

#linux #bash #regex

текст с допустимыми и не допустимыми адресами email'ов

Внезапно name1@ya.ru оказался в пустоте.
 Но затем na-me2@ya.ru появился из ниоткуда.
 В письме было указано: name3@ya1.ya-2.ru
 (spb@edu.ru) – это контакты университета.
 Кто-то сказал a@b.cd и все замолчали.
 Адрес для ответа: first.last@sub.domain.co.uk
 Он оставил email123@test-mail.com однако не ответил.
 x@example.info был на экране.
 (other.email-with-dash@example.com) – еще один вариант.

А вот недопустимые: привет1not@ya.гумир и словопо.-t2@ya.руеще.
 Пишитенаabc@cde.efсразу если что.
 А этоnot@.comнеправильно.
 Сообщение-not-allowed@example.comне доходит.
 В not@allowed..comслучае ошибки.
 Для not@-allowed.comтеста.
 Not@allowed_.com это не работает.
 Abc.example.com просто текст.
 Not@allowedдомен не существует.

regular expression (deepseek):

```
grep -E "(?<=^|[\s(:)]) [a-zA-Z0-9] [a-zA-Z0-9._-]*@[a-zA-Z0-9-]+\.[a-zA-Z0-9-]+\.[a-zA-Z]{2,} (?=[\s)\t\n]|$)"
```

2. Напишите команду `find`, которая используя `grep`, найдет с помощью регулярных выражений электронные адреса в текстовых файлах.

```
find . -name "*.txt" -exec grep -E "(?<=^|[\s(:)]) [a-zA-Z0-9] [a-zA-Z0-9._-]*@[a-zA-Z0-9-]+\.[a-zA-Z0-9-]+\.[a-zA-Z]{2,} (?=[\s)\t\n]|$)" {} \;
```

Задания

1. Выбрать существующие даты между 1000 и 2012 годом. Секунды могут быть опущены. В каждом месяце 30 дней.

```
grep -E "([12]0[01] [012])\|([01-9] [112])\|([01-9] [11-9] [21-9] |30) (([01] [0-9] |2[0-3]):[0-5] [0-9])"
```

test: <https://regex101.com/r/sVXSAp/1>

2. Преобразовать текст, обранный в звездочки, в курсив. Не трогать текст в двойных звездочках (жирный).
Использовать команду `sed`.