



Palladian

KNIME User Group Meeting,
Zürich, 07.03.2013

Miranda Ackerman, David Urbansky, Philipp Katz
TU Dresden

About us

- Information retrieval team at Lehrstuhl Rechnernetze, TU Dresden
- Research focus
 - Efficient feed polling strategies
 - Forum Q/A detection
 - Clustering heterogeneous content
 - Entity and fact extraction

Palladian?

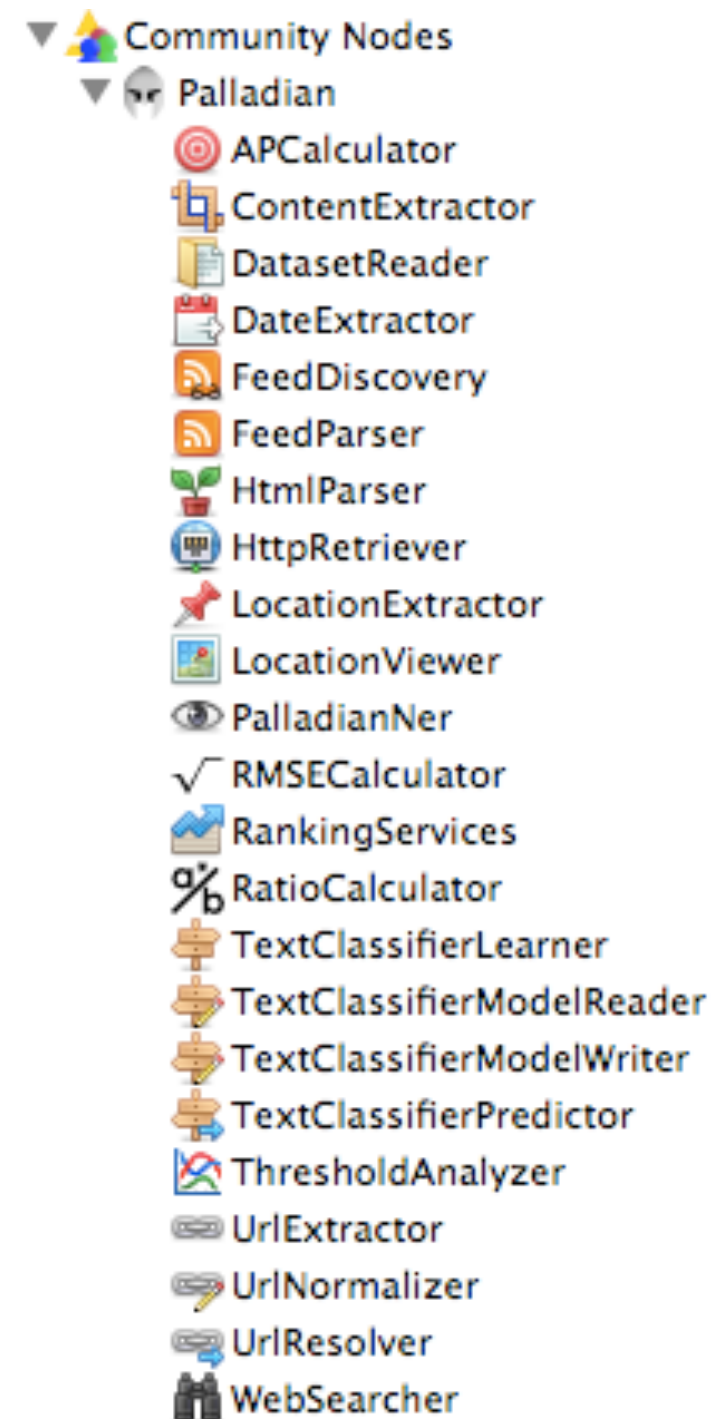
- Java-based toolkit for information retrieval
- Founded in 2009, still growing
 - Palladian KNIME nodes since 2011
- Provide users with a basic set of tools
- Used in commercial and academic projects

Integration with others



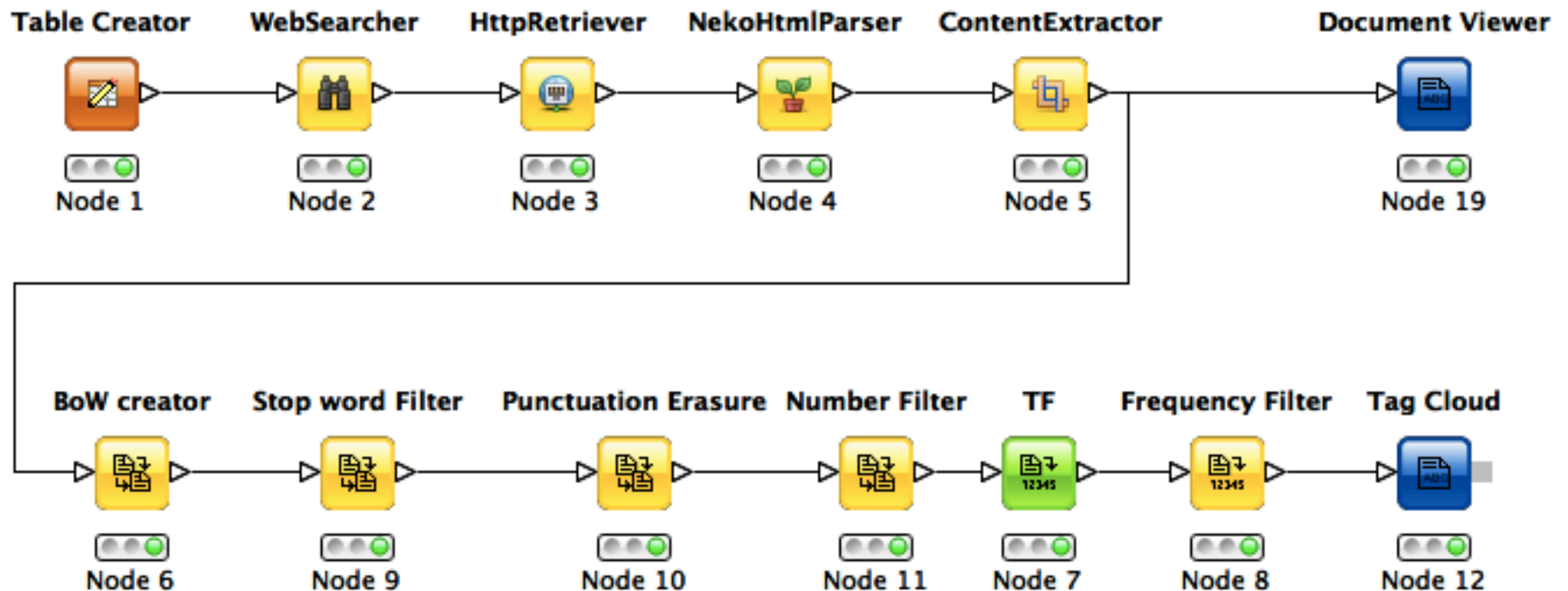
Palladian KNIME Nodes

- Text classification
- Content extraction
- Date extraction
- Location extraction
- Named entity recognition
- Web page, image, news search
- HTML, RSS, Atom parsing
- Ranking value retrieval
- Evaluation metrics



Examples

- Palladian + KNIME Text Processing



Examples

- Check news feeds, retrieve ranking values from Facebook, Bit.ly, Google, ...



Examples

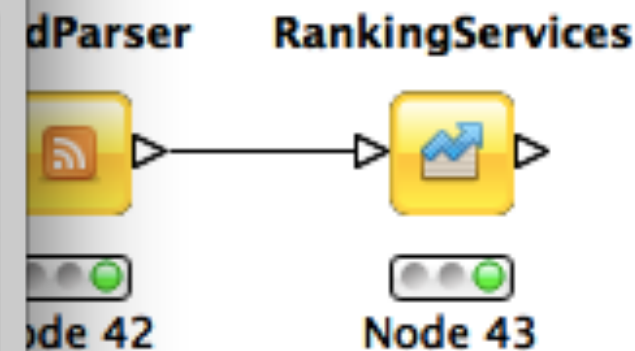
- Check news feeds, retrieve ranking values from Facebook, Bit.ly, Google, ...

Manually created table - 4:40 - Table Creator

File

Table "default" - Rows: 3 Spec - Column: 1

Row ID	S column1
Row1	http://rss.cnn.com/rss/edition.rss
Row2	http://feeds.reuters.com/reuters/topNews
Row3	http://feeds.bbc.co.uk/news/rss.xml



Examples

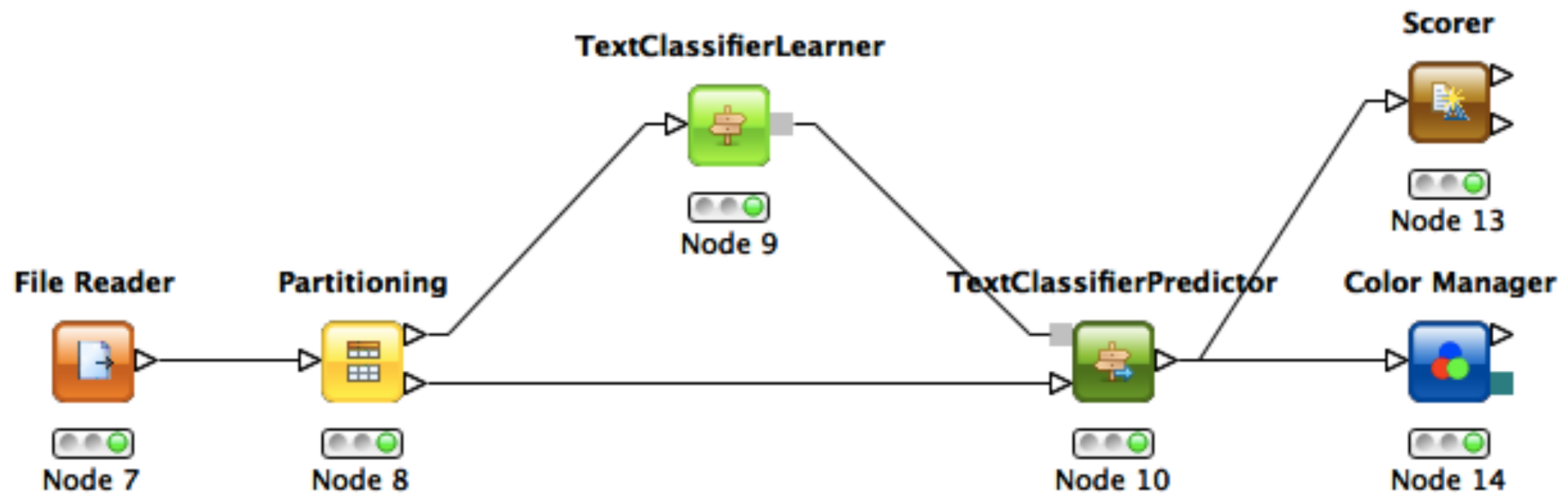
- Check news feeds, retrieve ranking values from Facebook, Bit.ly, Google, ...

The screenshot shows a software interface with two windows. The top window is titled "Manually created table - 4:40 - Table Creator". The bottom window is titled "Ranking values - 4:43 - RankingServices" and contains a table with 96 rows. The table has columns for Row ID, title, Bit.ly Clicks, Facebook Shares, and Facebook Likes. The first row, Row91, is highlighted and shows the title "Living in the world's most expensive city" with 2,725 Bit.ly Clicks, 0 Facebook Shares, and 0 Facebook Likes. The other rows show various news titles and their corresponding ranking values.

Row ID	title	Bit.ly Clicks	Facebook Shares	Facebook Likes
Row91	Living in the world's most expensive city	2,725	0	0
Row51	'Big cat' theory ruled out by DNA	1,989	0	0
Row35	Terry 'will not resign captaincy'	1,261	0	76
Row61	Scores saved off Papua New Guinea	860	0	2
Row20	Tensions rise in Cairo over riots	829	0	10
Row74	In pictures: Sony World Photography Awards sh...	791	0	1
Row28	Covert policeman 'defied' bosses	664	0	0
Row83	VIDEO: Europe's cold spell to last for days	639	0	1
Row43	UK download speed gains 'uneven'	581	0	1
Row89	Meet the galanthophiles	562	0	1
Row84	VIDEO: British men trafficked into slavery	551	0	0
Row65	Deadly attack on Colombia police	540	0	0
Row66	Chile arrest in glacier ice theft	418	0	1

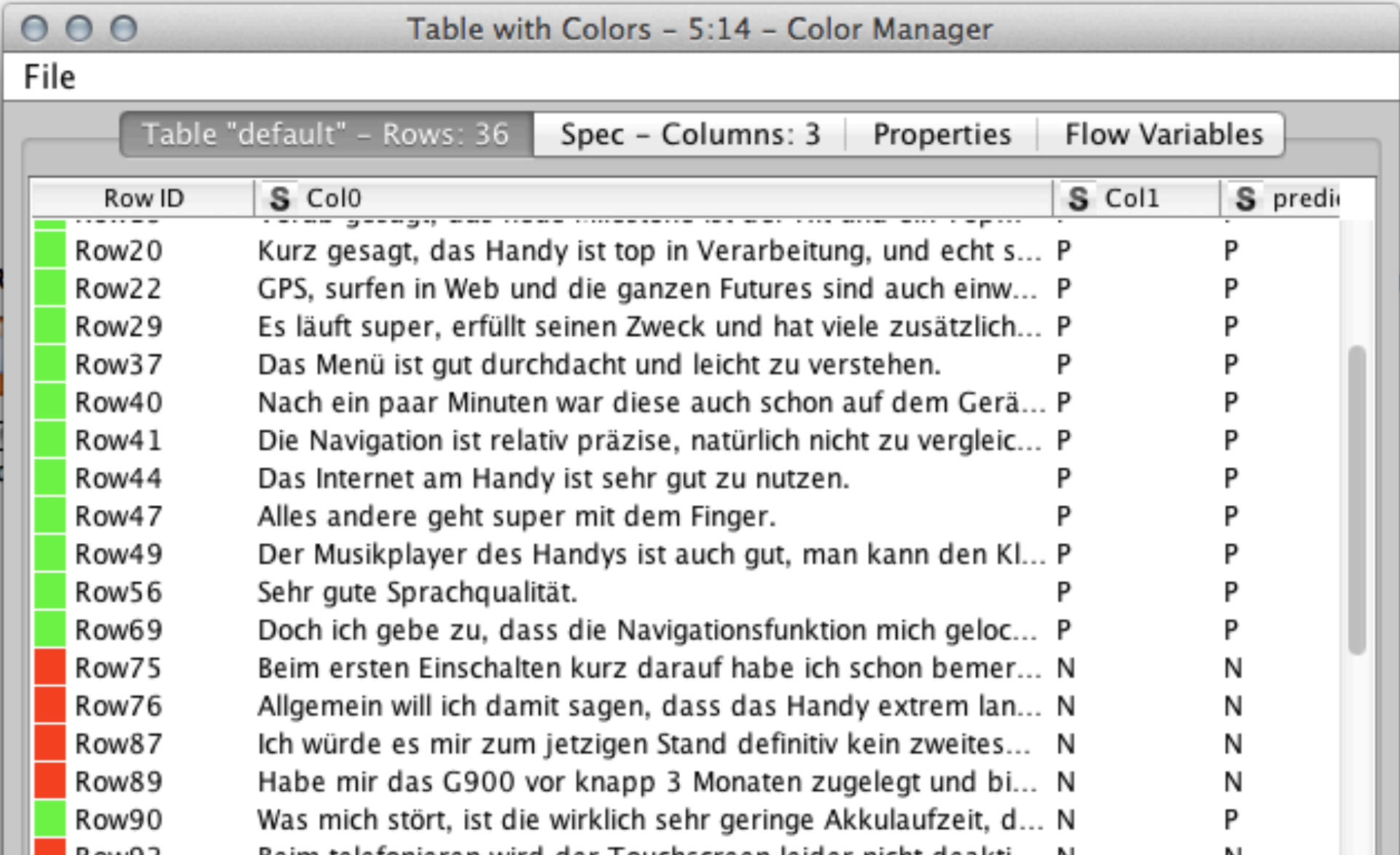
Examples

- Palladian text classifier for sentiment analysis



Examples

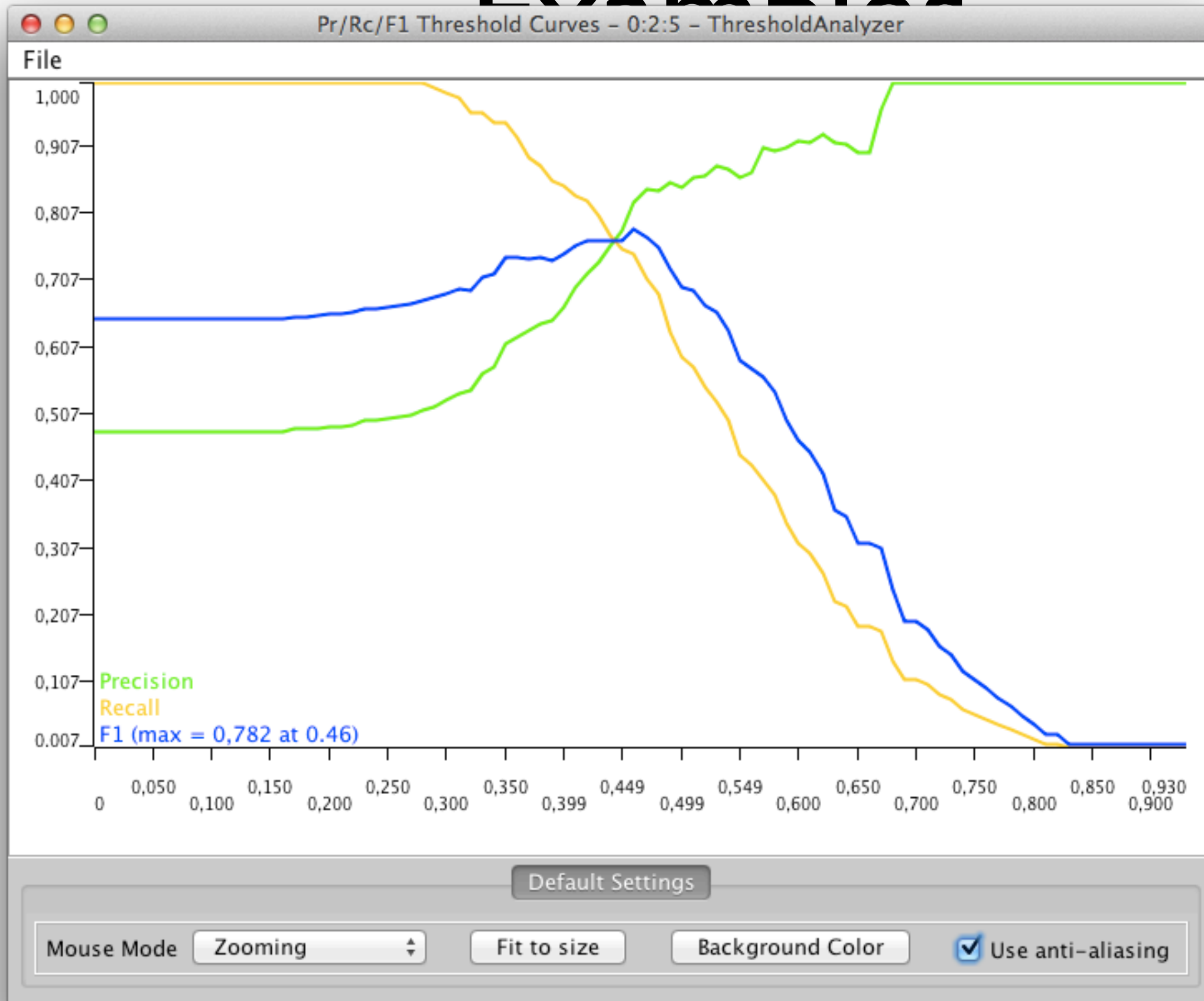
- Palladian text classifier for sentiment analysis



The screenshot shows a software window titled "Table with Colors - 5:14 - Color Manager". Inside, there's a tabbed interface with "Table 'default' - Rows: 36" selected. The table has four columns: "Row ID", "S Col0", "S Col1", and "S predi". The rows are color-coded: green for positive sentiment and red for negative sentiment. The text in the "S Col0" column represents user reviews of a mobile phone, and the "S predi" column shows the classifier's prediction.

Row ID	S Col0	S Col1	S predi
Row20	Kurz gesagt, das Handy ist top in Verarbeitung, und echt s...	P	P
Row22	GPS, surfen in Web und die ganzen Futures sind auch einw...	P	P
Row29	Es läuft super, erfüllt seinen Zweck und hat viele zusätzlic...	P	P
Row37	Das Menü ist gut durchdacht und leicht zu verstehen.	P	P
Row40	Nach ein paar Minuten war diese auch schon auf dem Gerä...	P	P
Row41	Die Navigation ist relativ präzise, natürlich nicht zu verglic...	P	P
Row44	Das Internet am Handy ist sehr gut zu nutzen.	P	P
Row47	Alles andere geht super mit dem Finger.	P	P
Row49	Der Musikplayer des Handys ist auch gut, man kann den Kl...	P	P
Row56	Sehr gute Sprachqualität.	P	P
Row69	Doch ich gebe zu, dass die Navigationsfunktion mich geloc...	P	P
Row75	Beim ersten Einschalten kurz darauf habe ich schon bemer...	N	N
Row76	Allgemein will ich damit sagen, dass das Handy extrem lan...	N	N
Row87	Ich würde es mir zum jetzigen Stand definitiv kein zweites...	N	N
Row89	Habe mir das G900 vor knapp 3 Monaten zugelegt und bi...	N	N
Row90	Was mich stört, ist die wirklich sehr geringe Akkulaufzeit, d...	N	P
Row92	Beim telefonieren wird der Touchscreen leider nicht deakti...	N	N

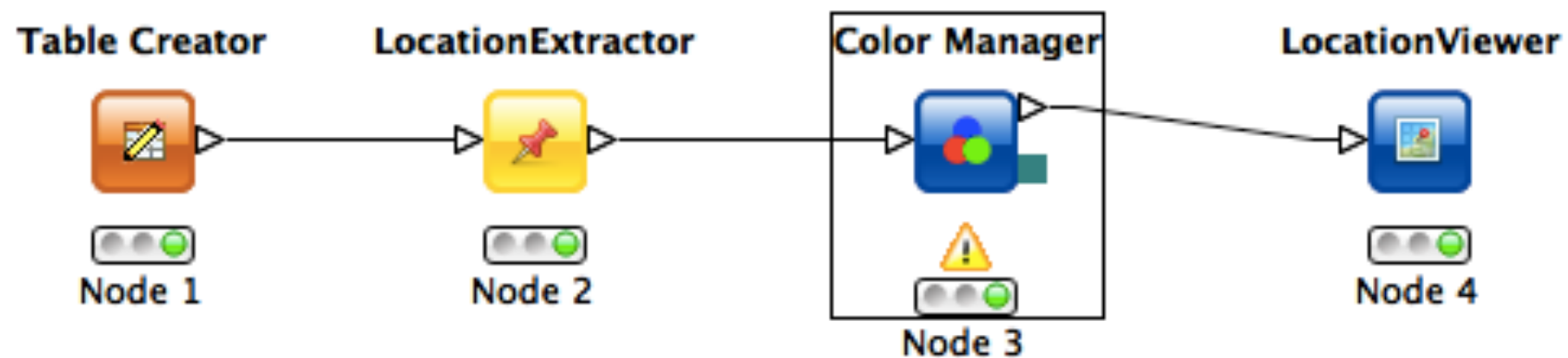
Example





Examples

- Extract and disambiguate locations, visualize them on the map



Examples

- Extract and disambiguate locations, visualize them on the map

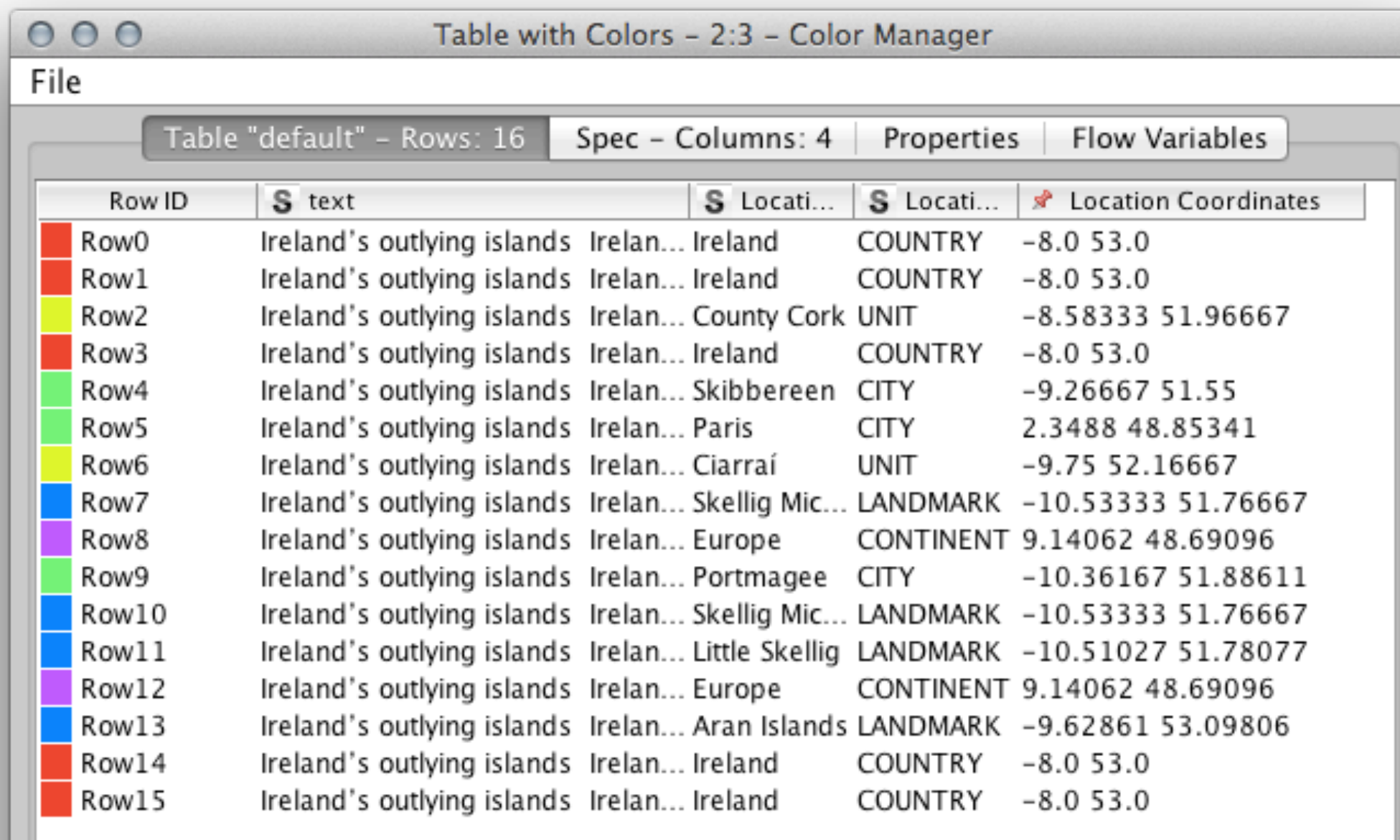


Table with Colors - 2:3 - Color Manager

File

Table "default" - Rows: 16 | Spec - Columns: 4 | Properties | Flow Variables

Row ID	S text	S Locati...	S Locati...	Location Coordinates
Row0	Ireland's outlying islands	Irelan...	Ireland	COUNTRY -8.0 53.0
Row1	Ireland's outlying islands	Irelan...	Ireland	COUNTRY -8.0 53.0
Row2	Ireland's outlying islands	Irelan...	County Cork	UNIT -8.58333 51.96667
Row3	Ireland's outlying islands	Irelan...	Ireland	COUNTRY -8.0 53.0
Row4	Ireland's outlying islands	Irelan...	Skibbereen	CITY -9.26667 51.55
Row5	Ireland's outlying islands	Irelan...	Paris	CITY 2.3488 48.85341
Row6	Ireland's outlying islands	Irelan...	Ciarraí	UNIT -9.75 52.16667
Row7	Ireland's outlying islands	Irelan...	Skellig Mic...	LANDMARK -10.53333 51.76667
Row8	Ireland's outlying islands	Irelan...	Europe	CONTINENT 9.14062 48.69096
Row9	Ireland's outlying islands	Irelan...	Portmagee	CITY -10.36167 51.88611
Row10	Ireland's outlying islands	Irelan...	Skellig Mic...	LANDMARK -10.53333 51.76667
Row11	Ireland's outlying islands	Irelan...	Little Skellig	LANDMARK -10.51027 51.78077
Row12	Ireland's outlying islands	Irelan...	Europe	CONTINENT 9.14062 48.69096
Row13	Ireland's outlying islands	Irelan...	Aran Islands	LANDMARK -9.62861 53.09806
Row14	Ireland's outlying islands	Irelan...	Ireland	COUNTRY -8.0 53.0
Row15	Ireland's outlying islands	Irelan...	Ireland	COUNTRY -8.0 53.0

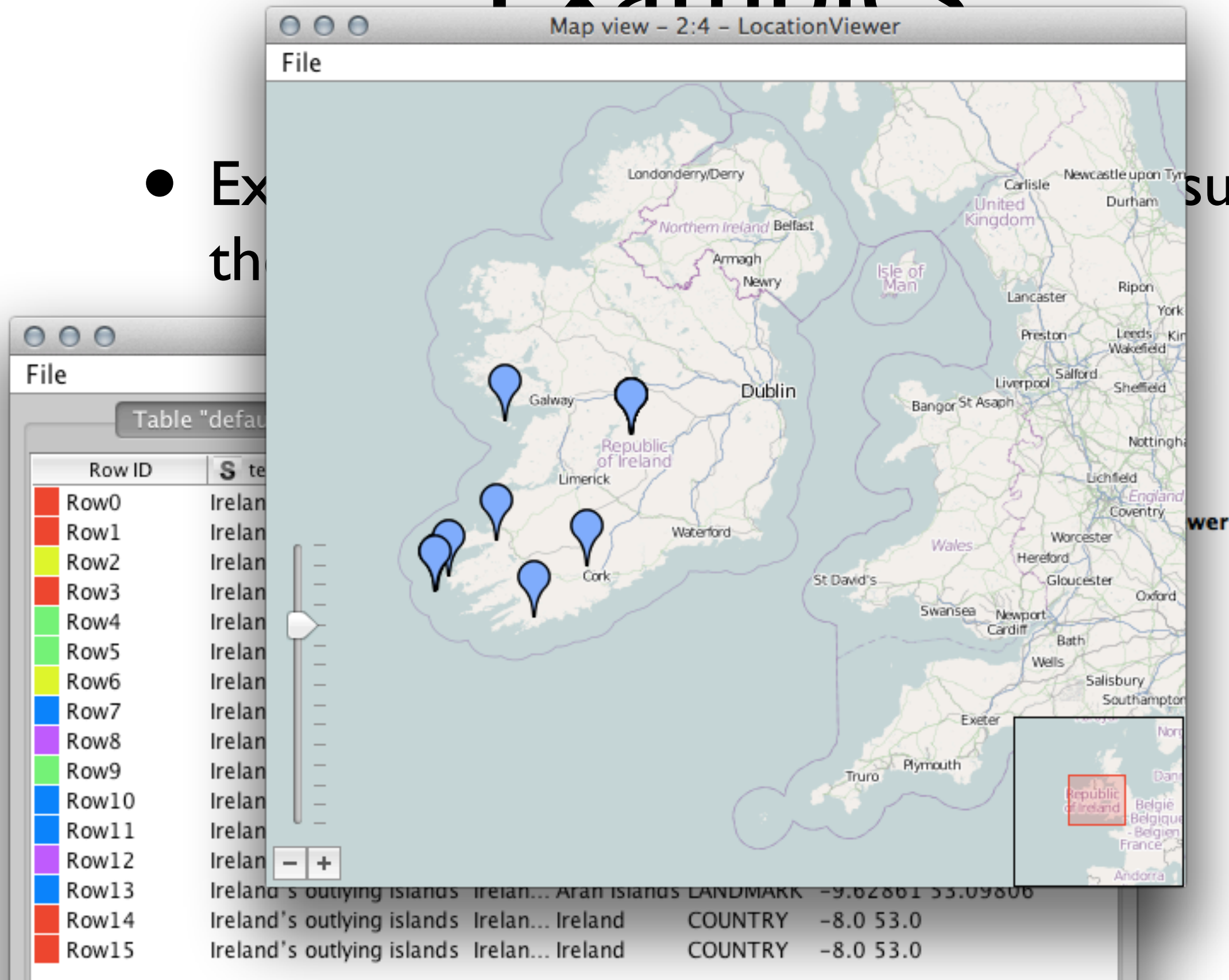
nViewer

de 4

Examples

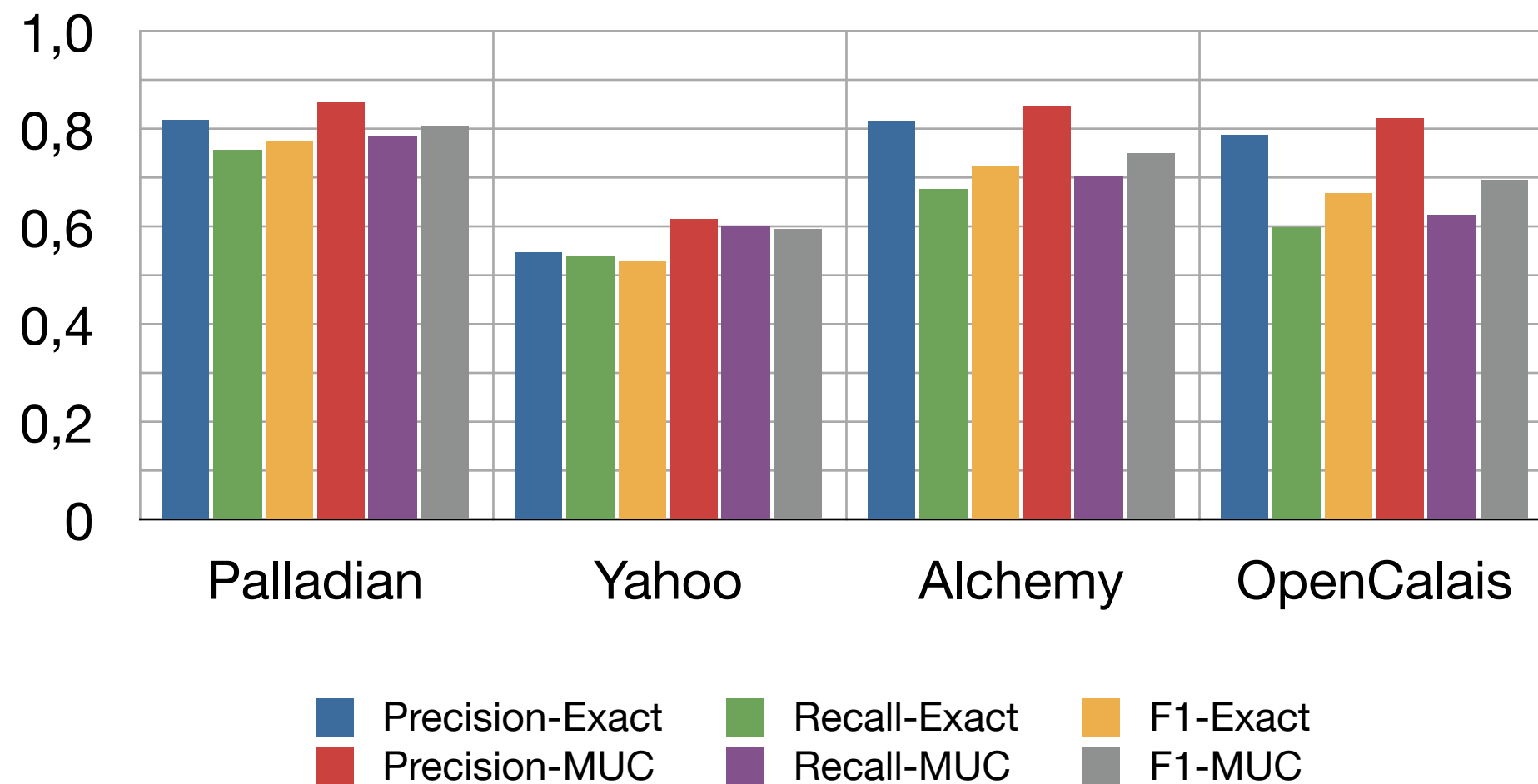
- Example 1

Visualize



Location Extraction

- Preliminary evaluation results



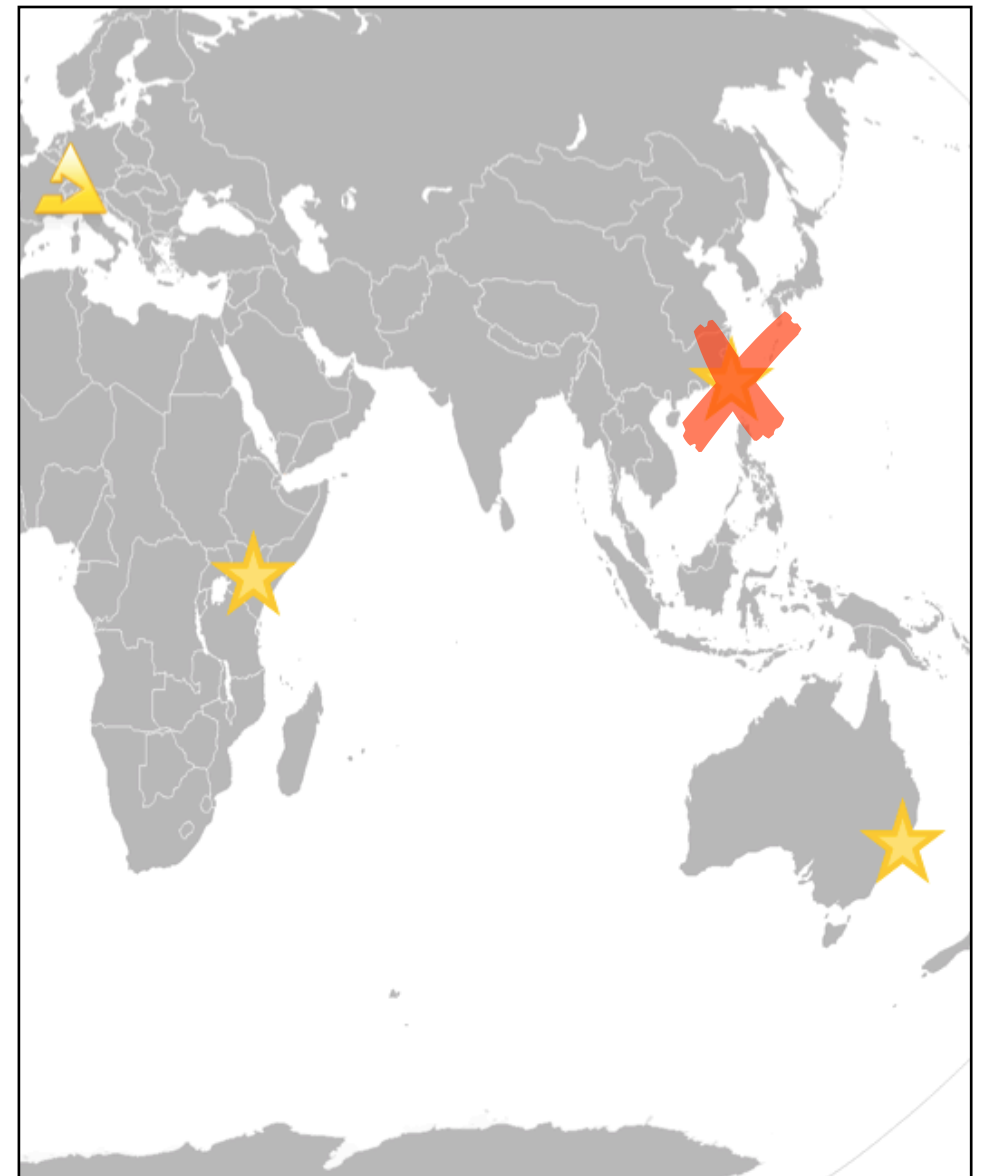
Tomorrow's Workshop

- Learn about
 - HTTP, feed retrieval
 - Text classification for sentiment analysis
 - Location extraction



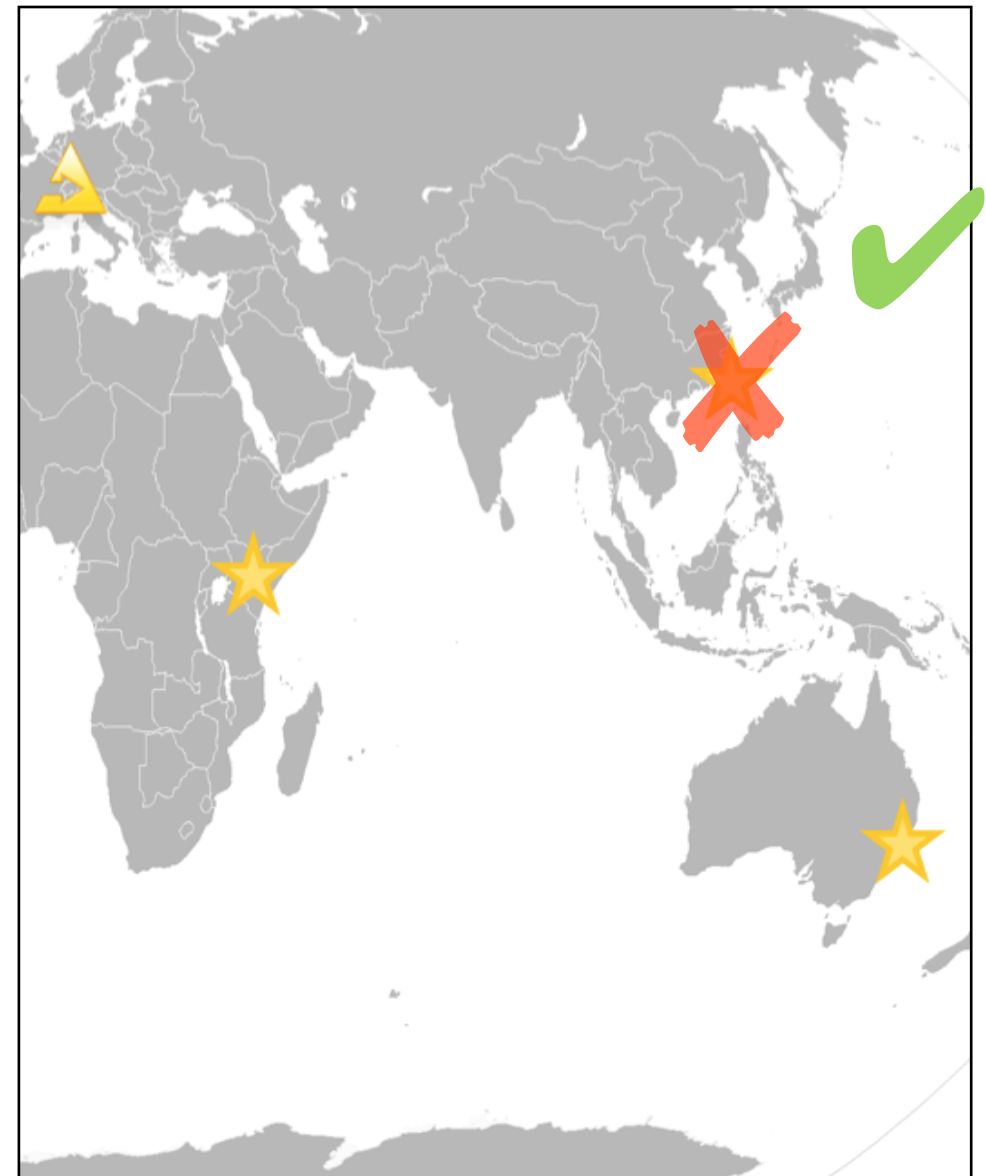
Tomorrow's Workshop

- Learn about
 - HTTP, feed retrieval
 - Text classification for sentiment analysis
 - Location extraction



Tomorrow's Workshop

- Learn about
 - HTTP, feed retrieval
 - Text classification for sentiment analysis
 - Location extraction





Thank you!

Any questions?

<http://palladian.ws>
mail@palladian.ws

The Palladian nodes are available via
KNIME community contributions.