authentication, including a text box for the username, a password field, and a **LOGIN** button. It also provides instructions on using an Athens ID and a link to register for Athens.
- SEARCH RESOURCES:** A search interface with a search box and a **Search** button. It offers two search options: Search title only and Search title and text. Below the search box, there are checkboxes to limit the search to: All resources, Evidence Based Reviews, Guidance, Specialist Libraries, Books and Journals, Images, For Patients, and Medical Dictionaries.
- RESOURCES:** A section listing various resources, including **Evidence Based Reviews** (Bandolier, Cochrane Library, DARE, HTA Database, NHS EED), **Guidance** (CKS, National Library of Guidelines, NICE Guidance, etc.), **Specialist Libraries** (collections for different communities of practice), and **Books, Journals and Healthcare Databases** (AMED, British Nursing Index, CINAHL, etc.).
- EVIDENCE UPDATE:** A section with an RSS icon, featuring a **Health Management Specialist Library - 31/10/08** update. The update text includes: "NHS foundation trusts: review and consolidated accounts 2007-08", "Developing the Quality and Outcomes Framework: proposals for a new, independent process", and "Real involvement: working with people to improve health services".
- NLH HEALTH NEWS:** A section for news, featuring a **Health News from NHS Choices - 03/11/2008** update at 09:14:31, with a link to "Exercise and breast cancer".

A red circular stamp on the right side of the page reads "NLH 2.0 SEARCH INSIDE" and includes the website URL www.library.nhs.uk/. Below the stamp is a link to [Learn more about NLH 2.0 >](#).

The Windows taskbar at the bottom shows the Start button, several open applications (Microsoft Office, Outlook, AGENDA, MELODIA, National Library for Health, peer_review_age...), and system tray information including the date and time (11:11).

Inputs/ outputs: Clean Hands Model



Central content collection and distribution

* Flow of licences directly between the users

Local clearance point

Hybrid Model



[C]

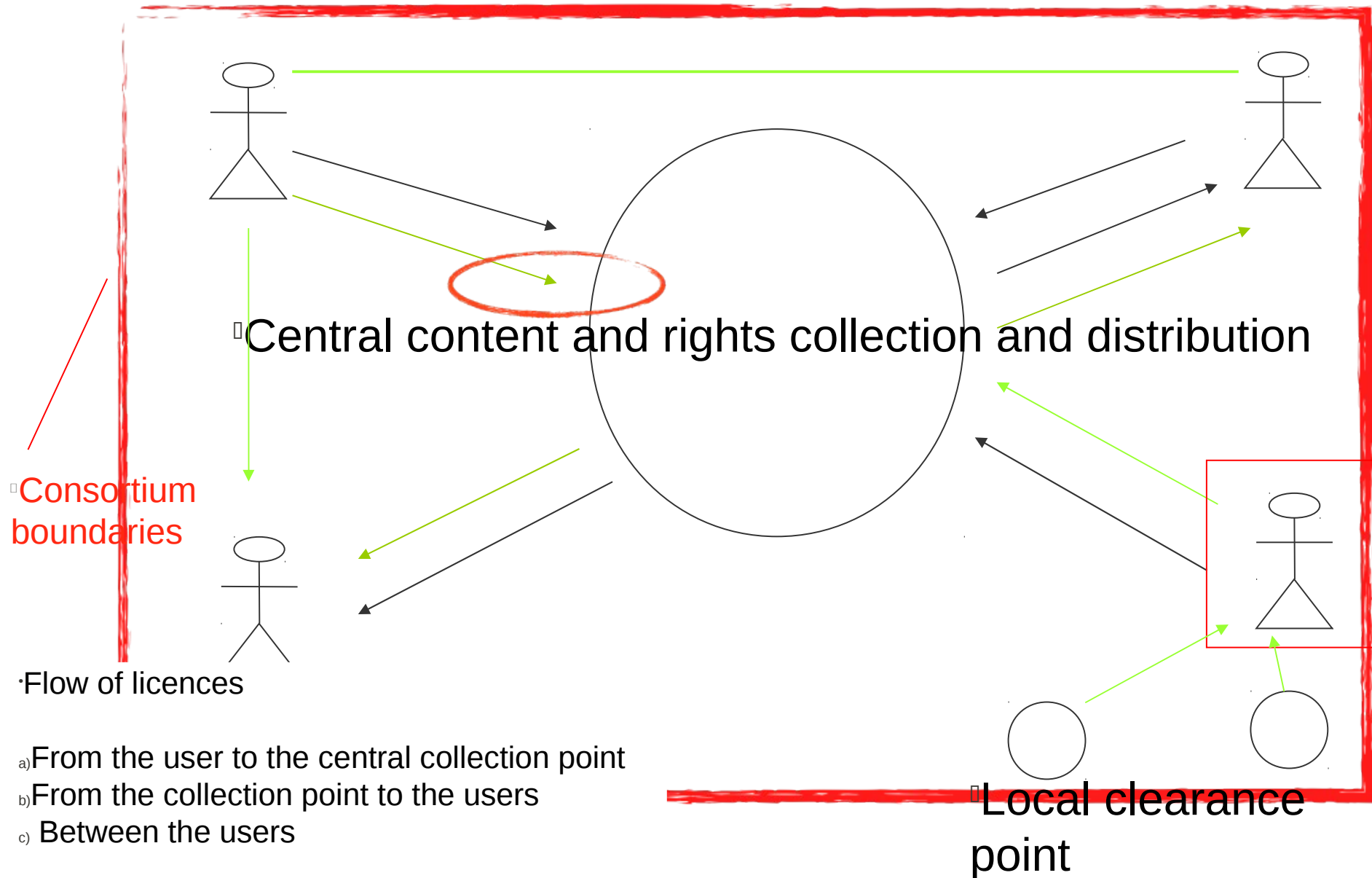


National Education Network (Gallery + Repurpose projects)

The screenshot shows the National Education Network (NEN) website in a Mozilla Firefox browser window. The browser's address bar displays the URL <http://www.nen.gov.uk/>. The website's navigation menu includes links for Home, About us, E-Safety, Learning & Teaching, Media Centre, and FAQ & Contacts, along with a search bar. The main content area features the NEN logo, a photograph of two young girls in school uniforms, and a descriptive text: "The NEN is the UK collaborative network for education, providing schools with a safe, secure and reliable learning environment and direct access to a growing range of online services and content." Below this is a circular diagram of network providers, including CLEOD, glow, YH GfL, embc, EBN, WMnet, and SEG/L. A "News" section on the right lists several articles, such as "Computers and Home Access for All" (22nd October 2008), "London - Beijing" (10th October 2008), and "Latest LMLE Case Study Videos" (24th September 2008). The browser's taskbar at the bottom shows various open applications and the system clock set to 11:10.

Inputs/ outputs: Hybrid Model

Decentralised rights distribution



6. The Role of an MoU

-
- * An internal document that sets the rules of the members of the group in greater detail
 - * The MoU is used when legal personality is not required
 - * it allows the standardisation of the custom made licences
 - * it makes possible the non-transfer of rights, yet the lowering of transaction costs

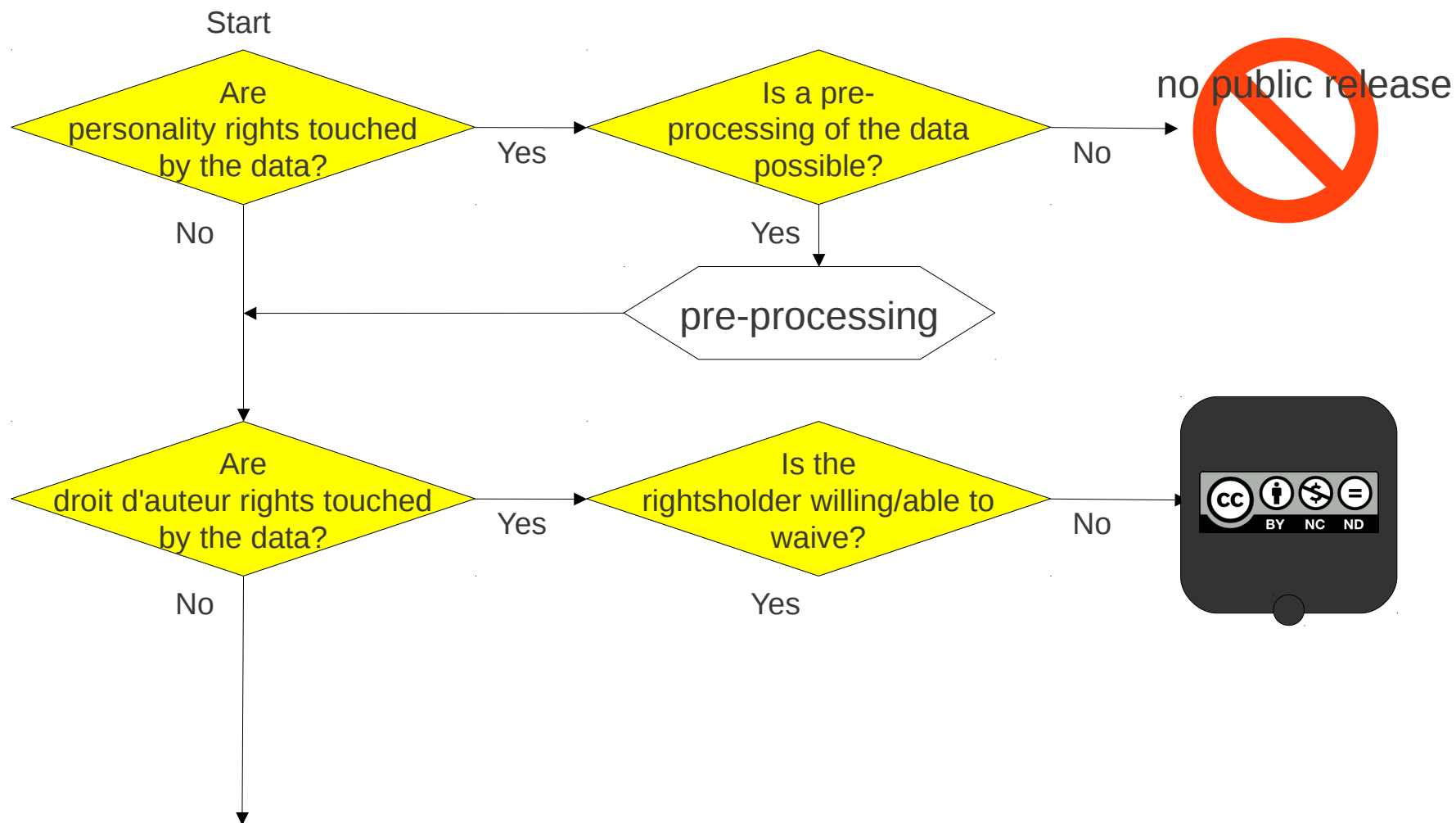
7. Technology

- * To complement the legal instruments and lower transaction costs (keep the honest honest)
- * Repositories
- * Semantic Representation of licensing information – links with OpenBiblio
- * Indexes and meta-data
- * Interoperability

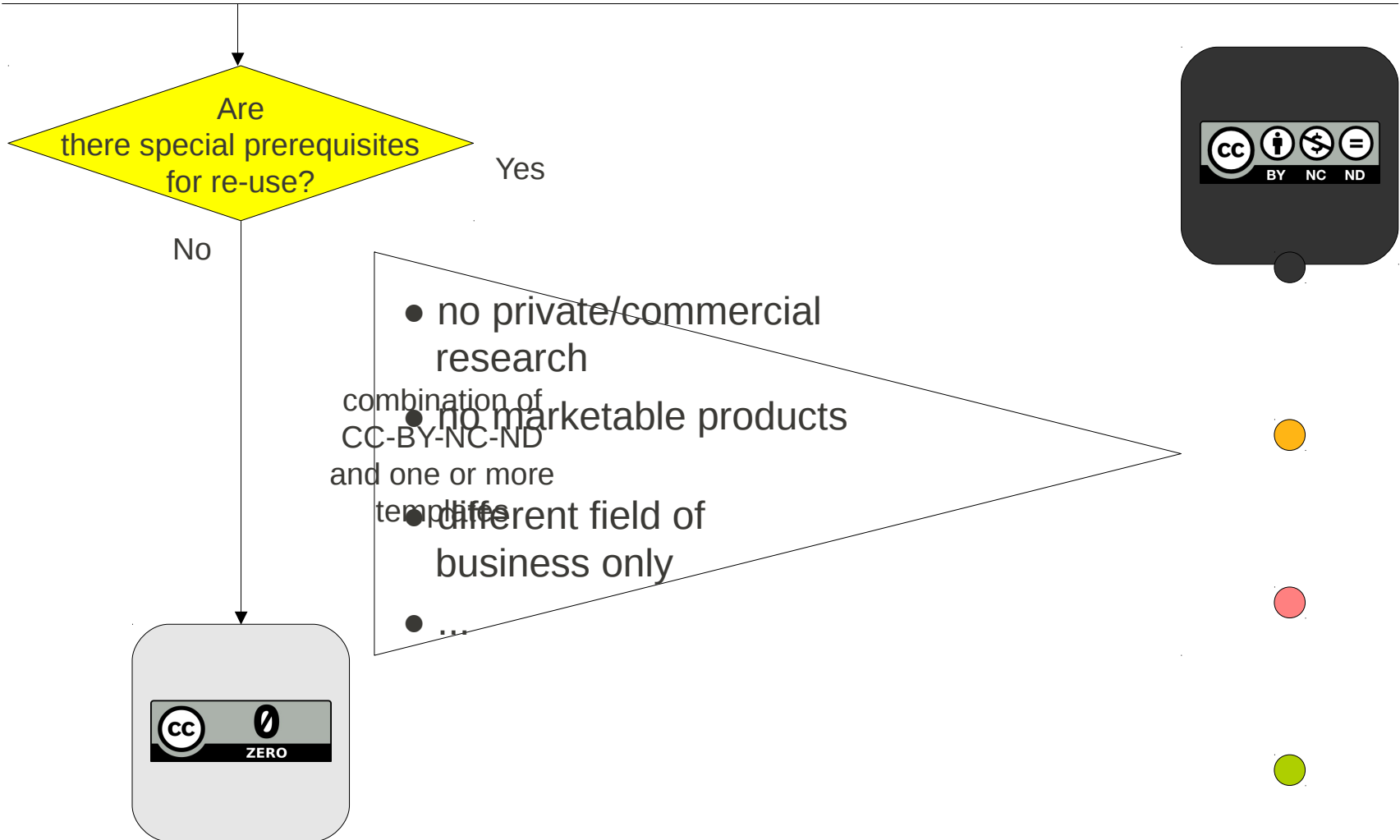
8. Workflows

-
- * Workflows ensure that all necessary rights enter the network
 - * $\text{RightsInput} \geq \text{RightsOutput}$
 - * Clearance and Consent at the points of concentration
 - * Adopt precedent from other areas (e.g. Open Educational Resources)

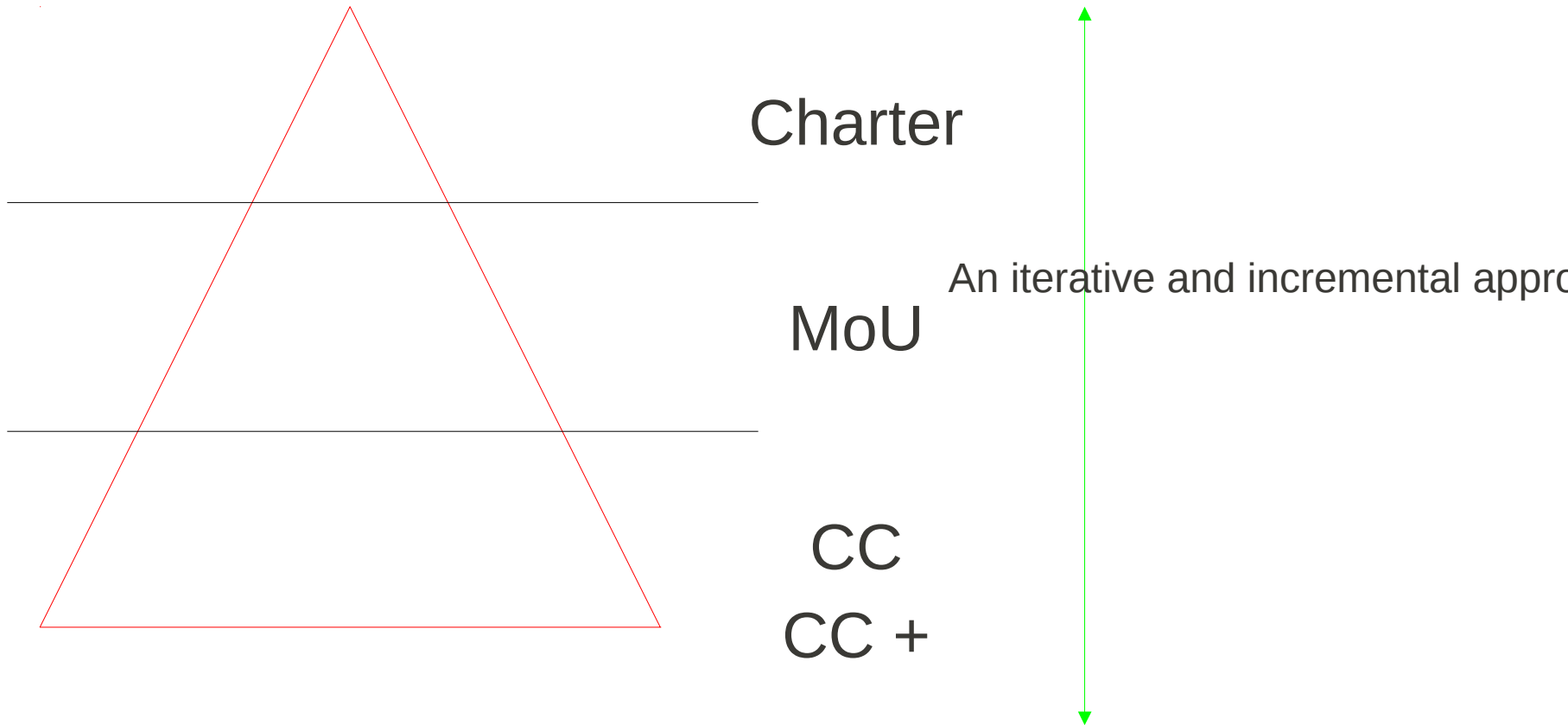
Public release workflow I



Public release workflow II



9. An iterative approach



Thank you very much.

office@meta-net.eu

<http://www.meta-net.eu>

<http://www.facebook.com/META.Alliance>