

AE4B99RPH: Problem Solving and Games

Task 2: Spam filter

Petr Pošík

Dept. of Cybernetics
CTU FEE

Spam filter

SPAM?

OK, seriously, what is

SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is

SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

Spam filter

SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

When?

✓ 1937

Who?

✓ Hormel Foods Corporation

What?

SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

When?

✓ 1937

Who?

✓ Hormel Foods Corporation

What?



SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

When?

✓ 1937

Who?

✓ Hormel Foods Corporation

What?



The meaning of SPAM acronym:

SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

When?

- ✓ 1937

Who?

- ✓ Hormel Foods Corporation

What?



The meaning of SPAM acronym:

- ✓ *Shoulder of Pork And Ham or Spiced Ham*

SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

When?

- ✓ 1937

Who?

- ✓ Hormel Foods Corporation

What?



The meaning of SPAM acronym:

- ✓ *Shoulder of Pork And Ham or Spiced Ham*
- ✓ “Specially Processed American Meats” (USA offered SPAM to England during WW2)
- ✓ “Something Posing As Meat”
- ✓ “Specially Processed Artificial Meat”
- ✓ “Stuff, Pork and Ham”
- ✓ “Special Product of Austin Minnesota”

OK, seriously, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

In 1970s:

- ✓ in USENET groups, messages that nobody was interested in started to show up

OK, seriously, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

In 1970s:

- ✓ in USENET groups, messages that nobody was interested in started to show up
- ✓ the situation resembled a scene from **Monty Python's Flying Circus**:
 - ✗ see it e.g. on youtube.com
 - ✗ a married couple wants to have a breakfast in a snack bar
 - ✗ the snack-bar staff reads the menu:
 - ✓ egg and bacon
 - ✓ egg, sausage and bacon
 - ✓ egg and spam
 - ✓ egg, bacon and spam
 - ✓ egg, bacon, sausage and spam
 - ✓ spam, bacon, sausage and spam
 - ✓ spam, egg, spam, spam, bacon and spam
 - ✓ spam, spam, spam, egg and spam
 - ✓ spam, spam, spam, spam, spam, spam, baked beans, spam, spam, spam and spam
 - ✓ lobster thermidor aux crevettes, with a mornay sauce garnished with truffle paté, brandy and a fried egg on top and spam
 - ✗ the wife does not want any meal containing SPAM
 - ✗ the staff and her husband keep forcing her to get SPAM
 - ✗ "I DON'T LIKE SPAM!!!"

OK, seriously, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

In 1970s:

- ✓ in USENET groups, messages that nobody was interested in started to show up
- ✓ the situation resembled a scene from **Monty Python's Flying Circus**:
 - ✗ see it e.g. on youtube.com
 - ✗ a married couple wants to have a breakfast in a snack bar
 - ✗ the snack-bar staff reads the menu:
 - ✓ egg and bacon
 - ✓ egg, sausage and bacon
 - ✓ egg and spam
 - ✓ egg, bacon and spam
 - ✓ egg, bacon, sausage and spam
 - ✓ spam, bacon, sausage and spam
 - ✓ spam, egg, spam, spam, bacon and spam
 - ✓ spam, spam, spam, egg and spam
 - ✓ spam, spam, spam, spam, spam, spam, baked beans, spam, spam, spam and spam
 - ✓ lobster thermidor aux crevettes, with a mornay sauce garnished with truffle paté, brandy and a fried egg on top and spam
 - ✗ the wife does not want any meal containing SPAM
 - ✗ the staff and her husband keep forcing her to get SPAM
 - ✗ "I DON'T LIKE SPAM!!!"
- ✓ SPAM - uninteresting, unwanted, obtrusive, repeating messages
- ✓ HAM - normal correct messages on topic

SPAM or HAM? (1)

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

As seen on NBC, CBS, CNN, and even Oprah! The health discovery that actually reverses aging while burning fat, without dieting or exercise! This proven discovery has even been reported on by the New England Journal of Medicine. Forget aging and dieting forever! And it's Guaranteed!

Click here: <http://web.kuhleersparnis.ch/hgh/index.html>

Would you like to lose weight while you sleep!?

No dieting!

No hunger pains!

No Cravings!

No strenuous exercise!

Change your life forever!

100% GUARANTEED!

1.Body Fat Loss	82% improvement.
2.Wrinkle Reduction	61% improvement.
3.Energy Level	84% improvement.
4.Muscle Strength	88% improvement.
5.Sexual Potency	75% improvement.
6.Emotional Stability	67% improvement.
7.Memory	62% improvement.

SPAM or HAM? (2)

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

Filled with useful examples and the depth, clarity, and attention to detail that made the first edition so popular with web developers, the just-released "JavaServer Pages, 2nd Edition" (Bergsten, \ \$44.95) is completely revised and updated to cover the substantial changes in the 1.2 version of the JSP specifications, and includes coverage of the new JSTL Tag libraries--an eagerly anticipated standard set of JSP elements for the tasks needed in most JSP applications, as well as thorough coverage of Custom Tag Libraries.

What people said about the first edition:

"an excellent printed resource on JSPs...I have been extremely impressed by its depth, clarity, and attention to detail. "

--Reuven M. Lerner, Linux Journal

"This is a great book: it was written by a key contributor not only to the JSP specification, but also to the JSP and Servlet reference implementations. Filled with useful examples, it stands as an important text in the adoption of JSP in the market."

--Eduardo Pelegri-Llopart, lead JSP Specification Engineer.

To order your copy or for more information, see:

<http://www.oreilly.com/catalog/jserverpages2/?CMP=7337>

or call 1-800-998-9938

or email orders@oreilly.com

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

**Third time lucky, what is
SPAM?**

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

**Third time lucky, what is
SPAM?**

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

**Third time lucky, what is
SPAM?**

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?
- ✓ Unsolicited commercial email?

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?
- ✓ Unsolicited commercial email?
- ✓ Bulk unsolicited commercial email?

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?
- ✓ Unsolicited commercial email?
- ✓ Bulk unsolicited commercial email?
- ✓ **Spam is an *unsolicited* e-mail sent *indiscriminately* to multiple mailing lists, individuals, or newsgroups, often of a *commercial nature*.**

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?
- ✓ Unsolicited commercial email?
- ✓ Bulk unsolicited commercial email?
- ✓ **Spam is an *unsolicited* e-mail sent *indiscriminately* to multiple mailing lists, individuals, or newsgroups, often of a *commercial nature*.**
 - ✗ vague definition, may not cover all cases, may cover cases which are not spam

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?
- ✓ Unsolicited commercial email?
- ✓ Bulk unsolicited commercial email?
- ✓ **Spam is an *unsolicited* e-mail sent *indiscriminately* to multiple mailing lists, individuals, or newsgroups, often of a *commercial nature*.**
 - ✗ vague definition, may not cover all cases, may cover cases which are not spam

Further questions about spam:

1. Would all people describe as spam the same set of messages?

Third time lucky, what is SPAM?

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is spam? How would you define it?

- ✓ Unsolicited email?
- ✓ Unsolicited commercial email?
- ✓ Bulk unsolicited commercial email?
- ✓ **Spam is an *unsolicited* e-mail sent *indiscriminately* to multiple mailing lists, individuals, or newsgroups, often of a *commercial nature*.**
 - ✗ vague definition, may not cover all cases, may cover cases which are not spam

Further questions about spam:

1. Would all people describe as spam the same set of messages?
2. If the same man got the same set of messages now and five years in the future, would he describe as spam the same emails?

Spam filter

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is a *Spam filter*?

Spam filter

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is a *Spam filter*?

- ✓ A machine that can label each message either as SPAM or as HAM.

Spam filter

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is a *Spam filter*?

- ✓ A machine that can label each message either as SPAM or as HAM.

What are its inputs and outputs?

Spam filter

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is a *Spam filter*?

- ✓ A machine that can label each message either as SPAM or as HAM.

What are its inputs and outputs?

- ✓ Input: all information about the email that we can get (message body, its formatting, headers, attachments, ...). Very often unstructured text.

Spam filter

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is a *Spam filter*?

- ✓ A machine that can label each message either as SPAM or as HAM.

What are its inputs and outputs?

- ✓ Input: all information about the email that we can get (message body, its formatting, headers, attachments, ...). Very often unstructured text.
- ✓ Output: label SPAM or HAM

Spam filter

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

What is a *Spam filter*?

- ✓ A machine that can label each message either as SPAM or as HAM.

What are its inputs and outputs?

- ✓ Input: all information about the email that we can get (message body, its formatting, headers, attachments, ...). Very often unstructured text.
- ✓ Output: label SPAM or HAM

Further questions about spam filter:

1. Can a single spam filter satisfy all potential users?
2. What feature must the spam filter have so that
 - ✓ it can preserve its performance in time?
 - ✓ it can be equally useful for different people?

Spam filtering

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

Other things to think about:

- ✓ Will the definition help us in recognizing the spam?

Spam filtering

Spam filter

SPAM?

OK, seriously, what is
SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is
SPAM?

Spam filter

Spam filtering

Goals and organization

Questions?

Other things to think about:

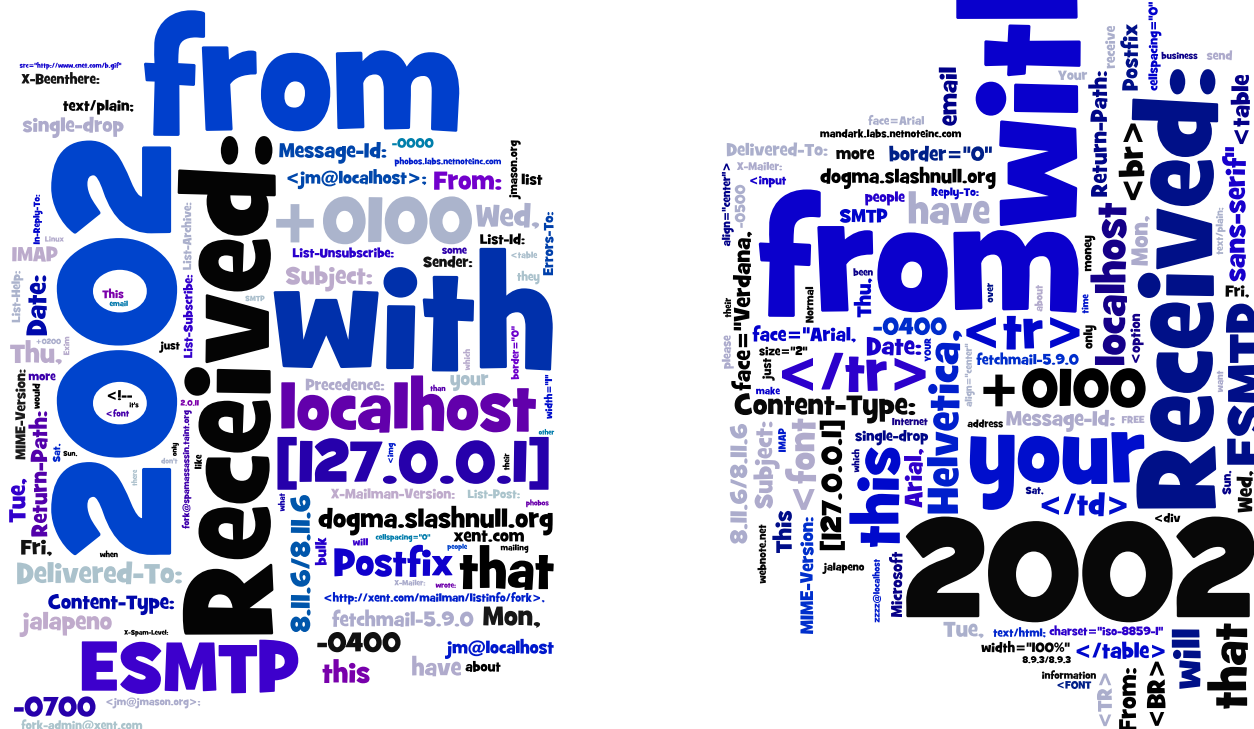
- ✓ Will the definition help us in recognizing the spam?
- ✓ What are the features of spam? What characteristics can the filter use to decide whether a mail is spam or not?

Spam filtering

- Spam filter
- SPAM?
- OK, seriously, what is SPAM?
- SPAM or HAM? (1)
- SPAM or HAM? (2)
- Third time lucky, what is SPAM?
- Spam filter
- Spam filtering
- Goals and organization
- Questions?

Other things to think about:

- ✓ Will the definition help us in recognizing the spam?
- ✓ What are the features of spam? What characteristics can the filter use to decide whether a mail is spam or not?

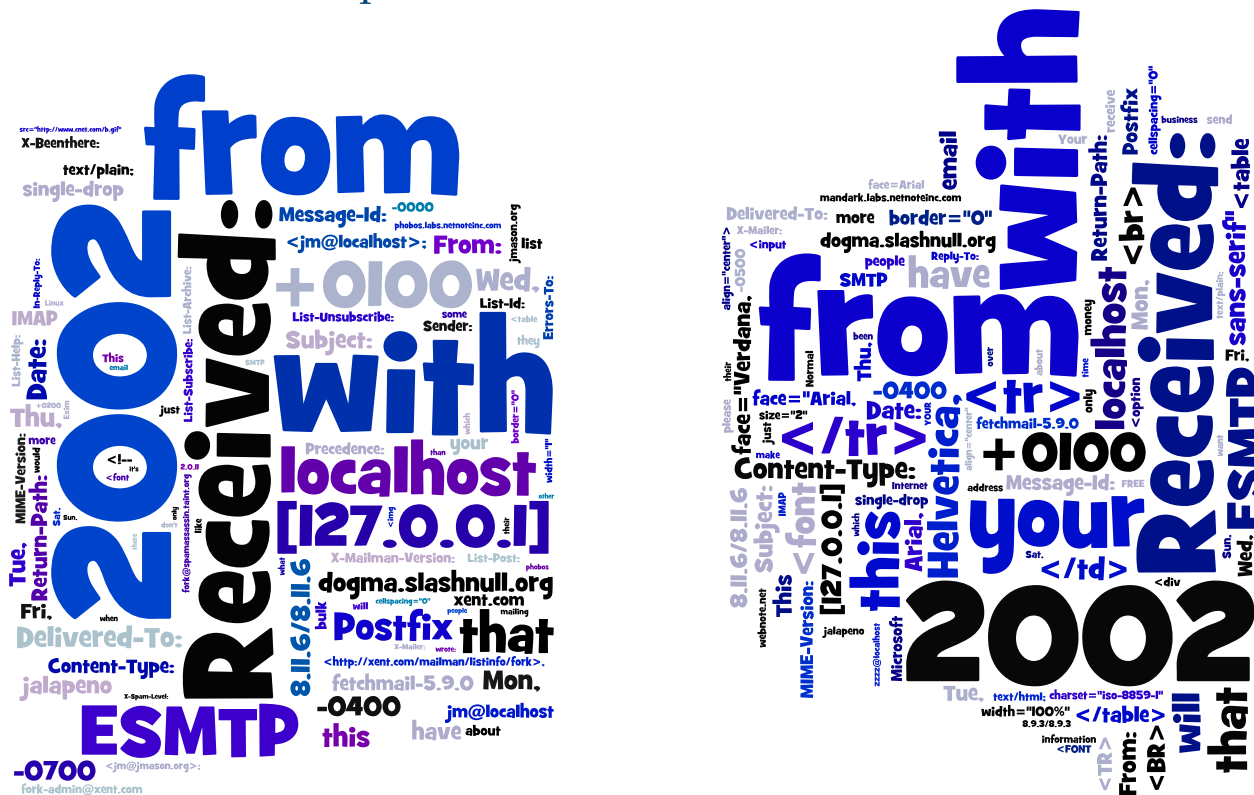


Spam filtering

- Spam filter
- SPAM?
- OK, seriously, what is SPAM?
- SPAM or HAM? (1)
- SPAM or HAM? (2)
- Third time lucky, what is SPAM?
- Spam filter
- Spam filtering
- Goals and organization
- Questions?

Other things to think about:

- ✓ Will the definition help us in recognizing the spam?
- ✓ What are the features of spam? What characteristics can the filter use to decide whether a mail is spam or not?



- ✓ Can the filter make an error? What types of errors?

Spam filtering

Spam filter

SPAM?

OK, seriously, what is SPAM?

SPAM or HAM? (1)

SPAM or HAM? (2)

Third time lucky, what is SPAM?

Spam filter

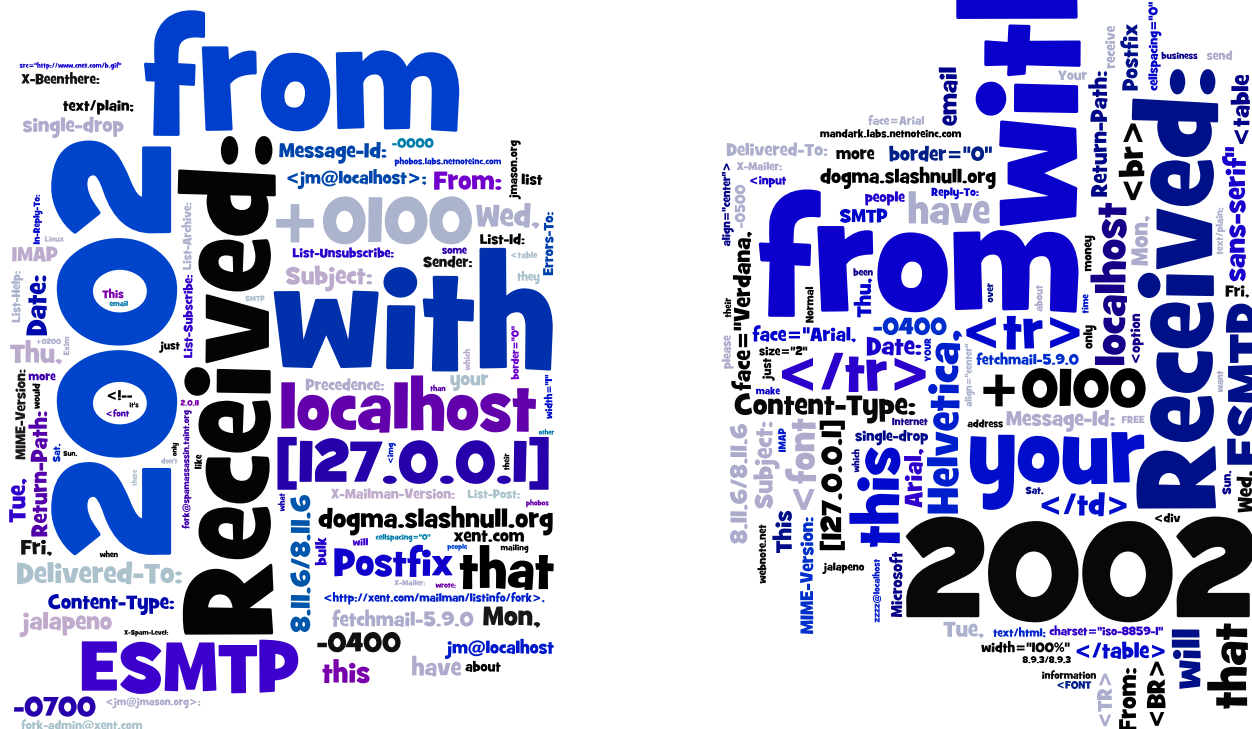
Spam filtering

Goals and organization

Questions?

Other things to think about:

- ✓ Will the definition help us in recognizing the spam?
- ✓ What are the features of spam? What characteristics can the filter use to decide whether a mail is spam or not?



- ✓ Can the filter make an error? What types of errors?
- ✓ Are all the errors equally serious for us?

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Goals and organization

Spam filtering as task 2?

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Why was spam filtering chosen as a task?

Spam filtering as task 2?

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Why was spam filtering chosen as a task?

- ✓ Practical, non-trivial task, yet solvable with the knowledge we have.
- ✓ Showcase of the basic principles and problems in machine learning:
 - ✗ dataset and its characteristics
 - ✗ training and testing of a classifier
 - ✗ overfitting

Spam filtering as task 2?

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Why was spam filtering chosen as a task?

- ✓ Practical, non-trivial task, yet solvable with the knowledge we have.
- ✓ Showcase of the basic principles and problems in machine learning:
 - ✗ dataset and its characteristics
 - ✗ training and testing of a classifier
 - ✗ overfitting

What will I gain from this task?

Spam filtering as task 2?

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Why was spam filtering chosen as a task?

- ✓ Practical, non-trivial task, yet solvable with the knowledge we have.
- ✓ Showcase of the basic principles and problems in machine learning:
 - ✗ dataset and its characteristics
 - ✗ training and testing of a classifier
 - ✗ overfitting

What will I gain from this task?

- ✓ I am interested in the ways how a computer program can learn from examples.
- ✓ I am interested in the principles of spam filtering.
- ✓ I want practice my knowledge of Python.
- ✓ I am interested in automatic text processing.

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them and
 - ✗ clarify how the spam filter should be implemented.

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them and
 - ✗ clarify how the spam filter should be implemented.
- ✓ Spend another 2 weeks working on the actual filter:

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them and
 - ✗ clarify how the spam filter should be implemented.
- ✓ Spend another 2 weeks working on the actual filter:
 - ✗ we do not require any super-anti-contra-multi-extra-unique filter; you do not have the needed background for them

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them and
 - ✗ clarify how the spam filter should be implemented.
- ✓ Spend another 2 weeks working on the actual filter:
 - ✗ we do not require any super-anti-contra-multi-extra-unique filter; you do not have the needed background for them
 - ✗ at least, try out some relatively simple method that - in your opinion - should work for spam filtering

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them and
 - ✗ clarify how the spam filter should be implemented.
- ✓ Spend another 2 weeks working on the actual filter:
 - ✗ we do not require any super-anti-contra-multi-extra-unique filter; you do not have the needed background for them
 - ✗ at least, try out some relatively simple method that - in your opinion - should work for spam filtering

What you will get from us:

- ✓ Automated tests for individual solution steps

The work plan

Spam filter

Goals and organization

Spam filtering as task 2?

The work plan

Questions?

Plan:

- ✓ We have about 4 weeks for this task.
- ✓ Spend 2 weeks working on the foundations allowing us to
 - ✗ easily create new filters,
 - ✗ easily test them and
 - ✗ compare them and
 - ✗ clarify how the spam filter should be implemented.
- ✓ Spend another 2 weeks working on the actual filter:
 - ✗ we do not require any super-anti-contra-multi-extra-unique filter; you do not have the needed background for them
 - ✗ at least, try out some relatively simple method that - in your opinion - should work for spam filtering

What you will get from us:

- ✓ Automated tests for individual solution steps

What shall be handed in?

- ✓ Implementation of the evaluation function for spam filters
- ✓ Implementation of the spam filter

Spam filter

Goals and organization

Questions?

Questions?