

# *Clariisa, uma Plataforma para a Integração de Dados para um Sistema de Governança em Saúde*

Odara Sena, Mariano Roberval, Martinho Vieira, Regis Magalhães, Vania Vidal, Mauro Oliveira

**Resumo**—Lariisa é um sistema inteligente para tomada de decisão em gestão da saúde. Este artigo apresenta o Clariisa, uma plataforma de integração de dados para o Lariisa. O objetivo é permitir a integração de várias bases de dados em saúde, com diferentes questões de governança envolvidas, possibilitando a interoperabilidade entre estas diversas fontes de dados. Um protótipo está sendo desenvolvido utilizando os conceitos do Framework Linked Data Mashup - LDMF, com Linked Data Mashup Service - LIDMS. Com isso, o Clariisa possibilitará a construção de mashups Linked Data com integração de dados de saúde de entidades públicas e privadas.

**Palavras-chave:** *Linked Data Mashup, Ontology, Context-Aware, Health-Governance System.*

**Abstract**—Lariisa is an intelligent system for decision-making in health management. This article introduces the Clariisa, a data integration platform for the Lariisa. The goal is to enable the integration of various databases in health, with different governance issues involved, enabling interoperability between these multiple sources of data. A prototype is being developed using the Framework concepts Linked Data Mashup-LDMF, with Linked Data Mashup Service-LIDMS. With this, the Clariisa will enable the construction of Linked Data mashups with health data integration of public and private entities.

**Keywords**— *Linked Data Mashup, Ontology, Context-Aware, Health-Governance System.*

## I. INTRODUÇÃO

Governança é um termo geral que conota uma série de tendências relacionadas na administração pública e gestão de políticas públicas cujo objetivo é alavancar o conhecimento disponível (em uma comunidade, por exemplo) para melhorar o desempenho administrativo e a democratização nos processos de tomada de decisão local. Por exemplo, a governança urbana visa promover uma relação mais estreita da sociedade civil com as organizações públicas para a melhoria do bem-estar em grandes cidades [1] [2] [3]. O conceito de governança também está sendo usado por organizações internacionais como o Banco Mundial e as Nações Unidas.

O uso da tecnologia da informação aplicada à governança da saúde tem sido alvo de várias discussões atualmente. Gestores de saúde sentem grande necessidade de soluções inteligentes que podem ajudar na tomada de decisões. As TIC's (Tecnologias de Informação e Comunicação) vem mostrando ser um pilar importante para o estabelecimento de governança, tanto pelas possibilidades que o uso de

dispositivos móveis, pequenos e mais robustos, traz para o sistema de saúde, quanto pelos sistemas inteligentes que podem ser criados para auxiliar diretamente os gestores de saúde em sua tomada de decisão.

Ainda relacionado a governança de saúde, é notório que um dos grandes problemas da gestão da informação em saúde é a fragmentação dos serviços de saúde, aumentando as dificuldades das autoridades de manter a integração do sistema, com danos para a sua governança, que se dissocia e se pulveriza. É sabido, por exemplo, que uma das grandes buscas do DATASUS é a integração de seus dados, espalhados em diversas fontes e sistemas [19].

O projeto Lariisa [21] [5] [6], um *framework* sensível ao contexto para a tomada de decisão em governança de saúde pública, atende aos requisitos buscados pelo DATASUS. O Lariisa propõe realizar inferência de informações a partir de cinco domínios de governança: (i) conhecimento, (ii) normativo, (iii) clínico-epidemiológico, (iv) administrativo e (v) gerenciamento compartilhado, a fim de permitir que o Gerente de Saúde possa tomar as melhores decisões possíveis [7].

A Web semântica destaca-se na evolução do projeto Lariisa. Ela fornece tecnologias para efetivamente publicar, recuperar e descrever dados distribuídos na Web. A integração de dados em grande escala é provavelmente um dos melhores casos de uso para as tecnologias de Web semântica.

Essa integração de dados na Web Semântica pode ser alcançada através do uso de *Linked Data*. *Linked Data* é um conjunto de princípios que se baseiam nas tecnologias da Web Semântica, como o RDF, *Remote Description Framework*, e permite reduzir a complexidade de integração de dados devido às ligações corretamente estabelecidas e descritas entre fontes [20].

Diante desse contexto, a arquitetura inicial do Lariisa está evoluindo para uma versão chamada Clariisa, onde passa a considerar as tecnologias abordados anteriormente: Web Semântica e *Linked Data*. Acredita-se que a aplicação de técnicas e métodos de integração de dados pode trazer grandes benefícios à integração sistêmica para as diversas aplicações do domínio de governança eletrônica.

Conforme a tabela I, disponível no site da *LOD Cloud* - <http://lod-cloud.net/state/> -, a área governamental responde por

mais de 42% dos dados linkados abertos e os das Ciências da Vida, no que se inserem os da Saúde, corresponde a mais de 9%, totalizando mais de 50% dos dados existentes.

Apesar de toda esta representatividade e os esforços realizados por universidades, pelo GT-INDA (Grupo de Trabalho da Infraestrutura Nacional dos Dados Abertos), pelo Data Gov do Brasil e ações de instituições, Estados e Municípios, o Brasil ainda não tem disponibilizado de forma massiva os seus dados abertos, na forma de *Linking Open Government Data* – LOGD. Por este aspecto já seria importante a disponibilização dos dados de saúde, como *Linked Data*.

Analisando os dados disponibilizados relacionados à saúde e as atividades desenvolvidas pelo LARISSA observa-se que a disponibilização dos dados de saúde como LOGD possibilita a adição desses dados na *LOD Cloud*, possibilitando a integração das diversas ações relacionadas à Saúde, em diferentes esferas de governos.

TABELA I. REPRESENTATIVIDADE DOS LINKED DATA BY DOMAIN

Domain	Number of datasets	Triples	%	(Out-)Links	%
Media	25	1,841,852,061	5.82 %	50,440,705	10.01 %
Geographic	31	6,145,532,484	19.43 %	35,812,328	7.11 %
Government	49	13,315,009,400	42.09 %	19,343,519	3.84 %
Publications	87	2,950,720,693	9.33 %	139,925,218	27.76 %
Cross-domain	41	4,184,635,715	13.23 %	63,183,065	12.54 %
Life sciences	41	3,036,336,004	9.60 %	191,844,090	38.06 %
User-generated content	20	134,127,413	0.42 %	3,449,143	0.68 %
	295	31,634,213,770		503,998,829	

Sendo assim, este artigo especifica uma plataforma para publicar dados de saúde do governo, originados de várias fontes e padrões, na Web, utilizando *Linked Data Mashups*. Tais informações vinculadas poderão ser consumidas, manipuladas e usadas pelo Lariisa ou outros aplicativos da Web.

Este trabalho esta organizado da seguinte forma: a Seção II apresenta a plataforma Lariisa; a Seção III apresenta informações sobre *Linked Data Mashups*; a Seção IV apresenta trabalhos relacionados relevantes; a Seção V apresenta a proposta do Clariisa. Por fim, a Seção VI apresenta as conclusões e trabalhos futuros.

II. PLATAFORMA LARISSA

O LARISSA está centrado no conceito de informação de contexto de saúde [4] [11] [12] [13], caracterizando situações de entidades (membro da família, um agente de saúde, gestor da saúde, entre outros, que são considerados relevantes para as interações entre um usuário e um sistema de saúde em um sistema de saúde).

Este contexto é formalmente definido a fim de facilitar sua representação, o compartilhamento e a interoperabilidade semântica no sistema de governança da saúde.

Para este fim, o LARISSA define duas ontologias OWL-DL para a modelagem de informações de contexto de saúde local e global. Contexto de saúde local descreve a situação de qualquer entidade interagindo com o sistema de governança, tais como usuários finais (pacientes), gestores de saúde, agentes de saúde, etc. Essas informações são utilizadas para a definição de regras de decisão locais de saúde e para construir o contexto de saúde

global que descreve informações de alto nível, derivado do contexto de saúde local, e é utilizado para tomada de decisão em governança de saúde. Por exemplo, o contexto de saúde global descreve o número de casos de dengue confirmados em uma região (ex: bairro, cidade, comunidade), durante um determinado período de tempo (ex: um dia, uma semana). Portanto, essas informações podem ser vistas como indicadores globais utilizados para melhorar as decisões de governança.

Os contextos de saúde local e global são classificados em seis dimensões:

- Espacial – quaisquer informações que caracterizem a situação da dimensão espacial (ex: localização, local, coordenadas GPS).
- Temporal – quaisquer informações que caracterizem a situação da dimensão do tempo (ex: instante, intervalo, período do dia, período do mês, período do ano, estação).
- Espaço-Temporal – quaisquer informações que caracterizam a situação que é dependente tanto da dimensão espacial quanto da dimensão temporal (ex: condições climáticas, temperatura, ruído, luminosidade).
- Social – quaisquer informações que caracterizem a situação dos relacionamentos sociais.
- Computacional – quaisquer informações que descrevem a situação das características computacionais (ex: configuração de dispositivos do usuário).
- Elemento de saúde – classifica o contexto da informação a partir do ponto de vista da saúde (ex: batimento cardíaco, pulso, pressão sanguínea).

O LARISSA reutiliza conceitos do GeographicallyEncoded Objects for ReallySimpleSyndicationfeeds (GeoRSS), uma simples marcação com informação de localização para descrição de coordenadas e relações geo-espaciais, assim como o OWL-Time que é utilizado para representar conteúdo temporal. Tais conceitos foram baseados no trabalho de [8].

O Lariisa define as classes Local\_Health\_Context (ver modelo da Figura 1) e Global\_Health\_Context (ver modelo da Figura 2) para representar os conceitos do contexto.

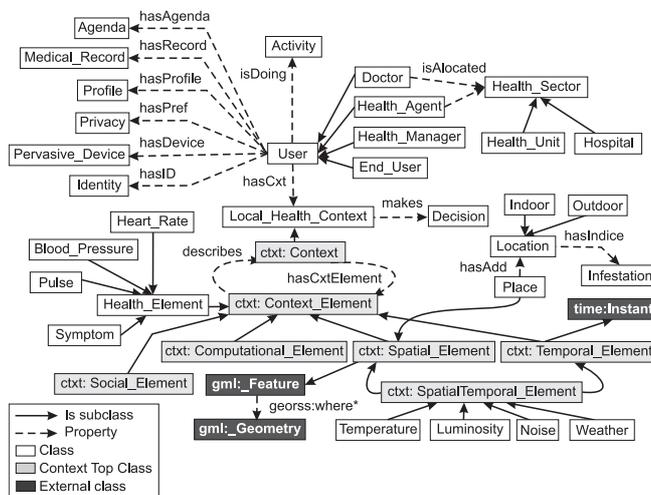


Fig. 1. Modelo de Saúde Local do LARISSA

Estes conceitos capturam do contexto quaisquer informações para caracterizar uma situação que é relevante para contribuir com decisões em governança de saúde, isto é, que podem ser utilizadas para definir regras de decisão local e global.

O *framework* utiliza a base do modelo ECA (*Event-Condition-Action*) para descrever regras de decisão local e global que são traduzidas dentro de regras utilizando a *Semantic Web Rule Language* (SWRL), “uma linguagem com sintaxe de abstração de alto nível para regras da OWL.” [Horrocks, 2004]. Um evento representa a identificação de mudanças no contexto. Uma condição descreve um conjunto válido de restrições de contexto, e uma ação descreve uma decisão.

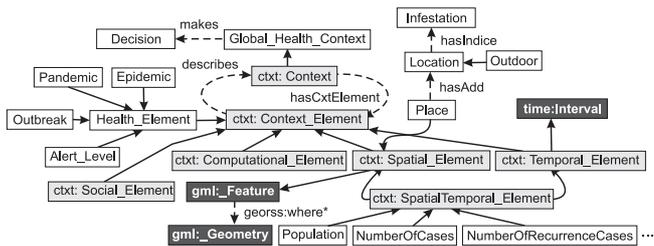


Fig. 2. Modelo de Saúde Global do LARIISA

Algumas aplicações são resultantes do framework Lariisa, como por exemplo, as especificadas em [14] e [15].

### III. LINKED DATA MASHUP

*Mashup* é uma aplicação Web que interativamente combina conteúdo de vários serviços ou fontes gerando uma nova fonte de dados ou serviço. Existe uma dificuldade em projetar um *mashup* da forma convencional, já que seria necessário uma linguagem de alto nível capaz de consultar fontes heterogêneas e conhecimento do usuário sobre essa linguagem de consulta, das URIs das fontes de dados e seus vocabulários. Ainda, de acordo com [17], o usuário fará uso de diferentes APIs Web para acessar diferentes fontes e serviços.

*Linked Data Mashup* surge como uma solução para a integração de dados de forma semântica de acordo com o que se segue: (i) facilidade de criação e manutenção em relação ao método de um *mashup* convencional. *Linked Data Mashup* identifica as relações entre os vocabulários das fontes e uma linguagem de consulta de alto nível (por exemplo: SPARQL) poderá ser usada para consultar um conjunto de dados heterogêneos. (ii) quantidade crescente de dados RDF publicados na Web de acordo com *Linked Data*.

A arquitetura de *Linked Data Mashup*, mostrada na Figura 3, define: (i) como os dados serão acessados, (ii) como são definidos para descoberta e reuso, (iii) como são executadas as consultas para recuperação de dados e (iv) como o conjunto de dados são acessados.

### IV. TRABALHOS RELACIONADOS

Alguns trabalhos anteriores trataram da integração de dados como *Linked Data*. Em [16], é proposto o DIGO–Delivery Information of Government. O DIGO permite acesso a dados governamentais primários por máquinas na forma de dados abertos de forma que os cidadãos interessados em ter acesso a esses dados possam combiná-los e produzir novas

informações e aplicações *mashup*, consequentemente habilitando o OGD e a fusão de dados no LOD Cloud.

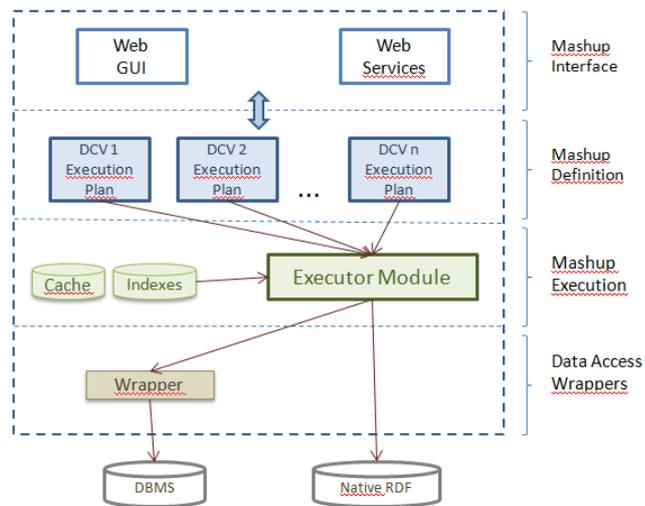


Fig. 3. Arquitetura *Linked Data Mashup*.

A arquitetura DIGO é dividida em 5 camadas. Figura 4 ilustra a arquitetura e as camadas.

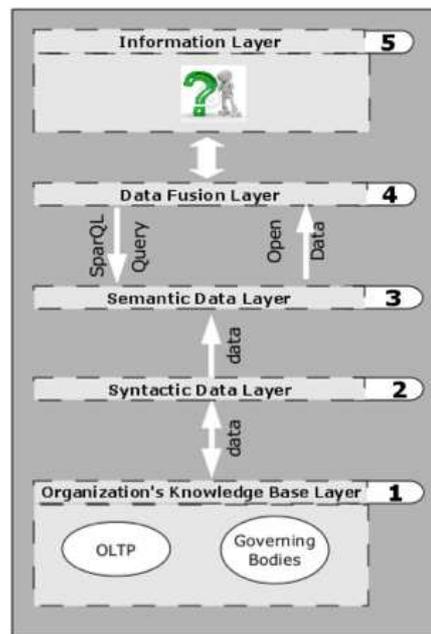


Fig. 4. Camadas da Arquitetura DIGO

Em especial, o trabalho desenvolvido na camada 3 dessa arquitetura foi norteadora para chegarmos a proposta do Clariisa.

Um processo comum e bem definido para a criação de um projeto de *Linked Data Mashup* é proposto por [18], através de uma especificação de alto nível do conteúdo dinâmico de dados, sobre uma visão unificada dos conjuntos de dados que serão consultados. Com base nestas especificações, é possível gerar automaticamente a consulta que recuperará estes dados, bem como os planos de execução de tais consultas sobre as diversas fontes, no contexto de *Linked Data*, do qual obtemos o entendimento das relações entre as fontes e na integração

semântica entre estas. É com base nessas características de recuperação de dados que propomos, na seção V, uma forma eficaz de integrar ao LARIISA serviços de *context-aware*, base de dados de instituições de saúde e disponibilização de dados de forma semântica.

V. PROPOSTA DO CLARIISA

LARIISA será usado como uma plataforma de software que contém muitos serviços orientados para a publicação de dados abertos, que permitem a sua integração com dados de outras fontes de dados. O módulo de integração descrito abaixo terá a função de consultar e extrair dados específicos de sistemas *Context-Aware* voltados para saúde, informações de Ministério e Secretarias de Saúde, ANVISA ou qualquer outro sistema que tenha informações contundentes para a Governança em Saúde.

Linked Data Mashups foi a tecnologia escolhida para integrar dados de fontes heterogêneas ao LARIISA, devido as características de facilidade de consulta, através de uma linguagem de alto nível (SPARQL) e publicação de dados abertos.

Seguindo o processo proposto de criação de aplicações Linked Data Mashup em [18] e focando nas necessidades e ambiente da plataforma de governança em saúde, LARIISA, é apresentado na Figura 5 a arquitetura do módulo de integração proposto.

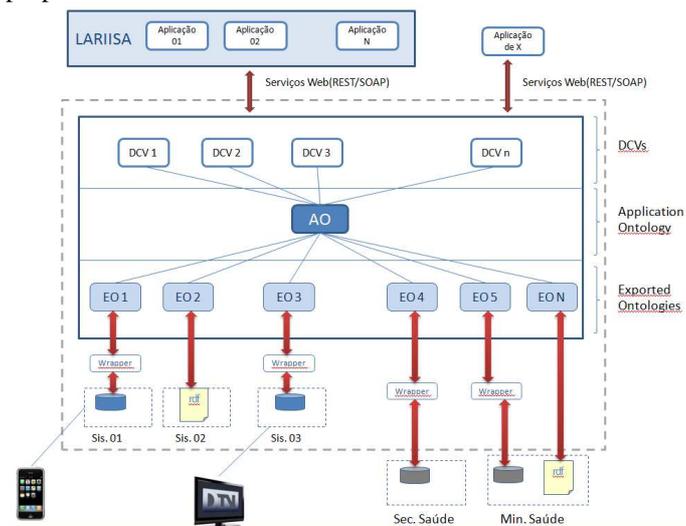


Fig. 5. Módulo de Integração do Clariisa.

A partir da Ontologia de Aplicação (OA) é definido o esquema conceitual global do mashup, que é elaborado através do PROTÈGE, de acordo com as necessidades das aplicações do LARIISA. Serão usados vocabulários padrão, como FOAF, GeoNames, DBpedia, Diseaseome, DailyMed, DrugBank, para a criação dessa ontologia.

Quando os dados da OA não estão relacionados a um vocabulário comum publicado em Linked Data, se faz necessário um mapeamento entre as estruturas de dados RDF ou DBMS (neste caso usa-se um wrapper) dos sistemas periféricos para a OA. Entende-se sistemas periféricos, sistemas Context-Aware, que usam serviços mobile e da TV Digital Brasileira [9] [10] ou base de dados de instituições.

Para cada um desses será gerado uma Ontologia Exportada que é a ontologia que representa o próprio sistema e descreve um conjunto de dados de acordo com o vocabulário definido na OA.

Os Dynamic Content Vision (DCV) consiste de uma quádrupla  $\langle S, P, Q, E \rangle$ , onde  $S$  é o esquema XML que define os atributos e o resultado esperado,  $P$  é uma lista de parâmetros definidos com base na troca de dados entre as páginas do mashup,  $Q$  é uma consulta SPARQL parametrizada,  $E$  é um conjunto de links especificando a navegação com outras DCVs, se houver. Desta forma, um DCV define a consulta sobre a OA obtendo assim os dados que será passado a aplicação requisitante do LARIISA através de web services, com REST ou SOAP, criando segurança e sigilo ao acesso dos dados.

Na arquitetura exposta fica claro a facilidade de integração de dados em um ambiente tão heterogêneo. O uso de Linked Data Mashup também facilita a divulgação de um documento RDF para outra aplicação ou usuário. Uma DCV poderá ser criada para disponibilizar Dados Abertos.

O módulo integrará serviços disponibilizado pelos Dados Abertos Governamentais, podendo usá-los e contribuir também com este projeto. Um exemplo disto é o uso de dados disponibilizados pelo Ministério da Saúde, como mostra a Figura 5.

O uso dessa tecnologia permite uma aplicação de Linked Data Mashup usar outra aplicação LDM. Nesse sentido, o Clariisa poderia utilizar mashups existentes, como o Mashup de Doenças e Drogas [22] e o de compras municipais para informações sobre as compras municipais de remédios, e possibilitar para o gestor de saúde verificar alguma irregularidade na compra de remédios efetuada pelos municípios. Ver Figura 6.

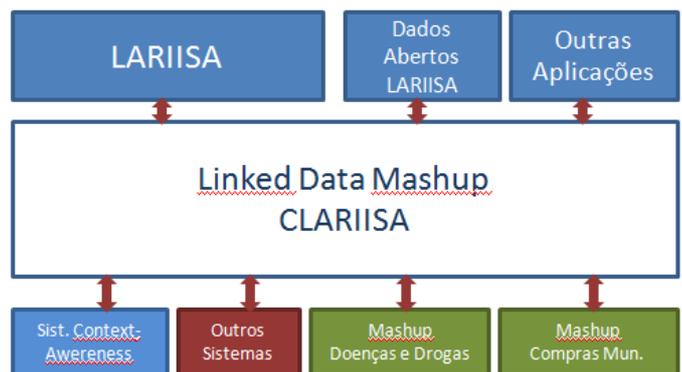


Fig. 6. Exemplo de Aplicação do uso do Linked Data Mashup Lariisa

VI. CONCLUSÃO

Este artigo apresentou a plataforma Clariisa, uma evolução do Lariisa, um projeto que busca realizar inferências que auxiliem a tomada de decisão pelo gestor de saúde, baseado no contexto dos atores envolvidos (dados coletados usuários do sistema de saúde, perfil dos agentes de saúde, localização destes atores, etc.), conhecimentos em diversas bases de informações integradas de saúde.

Esses dados, conhecimentos e informações serão alcançados a partir da arquitetura de integração de dados de saúde proposto. Com o CLariisa poderemos capturar dados abertos pulverizados em diversas fontes e integrá-los de forma a se conseguir informações linkadas que poderão ser consumidas tanto por aplicações do próprio Lariisa, como por outras aplicações.

O *Linked Data Mashup Lariisa* poderá ser utilizado, por exemplo, pelo DATASUS para consumir informações integradas de saúde. Na verdade, o CLariisa é um dos eixos estratégicos do NextSaúde, um projeto de utilização de TICs em saúde que deverá apoiar o DATASUS na pesquisa e desenvolvimento de soluções para o setor.

Como trabalhos futuros temos o desenvolvimento de aplicações, *endpoints* de consulta, com domínios específicos de informações de saúde como forma de servir de estudos de caso para a plataforma proposta. Por exemplo, poderíamos unir os *mashups* de Doenças & Drogas e de compras municipais para possibilitar para o gestor de saúde verificar alguma irregularidade na compra de remédios pelos municípios, como mencionado na sessão anterior.

## VII. REFERÊNCIAS

- [1] Scherer-Warren, I. Cidadania sem fronteiras: ações coletivas na era da globalização. Rio de Janeiro: Hucitec, 1999.
- [2] Castells, M. A sociedade em rede. A era da informação: economia, sociedade e cultura, v.1, São Paulo: Paz e Terra, 1999.
- [3] Putnam, R. Bowling Alone: the Collapse and Revival of American Community. *New York: Simon & Schuster, First Touchstone Edition, 2001*
- [4] Dey, A.; Abowd, G. Towards a Better Understanding of Context and Context-Awareness. In: *Workshop on the what, who, where, when and how of context-awareness*, CHI, Abril 2000.
- [5] Oliveira M., Andrade O.M., Hairon C.G., Moura R.C, Fernandes S., Bringel J., Gensel J., Martin H., Sicotte C., Denis J-L. “A Context-Aware Framework for Health Care Governance Decision-Making Systems: A model based on the Brazilian Digital TV”, in *Second IEEE Workshop on Interdisciplinary Research on E-health Services and Systems (IREHSS)*.
- [6] Oliveira, A.M.B., Et al. “A Context-Aware Framework for Healthcare Governance Decision-Making Systems: A model based on the Brazilian Digital TV”. In. *CRIFPE, centre de recherche interuniversitaire sur la formation et la profession enseignante. Université de Montréal, Canadá, 2010.*
- [7] Antunes, Franciano; Oliveira, A. M. B. ; Andrade, O. M. ; Bringel, J. R.; Barreto, I. “Um Protótipo Sensível ao Contexto para Governança de Saúde”. In: *2011 VI Congresso de Pesquisa e Inovação da Rede Norte e Nordeste de Educação Tecnológica.*, 2011, Aracaju. *VI Congresso de Pesquisa e Inovação da Rede Norte e Nordeste de Educação Tecnológica – CONNEPI*, 2011. Aracaju: IFS, 2011
- [8] Braga, R.B., Martin, H.: “Captain: A context-aware system based on personal tracking”, in: *The 17th International Conference on Distributed Multimedia Systems/DMS 2011*, KSI, Florence, 2011.
- [9] Oliveira, A.M.B., and CUNHA. P.R.F. Implementing Home Care Application in Brazilian Digital TV”, in *IEEE GIS Global Information Infrastructure Symposium*, Tunisia, 2009.
- [10] L.F.G. Soares, R.M. Rodrigues, and M.F. Moreno. NCL: the Declarative Environment of the Brazilian Digital TV System”, in *Journal of the Brazilian Computer Society - Vol 13 - Number 1, 2006.*
- [11] J. H. Jahnke, Y. Bychkov, D. Dahlem and L. Kawasme, “Implicit, Context-Aware Computing for HealthCare”, <http://www.ics.uci.edu/lopes/bspc04documents/Jahnke.pdf>, 2004.
- [12] T. Gu, Z. Kwok, K. K. Koh, and H. K. Pung. A Mobile Framework Supporting Ontology Processing and Reasoning, Proc. of the 2nd Workshop RSPSI - Ubicomp '07, 2007, Austria.
- [13] O'Hare, N., Smeaton, A.F.: Context-aware person identification in personal photo collections. *Trans. Multi.* 11 (2009) 220{228
- [14] J. B. B. Frota, M. Oliveira, L. O. M. Andrade, I. Barreto, and C. O. M., “Integrating Mobile Devices In a Brazilian Health Governance Framework,” in *International Conference on Advances of Information & Communication Technology in Health Care*, 2011.
- [15] T. P. D. A. Alcântara, “Paola: Uma Plataforma Para O Desenvolvimento De Aplicações Baseadas Em Ontologias Para O Projeto Lariisa,” Universidade Estadual Do Ceará, 2012.
- [16] A. L. Machado and J. M. Parente de Oliveira, “DIGO: An Open Data Architecture for e-Government,” *2011 IEEE 15th International Enterprise Distributed Object Computing Conference Workshops*, pp. 448–456, Aug. 2011.
- [17] Lorenzo, G. D., Hacid, H., Paik, H.-Y., And Benatallah, B. Data integration in mashups. *SIGMOD Record* 38 (1):59-66, 2009.
- [18] Ribeiro, T. G., Magalhães, Régis P., Vidal, V. M. P., Macêdo, J. A. F., *Uma Abordagem Baseada em Ontologias para o Projeto de Linked Data*
- [19] B. de Pierro, “Dados sobre saúde precisam de integração | Brasilianas.Org,” 2011. [Online]. Available: <http://www.advivo.com.br/materia-artigo/dados-sobre-saude-precisam-de-integracao>. [Accessed: 18-May-2013].
- [20] M. Hiatt, “About Mashups and Linked Data | Mashstream,” 2009. [Online]. Available: <http://mashstream.com/semantic-web/about-mashups-and-linked-data/>. [Accessed: 19-May-2013].
- [21] M. Oliveira, L. O. M. Andrade, F. Antunes, and L. M. Gardini, “Applying Ontology And Context Awareness Concepts On Health Management System: A Dengue Crisis Study Case,” in *IADIS International Conference e-Health 2013*, 2013.
- [22] R. P. Magalhães, “Um Ambiente Para Processamento De Consultas Federadas Em Linked Data Mashups,” Universidade Federal Do Ceará, 2012.