

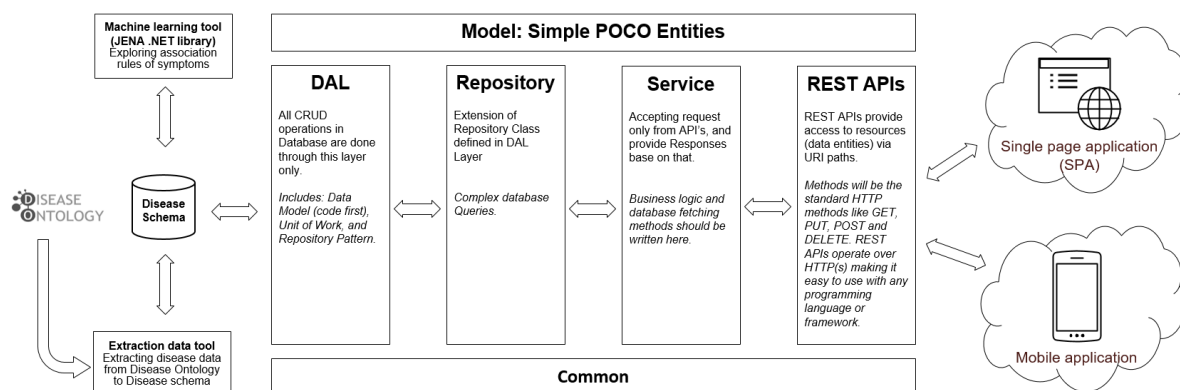
# Smart Doctor

## Architectural Design Description

# 1 Cấu trúc hệ thống

## 1.1 Tổng quan

Sơ đồ bên dưới mô tả những thành phần và mối quan hệ của chúng trong hệ thống:



Hình 1 – Cấu trúc khung của ứng dụng

Bảng diễn giải những thành phần trong cấu trúc:

STT	Thành phần	Diễn giải	Nguồn tham khảo
1	Disease Ontology	Được phát triển theo chuẩn bản thể học (Ontology) của những bệnh trên người.	<a href="http://disease-ontology.org/">http://disease-ontology.org/</a>
2	Extraction data tool	Công cụ dùng để trích xuất dữ liệu bệnh trong bản thể học về bệnh (Disease Ontology) sang cơ sở dữ liệu quan hệ.	
3	Machine learning tool	Công cụ dùng để khám quá luận quan hệ của những triệu chứng bệnh trong cơ sở dữ liệu bằng những phương thức của thư viện JENA .NET.	<a href="http://www.linkeddatatools.com/">http://www.linkeddatatools.com/</a>
4	Disease schema	Là cơ sở dữ liệu quan hệ, chứa thông tin bệnh.	
5	Common	Chứa những thực thể, phương thức chung được sử dụng ở những thành phần khác.	
6	Model: simple POCO entities	Chứa những mô hình thực thể của những bảng dữ liệu trong cơ sở dữ liệu bệnh. Công nghệ: Entity	<a href="https://www.asp.net/entity-framework">https://www.asp.net/entity-framework</a>

		framework	
7	Data access layer (DAL)	Tầng truy xuất cơ sở dữ liệu (thêm, xóa, sửa).	
8	Repository	Chứa những câu truy vấn phức tạp.	
9	Service	Hiện thức hóa quy trình nghiệp vụ	
10	REST APIs	Cung cấp những phương thức truy vấn dữ liệu của hệ thống thông qua URI cho những hệ thống khác.	<a href="https://www.asp.net/web-api">https://www.asp.net/web-api</a>
11	Single Page Application (SPA)	Ứng dụng website trên một trang với mục tiêu đưa trải nghiệm người dùng trên web như ứng dụng trên máy tính.  Công nghệ: HTML5, CSS3, Typescript, Angular.js	<a href="https://angularjs.org/">https://angularjs.org/</a>

## 1.2 Disease Ontology

Disease Ontology đã được phát triển như là một bản thể học chuẩn cho bệnh nhân với mục đích cung cấp cho cộng đồng y sinh thiếu phù hợp, tái sử dụng và bền vững về bệnh của con người, đặc điểm kiểu hình và khái niệm bệnh từ vựng y tế có liên quan thông qua các nỗ lực hợp tác của các nhà nghiên cứu tại Đại học Northwestern, trung tâm y học di truyền và Đại học Maryland School of Medicine, Viện Khoa học Di truyền.

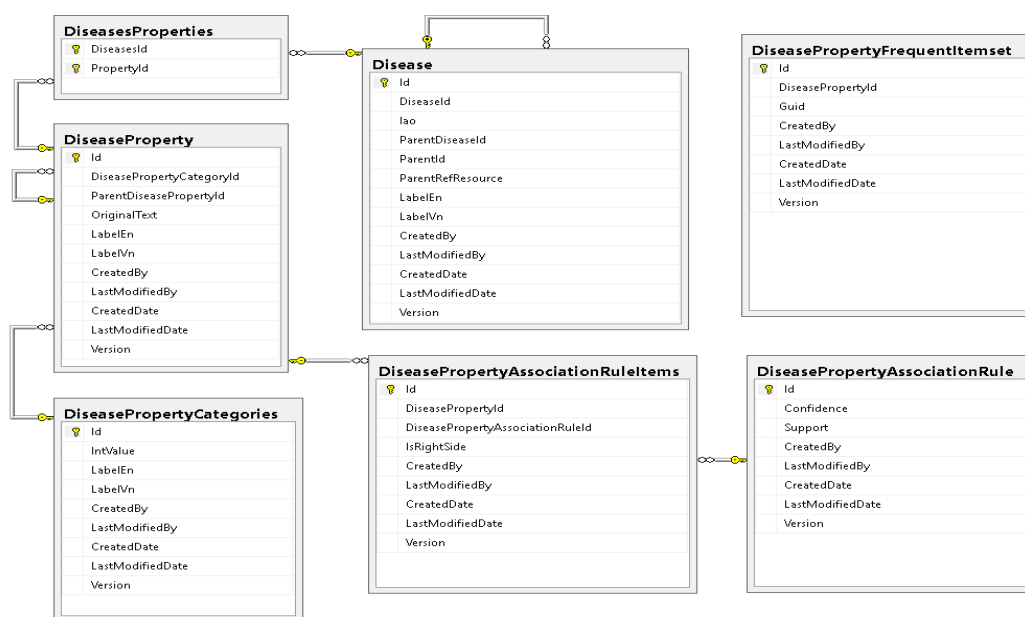
Ví dụ thông tin bệnh trong Disease Ontology:

Metadata		Visualize
DOID	DOID:14115	
Name	toxic shock syndrome	
Definition	<p>A commensal bacterial infectious disease that results in infection, has_material_basis_in <i>Streptococcus pyogenes</i> or has_material_basis_in <i>Staphylococcus aureus</i>, which produce toxins that are absorbed systemically and produce the systemic manifestations. The infection has_symptom fever, has_symptom rash, has_symptom hypotension, has_symptom multiorgan failure, has_symptom desquamation, has_symptom vomiting, has_symptom diarrhea, has_symptom headache, and has_symptom nonfocal neurologic abnormalities.</p> <p><a href="http://emedicine.medscape.com/article/169177-overview">http://emedicine.medscape.com/article/169177-overview</a>,  <a href="http://en.wikipedia.org/wiki/Toxic_shock_syndrome">http://en.wikipedia.org/wiki/Toxic_shock_syndrome</a></p>	
Xrefs	<p><a href="#">ICD10CM:A48.3</a>  <a href="#">ICD9CM:040.82</a>  <a href="#">MSH:D012772</a>  <a href="#">NCI:C35498</a>            SNOMEDCT_US_2016_03_01:18504008            UMLS_CUI:C0600327</p>	
Subsets	gram-positive_bacterial_infectious_disease	
Synonyms	<p>toxic shock [EXACT]            Toxic Shock syndrome [EXACT]            Toxic shock syndrome [EXACT]            Toxic shock syndrome (disorder) [EXACT]            TSS [EXACT]</p>	
Relationships	is_a <a href="#">commensal_bacterial_infectious_disease</a>	

Figure 2 - Mẫu ví dụ về một bệnh trong Disease Ontology

## 1.3 Disease database schema

### 1.3.1 Cấu trúc cơ sở dữ liệu



Hình 3 - Cấu trúc cơ sở dữ liệu bệnh

Bảng	Mô tả
Disease	Chứa thông tin bệnh được trích xuất từ Disease Ontology, như: <ul style="list-style-type: none"> <li>+ Bệnh cha (DiseaseParentId)</li> <li>+ Tên bệnh trong tiếng anh (LabelEn)</li> <li>+ Tên bệnh trong tiếng việt (LabelVn)</li> </ul>
DiseasePropertyCaterogies	Chứa những loại tính chất của bệnh, như: <ul style="list-style-type: none"> <li>• Phức tạp do (complicated by).</li> <li>• Gồm (composed of).</li> <li>• Xuất phát từ (derives from).</li> <li>• Có cơ sở vật chất (has material basis in).</li> <li>• Có triệu chứng (has symptom).</li> <li>• Thuộc về (inherits in).</li> <li>• Là một (is a).</li> </ul>

	<ul style="list-style-type: none"> <li>Nằm ở (located in).</li> <li>Xảy ra với (occurs with).</li> <li>Một phần của (part of).</li> <li>Thực hiện bởi (realized by).</li> <li>Thực hiện bởi loại bỏ với (realized by suppression with).</li> <li>Kết quả trong (results in).</li> <li>Kết quả trong sự hình thành của (results in formation of).</li> <li>Truyền qua (transmitted by).</li> </ul>
DiseaseProperty	Chứa những tính chất của bệnh, như: + Tính chất cha (DiseasePropertyId) + Tên tính chất bệnh trong tiếng anh (LabelEn) + Tên tính chất bệnh trong tiếng việt (LabelVn)
DiseaseProperties	Chứa mối quan hệ giữa bệnh và tính chất của bệnh.
DiseasePropertyFrequentItemset	Chứa thông tin những tính chất bệnh thường xuyên xuất hiện cùng nhau.
DiseasePropertyFrequentAssociationRule	Chứa thông tin của những luật quan hệ của những tính chất bệnh, như: + Độ hỗ trợ của luật (Support). + Độ tin cậy của luật (Confident).
DiseasePropertyFrequentAssociationRuleItems	+ Có liên kết với bảng "DiseasePropertyFrequentAssociationRule" và thể hiện những tính chất bệnh trong từng luật.

#### 1.4 Extraction data tool

Dựa vào cấu trúc thông tin với định dạng XML của bệnh trong "Disease Ontology", công cụ "Extraction data tool" mô hình hóa cấu trúc dữ liệu của "Disease Ontology" thành những lớp trong ngôn ngữ lập trình, và trích xuất dữ liệu đưa vào cơ sở dữ liệu bệnh (Disease schema) the những phương pháp sau:

### 1.4.1 Phương pháp rút thông tin bệnh:

```
<owl:Class rdf:about="http://purl.obolibrary.org/obo/DOID_0050004">
  <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">seminal vesicle acute gonorrhea</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://purl.obolibrary.org/obo/DOID_10400">
  <oboInOwl:id rdf:datatype="http://www.w3.org/2001/XMLSchema#string">DOID:0050004</oboInOwl:id>
  <oboInOwl:hasDbXref rdf:datatype="http://www.w3.org/2001/XMLSchema#string">ICD9CM:098.14</oboInOwl:hasDbXref>
  <oboInOwl:hasDbXref rdf:datatype="http://www.w3.org/2001/XMLSchema#string">SNOMEDCT_2010_1_31:301990003</oboInOwl:hasDbXref>
  <oboInOwl:hasDbXref rdf:datatype="http://www.w3.org/2001/XMLSchema#string">UMLS_CUI:C0578661</oboInOwl:hasDbXref>
  <oboInOwl:hasOBONamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#string">disease_ontology</oboInOwl:hasOBONamespace>
</owl:Class>
```

Figure 4 - Mẫu dữ liệu bệnh trong Disease Ontology

Mã số bệnh:

- Dữ liệu được chứa vào cột [DiseaseId] của bảng [Disease]
- Dữ liệu được rút từ tag "<owl:Class>", ví dụ từ mẫu trong hình 7:  
 <owl:Class rdf:about="http://purl.obolibrary.org/obo/DOID\_0050004">  
 => [DiseaseId] = "DOID:0050004"

Tên bệnh:

- Dữ liệu được chứa vào cột [LabelEn] của bảng [Disease]
- Dữ liệu được rút từ tag "<owl:label>", ví dụ từ mẫu trong hình 7:  
 <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">seminal vesicle acute gonorrhea</rdfs:label>  
 => [LabelEn] = "seminal vesicle acute gonorrhea"

Mã số bệnh cha:

- Dữ liệu được chứa vào cột [ParentDiseaseId] của bảng [Disease]
- Dữ liệu được rút từ tag "<owl:label>", ví dụ từ mẫu trong hình 7:  
 <rdfs:subClassOf rdf:resource="http://purl.obolibrary.org/obo/DOID\_10400"/>  
 => [LabelEn] = "DOID:10400"

### 1.4.2 Phương pháp rút loại thuộc tính bệnh:

```
<owl:ObjectProperty rdf:about="http://purl.obolibrary.org/obo/doi#has_symptom">
  <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">has_symptom</rdfs:label>
  <oboInOwl:hasOBONamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#string">disease_ontology
  </oboInOwl:hasOBONamespace>
  <oboInOwl:id rdf:datatype="http://www.w3.org/2001/XMLSchema#string">has_symptom</oboInOwl:id>
</owl:ObjectProperty>
```

Hình 5 - Mẫu dữ liệu loại thuộc tính bệnh trong Disease Ontology

Tên loại thuộc tính bệnh (tổng cộng 15 loại):

- Dữ liệu được chứa vào cột [LabelEn] của bảng [DiseasePropertyCategory]
- Dữ liệu được rút từ tag "<rdfs:label >", ví dụ từ mẫu trong hình 8:  
 "<rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">has\_symptom</rdfs:label>"  
 => Rút được [LabelEn] = "has\_symptom"

### 1.4.3 Phương pháp rút dữ liệu tính chất bệnh:

```
<owl:Class rdf:about="http://purl.obolibrary.org/obo/DOID_0050012">
  <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">chikungunya</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://purl.obolibrary.org/obo/DOID_934"/>
  <obo:IAO_0000115 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">A viral infectious disease
that results_in infection located_in joint, has_material_basis_in Chikungunya virus, which is
transmitted_by Aedes mosquito bite. The infection has_symptom fever, has_symptom arthralgia, and
has_symptom maculopapular rash.</obo:IAO_0000115>
  <oboInOwl:hasExactSynonym rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Chikungunya fever
</oboInOwl:hasExactSynonym>
  <oboInOwl:id rdf:datatype="http://www.w3.org/2001/XMLSchema#string">DOID:0050012</oboInOwl:id>
  <oboInOwl:hasOBONamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#string">disease_ontology
</oboInOwl:hasOBONamespace>
  <oboInOwl:inSubset rdf:resource="http://purl.obolibrary.org/obo/doi#zoonotic_infectious_disease"/>
</owl:Class>
<owl:Axiom>
  <owl:annotatedTarget rdf:datatype="http://www.w3.org/2001/XMLSchema#string">A viral infectious
disease that results_in infection located_in joint, has_material_basis_in Chikungunya virus, which
is transmitted_by Aedes mosquito bite. The infection has_symptom fever, has_symptom arthralgia,
and has_symptom maculopapular rash.</owl:annotatedTarget>
  <oboInOwl:hasDbXref rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
url:http://en.wikipedia.org/wiki/Chikungunya_virus</oboInOwl:hasDbXref>
  <owl:annotatedSource rdf:resource="http://purl.obolibrary.org/obo/DOID_0050012"/>
  <owl:annotatedProperty rdf:resource="http://purl.obolibrary.org/obo/IAO_0000115"/>
</owl:Axiom>
```

Hình 6- Mẫu dữ liệu chứa tính chất bệnh trong Disease Ontology

Tên tính của tính chất bệnh:

- Dữ liệu được chứa vào cột [LabelEn] của bảng [DiseaseProperty]
- Rút dữ liệu tính chất bệnh trong tag “<obo:IAO\_0000115 ...>” và “<owl:annotatedTarget...>”. Dữ liệu trong những tags này chứa nhiều tính chất bệnh, như:  
 “A viral infectious disease that results\_in infection located\_in joint, has\_material\_basis\_in Chikungunya virus, which is transmitted\_by Aedes mosquito bite. The infection has\_symptom fever”  
 => Trích xuất được thông tin của tính chất loại: “results\_in”, “has\_material\_basis\_in”, “transmitted\_by”, và “has\_symptom”.

Vì vậy, công cụ này sẽ dùng trích xuất dữ liệu bằng kỹ thuật “regular expression”, theo từng bước như sau:

Bước 1: chia nhỏ từng phần của đoạn dữ liệu lớn thành từng phần nhỏ, và đảm bảo một phần chỉ chứa duy nhất một tính chất của bệnh, như ví dụ trên thì có thể chia thành những phần sau:

```
“A viral infectious disease that”
“results_in infection”
“located_in joint,”
“has_material_basis_in Chikungunya virus, which is”
“transmitted_by Aedes mosquito bite. The infection”
“has_symptom fever”
```

Bước 2: sử dụng mẫu (pattern) của “regular expression” theo từng loại thuộc tính bệnh để rút chính thông tin, như ví dụ trên thì:

“has\_symptom fever” : dùng mẫu của loại “has\_symptom” là “has\_symptom (.\*)”, và có thể rút ra được dữ liệu “fever”.

Bảng dưới đây là những mẫu cho từng loại thuộc tính bệnh.

Loại thuộc tính bệnh	Bắt đầu bởi	Dùng mẫu (pattern)
Phức tạp do	complicated_by	@"complicated_by (.*)that{which{1}"
Gồm	composed_of	@"composed_of (.*)"
Xuất phát từ	derives_from;derived_from	@"derives_from (.*) and is", @"derives_from (.*) and/or", @"derives_from (.*) or (.*)", @"derives_from (.*) and", @"derives_from (.*) and (.*)", @"derives_from (.*) that", @"derives_from (.*){1}", @"derives_from (.*)"
Có cơ sở vật chất	has_material_basis_in	@"has_material_basis_in (.*)"
Có triệu chứng	has_symptom	@"has_symptom (.*) (.*) (.*) and (.*)", has_symptom (.*) and (.*) and (.*)", @"has_symptom (.*) (.*) and (.*) and\$", @"has_symptom (.*) or (.*) and\$", @"has_symptom (.*) and\$", @"has_symptom (.*) and\$", @"has_symptom (.*) (.*) or (.*)", @"has_symptom (.*) or\$", @"has_symptom (.*) or (.*)", @"has_symptom (.*)\$", @"has_symptom (.*)"
Thuộc về	inheres_in	@"inheres_in (.*)"
Là một	is_a	@"is_a (.*)"
Nằm ở	located_in	@"located_in (.*)"
Xảy ra với	occurs_with	@"occurs_with (.*)"
Một phần của	part_of	@"part_of (.*)"
Thực hiện bởi	realized_by	@"realized_by (.*)"
Thực hiện bởi loại bỏ với	realized_by supression_with	@"realized_by supression_with (.*)"
Kết quả trong	results_in	@"results_in (.*)"
Kết quả trong sự hình thành của	results_in_formation_of; resulting_in_formation_of	@"results_in_formation_of (.*)", @resulting_in_formation_of
Truyền qua	transmitted_by	@"transmitted_by (.*) and (.*)", @"transmitted_by (.*) (.*) and\$", @"transmitted_by (.*) and\$", @"transmitted_by (.*) and\$", @"transmitted_by (.*) (.*) or (.*) or\$", @"transmitted_by (.*) or (.*) or\$", @"transmitted_by (.*) or\$", @"transmitted_by (.*) or\$", @"transmitted_by (.*)\$", @"transmitted_by (.*) (.*)", @"transmitted_by (.*)"

#### 1.4.4 Phương pháp dịch dữ liệu sang tiếng việt:

Hệ thống hỗ trợ phương thức dịch từ ngữ từ tiếng anh tiếng việt, thông qua “Translation website” của Google:



```

private static string Translate(string text, string from, string to) {
    ClassTracer.Debug("Translate");
    var translateSuccess = false;
    var translatedText = string.Empty;
    while (!translateSuccess) {
        try {
            var wc = new WebClient();
            wc.Headers.Add(HttpRequestHeader.UserAgent, "Mozilla/5.0");
            wc.Headers.Add(HttpRequestHeader.AcceptCharset, "UTF-8");
            wc.Encoding = Encoding.UTF8;
            string url = string.Format(@"http://translate.google.vn/m?hl=en&sl={0}&tl={1}&ie=UTF-8&prev=_m&q={2}",
                @from, to, Uri.EscapeUriString(text));
            var page = wc.DownloadString(url);
            page = page.Remove(0, page.IndexOf("<div dir=\"ltr\" class=\"t0\">"));
            .Replace("<div dir=\"ltr\" class=\"t0\">", "");
            var last = page.IndexOf("</div>");
            translatedText = page.Remove(last, page.Length - last);
            translateSuccess = true;
        } catch (Exception exception) {
            Console.WriteLine(exception);
        }
    }
    return translatedText;
}

```

Figure 7 - Phương thức dịch ngôn ngữ

Dữ liệu sau khi được dịch sẽ được lưu vào cột “LabelVn” của những bảng như: “Disease”, “DiseaseProperty”, “DiseasePropertyCategory”.

#### 1.4.5 Kết quả trích xuất dữ liệu từ Disease Ontology

Bệnh: có 8796 bệnh trong bảng “Disease”

DiseaseId	ParentDiseaseId	LabelEn	LabelVn
1... DOID:10778	DOID:3951	fiedlers myocarditis	fiedlers viêm cơ tim
1... DOID:10779	DOID:3951	septic myocarditis	viêm cơ tim tị hoại
1... DOID:1078	NULL	Chromadorea infectious disease	bệnh truyền nhiễm chromadorea
1... DOID:10780	DOID:8432	primary polycythemia	đa hồng cầu chính
1... DOID:10782	DOID:1602	mesenteric lymphadenitis	viêm hạch mạc treo
1... DOID:10783	DOID:2860	methemoglobinemia	methemoglobinemia
1... DOID:10784	DOID:11104	Queensland tick typhus	Queensland tick sốt phát ban
1... DOID:10787	DOID:1414	premature menopause	mãn kinh sớm
1... DOID:10788	NULL	osteitis condensans	condensans viêm xương
1... DOID:1079	DOID:1080	setariasis	setariasis
1... DOID:10790	DOID:10791	chronic frontal sinusitis	viêm xoang trán mạn tính
1... DOID:10791	DOID:0050127	frontal sinusitis	viêm xoang trán
1... DOID:10792	DOID:2051	chronic maxillary sinusitis	viêm xoang hàm mạn tính
1... DOID:10793	DOID:10794	chronic sphenoidal sinusitis	viêm xoang mắt tĩnh sphenoidal
1... DOID:10794	DOID:0050127	sphenoid sinusitis	viêm xoang xương bướm
1... DOID:10795	NULL	congenital syphilitic encephalitis	viêm não syphilitic bẩm sinh
1... DOID:1080	DOID:883	filariasis	giun chỉ
1... DOID:10802	DOID:9401	acute gonococcal epididymo-orchitis	cấp tính do lậu epididymo-viêm tinh hoàn
1... DOID:10803	NULL	Haemophilus influenzae meningitis	Haemophilus influenzae viêm màng não

Query executed successfully. DESKTOP-GSD36K7 (12.0 RTM) | sa (52) | DiseaseOntologyDbNewDev | 00:00:00 | 8796 rows

Figure 8 - Mẫu dữ liệu trong bảng “Disease”

Loại thuộc tính bệnh: có 15 loại tính chất bệnh trong “DiseasePropertyCategories”

100 %

Results Messages

	Id	IntValue	LabelEn	LabelVn	CreatedBy	LastModifiedBy	CreatedDate
1	1	1	complicated_by	phức tạp do	1	1	2016-11-26 11:28:00.423
2	2	2	composed_of	gồm	1	1	2016-11-26 11:28:00.777
3	3	3	derives_from	xuất phát từ	1	1	2016-11-26 11:28:01.030
4	4	4	has_material_basis_in	có cơ sở vật chất	1	1	2016-11-26 11:28:01.220
5	5	5	has_symptom	có triệu chứng	1	1	2016-11-26 11:28:01.417
6	6	6	inheres_in	inheres trong	1	1	2016-11-26 11:28:01.650
7	7	7	is_a	là một	1	1	2016-11-26 11:28:01.850
8	8	8	located_in	nằm ở	1	1	2016-11-26 11:28:02.047
9	9	9	occurs_with	xảy ra với	1	1	2016-11-26 11:28:02.217
10	10	10	part_of	một phần của	1	1	2016-11-26 11:28:02.403
11	11	11	realized_by	thực hiện bởi	1	1	2016-11-26 11:28:02.577
12	12	12	realized_by_suppression_with	thực hiện bởi suppression với	1	1	2016-11-26 11:28:02.763
13	13	13	results_in	kết quả trong	1	1	2016-11-26 11:28:02.943
14	14	14	results_in_formation_of	kết quả trong sự hình thành của	1	1	2016-11-26 11:28:03.123
15	15	15	transmitted_by	truyền qua	1	1	2016-11-26 11:28:03.463

Query executed successfully. | DESKTOP-GSD36K7 (12.0 RTM) | sa (58) | DiseaseOntologyDbNewDev | 00:00:00 | 15 rows

Hình 9 - Mẫu dữ liệu trong bảng “DiseasePropertyCategories”

Thuộc tính bệnh: có 2444 thuộc tính bệnh trong bảng “DiseaseProperty”

Results Messages

	Id	DiseasePropertyCategoryId	LabelEn	LabelVn
295	295	5	muscle rigidity	cứng cơ
296	296	5	loss of conjugate eye movements	mất cử động mắt liên hợp
297	297	4	polioviruses	poliovirus
298	298	4	Human adenovirus 4	adenovirus Nhân 4
299	299	4	Human adenovirus 2	adenovirus Nhân 2
300	300	15	or talking from an infected person	hoặc nói chuyện của một người bệnh
301	301	15	droplet spread of oronasal secretions during ...	lây lan giọt dịch tiết oronasal trong khi ho, hắt hơi
302	302	5	ocular muscle palsies	liệt cơ mắt
303	303	5	reduced consciousness	giảm ý thức
304	304	4	Influenza B virus	Virus cúm B
305	305	4	Influenza A virus	Virus cúm A
306	306	15	or their contaminated tissues	hoặc bị ô nhiễm của họ mô
307	307	15	direct contact with sick person or animals	tiếp xúc trực tiếp với người hoặc động vật bị ốm
308	308	5	dizziness	chóng mặt
309	309	4	Nipah virus	virus Nipah
310	310	15	Mus musculus	Mus musculus
311	311	15	common house mouse	phổ biến chuột nhà
312	312	4	Lymphocytic choriomeningitis virus	Virus lymphocytic choriomeningitis

Query executed successfully. | DESKTOP-GSD36K7 (12.0 RTM) | sa (52) | DiseaseOntologyDbNewDev | 00:00:00 | 2444 rows

Hình 10 - Mẫu dữ liệu trong bảng "DiseaseProperty"

## 1.5 Machine learning tool

Dựa vào dữ liệu từ trích xuất từ “Disease Ontology”, thì những tính chất của bệnh có mối quan hệ với nhau, ví dụ:

DiseaseId	DiseaseLabelEn	DiseasePropertyId	DiseasePropertyLabelVn
1630	poliovirus type II nonparalytic poliomyelitis	7	fever
1630	poliovirus type II nonparalytic poliomyelitis	34	fatigue
1630	poliovirus type II nonparalytic poliomyelitis	36	headache
1630	poliovirus type II nonparalytic poliomyelitis	107	vomiting
1630	poliovirus type II nonparalytic poliomyelitis	226	neck stiffness
1630	poliovirus type II nonparalytic poliomyelitis	332	sore throat
1630	poliovirus type II nonparalytic poliomyelitis	561	direct contact with the oral secretions
1630	poliovirus type II nonparalytic poliomyelitis	562	water contaminated with feces
1630	poliovirus type II nonparalytic poliomyelitis	563	ingestion of food
1630	poliovirus type II nonparalytic poliomyelitis	564	destruction
1630	poliovirus type II nonparalytic poliomyelitis	569	Human poliovirus 2
1630	poliovirus type II nonparalytic poliomyelitis	574	muscle spasms
1630	poliovirus type II nonparalytic poliomyelitis	983	motor neurons

DiseaseId	DiseaseLabelEn	DiseasePropertyId	DiseasePropertyLabelVn
1627	poliovirus type III nonparalytic poliomyelitis	7	fever
1627	poliovirus type III nonparalytic poliomyelitis	34	fatigue
1627	poliovirus type III nonparalytic poliomyelitis	36	headache
1627	poliovirus type III nonparalytic poliomyelitis	107	vomiting
1627	poliovirus type III nonparalytic poliomyelitis	226	neck stiffness
1627	poliovirus type III nonparalytic poliomyelitis	332	sore throat
1627	poliovirus type III nonparalytic poliomyelitis	561	direct contact with the oral secretions
1627	poliovirus type III nonparalytic poliomyelitis	562	water contaminated with feces
1627	poliovirus type III nonparalytic poliomyelitis	563	ingestion of food
1627	poliovirus type III nonparalytic poliomyelitis	564	destruction
1627	poliovirus type III nonparalytic poliomyelitis	568	Human poliovirus 3
1627	poliovirus type III nonparalytic poliomyelitis	574	muscle spasms
1627	poliovirus type III nonparalytic poliomyelitis	983	motor neurons

Figure 11 - Mẫu dữ liệu bệnh trong CSDL

Trong 2 bệnh nêu trên có những tính chất bệnh cùng xuất hiện ở hai bệnh như: sốt (fever), đau đầu (headache), v....

Vì vậy, có thể sử dụng một số thuật toán của “machine learning” để khám phá ra mối quan hệ của những tính chất bệnh. Cụ thể, công cụ này sử dụng một số phương thức được hỗ trợ của thư viện JENA .NET để khám phá những quy luật này, và kết quả như sau:

Id	Confidence	Support	CreatedBy	LastModifiedBy	CreatedDate	LastModifiedDate	Version
1	100	0.258175559380379	1	1	2016-11-27 18:34:13.360	2016-11-27 18:34:13.360	0x00000000000016F31
2	60	0.258175559380379	1	1	2016-11-27 18:34:13.360	2016-11-27 18:34:13.360	0x00000000000016F32
3	100	0.258175559380379	1	1	2016-11-27 18:34:13.360	2016-11-27 18:34:13.360	0x00000000000016F33
4	100	0.344234079173838	1	1	2016-11-27 18:34:13.360	2016-11-27 18:34:13.360	0x00000000000016F34
5	60	0.258175559380379	1	1	2016-11-27 18:34:13.360	2016-11-27 18:34:13.360	0x00000000000016F35
6	60	0.258175559380379	1	1	2016-11-27 18:34:13.740	2016-11-27 18:34:13.740	0x00000000000016F36
7	100	0.258175559380379	1	1	2016-11-27 18:34:13.750	2016-11-27 18:34:13.750	0x00000000000016F37
8	50	0.344234079173838	1	1	2016-11-27 18:34:13.757	2016-11-27 18:34:13.757	0x00000000000016F38
9	100	0.258175559380379	1	1	2016-11-27 18:34:13.757	2016-11-27 18:34:13.757	0x00000000000016F39
10	100	0.344234079173838	1	1	2016-11-27 18:34:13.787	2016-11-27 18:34:13.787	0x00000000000016F3A
11	2.56410256410256	0.344234079173838	1	1	2016-11-27 18:34:13.787	2016-11-27 18:34:13.787	0x00000000000016F3B
12	60	0.258175559380379	1	1	2016-11-27 18:34:13.787	2016-11-27 18:34:13.787	0x00000000000016F3C
13	32.183908045977	4.81927710843374	1	1	2016-11-27 18:34:13.793	2016-11-27 18:34:13.793	0x00000000000016F3D
14	100	0.258175559380379	1	1	2016-11-27 18:34:13.787	2016-11-27 18:34:13.787	0x00000000000016F3E
15	100	0.258175559380379	1	1	2016-11-27 18:34:13.797	2016-11-27 18:34:13.797	0x00000000000016F3F

Query executed successfully. DESKTOP-GSD36K7 (12.0 RTM) sa (59) DiseaseOntologyDbNewDev 00:00:20 1355834 rows

Hình 12 - Khám phá 1.355.834 luật liên kết của tính chất bệnh

Id	Confidence	Support	DiseasePropertyId	LabelEn	IsRightSide
2	60	0.258175559380379	7	fever	0
2	60	0.258175559380379	107	vomiting	0
2	60	0.258175559380379	562	water contaminated with feces	0
2	60	0.258175559380379	574	muscle spasms	0
2	60	0.258175559380379	561	direct contact with the oral secretions	1
2	60	0.258175559380379	563	ingestion of food	1
2	60	0.258175559380379	568	Human poliovirus 3	1

Hình 13 - Mẫu dữ liệu về một luật liên kết

Theo hình bên trên, thì ta thông tin luật với mã số (ID) = 2 là:

- Độ hỗ trợ (Support) = 0.2581
- Độ tin cậy (Confidence) = 60
- Mối quan hệ:

{fever, vomiting, water contaminated with feces, muscle spasms} =>  
 {direct contact with the oral secretions, ingestion of food, human poliovirus 3}.

## 1.6 Common

Chứa những thực thể/phương thức chung được dùng trong hệ thống, như

- Thực thể chứa thông tin để chuyển qua mạng (data transfer object), nhằm làm giảm lượng dữ liệu được truyền qua mạng, tăng tốc độ xử lý.

```

1  L /
2
3  using Diagnosis.Common.Infrastructure;
4
5
6  namespace Diagnosis.Common.Dto {
7      -references
8      public class DiseaseViewDto : BaseViewDto {
9          #region IProperties
10
11          -references | 0 exceptions
12          public string DiseaseName {
13              get;
14              set;
15          }
16
17          1 reference | 0 exceptions
18          public string Link {
19              get;
20              set;
21          }
22          #endregion
23      }
24  }
  
```

Figure 14 – Lớp trao đổi dữ liệu của bệnh

- Những phương thức hỗ trợ (helpers), như: xử lý chuỗi, xử lý thực thể ...

```

public static class StringHelper {
    Constants

    #region SMethods
    0 references | 0 exceptions
    public static string AsUtf8(StringBuilder str) ...

    /// <summary> Returns the cropped string if it is longer than maxLength.
    0 references | 0 exceptions
    public static string CropString(this string input, int maxLength) ...

    /// <summary> Returns the cropped string if it is longer than maxLength. When useSuspensionPoints is true, the return st ...
    1 reference | 0 exceptions
    public static string CropString(this string input, int maxLength, bool useSuspensionPoints) ...

    /// <summary>
    /// Convert Decimals to string with current Culture.
    /// </summary>
    /// <param name="value">The value.</param>
    /// <returns></returns>
    0 references | 0 exceptions
    public static string DecimalToString(this decimal value) {
        return DecimalToString(value, CultureInfo.CurrentCulture);
    }
}
  
```

Figure 15 - Lớp hỗ trợ xử lý chuỗi (Stringhelper)

## 1.7 Model: Simple POCO entities

Mô hình thực thể của những bảng dữ liệu trong cơ sở dữ liệu bệnh.

```

32 references
public class DiseaseProperty : BaseDiagnosisObject {
    #region IProperties

    7 references | 0/1 passing | 0 exceptions
    public int DiseasePropertyCategoryId ...

    3 references | 0 exceptions
    public int? ParentDiseasePropertyId ...

    4 references | 0 exceptions
    public virtual DiseasePropertyCategory DiseasePropertyCategory ...

    1 reference | 0 exceptions
    public virtual ICollection<Disease> Diseases ...

    1 reference | 0 exceptions
    public virtual ICollection<DiseaseProperty> ChildrenDiseaseProperties ...

    1 reference | 0 exceptions
    public virtual DiseaseProperty ParentDiseaseProperty ...

    1 reference | 0 exceptions
    public virtual ICollection<DiseasePropertyAssociationRuleItem> DiseasePropertyAssociationRuleItems ...

    2 references | 0 exceptions
    public string OriginalText ...

    #endregion
}

```

Figure 16 -- Mô hình thực thể của bảng "DiseaseProperty"

```

2 references
public class DiseasePropertyMap : BaseDiagnosisObjectMap<DiseaseProperty> {
    #region IConstructors

    1 reference | 0 exceptions
    public DiseasePropertyMap() {
        // Table & Column Mappings
        ToTable("DiseaseProperty");
        Property(t => t.ParentDiseasePropertyId).HasColumnName("ParentDiseasePropertyId");
        Property(t => t.DiseasePropertyCategoryId).HasColumnName("DiseasePropertyCategoryId");
        Property(t => t.OriginalText).HasColumnName("OriginalText");

        // Relationships
        HasRequired(t => t.DiseasePropertyCategory)
            .WithMany(t => t.DiseaseProperties)
            .HasForeignKey(d => d.DiseasePropertyCategoryId);

        HasOptional(t => t.ParentDiseaseProperty)
            .WithMany(t => t.ChildrenDiseaseProperties)
            .HasForeignKey(d => d.ParentDiseasePropertyId);
    }

    #endregion
}

```

Figure 17 - Cấu hình cho đối tượng "DiseaseProperty" và bảng "DiseaseProperty"

## 1.8 Data access layer

Lớp cung cấp phương thức truy xuất cơ sở như sau:

```

public interface IGenericService<TEntity> where TEntity : IObjectState {
    #region IMethods
    1 reference | 0 exceptions
    void Delete(object id);
    2 references | 0 exceptions
    void Delete(TEntity entity);
    1 reference | 0 exceptions
    Task<bool> DeleteAsync(params object[] keyValues);
    2 references | 0 exceptions
    Task<bool> DeleteAsync(CancellationTokens cancellationTokens, params object[] keyValues);
    4 references | 0 exceptions
    TEntity Find(params object[] keyValues);
    1 reference | 0 exceptions
    Task<TEntity> FindAsync(params object[] keyValues);
    1 reference | 0 exceptions
    Task<TEntity> FindAsync(CancellationTokens cancellationTokens, params object[] keyValues);
    5 references | 0/1 passing | 0 exceptions
    void Insert(TEntity entity);
    1 reference | 0 exceptions
    void InsertGraphRange(IEnumerable<TEntity> entities);
    4 references | 0/1 passing | 0 exceptions
    void InsertOrUpdateGraph(TEntity entity);
    2 references | 0 exceptions
    void InsertRange(IEnumerable<TEntity> entities);
    1 reference | 0 exceptions
    IQueryable<TEntity> Query();
    1 reference | 0 exceptions
    IQueryable<TEntity> Query(IQueryObject<TEntity> queryObject);
    1 reference | 0 exceptions
    IQueryable<TEntity> Query(Expression<Func<TEntity, bool>> query);
    14 references | 0/1 passing | 0 exceptions
    IQueryable<TEntity> Queryable();
    1 reference | 0 exceptions
    IQueryable<TEntity> SelectQuery(string query, params object[] parameters);
    2 references | 0 exceptions
    void Update(TEntity entity);
    #endregion
}

```

Figure 18 - Giao tiếp truy xuất dữ liệu của những POCO

## 1.9 Repository layer

Tầng chứa những câu truy vấn phức tạp sử dụng “Lambda expression”

```

public class DiseaseQuery : QueryObject<Disease> {
    /// <summary>
    /// The tracer used for logging.
    /// </summary>
    [System.Diagnostics.CodeAnalysis.SuppressMessage("Microsoft.Security", "CA2104:DoNotDeclareReadOnlyMutableReferenceTypes")]
    public static readonly ITracer ClassTracer = TracerManager.GetLogger(typeof(DiseaseQuery));

    #region IMethods
    1 reference | 0 exceptions
    public DiseaseQuery HasSymptoms(params string[] symptoms) {
        ClassTracer.Debug("DiseaseQuery.HasSymptoms");
        foreach (var symptom in symptoms) {
            And(disease => disease.DiseaseProperties.Any(sym => sym.LabelVn.Equals(symptom, StringComparison.InvariantCultureIgnoreCase)
                && sym.DiseasePropertyCategoryId == (int)EDiseasePropertyCategoryId.HasSymptom));
        }
        return this;
    }
    1 reference | 0 exceptions
    public DiseaseQuery HasPropertyIds(params int[] symptomIds) {
        ClassTracer.Debug("DiseaseQuery.HasSymptoms");
        foreach (var symptomId in symptomIds) {
            And(disease => disease.DiseaseProperties.Any(sym => sym.Id == symptomId));
        }
        return this;
    }
    #endregion
}

```

Figure 19 – Ví dụ về câu truy vấn bệnh

## 1.10 Service layer

Tầng hiện thực những nghiệp vụ của hệ thống

```
public class DiseasePropertyServiceImpl : GenericService<DiseaseProperty>, IDiseasePropertyService {
    [SFields]
    [IConstructors]
    5 references | 0/1 passing | 0 exceptions
    public IList<DiseasePropertyViewDto> LoadRecommendDiseaseProperties(int[] diseasePropertyIds) {
        if (diseasePropertyIds.Any()) {
            var numberOfDiseasePropertyIds = diseasePropertyIds.Length;
            var smallestDiseasePropertyId = diseasePropertyIds.OrderBy(id => id).First();
            var diseasePropertyAssociationRuleItemRepository =
                m_repository.GetRepository<DiseasePropertyAssociationRuleItem>();
            return diseasePropertyAssociationRuleItemRepository.Queryable()
                .Where(item => item.DiseasePropertyId == smallestDiseasePropertyId &&
                    item.DiseasePropertyAssociationRule.DiseasePropertyAssociationRuleItems.Count(dpa => !dpa.IsRightSide) == numberOfDiseasePropertyIds &&
                    !item.DiseasePropertyAssociationRule.DiseasePropertyAssociationRuleItems.Any(
                        i => !i.IsRightSide && !diseasePropertyIds.Contains(i.DiseasePropertyId)))
                .OrderByDescending(item => item.DiseasePropertyAssociationRule.Confidence)
                .SelectMany(leftItem => leftItem.DiseasePropertyAssociationRule.DiseasePropertyAssociationRuleItems
                    .Select(rightItem => new DiseasePropertyViewDto {
                        Id = rightItem.DiseaseProperty.Id,
                        text = rightItem.DiseaseProperty.DiseasePropertyCategory.LabelVn + " : " +
                            rightItem.DiseaseProperty.LabelVn,
                        id = rightItem.DiseaseProperty.Id
                    })).ToList();
        }
        return new List<DiseasePropertyViewDto>();
    }
    4 references | 0/1 passing | 0 exceptions
    public IList<DiseasePropertyViewDto> LoadFrequentlyDiseaseProperties(int[] diseasePropertyIds) {...}
    4 references | 0/1 passing | 0 exceptions
    public IList<DiseasePropertyViewDto> LoadDiseaseSuggestionSearchTexts(string searchText) {...}
}
```

Figure 20 - Lớp hiện thực phương thức nghiệp vụ của bệnh

## 1.11 REST APIs

Cung cấp những phương thức truy vấn dữ liệu của hệ thống thông qua URI cho những hệ thống khách:

```

1 reference
public class DiseaseApiController : ApiController {
    [IConstructors]

    #region IMethods

    [HttpGet]
    0 references | 0 requests | 0 exceptions
    public IHttpActionResult GetSuggestionDiseaseProperties([FromUri] int[] diseasePropertyIds) {
        if (!diseasePropertyIds.Any()) {
            return BadRequest();
        }
        var recommendDiseaseProperties = Service.LoadRecommendDiseaseProperties(diseasePropertyIds);
        var frequentlyDiseaseProperties = Service.LoadFrequentlyDiseaseProperties(diseasePropertyIds);

        var suggestionDiseasePropertiesViewDto = new SuggestionDiseasePropertiesViewDto {
            RecommendDiseaseProperties = recommendDiseaseProperties,
            FrequentlyDiseaseProperties = frequentlyDiseaseProperties
        };

        return Ok(suggestionDiseasePropertiesViewDto);
    }

    [HttpGet]
    0 references | 0 requests | 0 exceptions
    public IHttpActionResult GetDiseaseSuggestionSearchTexts([FromUri] string searchText) {
        if (string.IsNullOrEmpty(searchText)) {
            return BadRequest();
        }
        var diseaseSuggestionSearchTextsResults = Service.LoadDiseaseSuggestionSearchTexts(searchText);
        return Ok(diseaseSuggestionSearchTextsResults);
    }

    #endregion
}

```

Figure 21 - DiseaseApiController

## 1.12 Single page application (SPA)

Sử dụng công nghệ web như: HTML5, CSS3, Typescripts, và Angular.js để làm ứng dụng website trên một trang với mục tiêu đưa trải nghiệm người dùng trên web như ứng dụng trên máy tính.

```

<div collapsible-ctx-panel same-as-container-height class="simple-search">
  <div class="row">
    <div class="col-md-12">
      <div class="search-box">
        <input class="form-control"
          type="text"
          maxlength="255"
          ng-trim="true"
          ng-model="searchText"
          id="inputSearch"
          name="inputSearch"
          ignore-security="true" />
        <span class="glyphicon glyphicon-search"></span>
      </div>
      <button type="submit"
        ng-click="performSearch()"
        class="btn btn-primary"
        ng-i18next="[title]person:search.simple.button.sucheStarten;disease:search.simple.button.sucheStarten"></button>
      </div>
    </div>
  </div>
</div>

```

Hình 22 - Ví dụ dùng Angular trong lập trình giao diện



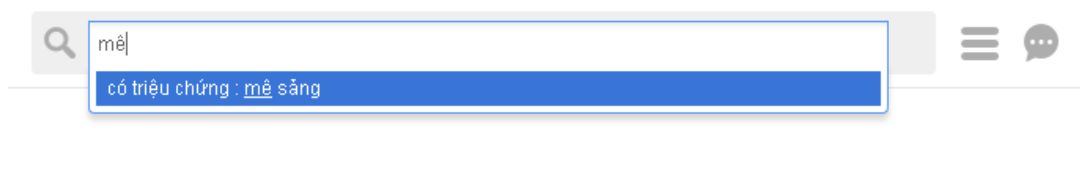
### **1.13 Ghi lại quá trình tìm kiếm của người dùng.**

Trong tìm kiếm của người sử dụng, hệ thống lưu lại các truy vấn của người dùng như các từ khóa tìm kiếm, bối cảnh của người sử dụng (thời gian, địa điểm); và cũng là tham khảo kết quả của người sử dụng như: kết quả mà bệnh được nhấp, bao lâu dùng sử dụng để đọc thông tin dịch bệnh, những nhận xét của người dùng về mỗi kết quả tìm kiếm của bệnh, và sự đánh giá của người dùng trên mỗi kết quả tìm kiếm của bệnh. Hệ thống sẽ lưu lại lịch sử tìm kiếm để khám phá các hành vi của người sử dụng, và điều chỉnh cơ sở dữ liệu bệnh trên đánh giá của người sử dụng.

## 2 Giao diện

### 2.1 Khung tìm kiếm

#### 2.1.1 Gợi ý từ khóa tìm kiếm



Hình 23 - Autocomplete trong khung tìm kiếm

Hệ thống này cung cấp tính năng gợi ý đề nghị để tăng tính tương tác giữa người sử dụng và hệ thống, trong một số tình huống dưới đây:

Khi người sử dụng chỉ cần nhập một số ký tự vào hộp tìm kiếm, hệ thống sẽ đề nghị truy vấn, mà là bắt đầu với những nhân vật này. Nó giúp người sử dụng điền các truy vấn nhanh hơn bằng cách lựa chọn câu hỏi từ gợi ý và sử dụng các tiêu chí tìm kiếm của chính xác hơn bằng cách đưa ra gợi ý chính tả.

Trong trường hợp, người sử dụng chỉ cần chỉ nhớ một từ khóa của các truy vấn, và họ có thể nhập các từ khóa vào ô tìm kiếm, hệ thống sẽ đề xuất các truy vấn, có chứa các từ khóa này. Nó giúp người sử dụng lựa chọn các truy vấn phù hợp với ý định của họ mà không cần nhớ toàn bộ truy vấn.

Ngoài ra, đề nghị gợi ý không chỉ giúp người dùng lựa chọn các truy vấn chính xác cho ý định của họ, mà còn giúp họ có các truy vấn thuộc vào từ điển của hệ thống. Điều này sẽ giúp sự tự tin của các truy vấn cao hơn và các kết quả tìm kiếm chính xác hơn.

#### 2.1.2 Gợi ý từ khóa liên quan

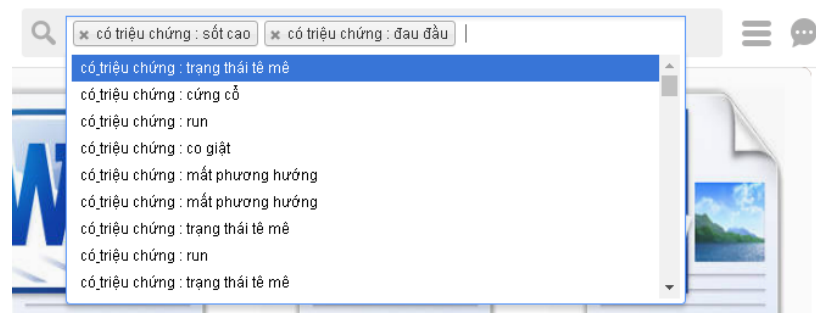


Figure 24 - Gợi ý từ khóa liên quan trong khung tìm kiếm

Hệ thống này cũng cung cấp đánh đề xuất tính năng để giới thiệu các triệu chứng có liên quan hoặc thuộc bệnh khác, ví dụ như khi người sử dụng lựa chọn một số từ khóa tìm kiếm là "sốt" (fever), "đau đầu" (headache), và họ dừng lại để nghĩ cho tìm kiếm tiếp theo từ khóa; lúc này hệ thống sẽ giới thiệu các triệu chứng liên quan với các triệu chứng bình chọn là "co giật" (convulsions). Các triệu chứng liên quan được khám phá từ các quy tắc quan hệ của các triệu chứng [1], và các triệu chứng liên quan được đặt hàng (trên xuống dưới) bởi sự tự tin của quy tắc của các triệu chứng liên quan.

### 2.1.3 Tìm kiếm kết quả bệnh

Khi người dùng thực hiện tìm kiếm thì hệ thống sẽ tìm bệnh tương ứng với những thuộc tính bệnh mà người dùng đã gõ vào, với những lựa chọn kết quả tìm kiếm như sau:

- Dữ liệu hệ thống chọn ra bệnh có thuộc tính đang tìm kiếm, bệnh được sắp xếp lượng đã xem và đánh giá đó, và nội dung của những bệnh này được cung cấp bởi những tài liệu trong hệ thống.
- Dữ liệu bên ngoài, ứng với mỗi bệnh từ dữ liệu hệ thống thì hệ thống sẽ tìm kiếm trên Google với từ khóa là tên bệnh, và lấy những kết quả (top 3) được đánh giá cao để gửi ý người dùng.

## 2.2 Hiện thị kết quả tìm kiếm

Kết quả hiện thị từng khối cho mỗi bệnh, và được sắp xếp từ trái sang phải, từ trên xuống dưới. Mỗi khối bệnh được hiện thị như sau:



Hình 25 – Hiện thị kết quả tìm kiếm bệnh

- Phần trên hiển thị ảnh đại diện của bệnh
- Phần giữa hiển thị tên bệnh, và số lượng người xem bệnh này.
- Phần cuối hiện đánh giá của người dùng về nội dung bệnh.

## 2.3 Hiện thị thông tin bệnh

Khi người dùng đã thấy kết quả tìm kiếm, người dùng có thể chọn (click) vào một bệnh thì thông tin về bệnh sẽ được hiện thị trên popup, như hình bên dưới:

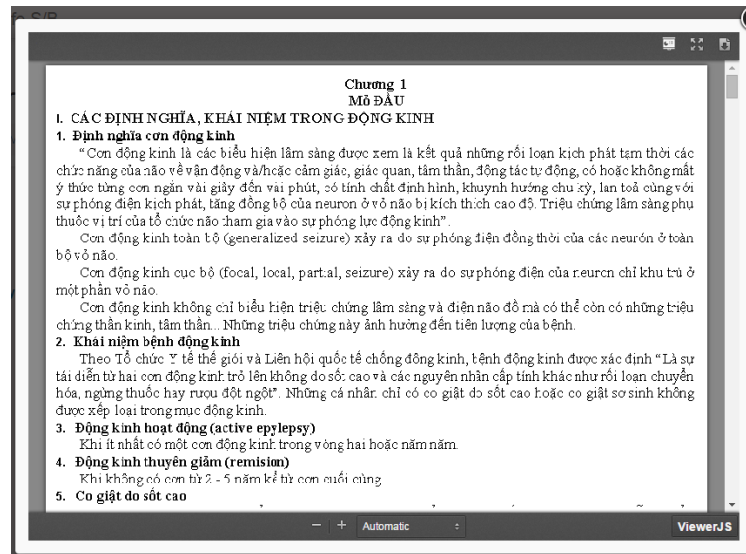


Figure 26 - Popup hiển thị thông tin bệnh

## 2.4 Người dùng phản hồi/đánh giá ứng dụng

Màn hình phản hồi/đánh giá sẽ giúp ứng dụng nhận được phản hồi, đánh giá của người dùng về ứng dụng nhằm cải thiện chất lượng ứng dụng:

Đánh giá

Ghi chú

Email

Rating

★ ★ ★ ★ ★

Gửi

Hình 27 - Màn hình phản hồi/đánh giá

## Nguồn tham khảo

Disease Ontology	<a href="http://disease-ontology.org/">http://disease-ontology.org/</a>
JENA .NET	<a href="http://www.linkeddatatools.com/">http://www.linkeddatatools.com/</a>
Entity Framework	<a href="https://www.asp.net/entity-framework">https://www.asp.net/entity-framework</a>
REST APIs	<a href="https://www.asp.net/web-api">https://www.asp.net/web-api</a>
Angular JS	<a href="https://angularjs.org/">https://angularjs.org/</a>

## **Các từ viết tắt**

ADD	Architectural Design Description
HTML5	Hyper Text Markup Language version 5
CSS3	Cascading Style Sheet version 3
API	Application Programming Interface
POCO	Plain Old CLR Object
SPA	Single page application

