

Semantische Annotationen (mit Plone)

Dr. Hendrik Bunke <h.bunke@zbw.eu>

ZBW IT-Entwicklung

Plonekonferenz 2012

München

23.02.2012

Überblick

- Vorstellung
- Semantic Web
- Semantische Annotationen: RDFa, schema.org
- Beispiel Umsetzung (mit Plone)
- Tools und Infoquellen
- Conclusio

ZBW

- Deutsche Zentralbibliothek für Wirtschaftswissenschaften - Leibniz-Informationszentrum Wirtschaft
- Weltweit größte Fachbibliothek für Wirtschaftswissenschaften mit Sitz in Kiel und Hamburg
- Neben gedruckten Publikationen seit einigen Jahren zunehmend Schwerpunkt auf digitalen Angeboten und informationswissenschaftlicher Forschung
- Angebote u.a.: [EconBiz](#) (Suchportal), [EconStor](#) (Open Access Repository), [Economics](#) (Open Access E-Journal)
- <http://zbw.eu>

Economics

- Open-Access, Open-Assessment E-Journal
- Kooperation ZBW mit [Institut für Weltwirtschaft \(IfW\)](#)
- etabliertes, internationales Journal mit sechs Nobelpreisträgern im Advisory Board und Listung im SSCI
- derzeit: 2700 'registrierte Leser' (ausgewiesene Wissenschaftler), 383 AutorInnen, 150 Journalarticle, 240 Discussionpaper
- Plone basiert (derzeit: 3.3.5) mit diversen Eigenentwicklungen
- <http://www.economics-ejournal.org>

Semantic Web

Worum geht es?

- **Daten** (*implizit* und *explizit*)
- **Bedeutung** (*Semantik*)
- **Verknüpfung** (*Linked Data*)



- **Web of data** (vs. *Web of documents*)

Journal Article

No. 2011-22 | December 22, 2011

Corresponding Discussion Paper

Jonathan Pycroft, Lucia Vergano, Chris Hope, Daniele Paci, and Juan Carlos Ciscar

A Tale of Tails: Uncertainty and the Social Cost of Carbon Dioxide

(Published in Special Issue [The Social Cost of Carbon](#))

[Download \(pdf, 413.1 kB\)](#) | [Comments \(1\)](#) | [Assessment](#) | [References \(37\)](#)

Abstract

Recent thinking about the economics of climate change has concerned the uncertainty about the upper bound of both climate sensitivity to greenhouse gases and the damages that might occur at high temperatures. This argument suggests that the appropriate probability distributions for these factors may be fat-tailed. The matter of tail shape has important implications for the calculation of the social cost of carbon dioxide (SCCO₂). In this paper a probabilistic integrated assessment model is adapted to allow for the possibility of a thin, intermediate or fat tail for both (i) the climate sensitivity parameter and (ii) the damage function exponent. Results show that depending on the tail shape of the climate sensitivity parameter the mean SCCO₂ rises by 29–85%. Changes in the mean SCCO₂ due to the adjustments to the damage function alone range from a reduction of 7% to a rise of 12%. The combination of both leads to rises of 33–15%. Greater rises occur for the upper percentiles of the SCCO₂ estimates. Given the uncertainties in both the science and the economics of climate change different tail shapes deserve consideration due to their important implications for the range of possible values for the SCCO₂.

JEL Classification

Q54

Citation

Jonathan Pycroft, Lucia Vergano, Chris Hope, Daniele Paci, and Juan Carlos Ciscar (2011). A Tale of Tails: Uncertainty and the Social Cost of Carbon Dioxide. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 5, 2011-22.
<http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22>

Reference: [BibTex](#) | [Dublin Core](#) | [RIS \(Endnote\)](#)

Document

Journal Article

No. 2011-22 | December 22, 2011

Corresponding Discussion Paper

Jonathan Pycroft, Lucia Vergano, Chris Hope, Daniele Paci, and Juan Carlos Ciscar

A Tale of Tails: Uncertainty and the Social Cost of Carbon Dioxide

(Published in Special Issue [The Social Cost of Carbon](#))

[Download \(pdf, 413.1 kB\)](#) | [Comments \(1\)](#) | [Assessment](#) | [References \(37\)](#)

Abstract

Recent thinking about the economics of climate change has concerned the uncertainty about the upper bound of both climate sensitivity to greenhouse gases and the damages that might occur at high temperatures. This argument suggests that the appropriate probability distributions for these factors may be fat-tailed. The matter of tail shape has important implications for the calculation of the social cost of carbon dioxide (SCCO₂). In this paper a probabilistic integrated assessment model is adapted to allow for the possibility of a thin, intermediate or fat tail for both (i) the climate sensitivity parameter and (ii) the damage function exponent. Results show that depending on the tail shape of the climate sensitivity parameter the mean SCCO₂ rises by 29–85%. Changes in the mean SCCO₂ due to the adjustments to the damage function alone range from a reduction of 7% to a rise of 12%. The combination of both leads to rises of 33–15%. Greater rises occur for the upper percentiles of the SCCO₂ estimates. Given the uncertainties in both the science and the economics of climate change different tail shapes deserve consideration due to their important implications for the range of possible values for the SCCO₂.

JEL Classification

Q54

Citation

Jonathan Pycroft, Lucia Vergano, Chris Hope, Daniele Paci, and Juan Carlos Ciscar (2011). A Tale of Tails: Uncertainty and the Social Cost of Carbon Dioxide. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 5, 2011-22.
<http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22>

Reference: [BibTex](#) | [Dublin Core](#) | [RIS \(Endnote\)](#)

Data

Semantic Web

W3C Standards (Auswahl)

RDF

RDFS

N-Triple

SPARQL

Turtle

Notation3

OWL

SKOS

RDFa

Semantic Web

W3C Standards (Auswahl)

RDF

RDFS

N-Triple

Turtle

SPARQL

Notation3

OWL

RDFa

SKOS

RDF

Die Basis des Semantic Web

- **Resource Description Framework**
- maschinen-lesbare Beschreibung beliebiger Ressourcen
- Modellierung als *Triple*:
Subjekt, Prädikat (Eigenschaft), Objekt
- beliebig erweiterbar mit Vokabularien (bspw. FOAF, DC, bibo, schema.org)
- Diverse Syntaxen und Serialisierungsformate: u.a. RDF/XML, Turtle, N3, JSON-LD, RDFa

Semantische Annotationen

RDFa

- RDF Einbettung in beliebige Markup-Sprachen als **A**tribute
- Vorteile:
 - einfache Syntax, nur wenige Attribute
 - Verwendung des HTML-Inhalts als RDF-Datenquelle (Vermeidung von Dopplungen)
 - Erweiterung durch Vokabularien genauso möglich wie bei RDF
- aktuell: RDFa 1.0 (14.10.2008)
- "Last Call Draft" 31.1.2012:
RDFa Core 1.1, XHTML+RDFa 1.1, RDFa Lite 1.1

Semantische Annotationen

RDFa: Beispiel

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/">
  <h2 property="dc:title">The trouble with Bob</h2>
  <h3 property="dc:creator">Alice</h3>
</div>
```

Semantische Annotationen

RDFa: Probleme

- bislang keine breite Adaption
- Syntax kann wegen Verwendung verschiedener Vokabulare immer noch komplex und unter Umständen mühsam zu implementieren sein
- Direkter Nutzen nicht sofort erkennbar
- Kein *Push* durch Suchmaschinen

Semantische Annotationen

schema.org

- gemeinsame Initiative (Juni 2011) von Google, Bing, Yahoo, später Yandex
- Ziel: Vereinfachung von Annotierungen und Verbesserung von Suchergebnissen
- Implementierung mit *Microdata* (HTML5)
- <http://schema.org/>

Semantische Annotationen

schema.org: Beispiel

```
<div itemscope itemtype="http://schema.org/Movie">
  <h1 itemprop="name">Avatar</h1>
  <span>Director:
    <span itemprop="director">James Cameron</span>
    (born August 16, 1954)
  </span>
  <span itemprop="genre">Science fiction</span>
  <a href="../../movies/avatar-theatrical-trailer.html"
    itemprop="trailer">Trailer</a>
</div>
```

Semantische Annotationen

RDFa vs. schema.org

- schema.org lässt sich mit RDFa implementieren (Vokabular)
- schema.org unterstützt RDFa
- Transformation von Microdata in RDF wird möglich sein (W3C Working Draft 12.01.2012)



Flexibilität:

- schema.org, wenn SEO und einfache Implementierung im Vordergrund stehen, **UND** HTML5 benutzt wird
- RDFa, wenn Erweiterbarkeit und Verlinkung im Vordergrund stehen

Semantische Annotationen

Usecase

- Implementierung: Beispiel *Economics E-Journal*
- In Discussion Papers und Journal Articles sollen einige wesentliche Objekte annotiert werden: Autor(en), Titel, Datum, Abstract, URL
- Umsetzung in:
 - HTML5 / Plone [**DANGER!**]
 - RDFa 1.1
 - Vokabular: vorerst nur schema.org

Semantische Annotationen

RDFa 1.1

```
<h1>Journal Article</h1>

<div id="ej_paperview"
  prefix="schema: http://schema.org/
         dcterms: http://purl.org/dc/terms/
         bibo: http://purl.org/ontology/bibo/
         owl: http://www.w3.org/2002/07/owl#"
  typeof="schema:ScholarlyArticle"
  about="http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22">
```

Semantische Annotationen

```
<div class="authors">
```

```
  <span rel="schema:author" typeof="schema:Person">
```

```
    <a property="schema:url"
```

```
      href="http://www.economics-ejournal.org/community/autho
```

```
    <span property="schema:name">Jonathan Pycroft</span></a> ,
```

```
  </span>
```

```
[...]
```

```
</div>
```

Semantische Annotationen

```
[<div class="title">  
  <span property="schema:headline">A Tale of Tails: Uncertainty and the  
    Social Cost of Carbon Dioxide</span>  
</div>
```

[...]

```
<h2>Abstract</h2>
```

```
<div id="ejournal-abstract"  
  property="schema:description">
```

```
<p>
```

```
Recent thinking about the economics of climate change has concerned the  
uncertainty about the upper bound of both climate sensitivity to  
greenhouse gases and the damages that might occur at high temperatures.  
This argument suggests that the appropriate probability distributions
```

Semantische Annotationen

RDFa Distiller: RDF/XML

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:schema="http://schema.org/"
>
  <schema:ScholarlyArticle rdf:about="http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22"
    <schema:headline>A Tale of Tails: Uncertainty and the Social Cost of Carbon Dioxide</schema:headline>
    <schema:keywords>Q54</schema:keywords>
    <schema:url rdf:resource="http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22"/>
    <schema:author>
      <schema:Person rdf:nodeID="UMEqAPzg12">
        <schema:name>Jonathan Pycroft</schema:name>
        <schema:url rdf:resource="http://www.economics-ejournal.org/community/authors/author">
        </schema:Person>
      </schema:author>
    <schema:description>Recent thinking about the economics of climate change has concerned t
    <schema:datePublished>December 22, 2011</schema:datePublished>
  </schema:ScholarlyArticle>
</rdf:RDF>
```

Semantische Annotationen

RDFa Distiller: Turtle

```
@prefix schema: <http://schema.org/> .

<http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22> a schema:ScholarlyArticle;
  schema:author [ a schema:Person;
    schema:name "Jonathan Pycroft";
    schema:url <http://www.economics-ejournal.org/community/authors/author.2011-22> .
  ]
  schema:datePublished "December 22, 2011";
  schema:description "Recent thinking about the economics of climate change has conc";
  schema:headline "A Tale of Tails: Uncertainty and the Social Cost of Carbon Dioxid";
  schema:keywords "Q54";
  schema:url <http://dx.doi.org/10.5018/economics-ejournal.ja.2011-22> .
```

Semantische Annotationen

schema.org Microdata

```
<h1>Journal Article</h1>

<div id="ej_paperview"
  itemscope
  itemtype="http://schema.org/ScholarlyArticle">
[... ]
  <div class="authors">
    <span itemprop="author" itemtype="http://schema.org/Person">
      <a itemprop="url"
        href="http://www.economics-ejournal.org/community/authors/aut
      <span itemprop="name">Jonathan Pycroft</span></a>,
    </span>
  [... ]
  </div>

  <div class="title">
    <span itemprop="headline">A Tale of Tails: Uncertainty and the
      Social Cost of Carbon Dioxide</span>
  </div>
</div>
```

sieht einfach aus...
wo ist der Haken?

Plone + Validierung

- HTML5 Validator beim W3C kennt (noch) keine RDFa 1.1 Attribute
- **NEU** (08.02.2012): <http://validator.w3.org/nu/>
- Doctypes (X)HTML+RDFa machen ebenfalls Validierungs-Probleme mit Plone
- Plone wird ab Version 4.2 standardmäßig HTML5 nutzen (<http://dev.plone.org/ticket/11300>)
- Mit früheren Plone Versionen sollte Doctype angepasst werden.
- Nicht-Validierung muss mit RDFa zum jetzigen Zeitpunkt in Kauf genommen werden. 'Fehlermeldungen' sollten sich aber lediglich auf die RDFa Attribute und gegebenenfalls kss-metatags beziehen.
- schema.org Attribute mit Microdata/HTML5 sollten validieren.

Plone Verankerung

Ansatzpunkte (please help)

- Templates Standard Contenttypes?
- Field Property / Mappings in AT oder Dexterity?
- Web Konfiguration?
- ???

Tools

- RDFa Distiller (Ivan Herman):
 - <http://www.w3.org/2007/08/pyRdfa/> (RDFa 1.0)
 - <http://www.w3.org/2007/08/pyRdfa/Shadow.html> (RDFa 1.1)
 - pypi: <http://pypi.python.org/pypi/pyRdfa/>
- Rich Snippets (Google): <http://www.google.com/webmasters/tools/richsnippets>
- Redland Library + Commandline tools + Python bindings: <http://librdf.org>

Conclusio

- Das *Semantic Web* wird Realität
- schema.org und RDFa bringen es in den *Massenmarkt*

Ihre Plone Site

- Die Implementierung in Plone Templates sollte spätestens ab der kommenden Version 4.2 und mit HTML5 einfach sein
- Plone Standard Contenttypes sollten zukünftig (zumindest) schema.org Attribute ausgeben.

Conclusio

Act now!

Danke!

- Für Ihre Aufmerksamkeit!
- An alle Plone-EntwicklerInnen!
- Und die KonferenzorganisatorInnen!