

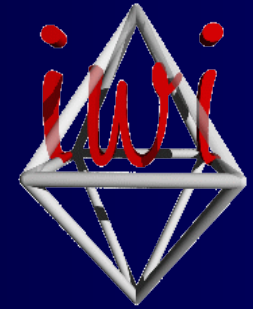
# MathNet, MPRESS

## and its search engines

**Judith Plümer**

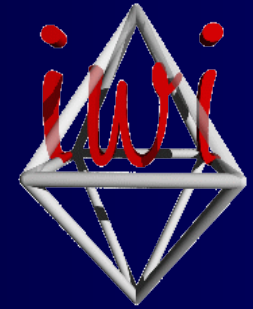
**Universität Osnabrück**

`Judith.Plumer@uos.de`



# Content

- **Math-Net and MPRESS**
  - Which problem is solved by Math-Net
  - Which methods are used
- **MetaData**
  - HTML META tag
  - RDF
- **Transition – methods and tools**



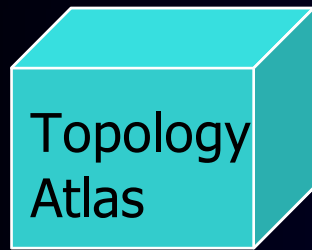
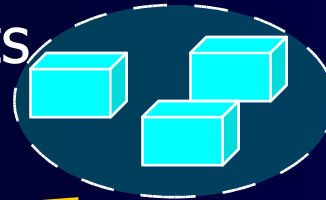
# Problem

- **How to find mathematical preprint**  
(teaching material, persons, institutions, projects, etc.)  
**on the internet?**
  - Centralized preprint servers
  - Departmental servers
- **MPRESS (Math-Net)**
  - User driven approach
  - Offering retrieval mechanisms
  - Worldwide initiative

# MPRESS – an example of a search or retrieval system



Math departments  
in Germany



*gathering*



*gathering*



*import*

France

*import*



Austria



*import*



In cooperation with



# MPRESS

# MathNet.preprints

## The Mathematics PREprint Search System

This is a searchable index of preprints from

[Index national des prépublications et thèses en mathématiques en France](#)

[JABaPub / Preprints from Austria](#)

[MathN / D-MathNet.preprints](#)

[Topology Atlas](#) (Preprints related to topology)

### [About MPRESS](#)

If you want MPRESS to index your preprints, please contact: [J. Plümer](#) or [R. Schwänzl](#)

Index home: U Osnabrück <http://MathNet.preprints.org>

Replica: EMIS <http://zblmath.fiz-karlsruhe.de/MPRESS>

The search interface will yield optimal results for servers providing [Dublin Core](#) MetaData.

A tool for creating MetaData is [MMM](#). (download [MMM source](#)).

[Help](#)

[Browse MSC](#)

Submit

Reset

Author:

Keywords:

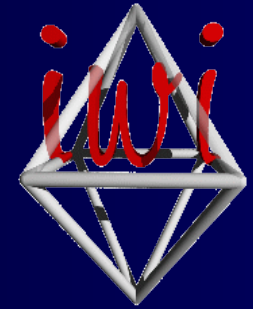
Title:

Free Text:

# Mathematical PREprint Search System



- **Decentralized system**
- **different methods to get data:**
  - import indices from remote brokers
  - import indices from remote gatherers
  - gather information by itself



# kind of content

- automatically generated summaries of mathematical preprints (PostScript)
- automatically generated summaries of mathematical preprints with additional metadata (HTML-Abstracts, RDF)
- no fulltext papers

hyperlinks to the original Files

1. More: [AlgGeo\\_I.ps.gz](#) on:ftp.mathematik.uni-osnabrueck.de
2. More: [iterated\\_monoidal.ps.gz](#) on:ftp.mathematik.uni-osnabrueck.de
3. More: [pr94013.ps.gz](#) on:www.mathematik.uni-bielefeld.de
4. More: [hollender.ps.gz](#) on:ftp.mathematik.uni-osnabrueck.de
5. More: [topologie.ps.gz](#) on:ftp.mathematik.uni-osnabrueck.de
6. More: [pr94028.ps.gz](#) on:ftp.mathematik.uni-osnabrueck.de
7. More: [algebra\\_I.ps.gz](#) on:ftp.mathematik.uni-osnabrueck.de
8. **Preprints of SFB 343**  
More: [sfb343.html](#) on:www.mathematik.uni-osnabrueck.de
9. More: [pr94028.ps.gz](#) on:www.mathematik.uni-bielefeld.de

[Help](#)[Browse MSC](#)

Submit

Reset

Author:

Keywords:

Title:

Free Text:

Fulltext search gives no precise results



# Searching in special fields

→ metadata



- **Author**
- **Title**
- **Keywords**
- **Classifikation (MSC)**
- **Dates**

### 1. Basic Differential Forms for Actions of Lie Groups

Peter W. Michor

Proc. AMS 124, 5 (1996) 1633–1642

[Source file](#)

Primary MSCClassification: 57S15 20F55

Upload: 27 Apr 1994

Update: 14 Nov 1994

More: [esi087.html](#) on:www.esi.ac.at

### 2. Bounds on the Multiplicity of Eigenvalues for Fixed Membranes

Thomas Hoffmann–Ostenhof, Peter W. Michor, Nikolai Nadirashvili

ESI preprints

[Source file](#)

Primary MSCClassification: 35B05 35P15 58G25

Upload: 15 Dez 1997

Update: 20 Jul 1998

More: [esi514.html](#) on:www.esi.ac.at

### 3. On the Geometry of the Virasoro–Bott Group

[Help](#)

[Browse MSC](#)

Submit

Reset

Author:

Keywords:

Title:

Free Text:



[Peter W. Michor](#)

[Basic Differential Forms for Actions of Lie Groups](#)

The paper is published: Proc. AMS 124, 5 (1996) 1633–1642

[MSC:](#)

57S15 Compact Lie groups of differentiable transformations

20F55 Coxeter groups, See also {[22E40](#)}

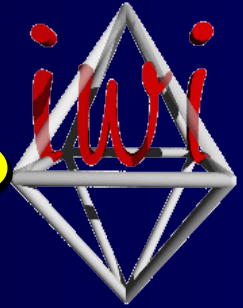
**Abstract:** A section of a Riemannian  $G$ -manifold  $M$  is a closed submanifold  $S$  which meets each orbit orthogonally. It is shown that the algebra of  $G$ -invariant differential forms on  $M$  which are horizontal in the sense that they kill every vector which is tangent to some orbit, is isomorphic to the algebra of those differential forms on  $S$  which are invariant with respect to the generalized Weyl group of  $S$ , under some condition.

**Keywords:** *Orbits, sections, basic differential forms*



34	Ordinary differential equations	54	<a href="#">search</a>	<a href="#">browse</a>
35	Partial differential equations	230	<a href="#">search</a>	<a href="#">browse</a>
39	Finite differences and functional equations	6	<a href="#">search</a>	<a href="#">browse</a>
40	Sequences, series, summability	1	<a href="#">search</a>	<a href="#">browse</a>
41	Approximations and expansions	48	<a href="#">search</a>	<a href="#">browse</a>
42	Fourier analysis	36	<a href="#">search</a>	<a href="#">browse</a>
43	Abstract harmonic analysis	3	<a href="#">search</a>	<a href="#">browse</a>
44	Integral transforms, operational calculus	9	<a href="#">search</a>	<a href="#">browse</a>
45	Integral equations	37	<a href="#">search</a>	<a href="#">browse</a>
46	Functional analysis	55	<a href="#">search</a>	<a href="#">browse</a>
47	Operator theory	60	<a href="#">search</a>	<a href="#">browse</a>
49	Calculus of variations and optimal control; optimization	40	<a href="#">search</a>	<a href="#">browse</a>
51	Geometry	28	<a href="#">search</a>	<a href="#">browse</a>
52	Convex and discrete geometry	42	<a href="#">search</a>	<a href="#">browse</a>
53	Differential geometry	67	<a href="#">search</a>	<a href="#">browse</a>
54	General topology	26	<a href="#">search</a>	<a href="#">browse</a>

<b>35Axx</b>	General theory	16	<a href="#">browse</a>
<b>35Bxx</b>	Qualitative properties of solutions	42	<a href="#">browse</a>
<b>35Cxx</b>	Representations of solutions	4	<a href="#">browse</a>
<b>35Dxx</b>	Generalized solutions of partial differential equations	24	<a href="#">browse</a>
<b>35Exx</b>	Equations and systems with constant coefficients	10	<a href="#">browse</a>
<b>35Fxx</b>	General first-order equations and systems	7	<a href="#">browse</a>
<b>35Gxx</b>	General higher-order equations and systems	5	<a href="#">browse</a>
<b>35H05</b>	Hypoelliptic equations and systems	1	<a href="#">browse</a>
<b>35Jxx</b>	Partial differential equations of elliptic type	77	<a href="#">browse</a>
<b>35Kxx</b>	Parabolic equations and systems	48	<a href="#">browse</a>
<b>35Lxx</b>	Partial differential equations of hyperbolic type	16	<a href="#">browse</a>
<b>35Mxx</b>	Partial differential equations of special type (mixed, composite, etc.)	5	<a href="#">browse</a>
<b>35Nxx</b>	Overdetermined systems	1	<a href="#">browse</a>
<b>35Pxx</b>	Spectral theory and eigenvalue problems for partial differential operators	24	<a href="#">browse</a>
<b>35Qxx</b>	Equations of mathematical physics and other areas of application	64	<a href="#">browse</a>



# Who produces metadata?

- Traditionally metadata are produced by librarians:
  - This takes time
  - Costs money
  - Therefore scientific knowledge in the respective area is needed
  - It is not realizable

**but then there is the author himself**

# Scientists are typically not experts in data formats



## Mathematics Metadata Markup 3.0

**Author(s)**

Please give me one more author field

**Lastname**

**Firstname**

**Emailaddress**

Schwänzl

Roland

roland@mathematik.uni-osnabrue

Staffeldt

Ross

ross@nmsu.edu

Waldhausen

Friedhelm

Friedhelm.Waldhausen@mathemati

**Title of paper:**

K-theory and generalized free products of S-algebras: Locali

**URL of paper:**

ftp://ftp.mathematik.uni-osnabrueck.de/pub/sfb343/pr99XXX.ps

The paper is

a preprint. ▾

It is written in

English ▾

and coded as

Postscript Document ▾





Open MSC2000 Prim. 19D10 Sec. 19D45 ZDM 1235 CR 67890 PACS: 12345



Keywords:

Stable K-theory, Topological

Abstract:

A generalized free product of algebras, a generalization and stabilization of the free product of rings arising from a Seifert manifold. Our eventual goal is to obtain a higher algebraic K-theory of the free product diagram in terms of smaller algebras. We first present a generalized free product diagram associated categories of May. We show that the categories are categories with cofibrations and mapping cylinders. In partic

Notes:

19Dxx | Prev: [19C](#) | Up: [19](#) | Next: [19E](#)

### Higher algebraic $SK$ -theory

- 19D06  $Q$ - and plus-constructions
- 19D10 Algebraic  $SK$ -theory of spaces
- 19D23 Symmetric monoidal categories [See also [18D10](#)]
- 19D25 Karoubi-Villamayor-Gersten  $SK$ -theory
- 19D35 Negative  $SK$ -theory, NK and Nil
- 19D45 Higher symbols, Milnor  $SK$ -theory
- 19D50 Computations of higher  $SK$ -theory of rings [See also [13D15](#), [16E20](#)]
- 19D55  $SK$ -theory and homology; cyclic homology and cohomology [See also [18G60](#)]
- 19D99 None of the above, but in this section

# Storing of data by Harvest

(SOLF=summary object interchange format)



```
<META NAME=„DC.Creator.Personalname“>
```

```
CONTENT=„Erwin Mustermann“>
```

```
dc.creator.personalname{16}: Erwin Mustermann
```

```
<META NAME=„DC.Title“>
```

```
CONTENT=„Theorems on Potatoes“>
```

```
dc.title{20}: Theorems on Potatoes
```

# Problem with HTML-META



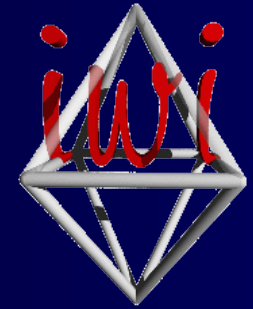
```
<META NAME="dc.creator"  
          CONTENT="Marie">
```

```
<META NAME="dc.creator"  
          CONTENT="Paul">
```

```
<META NAME="dc.creator.email"  
          CONTENT="marie@email.org">
```

```
<META NAME="dc.creator.email"  
          CONTENT="paul@email.org">
```

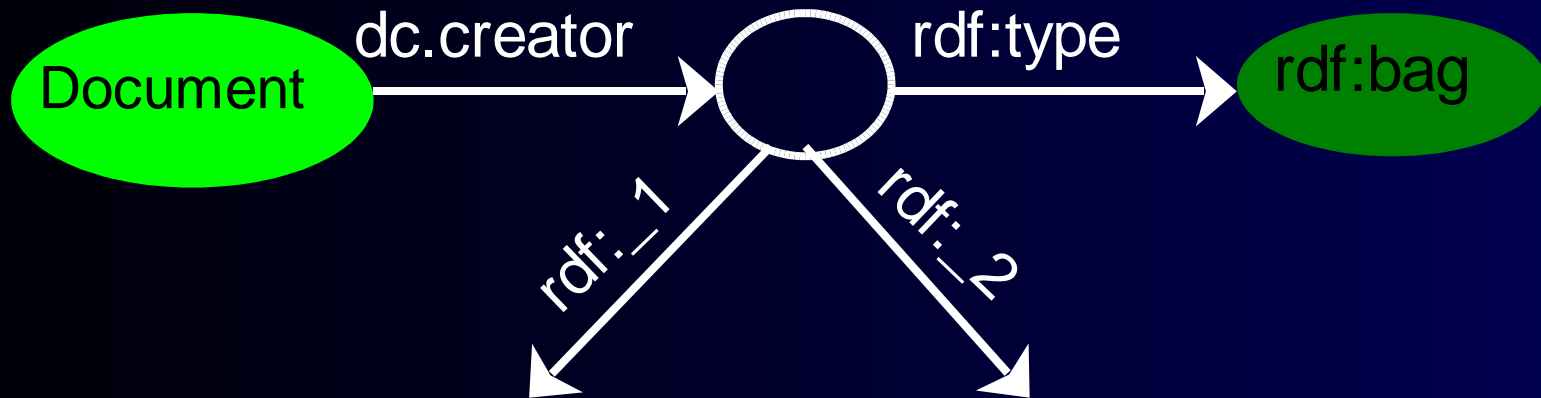
HTML-META is commutative and associative



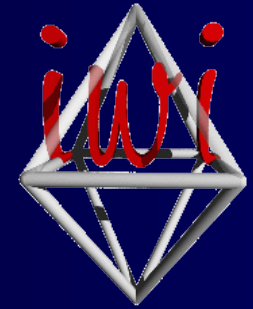
# RDF - What's that?

- **W3C:**  
***“The Resource Description Framework (RDF) integrates a variety of web-based metadata activities including sitemaps, content ratings, stream channel definitions, search engine data collection (web crawling), digital library collections, and distributed authoring, using XML as an interchange syntax.”***

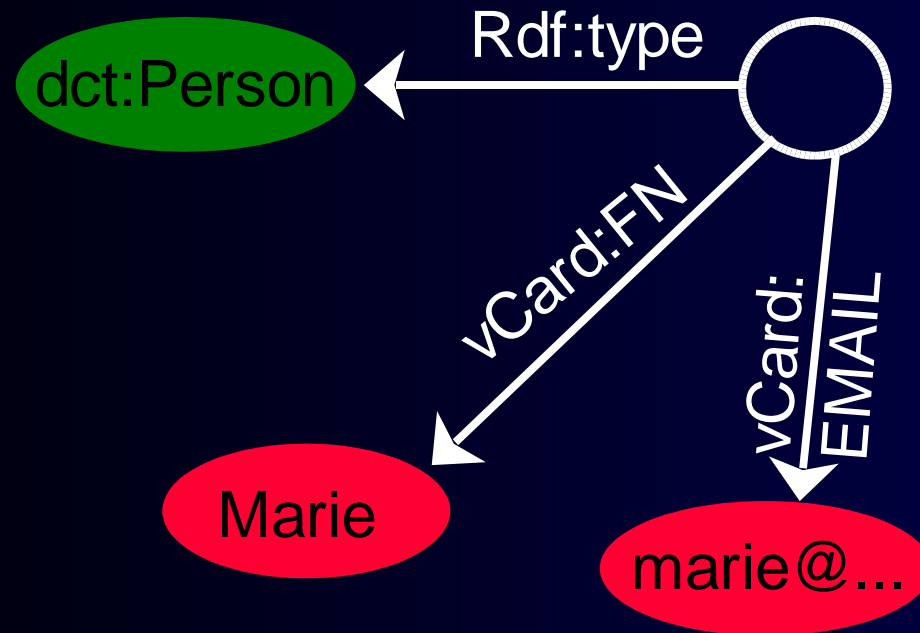
# RDF – first impression



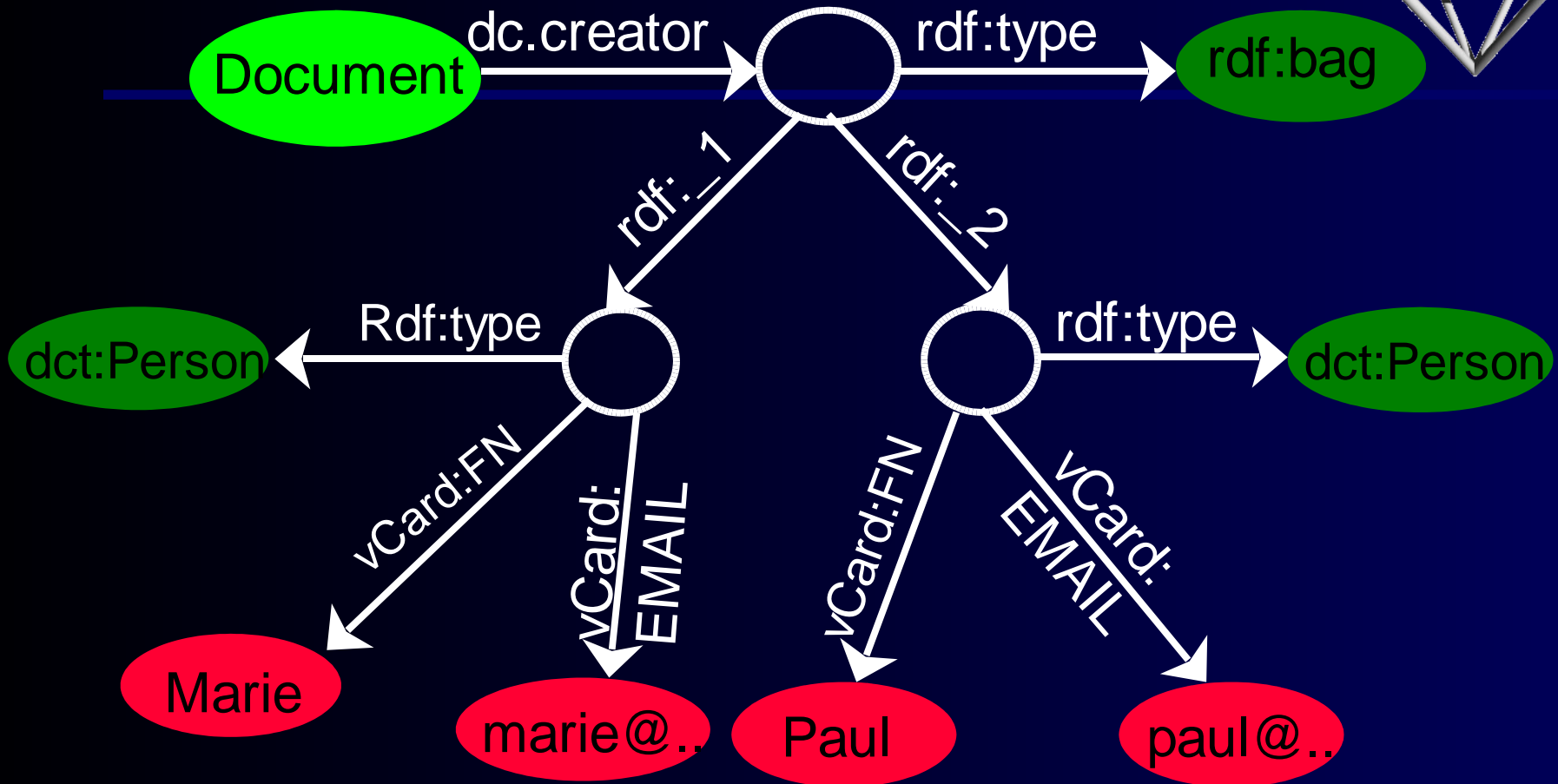
Document with two authors

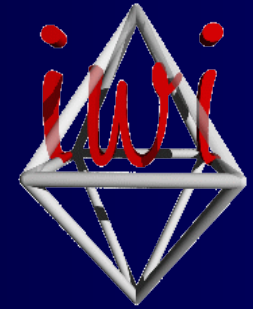


# RDF – first impression



# RDF – first impression





# HTML Meta $\leftrightarrow$ RDF

- HTML META consist of attribute – value pairs
- RDF consists of triples (Predicate, Subject, Object)
- RDF encodes statements





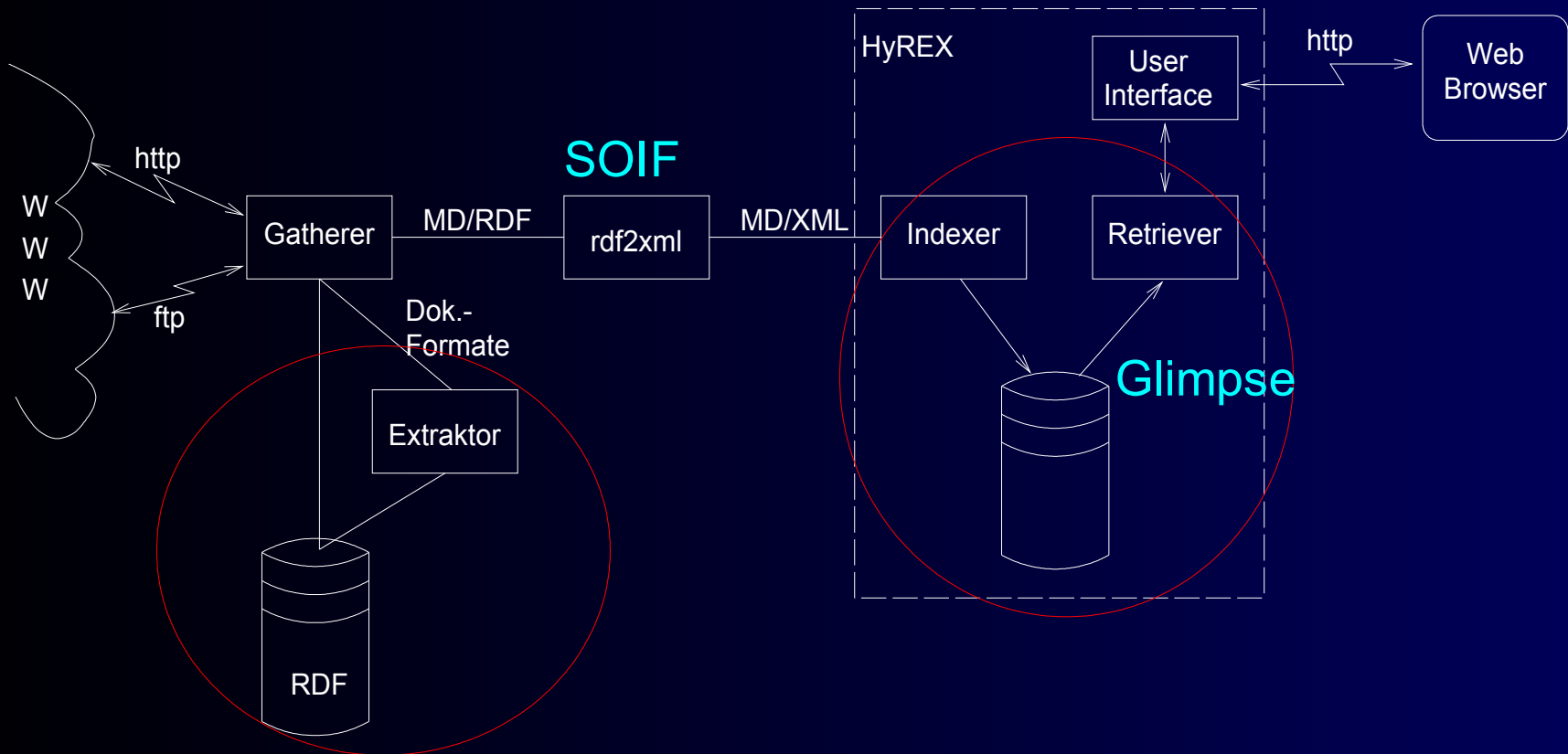
# CARMEN AP7



An integrated hypertext and retrieval system for digital libraries

- Development of a retrieval system on the base of harvest-ng including all functionalities of harvest.
- Gatherer generates RDF instead of SOIF
- Retrieval component designed for XML

# Architecture





# XIRQL: XML IR query language

- ⌘ Weights for documents with metadata that were generated by probabilistic methods/heuristics
- ⌘ Vague predicates



# Reality

60.000 records in MPRESS

- /// 45 % with MetaData
- /// 30 % with MSC classification

heterogeneous material

Which problems occure?



# Problems

Generation of RDF instead of HTML

- /// New versions of MMM are there for more than 2 years
- /// Changes of the RDF specification require changes in the software...

# Problems



- new summarizer for .PS, .PDF, .HTML
- Summarizers were developed which use heuristics and statistical methods for the extraction/generation of metadata
  - extraction of classification/title/abstract

# Problems

---



dump down of RDF metadata and  
conversion to XML with a predefined  
DTD

because currently the retrieval does not  
use the whole RDF structure of the  
metadata files



# Problems

---

OAI import und export

- /// Interfaces to common protocols are organized by the broker function





# Problems

modification of the userinterface

- /// open questions till now:
  - /// use of new functionalities without bothering the user
  - /// design of a new user interface?



# A lot to do...

---

- Test installations are already running on a part of the data

consult us at [MathNet.preprints.org](http://MathNet.preprints.org)