

شرایط دنیای واقعی، بسیار متفاوت است از طراحی‌های ساده‌ی اولیه‌ی ثبت نام. در طراحی‌های ساده، ایمیل، نام کاربری و بسیاری از اطلاعات دیگر باید منحصر بفرد باشند. ایندکس منحصر بفرد تعریف می‌کنید. قیود و اعتبار سنجی سمت سرور و سمت کاربر را اضافه می‌کنید. چقدر عالی! اما ... دنیای واقعی شکل دیگری را دارد!

یک روز با ایمیل `username@gmail.com` ثبت نام می‌کنند. فردا با ایمیل `user.name@gmail.com` ثبت نام خواهند کرد. پس فردا با ایمیل `us.er.name@gmail.com` و به همین ترتیب! امروز با نام «کاربر یک» ثبت نام می‌کنند. فردا با نام «کاربر یک»! امروز با نام «مجید» ثبت نام می‌کنند، فردا با نام «مجید»! همچنین علاقه‌ی شدیدی هم به استفاده از ایمیل‌های fake دارند ( [راه حل](#) ). بنابراین نیاز است اطلاعات کاربران را پیش از ثبت نام نرمال سازی کرد. برای مثال نقطه‌های ایمیل‌های جیمیل را حذف کرد؛ یا اگر اجازه داده‌اید که در بین کلمات نام کاربری، فاصله‌ای را وارد کنند، فقط یک فاصله مجاز باشد و یا اگر نامی را ثبت می‌کنید، به فکر حالت‌های کش آمده‌ی آن مانند «مجید» هم باید بود و آن‌را تبدیل به حالت اصلی‌اش کرد.

## نرمال سازی ایمیل‌های gmail

تا جایی که اطلاع دارم، حداقل فیس بوک و جی‌میل، بکارگیری نقاط را در ایمیل‌ها مجاز می‌دانند. برای مثال ترکیب‌های زیر از دید gmail تنها یک ایمیل محسوب می‌شوند:

`johndoe@gmail.com`

`john....doe@gmail.com`

`johndoe+spamsite@gmail.com`

راه حل پیشنهادی:

```
public static string FixGmailDots(string email)
{
    if (string.IsNullOrEmpty(email))
        return string.Empty;

    email = email.ToLowerInvariant().Trim();
    var emailParts = email.Split('@');
    var name = emailParts[0].Replace(".", string.Empty).Replace("+", string.Empty);
    var emailDomain = emailParts[1];

    string[] domainsAllowedDots =
    {
        "gmail.com",
        "facebook.com"
    };

    var isFromDomainsAllowedDots = domainsAllowedDots.Any(domain => emailDomain.Equals(domain));
    return !isFromDomainsAllowedDots ? email : string.Format("{0}@{1}", name, emailDomain);
}
```

در اینجا بررسی می‌شود که آیا دومین ایمیل دریافتی از سمت جیمیل یا فیس بوک است؟ اگر بله، آنگاه نقطه‌ها و +های آن‌ها حذف می‌شوند.

این بررسی باید در حین ثبت نام و همچنین ویرایش اطلاعات کاربری جهت نرمال سازی اطلاعات اعمال شود. اگر سایت‌های دیگری هم هستند که بکارگیری نقاط را مجاز می‌دانند، آرایه‌ی `domainsAllowedDots` را تکمیل کنید.

## نرمال سازی ورود حروف ویژه

نرمال سازی ابتدایی ثبت نام کاربران در سایت جاری به صورت ذیل است:

```
friendlyName =
```

```
friendlyName.ApplyCorrectYeKe().RemoveDiacritics().CleanUnderLines().RemovePunctuation();  
var trimmedFriendlyName = friendlyName.Trim().Replace(" ", "");
```

با [ApplyCorrectYeKe](#) آشنایی دارید؛ همان یک دست کردن ی و ک فارسی و عربی است.

RemoveDiacritics همان [حذف اعراب](#) از کلمات است است.

متد پاکسازی underlineهای ویژه یا همان نامهای کش آمده، به صورت زیر است:

```
public static string CleanUnderLines(string text)  
{  
    if (string.IsNullOrEmpty(text))  
        return string.Empty;  
  
    const char chr1600 = (char)1600; //1600=  
    const char chr8204 = (char)8204; //8204=  
  
    return text.Replace(chr1600.ToString(CultureInfo.InvariantCulture), "")  
        .Replace(chr8204.ToString(CultureInfo.InvariantCulture), "");  
}
```

و متد RemovePunctuation :

```
public static string RemovePunctuation(string text)  
{  
    return string.IsNullOrEmpty(text) ? string.Empty : new string(text.Where(c =>  
        !char.IsPunctuation(c)).ToArray());  
}
```

این موارد، «حداقل»هایی هستند که باید جهت نرمال سازی اطلاعات، در حین ثبت نام اعمال شوند.

## نظرات خوانندگان

نویسنده: مجتبی آزاد  
تاریخ: ۱۸:۴۳ ۱۳۹۴/۰۵/۰۷

سلام

طبق دستوری که برای حذف کردن علامت + نوشته شده، بعد از اجرای کد ایمیل johndoe+spamsite@gmail.com به johndoespamsite@gmail.com تبدیل می‌شود. آیا بهتر یا صحیح‌تر نیست که مقدار بعد از علامت + حذف شود و ایمیل به این شکل در بیاید: johndoe@gmail.com ؟