The Open Web Index
Crawling and Indexing the Web for Public Use

Gijs Hendriksen, Michael Dinzinger, Sheikh Mastura Farzana, Noor Afshan Fathima, Maik Fröbe, Sebastian Schmidt, Saber Zerhoudi, Michael Granitzer, Matthias Hagen, Djoerd Hiemstra, Martin Potthast, Benno Stein
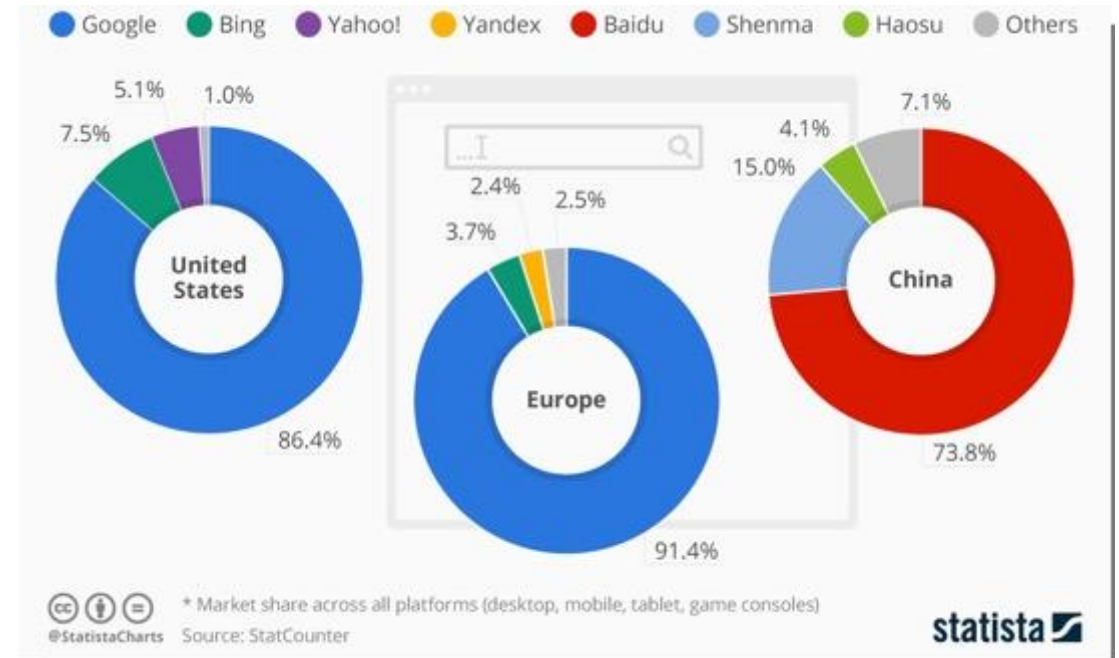
Open WebSearch

Open Web Search

Today's situation in Web Search

Two properties of Web Search that don't fit

- A critical infrastructure for society

- A market oligopoly
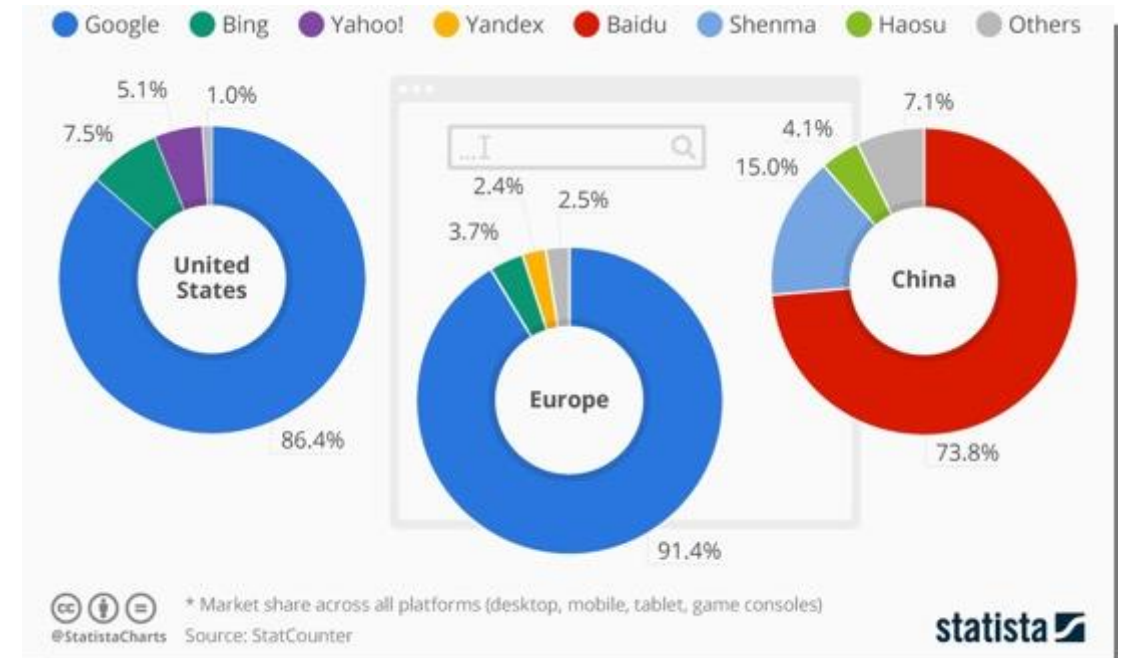
Today's situation in Web Search

Two properties of Web Search that don't fit

- A critical infrastructure for society

- A market oligopoly

Effects

- Reduced user choice

- Rich-get-richer
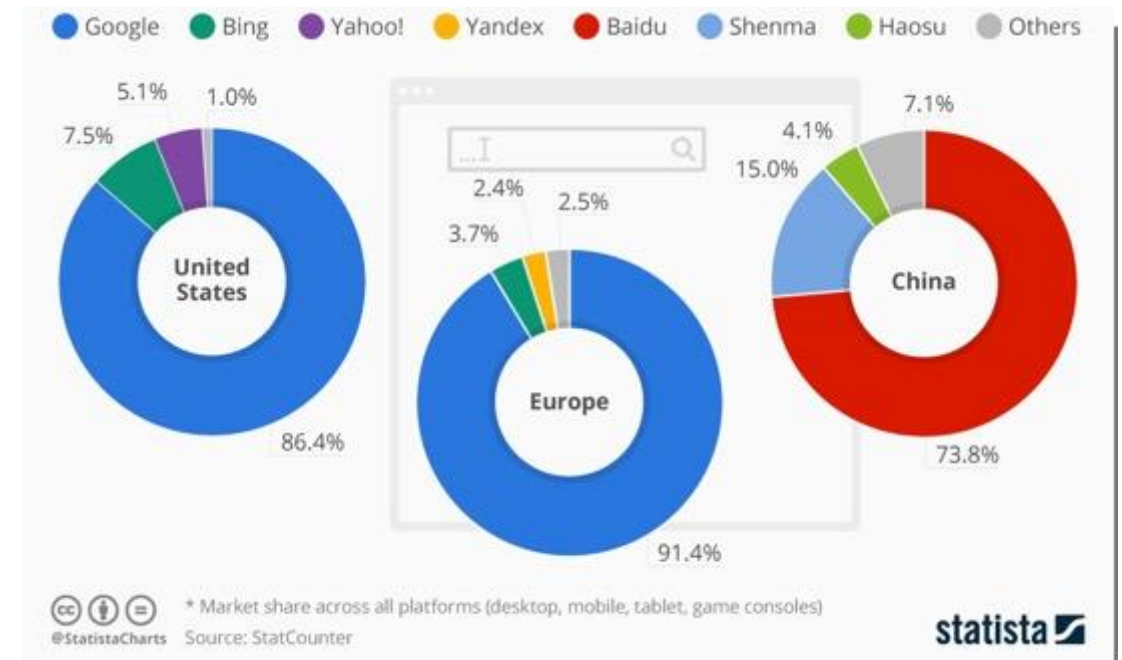
Today's situation in Web Search

Two properties of Web Search that don't fit

- A critical infrastructure for society

- A market oligopoly

Effects

- Reduced user choice

- Rich-get-richer
- Search Engine Manipulation Effect

- SEO optimized ranking vs. best information delivery

- Missing transparency in search

- User locked-in despite of "open" technologies

- Concerning market behaviour (e.g. Jedi Blue)

- Limited business models

OpenWebSearch.eu will create an open European infrastructure for internet search, based on European values and jurisdiction

What?

Restore an open search ecosystem / market as a basis for a new Internet Search

OpenWebSearch.eu will create an open European infrastructure for internet search, based on European values and jurisdiction

What?

Restore an open search ecosystem / market as a basis for a new Internet Search

Why?

1. Web search is dominated and limited by a few gatekeepers.

2. Tapping the Web as resource is challenging for innovators and researchers

OpenWebSearch.eu will create an open European infrastructure for internet search, based on European values and jurisdiction

What?

Restore an open search ecosystem / market as a basis for a new Internet Search

Why?

1. Web search is dominated and limited by a few gatekeepers.

2. Tapping the Web as resource is challenging for innovators and researchers

Who?

14 renowned European universities + institutions will pool their expertise and resources.
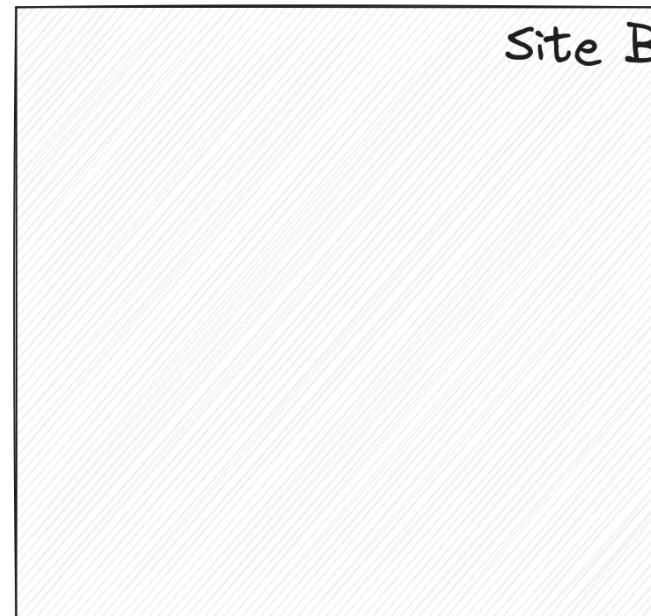
OpenWebSearch.eu will create an open European infrastructure for internet search, based on European values and jurisdiction

What?

Restore an open search ecosystem / market as a basis for a new Internet Search

Why?

1. Web search is dominated and limited by a few gatekeepers.

2. Tapping the Web as resource is challenging for innovators and researchers

Who?

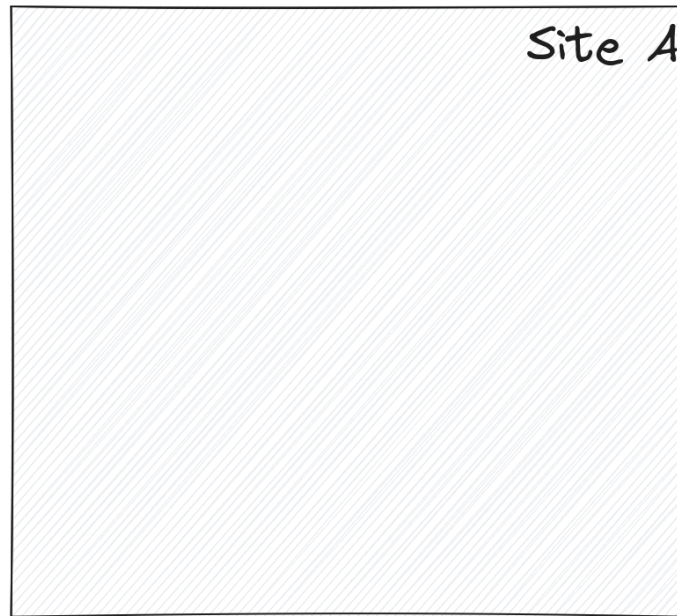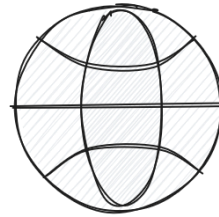14 renowned European universities + institutions will pool their expertise and resources.
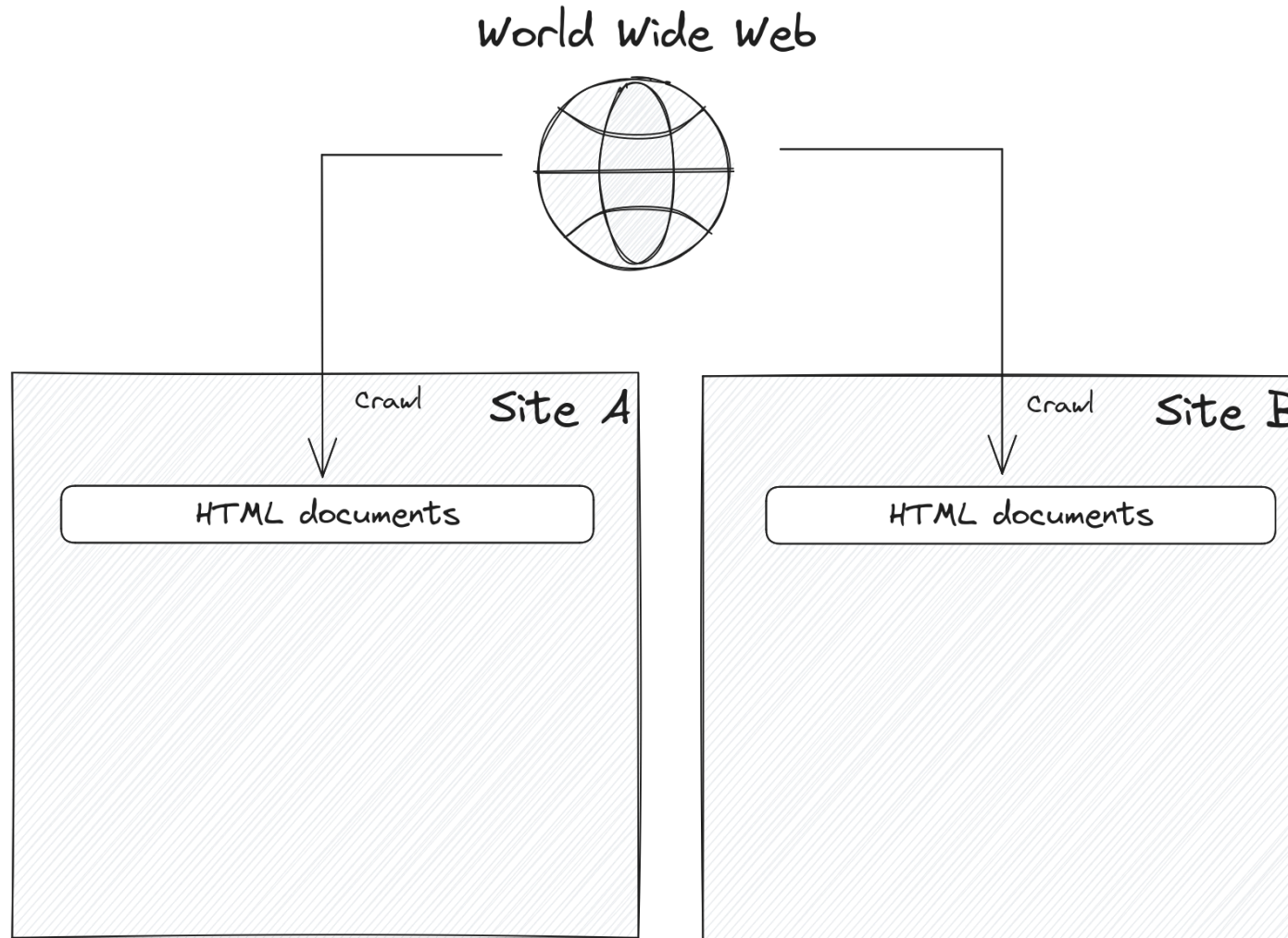
How?

Develop the core of a European Open Web Index

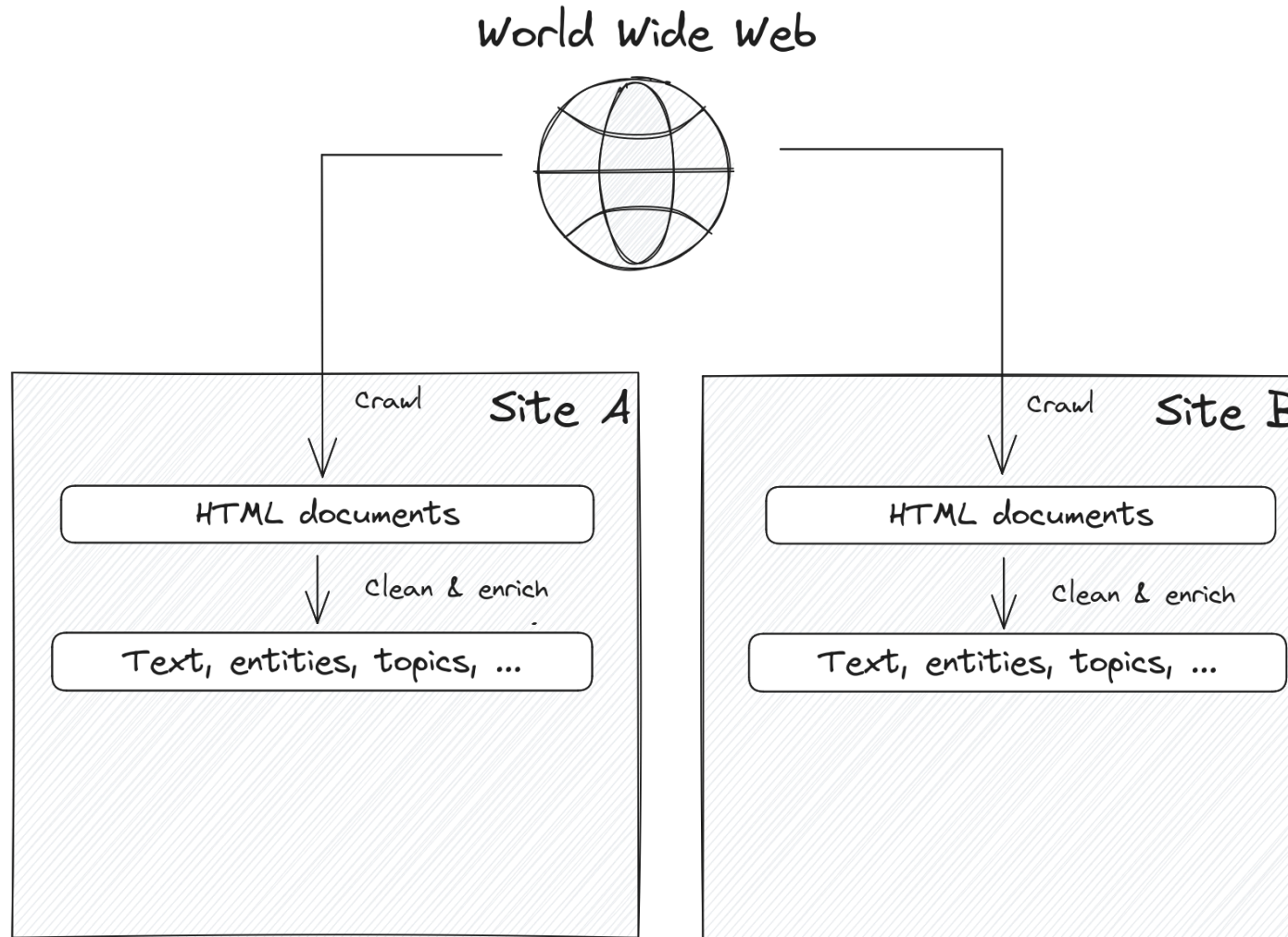# Infrastructure of the Open Web Index

World Wide Web

Site A

Site B

Infrastructure of the Open Web Index

Infrastructure of the Open Web Index

World Wide Web

Crawl | Site A

HTML documents

Clean & enrich

Text, entities, topics, ...

Crawl | Site B

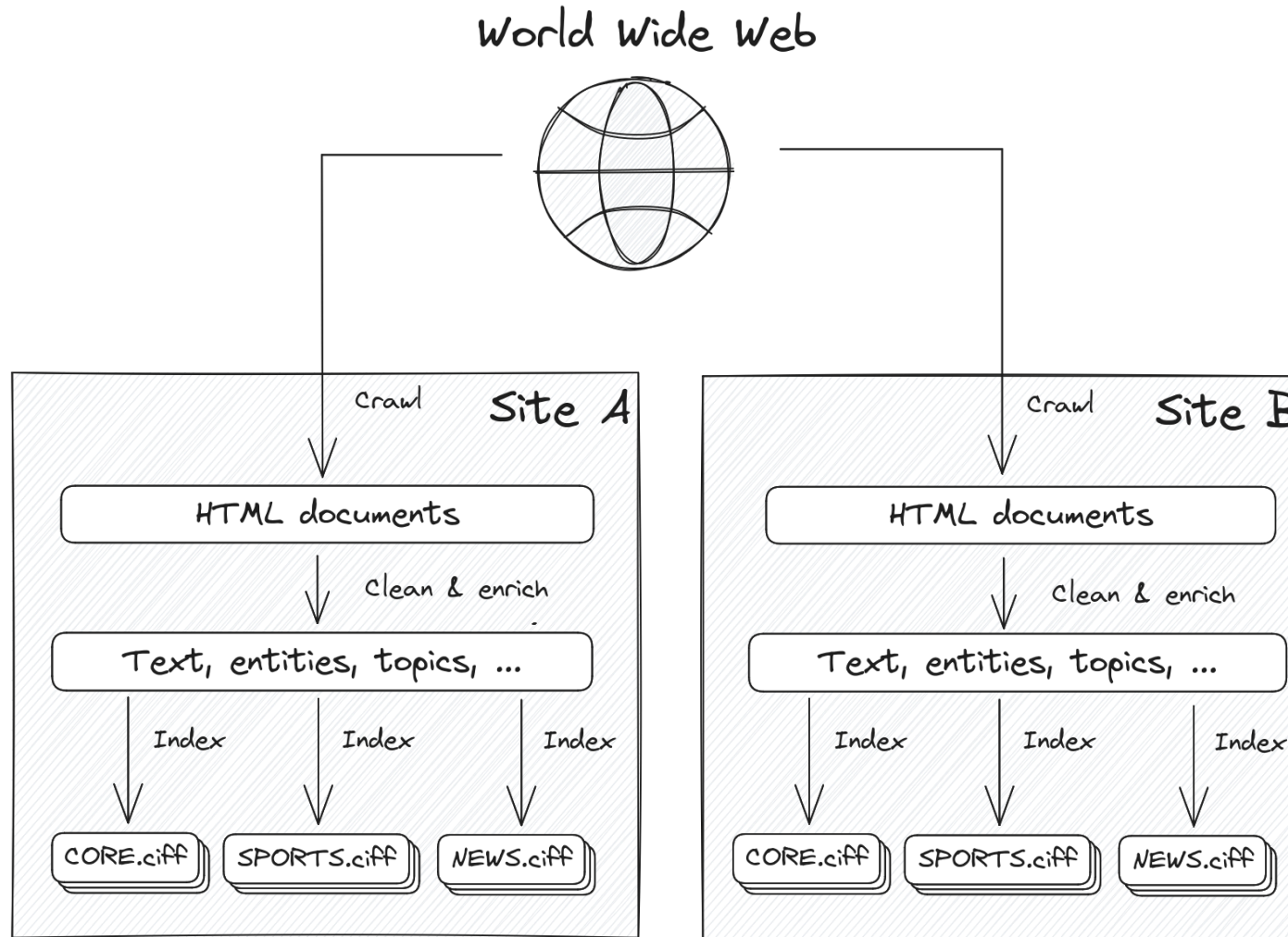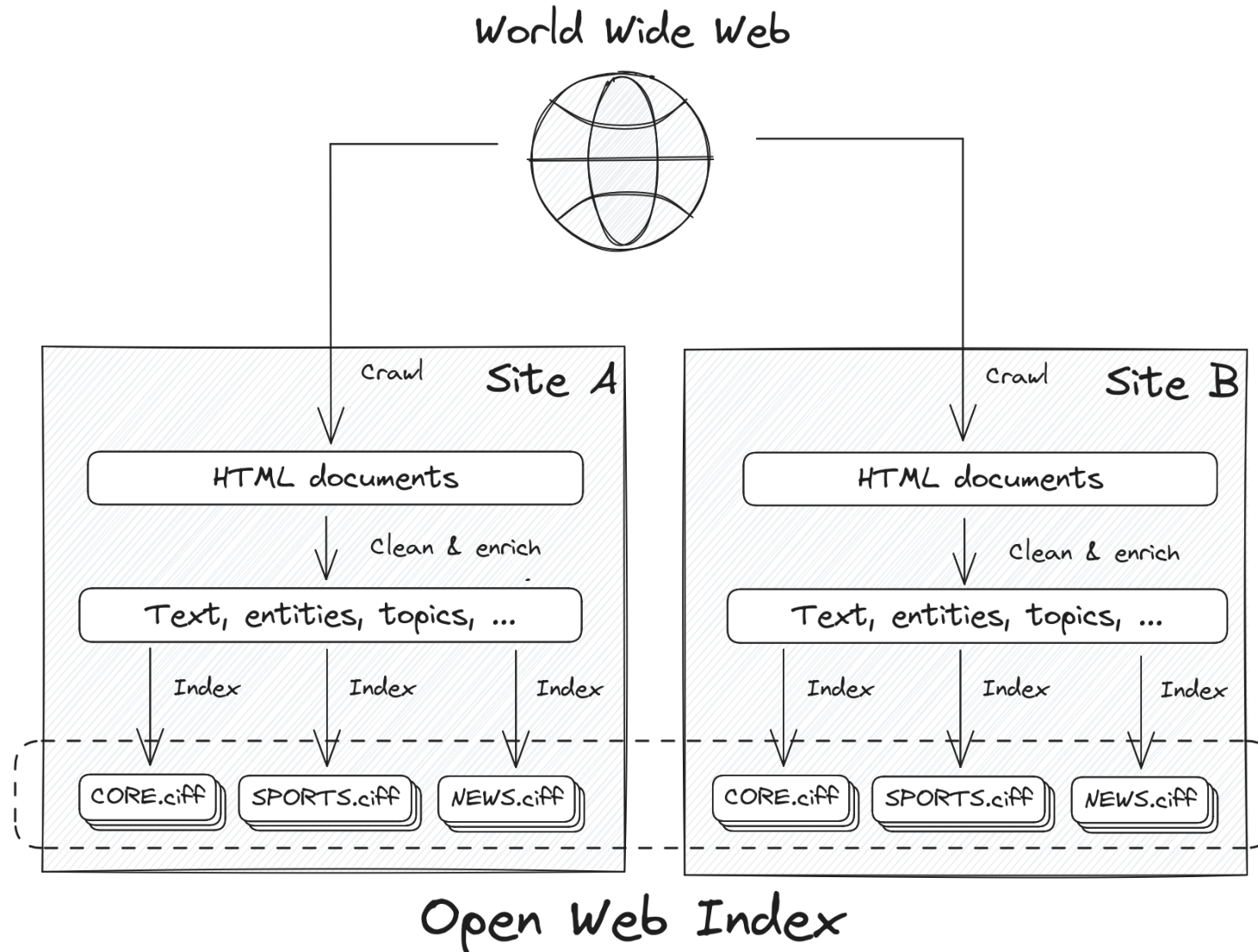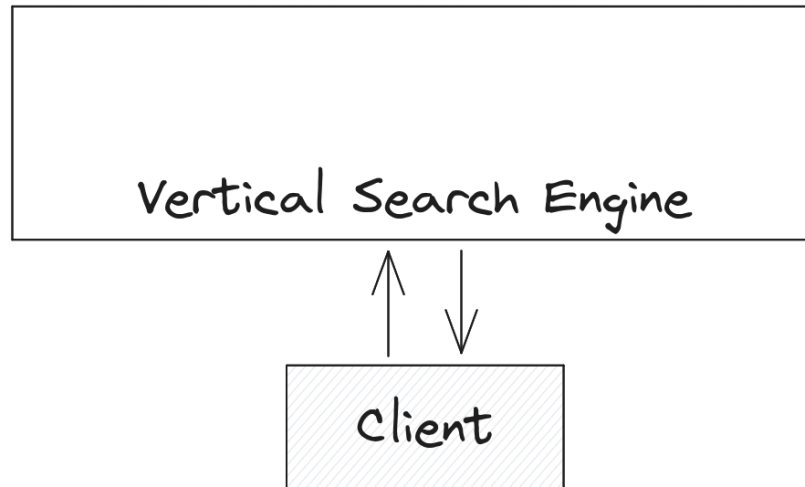HTML documents

Clean & enrich

Text, entities, topics, ...

Infrastructure of the Open Web Index
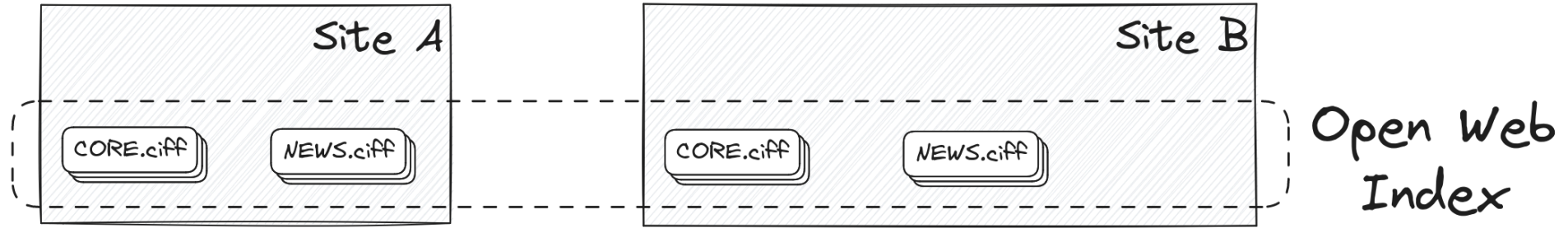
Infrastructure of the Open Web Index

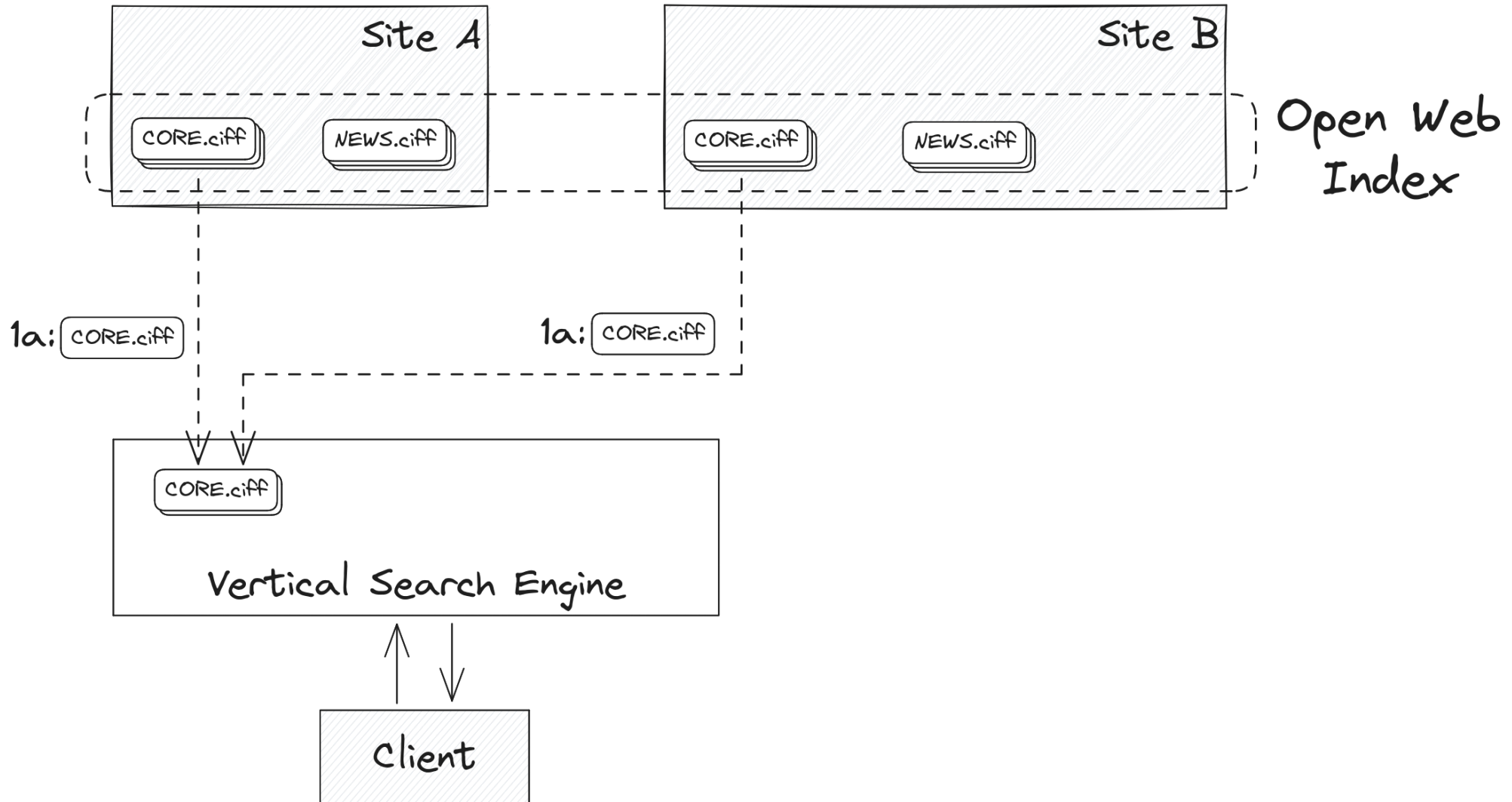World Wide Web

Crawl   **Site A**

HTML documents

Clean & enrich

Text, entities, topics, ...

Index   Index   Index

CORE.ciff   SPORTS.ciff   NEWS.ciff

Crawl   **Site B**

HTML documents

Clean & enrich

Text, entities, topics, ...

Index   Index   Index

CORE.ciff   SPORTS.ciff   NEWS.ciff

Open Web Index

Usage of the Open Web Index

Site A

CORE.ciff   NEWS.ciff

Site B
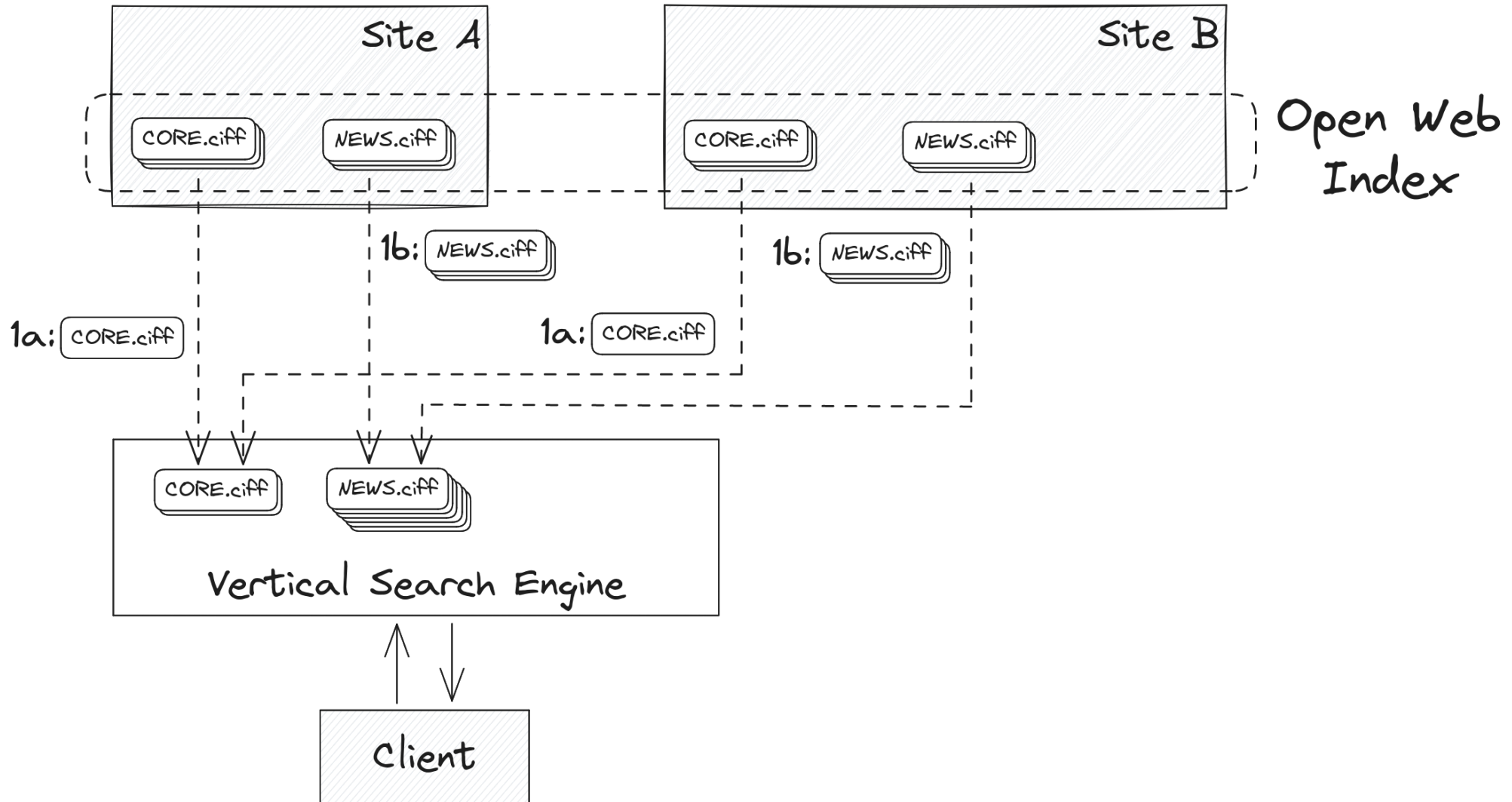
CORE.ciff   NEWS.ciff

Open Web Index
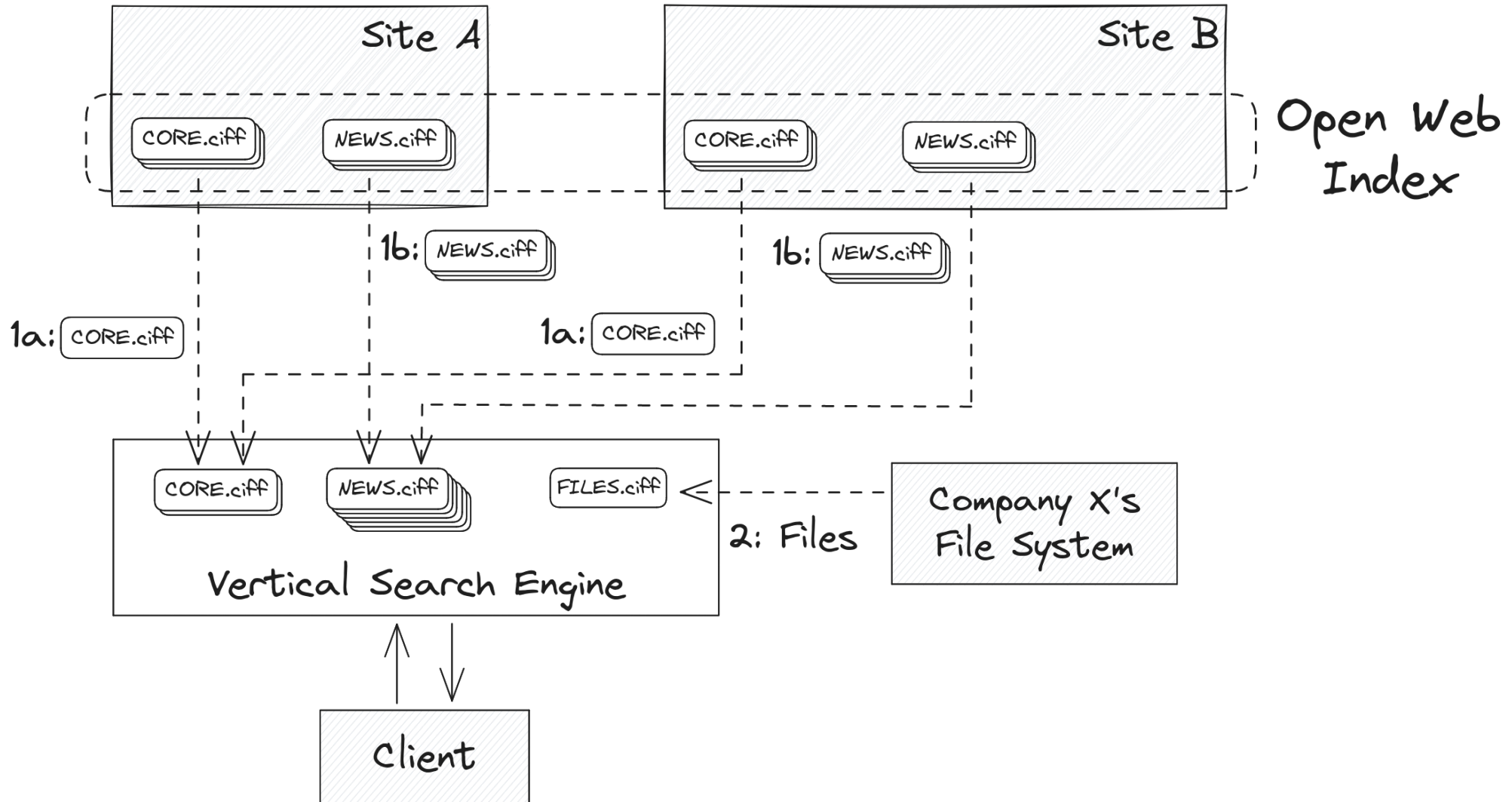
Vertical Search Engine

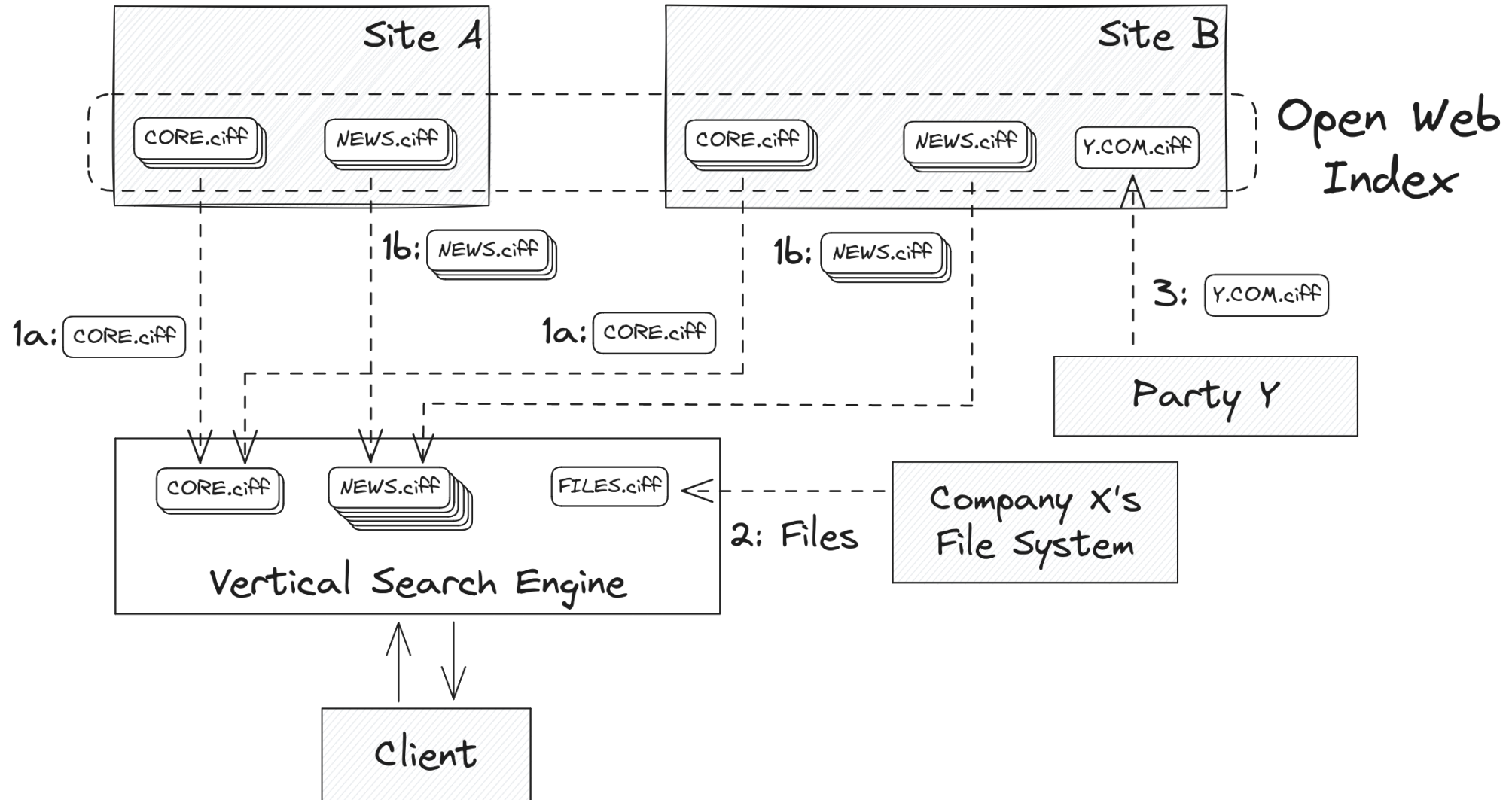Client

Usage of the Open Web Index

Usage of the Open Web Index

Usage of the Open Web Index

Usage of the Open Web Index

Contribute to Open Web Search!

Use the Open Web Index

- Search applications

- IR experiments & collections

Contribute to Open Web Search!

Use the Open Web Index

- Search applications

- IR experiments & collections

Join the community

- Third-party calls (deadline April 4th)

- Open Search Symposium #ossym2024

- OpenWebSearch.eu social media channels

Contribute to Open Web Search!

Use the Open Web Index

- Search applications

- IR experiments & collections

Join the community

- Third-party calls (deadline April 4th)

- Open Search Symposium #ossym2024

- OpenWebSearch.eu social media channels

Contribute

- Web Search components + evaluation

- Compute resources

Contribute to Open Web Search!

Use the Open Web Index

- Search applications

- IR experiments & collections

Join the community

- Third-party calls (deadline April 4th)

- Open Search Symposium #ossym2024

- OpenWebSearch.eu social media channels

Contribute

- Web Search components + evaluation

- Compute resources

Join us on Mattermost!

https://openwebsearch.eu/community/
ows-eu-community-on-mattermost/