# CLEF 2011 - Semi-automated Artificial Intelligence to assist editing: An opportunity for Wikimedia sites

To Aru Shiroi Neko とある白い猫

## What is Artificial Intelligence

# Artificial intelligence (AI)

- Is a branch of computer science
- Uses machines to process relationship of info
- Without explicit programming
- To predict how to handle future data
- Grown in its use particularly in the past decade
- Applications range from search engines to space exploration
- You use it all the time probably without knowing
  - US Postal service
  - Online shopping suggestions



#### What do I mean by Semi-automated Artificial Intelligence

- Wikimedia sites have relied on "crowdsourcing"
  - Wikimedia is growing much faster that the "crowd"
  - Wikimedians can use the help
- Common "fear" in regards to automation
  - Skynet-phobia or bot takeover
  - Bot making mistakes
  - These points are fair
- Initially (until confidence forms)
  - Program will give suggestions ONLY
  - Suggestions may be restricted to limited people

#### **CLEF: Cross-Language Evaluation Forum**

- CLEF: Cross-Language Evaluation Forum
  - 19 22 September 2011, Amsterdam, Netherlands
  - 17 20 September 2012, Rome, Italy
- Artificial Intelligence classification tasks:
  - Text
  - Images
  - Text and Images
  - Audio
- http://clef2011.org
- http://clef2012.org





#### Wikimedia projects

- Wikimedia wikis are Maintained by the Wikimedia Foundation
- Content hosted on Wikimedia projects include
  - Wikipedia, Wikimedia Commons,
  - Wikibooks, Wikisource,
  - Wikinews, Wikiquote,
  - Wiktionary, Wikispecies
- Content is used
  - Wikimedia-wide
  - Internet-wide
    - Mirrors, forum posts, social media
  - **Downloaded for offline use** 
    - Wikireader (Commercial)
    - OLPC (Non-Profit)
    - Research (Scientific)



#### Free licenses

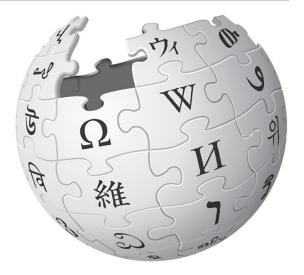
- Freely licensed
- Reusable, freely distributable
- Allows derivatives of any kind
- Commercially usable
- Used in very large projects like Wikipedia
- Used in research as demonstrated by ImageCLEF already
- Compels people to add to the existing repository of images







# Wikipedia



• Statistical information will be posted here

#### Wikipedia – Proposed task 1: Copyright/Plagiarism detection

## Classification task:

- Identify content copied to and from Wikipedia
  - Since the start this has remained to be a problem
  - Currently, tools such as Copyscape is used
  - Incorrectly identifiable as copyvio by tools like Copyscape
  - Make use of page histories to help determine
- Detect inclusion of previously oversigted content
  - Only users with oversight access would be able to read the reports

#### CLEF

Similar track on PAN workshop of CLEF 2011

#### Wikipedia – Proposed task 2: Author identification

## Classification task:

- Sockpuppets
  - Returning users that are previously blocked
- Meatpuppets & Lobbyists
  - Forum driven effort to introduce bias
  - List of blocks in regards to CAMERA is in light years
- Political edits
  - US Congressional staff edits to Wikipedia
- CLEF
  - Similar track on PAN workshop of CLEF 2011

#### Wikipedia – Proposed task 3: Vandalism detection

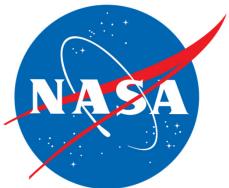
## Classification task:

- Identify Vandalism
  - Blanking, Gibberish, Sneaky
    - RC feed is full of these
  - Slander/Defamation
    - In other words: fact checking
    - John Seigenthaler controversy
  - Spam
    - Links (page rank boost)
    - Bot generated spam articles
- CLEF
  - Similar track on PAN workshop of CLEF 2011

#### Wikipedia – Proposed task 4: Disambiguation & Category Identification

- Classification task:
  - People
  - Places
  - Abbreviations
  - Incorrectly linked pages
- Disambiguation pages can be...
  - Confusing
  - Less than helpful
  - Outright misleading
- CLEF
  - Nothing similar on CLEF, yet









#### Wikipedia – Proposed task 5: Correlate real life events

#### Source?

- News, Facebook, Google+, Twitter trends, Blogs
- Wikipedia article on topic
  - Is it controversial? Sentiment analysis?

# Why?

- Assist admins by directing them to what is popular
- Help identify content for ITN section of main page
- Identify pages that can be updated

#### CLEF

Nothing similar on CLEF, yet

#### Wikisource – Proposed task: OCR for wiki

- OCR Object Character Recognition
  - Technology already exists for raw text
  - OCR then correlate with Wikipedia
    - Break up, categorize, link and tag
  - Help establish copyright
    - Text known to have copyright?
    - Deleted before as a copyright violation?
- CLEF
  - Nothing on OCR



#### Wikipedia – Proposed task 1

- Copyright identification (detection of text copied from elsewhere)
- Copyright identification (deleted images)
- OCR for Wiki
- Vandalism detection
- Category identification
  - Automated content identification (what images and/or text contain)
  - Automated plant identification (images)
- Disambiguation
- Correlate real life events, twitter trends and etc. to determine potential articles of attention for administrators

#### Wikimedia Commons

 Wikimedia Commons (or simply Commons) is an online repository of free-use

- Images,
- Audio,
- Video and
- Other files



http://commons.wikimedia.org

• midi	1,605	audio
<ul><li>wav</li></ul>	4	audio
• ogg	142,193	audio/video
• mp4	1	audio/video
<ul><li>gif</li></ul>	121,155	image/animation
<ul><li>jpeg</li></ul>	9,400,557	image
• png	835,988	image
<ul><li>svg+xml</li></ul>	485,356	images, vector
<ul><li>tiff</li></ul>	64,121	image
<ul><li>vnd.djvu</li></ul>	18,558	image
<ul><li>x-xcf</li></ul>	263	image
<ul><li>pdf</li></ul>	17,204	mixed text & images

Statistics by Benutzer:DaBPunkt

#### Wikimedia Commons – Statistics 2

- There are 103,112 galleries on Commons
- There are 11,085,316 files on Commons
- There are 1,259,001 deleted files on Commons
  - %11.342 of the existing files are deleted

#### Wikimedia Commons – Proposed task 1: Unwanted and/or controversial

# Examples of unwanted or controversial content

- Offensive/Vandalism
  - Genitalia (I will spare you the images)
  - Troll use of certain words or block of text within the images
  - Troll use of certain symbols within images: 🙈 🔌 📙 🌣
- Controversial
  - Nudity
  - Violence
- Copyright violations
  - Screenshots of movies
  - Photos of sculptures (non-FOP)
  - Special cases such as the Eiffel tower at night
  - Other stuff used without permission (frequently)





#### Wikimedia Commons – Proposed task 2: Categorization - 1

#### Category:Topics





This meta category should only contain other categories. Files should either be in the relevant subcategory or in the parent category.

English: This category is the global common root of the media files categorized by the topic, similar to en:Category:Main topic classifications. All media files should be categorized under this category for the sake of allowing others to find them by topic. For an explanation of Wikimedia's category system, see Commons:Categories, Help:Category and Category:CommonsRoot.

#### Subcategories

This category has the following 22 subcategories, out of 22 total.

- [+] <u>Categories by topic</u> (31 C, 1 F)
- [+] Academia (14 C, 66 F)
- [+] Activities (69 C, 3 F)
- В
- [+] Belief (7 C, 4 F)
- [+] Business (46 C, 271 F)
- [+] Causes (6 C, 23 F)
- [+] Communication (42 C, 1 P, 165 F)
- [+] Concepts (7 C, 15 F)

#### C cont.

- [+] <u>Creatures</u> (7 C, 3 F)
- [+] Culture (50 C, 1 P, 238 F)
- [+] Education (30 C, 2 P, 890 F)
- [+] Effects (10 C, 11 F)
- [+] Function (5 C, 2 F)
- [+] History (59 C, 4 P, 438 F)
- [+] Knowledge (14 C, 33 F)
- [+] Materials (56 C, 45 F)

- [-] Nature (30 C, 113 F)
  - [+] Nature by city (17 C, 1 F)
  - [+] Nature by country (226 C)
  - [+] Pronunciation of words relating to nature (7 C)
  - [+] Nature by region (19 C)
  - [+] Black and white photographs of nature (4 C, 3 F)

  - [+] Nature in art (41 C, 26 F)
  - [+] Colors in nature (2 C)
- [+] Countryside (10 C, 40 F)
- [+] Earth (30 C, 8 P, 178 F)
- [+] Environment (24 C, 1 P, 298 F)
- [+] Natural habitats (7 C, 1 P, 2 F)
- [+] Landscapes (31 C, 1 P, 201 F)
- [+] Matter (9 C, 3 P, 4 F)
- [+] Natural monuments (18 C, 63 F)
- [+] Natural history (8 C, 56 F)
- [x] Naturmentoring (5 F)
- [+] Nature at night (2 C, 4 F)
- [-] Organisms (15 C, 3 F)
- [x] Anaglyphs of organisms (29 F)
- [-] Animals (76 C, 3 P, 25 F)
- [+] Quality images of animals (7 C, 31 F)
- [+] Categories by animal (2 C)
- [+] Categories by zoo (6 C)
- [+] Animals by color (32 C, 1 F)
- Animals by common named groups (33 C)
- (9 C)
- [+] Badgers (1 C, 4 F)
- [+] Birds (40 C, 27 F)
- Butternies (16 C, ZP, ZF)
- [+] Calves (10 C, 1 P, 223 F)
- [+] Camels (9 C, 90 F)
- [+] Cattle (19 C, 1 P, 314 F)
- [+] Clams (3 C, 7 F)
- [+] Coral (7 C, 120 F)



#### Wikimedia Commons – Proposed task 2: Categorization - 2

#### Category:Birds

For the taxonomic classification of birds, see Category: Aves

This is a main category or subcategory requiring frequent diffusion and maybe maintenance. As many pictures and media files as possible should be woved into the right subcategories. (more information on categories)

| Boarisch | Беларуская (тарашкевіца) | Català | Česky | Deutsch | Deutsch (Sie-Form) | Еλληνικά | English | Español | Euskara | فريس | Suomi | Français | Galego | Alemannisch | أهريتها | Magyar | Interlingua | Italiano | 日本語 | პრთული | 한국어 | Kurdî | Lietuviu | Македонски | ஹமல் | Plattdüütsch | Nederlands | Occitan | Polski | Português | Português ob Brasil | Română | Русский | Slovenčina | <u> Svenska | Точикй | Türkçe | 中文(简体) | 中文(繁體) | +/-</u>

#### Vernacular names

English: Birds · Afrikaans: Voël · العربية: مائي · Asturianu: Páaru · Bahasa Indonesia: Burung · Bahasa Melayu: Burung · கு 😹 😹 Български: Птици · Català: Ocell · Česky: Ptáci · • Français: Oiseau • Frysk: Fûgels • Dererin Naoero: Tototl • Eesti: Linnud • Español: Ave • Esperanto: Birdoj • فارسي: برنده • Français: Oiseau • Frysk: Fûgels Gaeilge: Éan · Galego: Paxaro · 한국어: 새 · Hrvatski: Ptice · Ido: Ucelo · Íslenska: Fugl · Italiano: Uccelli · Italiano: עברית: עופות יופות יופו Kiswahili: Ndege·Kurdi: Cûk·Latina: Avis·Lëtzebuergesch: Vullen·Lietuvių: Paukščiai·Limburgs: Veugel·Magyar: Madár·Nederlands: Vogels·日本語: 鳥類·Norsk (bokmål): Fugler · Norsk (nynorsk): Fugl · Occitan: Aves · Plattdüütsch: Vagel · Polski: Ptaki · Português: Aves · Русский: Птицы · Slovenčina: Vtáky · Slovenščina: Ptič · Српски / Srpski: Птице · Suomi: Linnut · Svenska: Fåglar · ไทย: นก · Tiếng Việt: Chim · Türkçe: Kuşlar · Українська: Птахи · Walon: oujheas · 中文: 岛 ·

#### Subcategories

This category has the following 40 subcategories, out of 40 total.

```
• [+] Birds by color (11 C)

    [+] <u>Abused birds</u> (1 C, 1 F)

• [+] Birds by common name (13 C)

    [+] <u>Albinism in birds</u> (4 C, 32 F)

    [+] Birds by country (162 C, 1 F)

    [+] <u>Bird anatomy</u> (19 C, 20 F)

    [+] Birds by lake (3 C)

                                                            • [+] Audio files of Aves (16 C, 6 F)

    [+] Birds by region (10 C, 1 F)

 [+] Birds by setting (12 C)

    [+] Bird behavior (19 C, 23 F)

• [+] Featured pictures of birds (2 C, 278 F)

 [x] Bird harness (6 F)

• [+] Quality images of birds (3 C, 212 F)
                                                             • [+] Bird hunting (3 C, 14 F)
                                                            • [+] Bird repellents (5 C, 26 F)
                                                            • [+] Birdhouses (3 C, 1 P, 65 F)
• [+] Birds in art (74 C, 2 P, 383 F)
• [+] Aviculture (9 C, 3 F)
                                                            • [+] Bird conservation (9 C, 7 F)

    [+] Birdwatching (6 C, 1 P, 29 F)

• [+] Bird distribution maps (23 C, 1 P, 41 F)
• [+] Ornithology (7 C, 40 F)

    [+] <u>Bird feces</u> (1 C, 33 F)

    [x] <u>Feral birds</u> (58 F)

• [+] Unidentified birds (29 C, 107 F)
                                                            • [+] Bird flight (2 C, 4 F)

    [+] Flightless birds (27 C, 1 P, 3 F)
```

```
• [x] Invasive birds (5 F)

    [+] Bird reproduction (9 C)

• [+] Bird parks (12 C, 7 F)

 [x] Bird pellets (4 F)

• [+] People with birds (12 C, 76 F)
• [+] Pronunciation of words relating to birds (2 C, 29 F)

    [x] <u>Bird seed</u> (9 F)

• [+] Bird-related signs (2 C, 14 F)

    [+] Bird toys (1 C, 7 F)

    [+] Bird tracks (1 C, 36 F)

• [+] Veterinary medicine for birds (1 C)

    [+] Extinct birds (49 C, 8 F)
```

Q

## Wikimedia Commons – Proposed task 2: Categorization - 3

#### Category:Aircraft

The article Aircraft on Wikipedia projects:

af.Vliegtuig | articlinate, legical | bg:Летателен anapar | bs:Zrakoplov | ca:Aeronau | cs:Letadlo | cy:Awyren | da:Luftfarhzeug | en:Aircraft | eo:aviadilo | es:Aeronave | et:Õhusõiduk | fa:المنظرة المعالمة udara | nl:Luchtvaartuig | no:Luftfartøy | pl:Statek powietrzny | pt:Aeronave | ro:Aerodină | ru:Воздушное судно | simple:Aircraft | sk:Lietadlo | sl:Zrakoploy | sr:Ваздухоплов | sv:Flyqmaskin | ta:வானூர்கி | th:อากาศยาน | tr:Hava taṣiti | vi:Khí cu bay | vo:Oko ofurufu | zh:航空器

An aircraft is a vehicle which is able to fly by being supported by the air, or in general, the atmosphere, of a planet. Examples include balloons, airplanes and

Editors please ensure that the subcategory hierarchy is similar to the corresponding Wikipedia category hierarchy.

#### Subcategories

Category Aircraft on sister projects:



This category has the following 49 subcategories, out of 49 total.

• [+] Aircraft by color (14 C)	A cont.	O cont.
• [+] Aircraft by condition (6 C)	• [+] Anti-aircraft weapons (12 C, 1 P, 22 F)	• [+] Aircraft operators maps (1 C, 81 F)
• [+] Aircraft by country (119 C, 1 F)	• [+] Aircraft in art (15 C, 102 F)	P
• [-] Aircraft by function (15 C)	В	
<ul><li>[+] <u>Helicopters by function</u> (6 C)</li></ul>		• [+] <u>Aircraft parts</u> (34 C, 1 P, 132 F)
[+] Aerobatic aircraft (26 C, 14 F)	• [x] Black and white photographs of aircraft (153 F)	• [+] People with aircraft (8 C, 85 F)
[+] Agricultural aircraft (14 C, 40 F)	C	• [x] Pronunciation of words relating to aircraft (18 F)
[+] Airborne aircraft carriers (6 C, 12 F)	• [+] Aircraft carriers (23 C, 3 F)	R
[+] <u>Bush planes</u> (1 C, 8 F)	• [+] Cutaways of aircraft (1 C, 9 F)	• [+] Re-used aircraft registrations (11 C)
[+] <u>Civil aircraft</u> (61 C, 13 P, 27 F) [+] <u>Firefighting aircraft</u> (11 C, 41 F)		• [x] Reflections on aircraft (1 F)
[+] General aviation aircraft (25 C, 28	D	• [+] Aircraft-related road signs (1 C)
F)	• [+] <u>Drawings of aircraft</u> (8 C, 107 F)	S
[+] Military aircraft (19 C, 16 F)	F	
[+] Police aircraft (3 C, 1 F)	• [+] Aircraft in fiction (3 C, 20 F)	• [x] Sunrises taken from aircraft (6 F)
[+] Rescue aircraft (8 C, 7 F)	• [x] Fuel dumping (7 F)	• [x] Sunsets taken from aircraft (15 F)
[+] <u>Tanker aircraft</u> (10 C, 1 P, 4 F)		• [+] Symbols of aircraft (2 C, 1 F)
[+] <u>Trainer aircraft</u> (14 C, 1 P, 23 F)	L	T
[+] Transport aircraft (28 C, 12 F)	• [+] Aircraft liveries (3 C)	• [+] Transport of aircraft (4 C, 1 F)
[+] <u>Utility aircraft</u> (7 C, 1 F)	• [+] Locomotive nameplates featuring aircraft (1 C, 5 F)	V
<ul> <li>[+] Aircraft by location (27 C)</li> <li>[+] Aircraft by manufacturer (397 C, 1 F)</li> </ul>	M	AND THE PROPERTY OF THE PROPER
• [+] Aircraft by manuacturer (357 C, 117) • [+] Aircraft by motive power (6 C)	• [+] Maiden flights (2 C, 1 F)	• [+] Videos of aircraft (24 C, 5 F)
• [+] Aircraft by museum (2 C)	• [+] Aircraft maintenance (1 C, 84 F)	• [+] Views of aircraft (7 C, 1 F)
• [+] Aircraft by operator (4 C)	• [+] Aircraft manufacturers (293 C)	W
• [+] Aircraft by period (6 C)	• [+] Mechanics of Aircraft Structures - University of Florida course EAS4200C (1 C, 174 F	• [x] Water reflections of aircraft (6 F)
• [+] Aircraft by registration (9739 C)	• [+] Models of aircraft (16 C, 3 P, 168 F)	
• [+] Aircraft by type (33 C, 2 F)	• [+] Aircraft in motion (1 C, 62 F)	_
A	0	• [+] Unidentified aircraft (1 C, 47 F)
<ul> <li>[+] Aircraft with advertisement (3 C, 9 F)</li> </ul>	• [+] Aircraft operations (6 G, 1 F)	

Q

## Wikimedia Commons – Proposed task 2: Categorization - 4

#### Category:Aircraft

The article Aircraft on Wikipedia projects:

af.Vlieqtuiq | ar. كان اعظر المعربة | bg:Летателен апарат | bs:Zrakoploy | ca:Aeronau | cs:Letadlo | cy:Awyren | da:Luftfartøj | de:Luftfahrzeuq | en:Aircraft | eo:aviadilo | es:Aeronave | et:Õhusõiduk | fa: ومكودة | fi:llma-alus | fr.Aéronef | fy.Loftfartúch | gl.Aeronave | heːoːʊ ウン | hiːवाय्यान | hr.Zrakoplov | hu:Repülőgép | id:Pesawat terbang | io:Aeronavo | is:Loftfar | it.Aeromobile | ja:航空機 | koː항공기 | It.Orlaivis | ms:Pesawat udara | nl:Luchtvaartuig | no:Luftfartøy | pl:Statek powietrzny | pt:Aeronave | ro:Aerodină | ru:Воздушное судно | simple:Aircraft | sk:Lietadlo | sl:Zrakoploy | sr:Ваздухоплов | sv:Flyqmaskin | ta:வானூர்கி | th:อากาศยาน | tr:Hava taṣiti | vi:Khí cu bay | vo:Oko ofurufu | zh:航空器

An aircraft is a vehicle which is able to fly by being supported by the air, or in general, the atmosphere, of a planet. Examples include balloons, airplanes and

Editors please ensure that the subcategory hierarchy is similar to the corresponding Wikipedia category hierarchy.

#### Subcategories

Category Aircraft on sister projects:

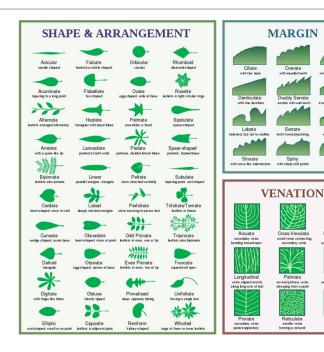


This category has the following 49 subcategories, out of 49 total.

• [+] Aircraft by color (14 C)	A cont.	O cont.
• [+] Aircraft by condition (6 C)	• [+] Anti-aircraft weapons (12 C, 1 P, 22 F)	• [+] Aircraft operators maps (1 C, 81 F)
• [+] Aircraft by country (119 C, 1 F)	• [+] Aircraft in art (15 C, 102 F)	P
<ul> <li>[-] Aircraft by function (15 C)</li> </ul>	В	
[+] Helicopters by function (6 C)		• [+] <u>Aircraft parts</u> (34 C, 1 P, 132 F)
[+] Aerobatic aircraft (26 C, 14 F)	• [x] Black and white photographs of aircraft (153 F)	• [+] People with aircraft (8 C, 85 F)
[+] Agricultural aircraft (14 C, 40 F)	C	• [x] Pronunciation of words relating to aircraft (18 F)
[+] Airborne aircraft carriers (6 C, 12 F)	• [+] Aircraft carriers (23 C, 3 F)	R
[+] <u>Bush planes</u> (1 C, 8 F)	• [+] Cutaways of aircraft (1 C, 9 F)	• [+] Re-used aircraft registrations (11 C)
[+] <u>Civil aircraft</u> (61 C, 13 P, 27 F)		• [x] Reflections on aircraft (1 F)
[+] <u>Firefighting aircraft</u> (11 C, 41 F) [+] <u>General aviation aircraft</u> (25 C, 28	D	• [+] Aircraft-related road signs (1 C)
F)	• [+] Drawings of aircraft (8 C, 107 F)	.,
[+] Military aircraft (19 C, 16 F)	F	S
[+] Police aircraft (3 C, 1 F)	• [+] Aircraft in fiction (3 C, 20 F)	<ul> <li>[x] Sunrises taken from aircraft (6 F)</li> </ul>
[+] Rescue aircraft (8 C, 7 F)	• [+] Aircraft in liction (3 C, 20 F) • [x] Fuel dumping (7 F)	<ul> <li>[x] Sunsets taken from aircraft (15 F)</li> </ul>
[+] Tanker aircraft (10 C, 1 P, 4 F)	• [*] Fuer dumping (7 F)	<ul> <li>[+] Symbols of aircraft (2 C, 1 F)</li> </ul>
[+] Trainer aircraft (14 C, 1 P, 23 F)	L	Т
[+] Transport aircraft (28 C, 12 F)	• [+] Aircraft liveries (3 C)	• [+] Transport of aircraft (4 C, 1 F)
[+] Utility aircraft (7 C, 1 F)	• [+] Locomotive nameplates featuring aircraft (1 C, 5 F)	
• [+] Aircraft by location (27 C)	M	V
• [+] Aircraft by manufacturer (397 C, 1 F)		<ul> <li>[+] Videos of aircraft (24 C, 5 F)</li> </ul>
• [+] Aircraft by motive power (6 C)	• [+] Maiden flights (2 C, 1 F)	<ul> <li>[+] Views of aircraft (7 C, 1 F)</li> </ul>
• [+] Aircraft by museum (2 C)	• [+] Aircraft maintenance (1 C, 84 F)	W
• [+] Aircraft by operator (4 C)	• [+] Aircraft manufacturers (293 C)	
• [+] Aircraft by period (6 C)	• [+] Mechanics of Aircraft Structures - University of Florida course EAS4200C (1 C, 174 F)	• [*] <u>water reflections of aircraft</u> (6 F)
• [+] Aircraft by registration (9739 C)	• [+] Models of aircraft (16 C, 3 P, 168 F) • [+] Aircraft in motion (1 C, 62 F)	~
• [+] Aircraft by type (33 C, 2 F)	• [T] Alicial III Mouoli (1 C, 02 F)	• [+] Unidentified aircraft (1 C, 47 F)
A	0	
• [+] Aircraft with advertisement (3 C, 9 F)	• [+] Aircraft operations (6 C, 1 F)	

#### Wikimedia Commons – Proposed task 3: Plant identification

- Plant identification
  - Is it a plant
  - Type of image
    - Scan of leaf
    - Photo of individual leaf
    - Photo in nature
  - What type of leaf does it have?
    - Shape and Arrangement
    - Margin
    - Venation
  - Identification isn't necessarily down to the species
- A human botanist can disambiguate further



#### Wikimedia Foundation – Proposed tasks

- Wikimedia servers
  - Performance
    - More efficient distribution of load
    - Detection of trends leading to failure
    - Help determine deadlines for upgrades
  - Cyber Defense
    - Anomaly detection
      - Detect botnets
      - Detect sneaky RATs
      - Consider Stuxnet and DuQu
    - Detection of manipulation to database
    - Does Wikipedia count as "critical infrastructure"?
- Sentiment analysis of social media and the web
  - Improve relations by not being deaf to the concerns of readers
  - Handle false information about the WMF
  - Identify impersonation attempts of WMF



## Wikimedia Commons – Questions?



## Wikimedia Commons

• Thank You!