



1996 – 2011: 15 years in the market

Opoint introduction



- Founded in 1996 – pioneer in the web monitoring business
- Acquired by Edda Media/Mecom in 2005
 - Edda Media is a leading media group and content provider in Norway
 - Mecom owns over 300 printed titles and over 200 websites in the Netherlands, Denmark, Norway and Poland.
- Rating "AAA" Dun & Bradstreet



1996 – 2011: 15 years in the market

Opoint introduction



- Opoint: 25 employees
- Innovative leader in web media monitoring and filtering
- Monitors radio/TV manually
- From January 2011 manual scanning of printed media; appr. 200 titles
- Technological platform developed in-house by own skilled IT staff



Opoint introduction



Supplier of premium quality web crawling, newsfeed and application MONITOR for internet media monitoring and information management

Target groups:

1. Medium to large companies and Government in Norway
(appr. 450 customers)
–7 Ministries, City of Oslo, Public and Private sector etc.

1996 – 2011: 15 years in the market

Opoint front-end: MONITOR2

Opoint Monitor 2.0 - Mozilla Firefox

File Rediger Vis Historikk Bokmerker Verktøy Hjelp

http://monitor2.opoint.com/

Opoint Webmail :: Innboks Google Oversetter Opoint Monitor 2.0

monitor by Opoint

Overvåkning

Rapporter

Varslinger

Statistikk

Analyse

Administrasjon

Brukerkonto

ski vm

Lagre som profil

MINE SNARVEIER

RESULTATLISTE

Merk

VISNING

Artikkel med bilder

JPG av artikkelen

Artikkelen i dag

Opprinnelig artikkel

Profiler + Lag ny

Merkelapper + Lag ny

Søkehistorikk Fjern alle søk


ski vm (2)

ski vm

Søk

▼ ski vm (2)

- 1. Re: Fototreff for Akam Oslo**
18:12 i dag Digital foto og video: Fotografering: Fotoaktiviteter
- 2. Bjørgen på VM-frimerker**
17:59 i dag Namdalsavisa - Førstesiden
13:31 i dag Bygg.no - Hovedsiden
13:53 i dag Trønder-Avisa
[+15 identiske artikler](#)
- 3. Disse var best i Sandefjordsrennet**
17:18 i dag mobil.sb.no - Sport
17:05 i dag Sandefjords Blad - Sport
- 4. Her er de beste gratisplassene under VM**
17:17 i dag Nettavisen mobil - Sport
17:02 i dag Nettavisen - Sport
- 5. Hyrer inn politistudenter og vektere**
17:15 i dag P4 - Nyheter
- 6. Grått og fuktig VM-vær**
17:02 i dag



VM i LØYPA. Det er gode muligheter for å få med seg flere VM-avbeiser ved å slå langs løypa ved Frognerseteren. Her kan du se løperne på nært hold og heie dem fram. Foto: Ragnhild I (Medehuset Nettavisen)

Her er de beste gratisplassene under VM

Fullført

start

All Mail Items - Micros...

Opoint Monitor 2.0 - ...

Opoint presentasjon...

2 Windows Explorer


Search Desktop

19:07

Opoint introduction

2. Portals (major portals in Nordic)

–Startsiden.no and kvasir.no



The screenshot shows a Mozilla Firefox browser window displaying the Kvasir news portal. The search results are for the query 'stoltenberg'. A large green arrow points from the text 'Startsiden.no and kvasir.no' to the search bar. Another green arrow points from the search bar to the 'opoint' logo, which is circled in green. The page shows several news items, including 'Et ekskluderende arbeidsliv', '- Ja, vi bryr oss', 'Støre kan bli helseminister', and 'Erna håper Jens roser henne for å være saklig'. The 'opoint' logo is located in the top right corner of the search results area.

1996 – 2011: 15 years in the market

Opoint introduction

3. Other Media Monitoring companies in Europe

- Infopaq (S,FIN,DK, D)
- Newsline (S)
- Infomedia (DK)
- Netsprint (PL)
- APA-deFacto (A)
- Digital Media Research (I)
- Visual Box (I)
- PMG Presse-Monitor (D)
- NEWBASE (D)
- X-Ray (D)
- AuxiPress (B)
- Press Index (F)
- NLA (Newspaper Licensing Agency) (UK)
- Baltic and Eastern-Europe (through Netsprint)



1996 – 2011: 15 years in the market

Opoint strategy

- Opoint has added to the minimum all the sources that Moreover has in Europe:
 - Q2 – 11: Collects appr. 750.000 different art. per day from 29.000 sources
 - Q4 – 11: Collect appr. 1 mill different art. per day from 35.000 sources.

- Various benchmarks shows that we are the supplier of content from web in Europe, that collects most different articles and headlines = the leading provider in the European market.

- Our strategy is to become a content provider to a wide range of media monitoring companies in Europe

1996 – 2011: 15 years in the market

Opoint technology compared



- On average we find, index and filter articles within 10 min. after publishing, most competitor have substantially higher delays
- We spend less time on manual processes, but with superior quality output
- On average we find 15 - 20 % more news articles than our competitors



1996 – 2011: 15 years in the market

Finding news fast and reliably



- Each section/watchpoint is configured manually
- Configuration separates news links from other links
- Visual tool makes this transparent and reliable
- We tune the visit frequency for each section
- Asynchronous checking for new links
- We are 2-4 times faster than competitors



1996 – 2011: 15 years in the market

Maintaining site configurations



- Manual configuration makes it possible to generate accurate information about sites needing attention
- Many check lists are built into the system
- The visual tool makes it easy to spot errors
- Monitoring works well over time

1996 – 2011: 15 years in the market

Classification of contents on article pages

- We use a patched Firefox to download article pages
- Our crawler acts almost like a browser
- The crawler can do «human» operations
- Rendering pages helps classifying contents
- Visual information is used in classification
- Firefox can block ads
- Automatic learning is used in classification

Delivery

- An XML Delivery from Opoint will include the following:

- o URL

```
http://redir.opoint.com/?url=http://news.bbc.co.uk/2/hi/business/8255482.stm&OpointData=990645113e16a459b60697fe6d57cc8fJmlkX3NpdGU9MTMxJmlkX2FydGljbGU9NjgyMDk0JmlkX3VzZXI9ODc4OCZpZF9hcHBsaWNhdGlvbj0xMDAwMzU5Jmxhbmc9ZW4=
```

- o Header

- o Summary

- o Text

```
</url>  
<direct_url>
```

```
http://news.bbc.co.uk/2/hi/business/8255482.stm
```

```
</direct_url>
```

```
<main_source id="429">
```

```
<name>BBC News - Business</name>
```

```
<url>http://news.bbc.co.uk/hi/english/business/default.stm</url>
```

- o Image text

- o Image URL and image text if found

- o Information about visibility on the section pages

where the articles was found → **Analysis!**

How we classify contents on article pages

- A patched Firefox downloads a given URL and renders it
- Firefox writes a modified version of the original HTML code containing
 - all original HTML code
 - Visual information from the rendered page for each tag, like position, size, font size, font weight etc.
 - Absolute links
 - HTML code for all frames and iframes on the web page.
- The use of Firefox export many hard parts (downloading, redirects, Javascript evaluation etc) to the Firefox team.
- The browser acts more as a normal user than a standard crawler, and can be programmed to preform user operations. This makes it easier to fool unfriendly sites. The system can also use anonymous proxies to get around IP blocking.

How we classify contents on article pages

- Descriptive vectors is generated for each text and image element, almost entirely based on visual information from Firefox. This is a good thing since visual properties are more stable than HTML code across sites.
- The elements are classified as header, summary, text, sub-header, publish date, image or image text based on automatic learning from a manually classified learning set containing 2500 articles.
- An XML version is generated from the web page.
- The special version of the HTML code is saved in our database.
- The articles are checked for uniqueness and quality before publishing.

Filtering and searching

- An article is published if the edit distance to the nearest article from that site is larger than 30% of its length.
- Our in-house developed filtering system can filter 50 articles each second to 10 000 different news lists.
- Each filter can contain hundreds of search terms and logical and mathematical relations between them, and they can depend on each other.

Filtering and searching

- The indexing system supports different access privileges for users, making it easy to include paid sites in the system.
- Our news portal can easily be adapted to a new brand and language.