



TECHNISCHE
UNIVERSITÄT
WIEN

Vienna University of Technology

CLEF-IP 2012: Retrieval in the Intellectual Property Domain

Florina Piroi, Mihai Lupu, Allan Hanbury ⁽¹⁾
Walid Magdy ⁽²⁾, Alan Sexton ⁽³⁾, Igor Filippov ⁽⁴⁾

IFS, Vienna University of Technology, Austria ⁽¹⁾

QCRI, Qatar Foundation, Doha, Qatar ⁽²⁾

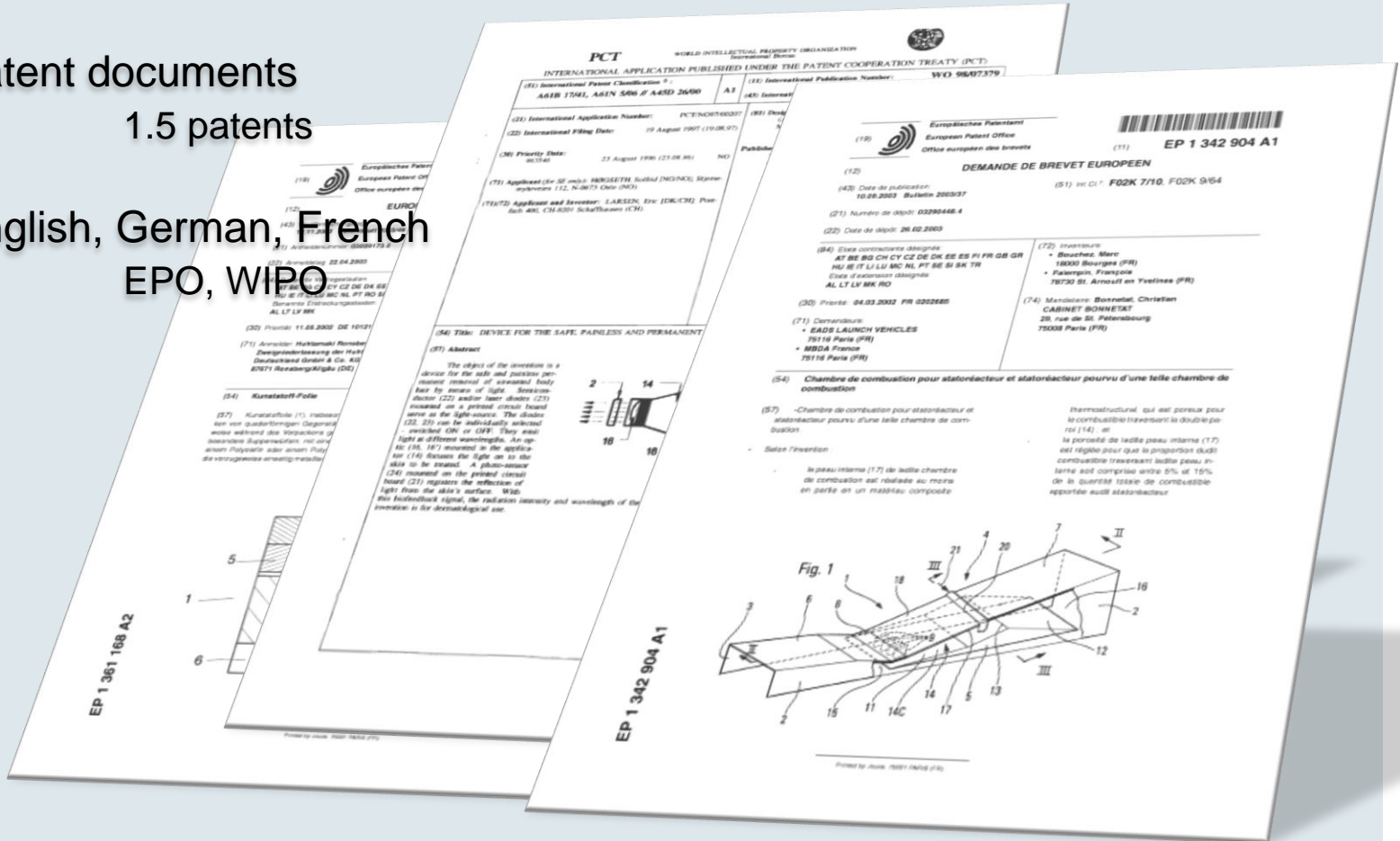
School of CS, University of Birmingham, UK ⁽³⁾

SAIC-Frederick Inc., USA ⁽⁴⁾

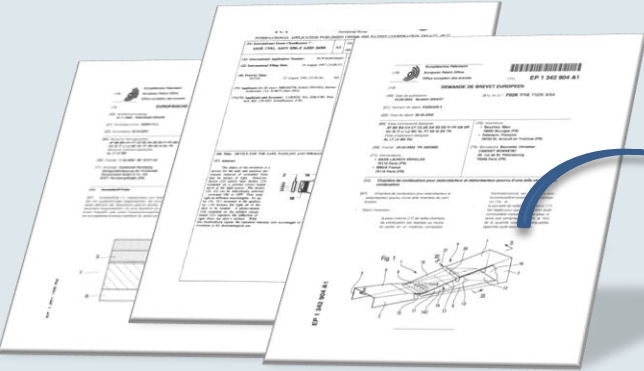
CLEF-IP Collection Content

Patent documents
1.5 patents

English, German, French
EPO, WIPO



CLEF-IP Collection Content

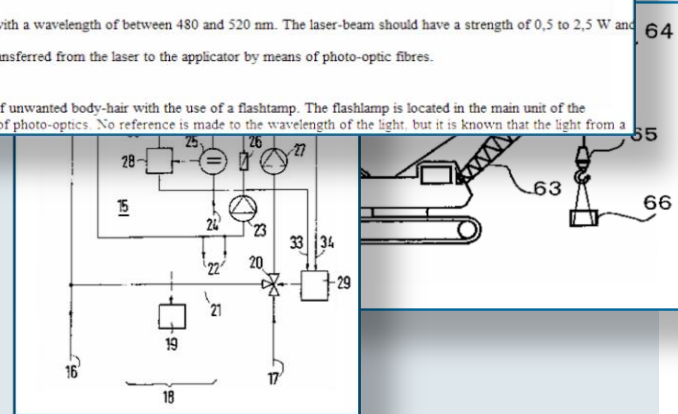
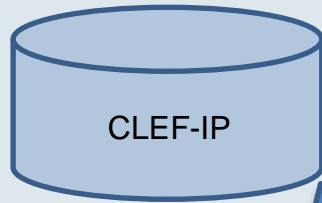


```

<patent-document uid="WO-1998007379-A1" country="WO" doc-number="1998007379" kind="A1" lang="EN" family-id="19899747" status="new"
date-produced="20090508" date="19980226">
  <bibliographic-data>< bibliographic-data>
  <abstract load-source="docdb" source="national office" status="new" lang="EN">
  <p>
    The object of the invention is a device for the safe and painless permanent removal of unwanted body hair by means of light. Semiconductor (22) and/or laser diodes (23) mounted on a printed circuit board serve as the light-source. The diodes (22, 23) can be individually selected - switched ON or OFF. They emit light at different wavelengths. An optic (16, 16') mounted in the applicator (14) focuses the light on to the skin to be treated. A photo-sensor (24) mounted on the printed circuit board (21) registers the reflection of light from the skin's surface. With this biofeedback signal, the radiation intensity and wavelength of the light-source (20) can be semi- or fully automatically adjusted. The invention is for dermatological use.
  </p>
  </abstract>
  <abstract load-source="docdb" source="national office" status="new" lang="FR"></abstract>
  <description load-source="pctxml" status="new" lang="EN">
  <!-- EPO <DP n="3"/>-->
  <p num="p0001">
    Device for the safe, painless and permanent removal of unwanted hairgrowth.
  </p>
  <p num="p0002">Background of the invention:</p>
  <p num="p0003">
    The invention concerns devices for the safe, painless and permanent removal of unwanted hairgrowth in accordance with the specification in claim 1.
  </p>
  <p num="p0004">State of the Art.</p>
  <p num="p0005">
    The removal of unwanted hairgrowth by means of light is already known.
  </p>
  <p num="p0006">
    U.S. Patent No. 5.226.907 proposes the use of a CO2 laser with a wavelength of ca. 10.000 nm, or a pulsed dye laser with wavelengths of 531 , 584, 587 or 632 nm. The applicator covers a treatment area of ca. 1 cm2, and the radiation intensity is from 5 to 10 J/cm2. Prior to treatment, however, the skin must be treated with a photosensitizer, such as a suspension of carbon in oil, which must penetrate through to the hair-roots via the pores.
  </p>
  <p num="p0007">
    WO 95/15725 proposes, for the same purpose, a ruby laser with a wavelength of 694 nm, a neodyme-YAG laser with a wavelength of 1064 nm, or the use of any other laser with a wavelength of between 600 and 1500 nm. The pulse-length should be between 1 µsec. and 1 millisecc.
  </p>
  <p num="p0008">
    U.S. Patent No. 5.059.192 describes the use of a ruby laser which emits a light-pulse of 30-40 nsec, treating an area of ca. 3 mm. with an intensity of 0,4 to 10 J/cm2. The irradiation time should be under 2 msec.
  </p>
  <p num="p0009">
    U.S. Patent No. 4.388.924 proposes the use of an argon laser with a wavelength of between 480 and 520 nm. The laser-beam should have a strength of 0,5 to 2,5 W and the treatment time be between 10 and 400 msec. The light is transferred from the laser to the applicator by means of photo-optic fibres.
  </p>
  <p num="p0010">
    U.S. Patent No. 4.608.978 describes a device for the removal of unwanted body-hair with the use of a flashtamp. The flashtamp is located in the main unit of the apparatus, the light being conveyed to the applicator by means of photo-optics. No reference is made to the wavelength of the light, but it is known that the light from a
  </p>
  </description>
  </patent-document>
  
```

3.5 million XML files

Multilingual content



CLEF-IP Tasks

Prior Art Retrieval (2009, 2010, 2011)

Patent Classification (IPC) (2010, 2011)

Image-based Patent Retrieval (2011)

Patent Image Classification (2011)

CLEF-IP Tasks

Prior Art Retrieval (2009, 2010, 2011)

Patent Classification (IPC) (2010, 2011)

Image-based Patent Retrieval (2011)

Patent Image Classification (2011)

Passage Retrieval based on Claims

Flow-chart Recognition

Chemical Structure Recognition

Passage Retrieval Based on Claims

Patentansprüche

1. Kunststoffolie(1.21), insbesondere zum Verpacken von quaderförmigen Gegenständen, die vorzugsweise während des Verpackens gefaltet werden, wie besonders Suppenwürfel, mit einer (2.22) aus einem Polyolefin oder einem Ienitrophtalat, die vorzugsweise einseitig sind, **dadurch gekennzeichnet, daß** folie(2.22) eine orientierte Polypropylen zugewisse eine biaxial orientierte Polypropylen oder eine PET-Folie vorgesehen ist, die Metallisierung(3) tragender Seite eintrag(5) vorgesehen ist, während auf der valseitigung(3) abgewandten Seite der (2.22) ein Kaltkleber(6) flächenartig, vorbandförmig, aufgetragen ist.

2. Kunststoffolie nach Anspruch 1, **dadurch zeichnet, daß** der Druckauftrag(5) mit conartig, auftragragmen Relieslack(7) abgedeckt ist, wobei die Auftragflächen des Kaltklebers(6) abgestimmt sind.

3. Kunststoffolie nach Anspruch 1 oder 2 **gekennzeichnet, daß** zwischen der Met(3) und dem Druckauftrag(5) eine Pfl Halftvermittler(4) vorgesehen ist.

4. Kunststoffolie nach Anspruch 1, 2 oder 3 **gekennzeichnet, daß** als Basisfolie(2) pakte Folie vorgesehen ist.

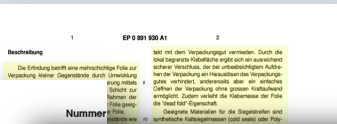
5. Kunststoffolie nach Anspruch 1, 2 oder 3 **gekennzeichnet, daß** als Basisfolie(2)



Europäisches Patentamt

EUROPÄISCHER RECHERCHENBERICHT

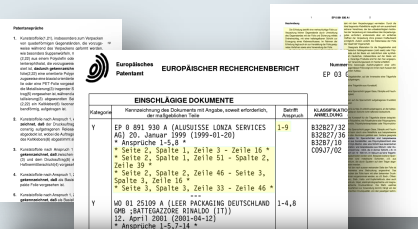
EINSCHLÄGIGE DOKUMENTE		
Kategorie	Kenzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrifft Anspruch
Y	EP 0 891 930 A (ALUSUISSE LONZA SERVICES AG) 20. Januar 1999 (1999-01-20) * Ansprüche 1-5, 8 * * Seite 2, Spalte 1, Zeile 3 - Zeile 16 * * Seite 2, Spalte 1, Zeile 51 - Spalte 2, Zeile 39 * * Seite 2, Spalte 2, Zeile 46 - Seite 3, Spalte 3, Zeile 16 * * Seite 3, Spalte 3, Zeile 33 - Zeile 46 *	1-9
Y	WO 01 25109 A (LEER PACKAGING DEUTSCHLAND GMB ;BATTEGAZZORE RINALDO (IT)) 12. April 2001 (2001-04-12) * Ansprüche 1-5, 7-14 * * Seite 4, Zeile 10 - Zeile 20 * * Seite 5, Zeile 1 - Seite 6, Zeile 3 *	1-4, 8
Y	EP 0 870 695 A (PECHINEE EMBALLAGE FLEXIBLE FU) 14. Oktober 1998 (1998-10-14)	1-9



- PSG2, EP-1172823-A1, /patent-document/claims/claim[6]
- PSG2, EP-1172823-A1, /patent-document/claims/claim[7]
- PSG2, EP-1172823-A1, /patent-document/claims/claim[8]
- PSG2, EP-1172823-A1, /patent-document/claims/claim[9]
- PSG2, EP-1172823-A1, /patent-document/description/heading[5]
- PSG2, EP-1172823-A1, /patent-document/description/p[78]
- PSG2, WO-2000031183-A1, /patent-document/claims/claim
- PSG2, WO-2000031183-A1, /patent-document/description/p[52]
- PSG2, WO-2000031183-A1, /patent-document/description/p[53]
- PSG2, WO-2000031183-A1, /patent-document/description/p[54]
- PSG2, WO-2000031183-A1, /patent-document/description/p[55]
- PSG2, WO-2000031183-A1, /patent-document/description/p[56]
- PSG2, WO-2000031183-A1, /patent-document/description/p[57]
- PSG2, WO-2000031183-A1, /patent-document/description/p[58]
- PSG2, WO-2000031183-A1, /patent-document/description/p[59]
- PSG2, WO-2000061834-A1, /patent-document/claims/claim
- PSG2, WO-2000061834-A1, /patent-document/description/p[151]
- PSG2, WO-2000061834-A1, /patent-document/description/p[152]
- PSG2, WO-2000061834-A1, /patent-document/description/p[153]
- PSG2, WO-2000061834-A1, /patent-document/description/p[154]
- PSG2, WO-2000061834-A1, /patent-document/description/p[155]
- PSG2, WO-2000061834-A1, /patent-document/description/p[156]
- PSG2, WO-2000061834-A1, /patent-document/description/p[157]
- PSG2, WO-2000061834-A1, /patent-document/description/p[158]
- PSG2, WO-2000061834-A1, /patent-document/description/p[159]
- PSG2, WO-2000061834-A1, /patent-document/description/p[160]
- PSG2, WO-2000061834-A1, /patent-document/description/p[161]
- PSG2, WO-2000061834-A1, /patent-document/description/p[162]
- PSG2, WO-2000061834-A1, /patent-document/description/p[163]
- PSG2, WO-2000061834-A1, /patent-document/description/p[88]
- PSG2, WO-2000061834-A1, /patent-document/description/p[89]
- PSG3, EP-0967251-A1, /patent-document/abstract[1]/p
- PSG3, EP-0967251-A1, /patent-document/abstract[2]/p
- PSG3, EP-0967251-A1, /patent-document/claims/claim[1]
- PSG3, EP-0967251-A1, /patent-document/claims/claim[2]
- PSG3, EP-0967251-A1, /patent-document/claims/claim[3]

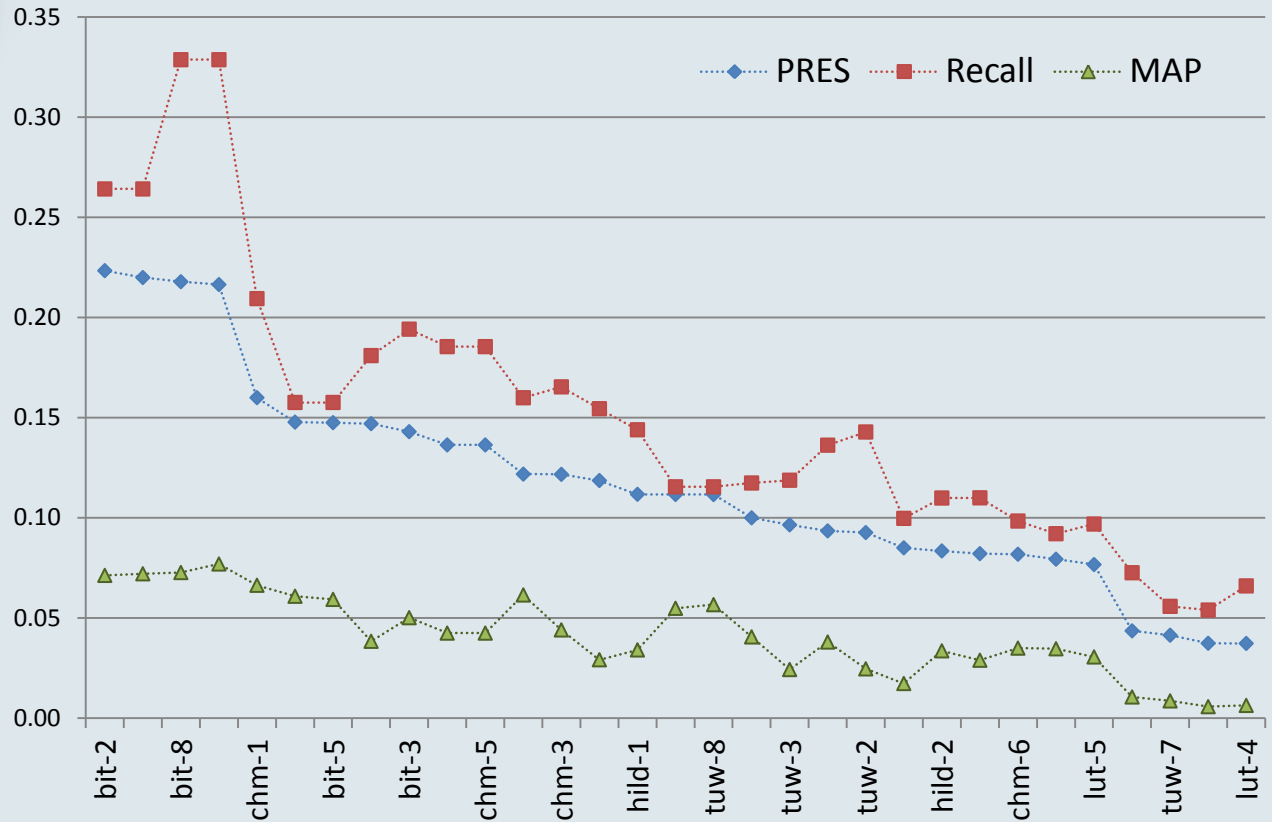
105 topics (3 x 35)
XPaths in qrels

Passage Retrieval Based on Claims



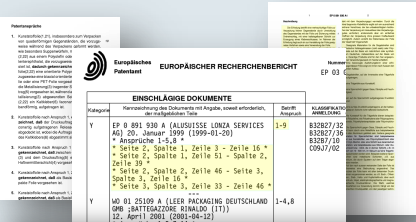
```

PSG2,EP-1172823-A1,/patent-document/claims/claim[6]
PSG2,EP-1172823-A1,/patent-document/claims/claim[7]
PSG2,EP-1172823-A1,/patent-document/claims/claim[8]
PSG2,EP-1172823-A1,/patent-document/claims/claim[9]
PSG2,EP-1172823-A1,/patent-document/description/heading[5]
PSG2,EP-1172823-A1,/patent-document/description/p[78]
PSG2,WO-2000031183-A1,/patent-document/claims/claim
PSG2,WO-2000031183-A1,/patent-document/description/p[52]
PSG2,WO-2000031183-A1,/patent-document/description/p[53]
PSG2,WO-2000031183-A1,/patent-document/description/p[54]
PSG2,WO-2000031183-A1,/patent-document/description/p[55]
PSG2,WO-2000031183-A1,/patent-document/description/p[56]
PSG2,WO-2000031183-A1,/patent-document/description/p[57]
PSG2,WO-2000031183-A1,/patent-document/description/p[58]
PSG2,WO-2000031183-A1,/patent-document/description/p[59]
PSG2,WO-2000061834-A1,/patent-document/claims/claim
PSG2,WO-2000061834-A1,/patent-document/description/p[151]
PSG2,WO-2000061834-A1,/patent-document/description/p[152]
PSG2,WO-2000061834-A1,/patent-document/description/p[153]
PSG2,WO-2000061834-A1,/patent-document/description/p[154]
PSG2,WO-2000061834-A1,/patent-document/description/p[155]
PSG2,WO-2000061834-A1,/patent-document/description/p[156]
PSG2,WO-2000061834-A1,/patent-document/description/p[157]
PSG2,WO-2000061834-A1,/patent-document/description/p[158]
PSG2,WO-2000061834-A1,/patent-document/description/p[159]
PSG2,WO-2000061834-A1,/patent-document/description/p[160]
PSG2,WO-2000061834-A1,/patent-document/description/p[161]
PSG2,WO-2000061834-A1,/patent-document/description/p[162]
PSG2,WO-2000061834-A1,/patent-document/description/p[163]
PSG2,WO-2000061834-A1,/patent-document/description/p[88]
PSG2,WO-2000061834-A1,/patent-document/description/p[89]
PSG3,EP-0967251-A1,/patent-document/abstract[1]/p
PSG3,EP-0967251-A1,/patent-document/abstract[2]/p
PSG3,EP-0967251-A1,/patent-document/claims/claim[1]
PSG3,EP-0967251-A1,/patent-document/claims/claim[2]
PSG3,EP-0967251-A1,/patent-document/claims/claim[3]
    
```



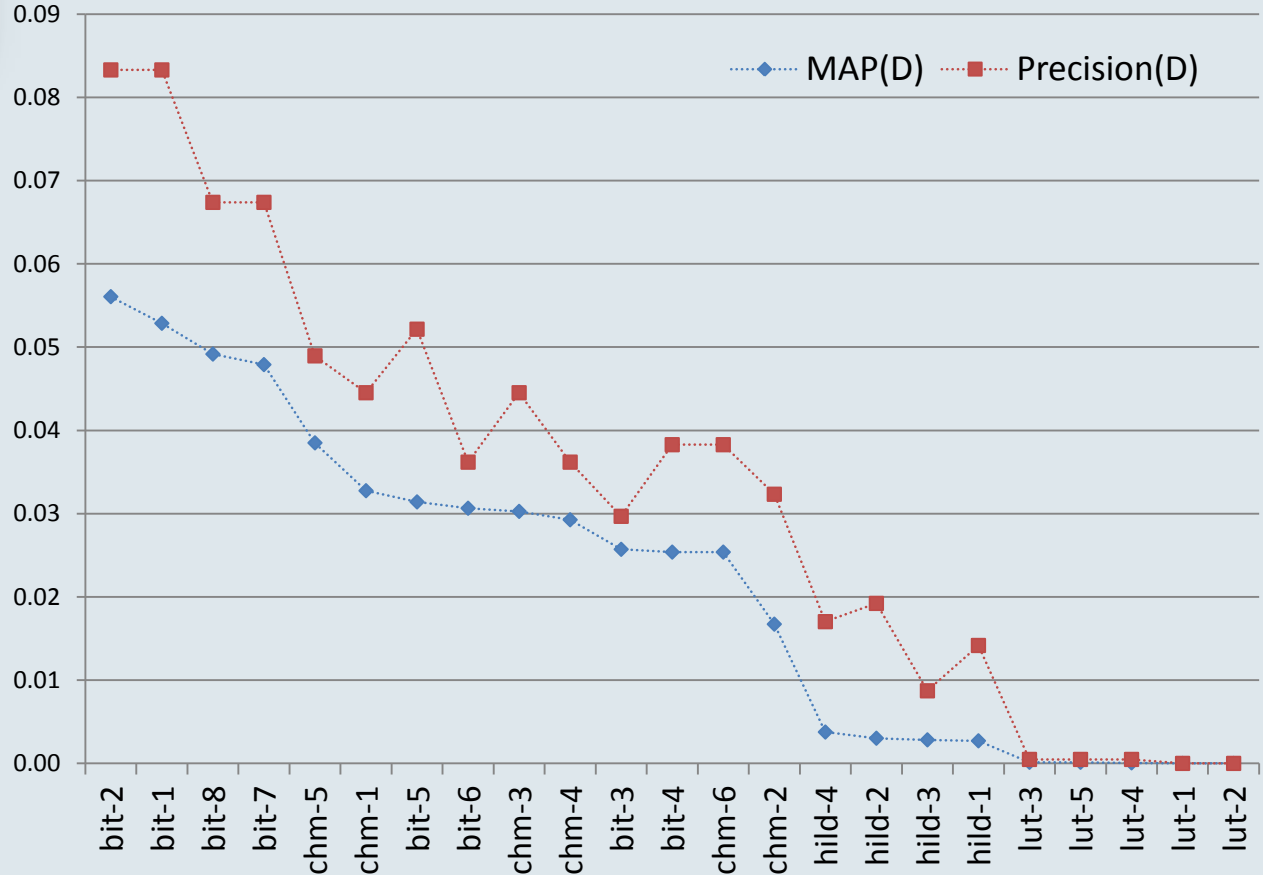
PRES
Recall
MAP

Passage Retrieval Based on Claims



```

FSG2,EP-1172823-A1,/patent-document/claims/claim[6]
FSG2,EP-1172823-A1,/patent-document/claims/claim[7]
FSG2,EP-1172823-A1,/patent-document/claims/claim[8]
FSG2,EP-1172823-A1,/patent-document/claims/claim[9]
FSG2,EP-1172823-A1,/patent-document/description/heading[5]
FSG2,EP-1172823-A1,/patent-document/description/p[78]
FSG2,WO-2000031183-A1,/patent-document/claims/claim
FSG2,WO-2000031183-A1,/patent-document/description/p[52]
FSG2,WO-2000031183-A1,/patent-document/description/p[53]
FSG2,WO-2000031183-A1,/patent-document/description/p[54]
FSG2,WO-2000031183-A1,/patent-document/description/p[55]
FSG2,WO-2000031183-A1,/patent-document/description/p[56]
FSG2,WO-2000031183-A1,/patent-document/description/p[57]
FSG2,WO-2000031183-A1,/patent-document/description/p[58]
FSG2,WO-2000031183-A1,/patent-document/description/p[59]
FSG2,WO-2000061834-A1,/patent-document/claims/claim
FSG2,WO-2000061834-A1,/patent-document/description/p[151]
FSG2,WO-2000061834-A1,/patent-document/description/p[152]
FSG2,WO-2000061834-A1,/patent-document/description/p[153]
FSG2,WO-2000061834-A1,/patent-document/description/p[154]
FSG2,WO-2000061834-A1,/patent-document/description/p[155]
FSG2,WO-2000061834-A1,/patent-document/description/p[156]
FSG2,WO-2000061834-A1,/patent-document/description/p[157]
FSG2,WO-2000061834-A1,/patent-document/description/p[158]
FSG2,WO-2000061834-A1,/patent-document/description/p[159]
FSG2,WO-2000061834-A1,/patent-document/description/p[160]
FSG2,WO-2000061834-A1,/patent-document/description/p[161]
FSG2,WO-2000061834-A1,/patent-document/description/p[162]
FSG2,WO-2000061834-A1,/patent-document/description/p[163]
FSG2,WO-2000061834-A1,/patent-document/description/p[88]
FSG2,WO-2000061834-A1,/patent-document/description/p[89]
FSG3,EP-0967251-A1,/patent-document/abstract[1]/p
FSG3,EP-0967251-A1,/patent-document/abstract[2]/p
FSG3,EP-0967251-A1,/patent-document/claims/claim[1]
FSG3,EP-0967251-A1,/patent-document/claims/claim[2]
FSG3,EP-0967251-A1,/patent-document/claims/claim[3]
    
```



Precision(D)
MAP(D)

Flow-chart Recognition

FIG.8

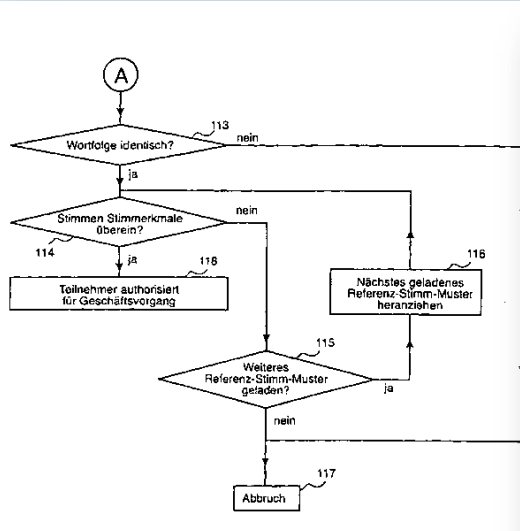
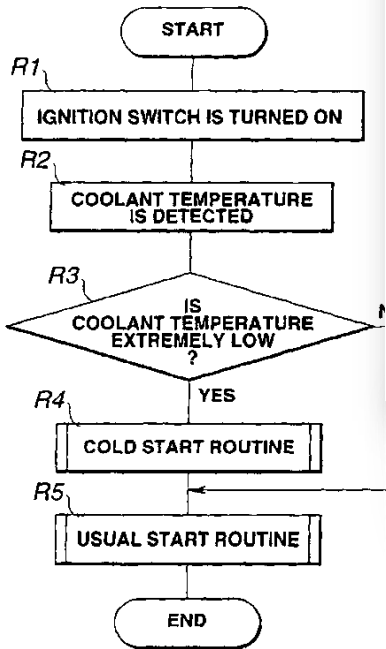
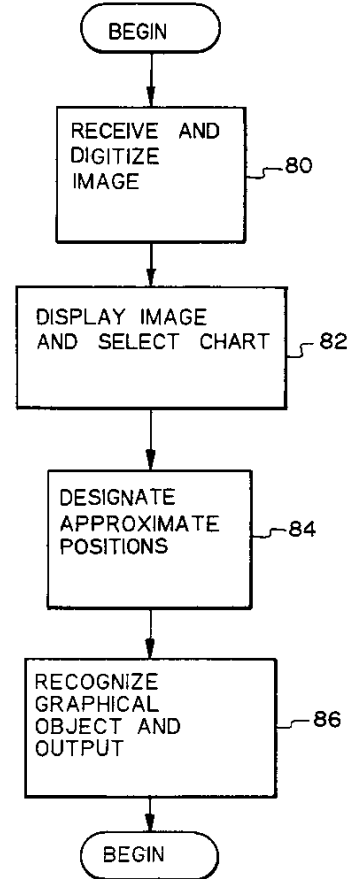


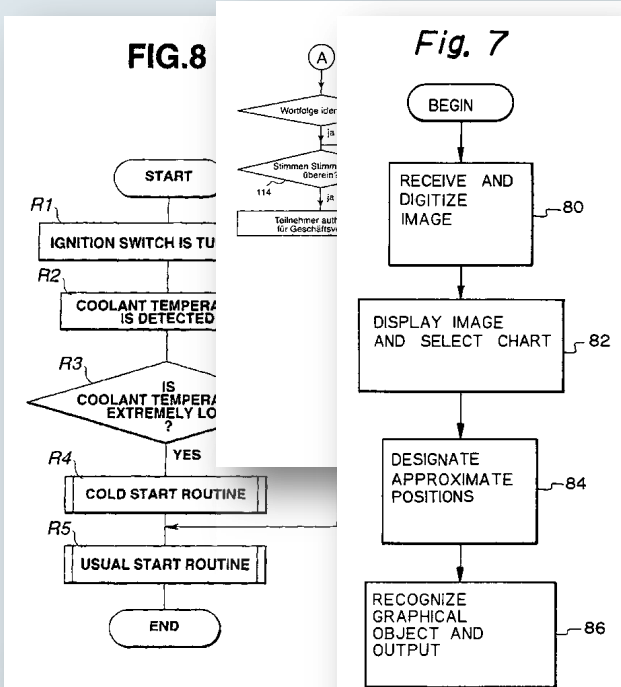
Fig. 2b

Fig. 7



Make content searchable

Flow-chart Recognition



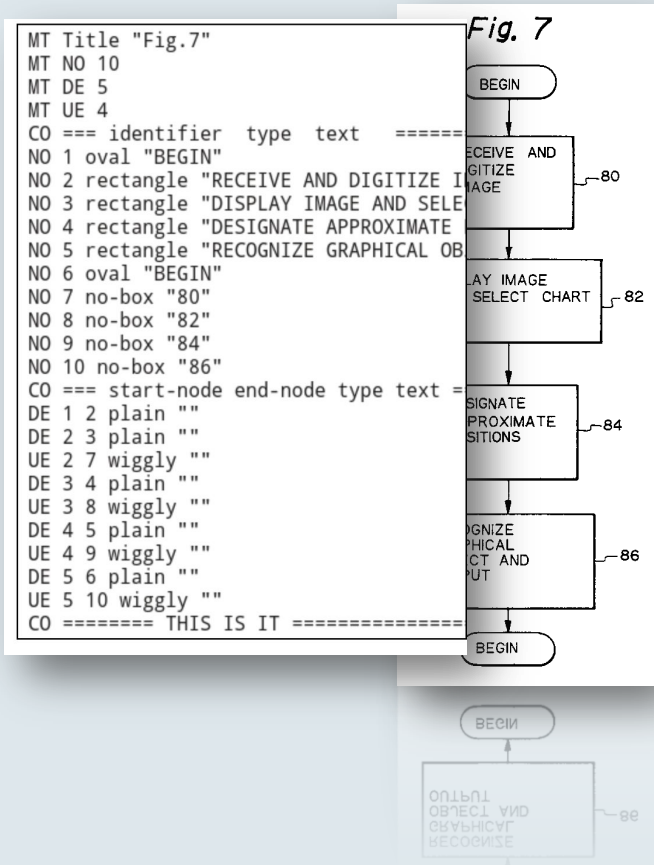
```

MT Title "Fig.7"
MT NO 10
MT DE 5
MT UE 4
CO === identifier type text =====
NO 1 oval "BEGIN"
NO 2 rectangle "RECEIVE AND DIGITIZE I
NO 3 rectangle "DISPLAY IMAGE AND SELE
NO 4 rectangle "DESIGNATE APPROXIMATE
NO 5 rectangle "RECOGNIZE GRAPHICAL OB
NO 6 oval "BEGIN"
NO 7 no-box "80"
NO 8 no-box "82"
NO 9 no-box "84"
NO 10 no-box "86"
CO === start-node end-node type text =
DE 1 2 plain ""
DE 2 3 plain ""
UE 2 7 wiggly ""
DE 3 4 plain ""
UE 3 8 wiggly ""
DE 4 5 plain ""
UE 4 9 wiggly ""
DE 5 6 plain ""
UE 5 10 wiggly ""
CO ===== THIS IS IT =====
  
```

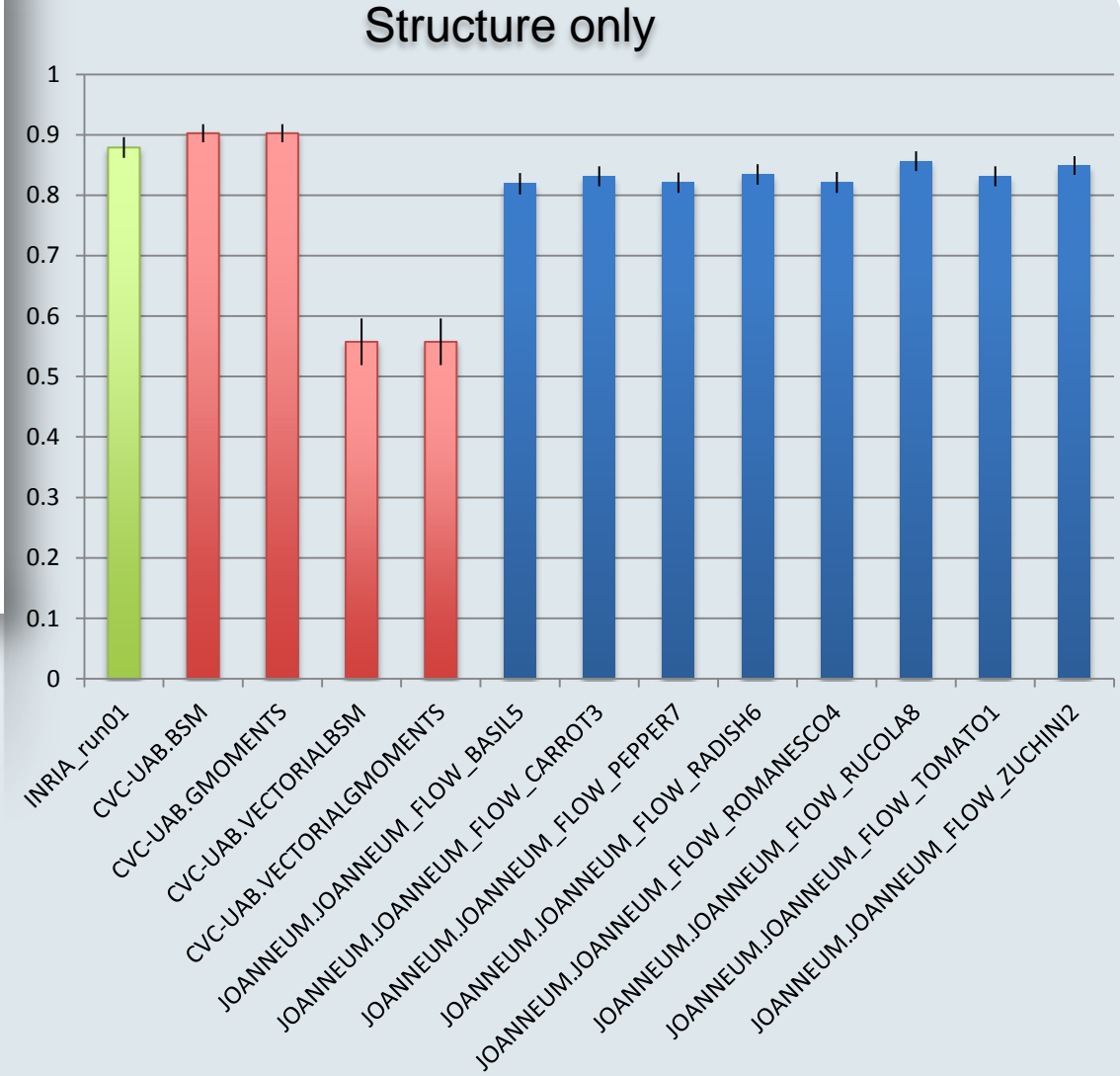
100 topics

Graph distance measure

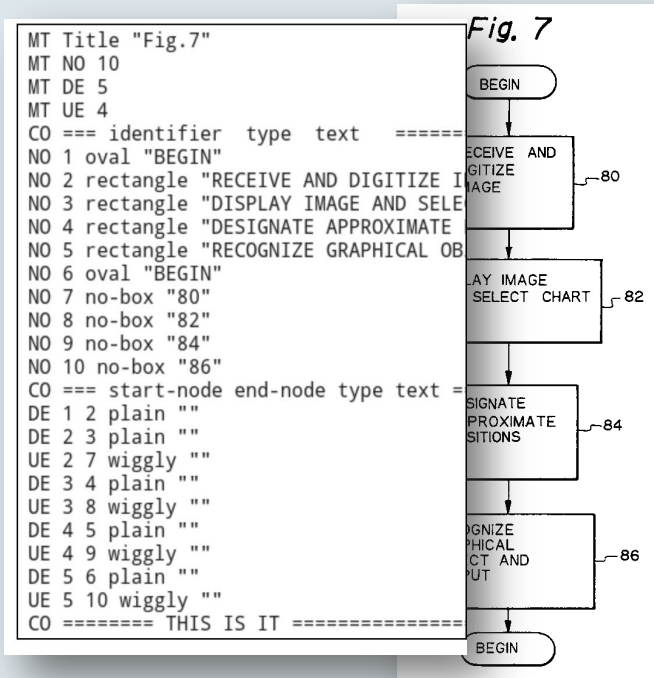
Flow-chart Recognition



40 topics evaluated

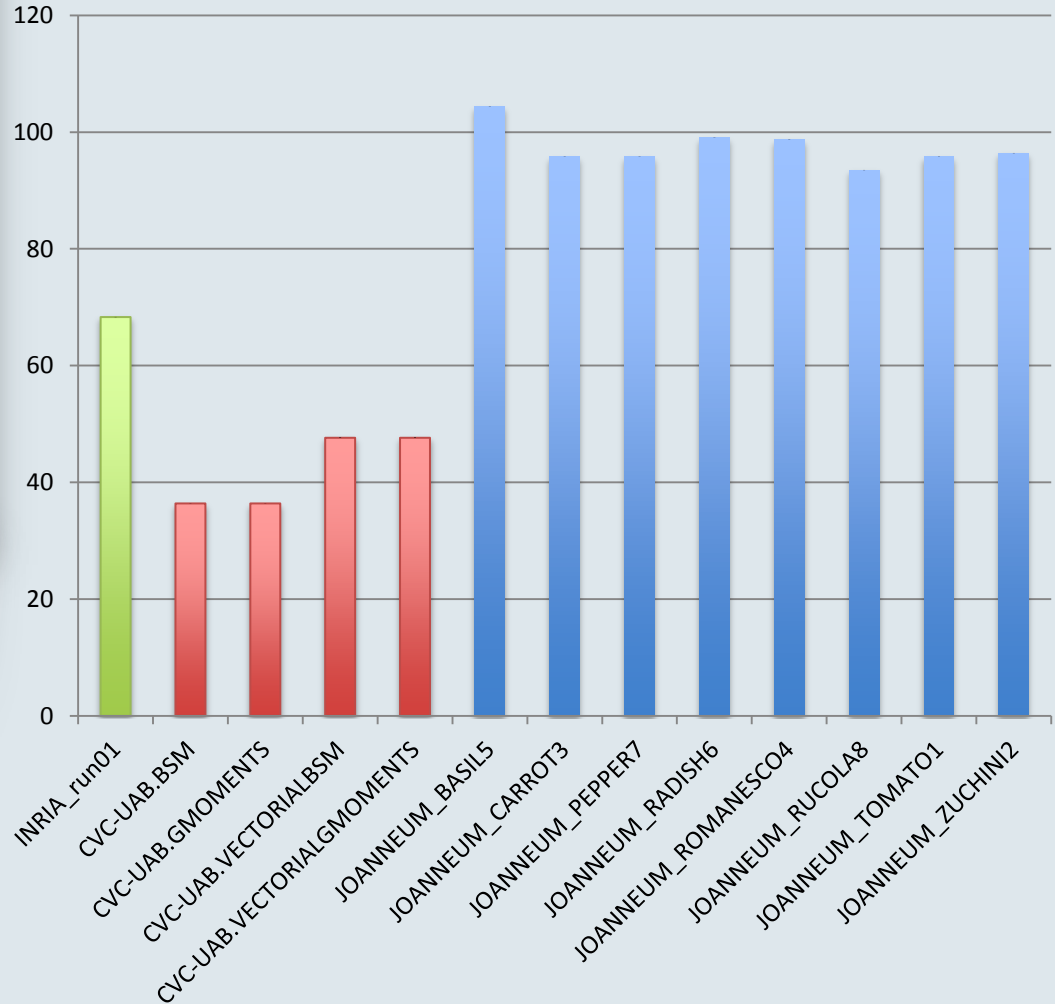


Flow-chart Recognition

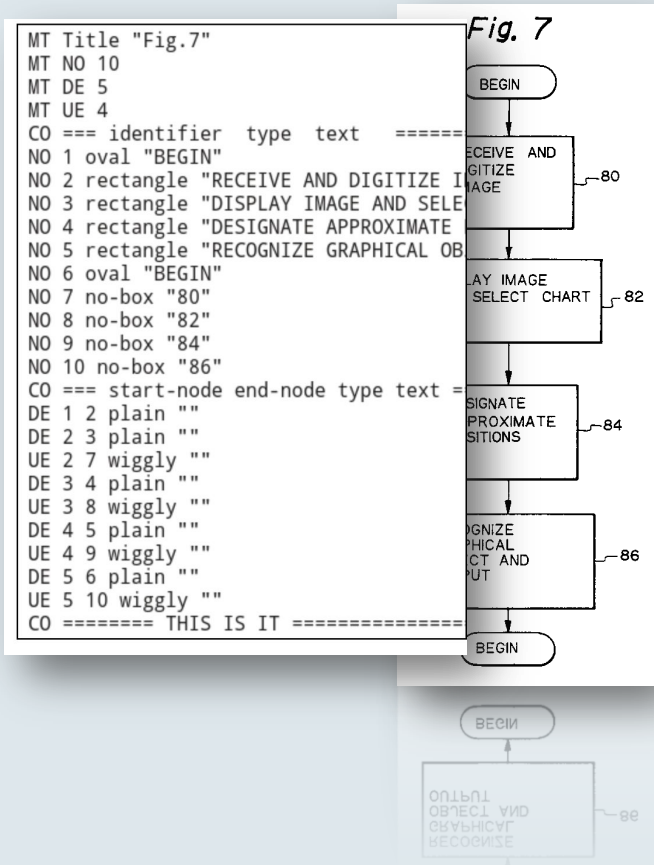


40 topics evaluated

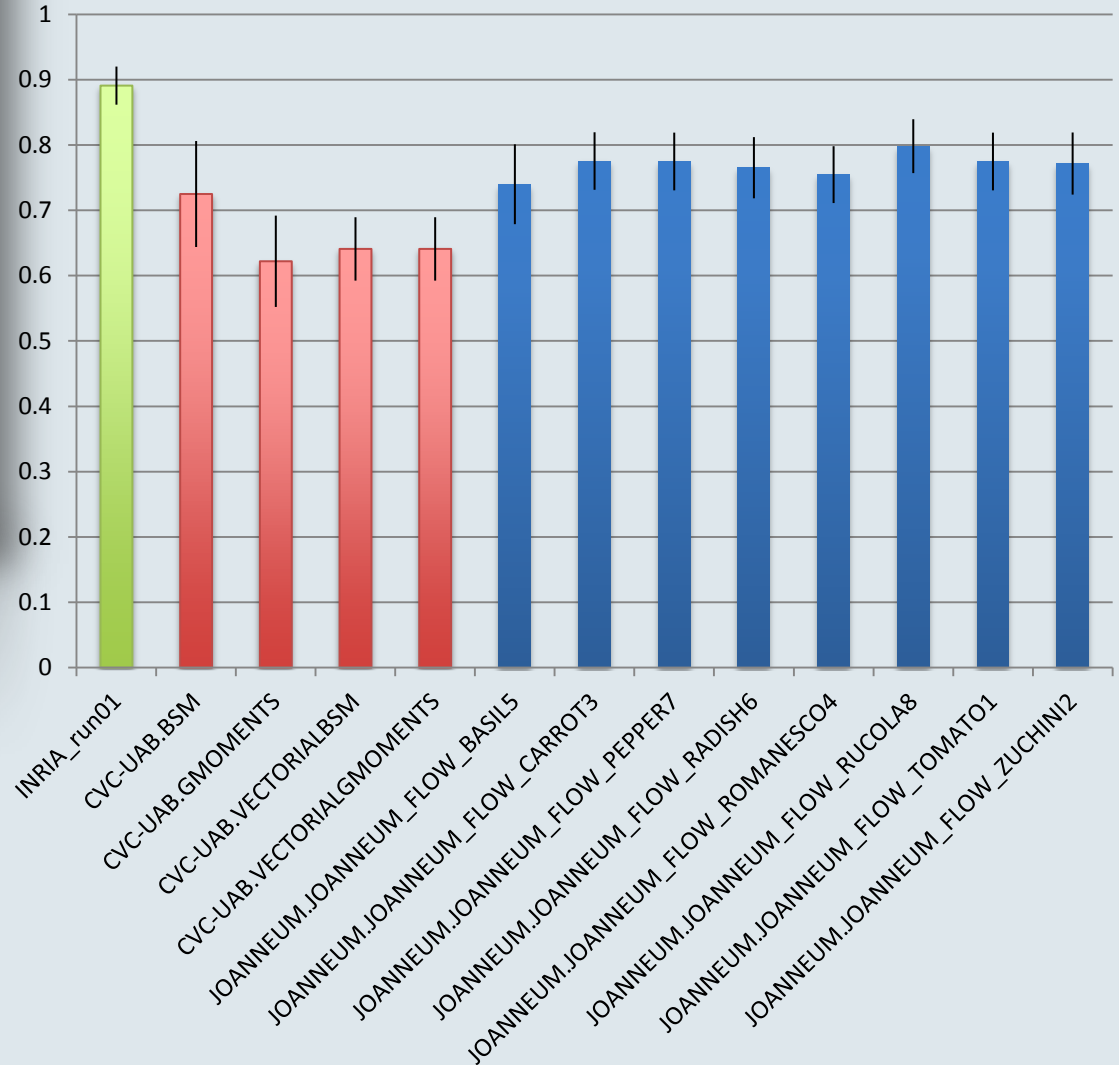
Text Recognition (Edit distance) - average



Flow-chart Recognition



Node type match



40 topics evaluated

Chemical Structure Recognition

Image segmentation

Tolerance	Precision	Recall	F ₁
0	0.70803	0.68622	0.69696
10	0.79311	0.76868	0.78070
20	0.82071	0.79543	0.80787
40	0.86696	0.84025	0.85340
55	0.88694	0.85962	0.87307

Structure recognition

	Automatic Set			Manual Set			Total		
	#Structures	Recalled	%	#Structures	Recalled	%	#Structures	Recalled	%
saic	865	761	88%	95	38	40%	960	799	83%
uob-1	865	832	96%	95	44	46%	960	876	91%
uob-2	865	821	95%	95	56	59%	960	877	91%
uob-3	865	821	95%	95	44	46%	960	865	90%
uob-4	865	832	96%	95	54	57%	960	886	92%

www.ifs.tuwien.ac.at/~clef-ip

CLEF-IP Workshop

Wednesday and Thursday

Thank You!

Florina Piroj, Mihai Lupu, Allan Hanbury
Walid Magdy, Alan Sexton, Igor Filippov