Modern Education
and Computer Science
PRESS

# An Ontology-Based System for Cancer Diseases Knowledge Management

**Marco Alfonse, Mostafa M. Aref, Abdel-Badeeh M. Salem**
Computer Science Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
Email: marco_alfonse@cis.asu.edu.eg, aref_99@yahoo.com, absalem@cis.asu.edu.eg

*Abstract*—Cancer is a class of diseases characterized by out-of-control cell growth. There are over 200 different types of cancer, and each is classified by the type of cell that is initially affected. This paper discusses the technical aspects of some of the ontology-based medical systems for cancer diseases. It also proposes an ontology based system for cancer diseases knowledge management. The system can be used to help patients, students and physicians to decide what cancer type the patient has, what is the stage of the cancer and how it can be treated. The system performance and accuracy are acceptable, with a cancer diseases classification accuracy of 92%.

*Index Terms*—Cancer Informatics, Ontology, Medical Informatics, Machine Learning, Computer Aided Diagnosis (CAD), Knowledge Engineering.

## I. INTRODUCTION

An ontology is a formal, explicit specification of a shared conceptualization. This definition identifies four main concepts: an abstract model of a phenomenon termed "conceptualization", a precise mathematical description hints the word "formal", the precision of concepts and their relationships clearly defined are expressed by the term "explicit"' and the existence of an agreement between ontology users is hinted by the term "shared". Ontologies can be used to support a great variety of tasks in diverse research areas such as knowledge representation, natural language processing, information retrieval, databases, knowledge management, on line database integration, digital libraries, geographic information systems, visual information retrieval or multi agent systems [1].

Cancer is one of the most dangerous diseases known to human. The deadliness of this disease can be notified by the fact that cancer caused 13% of all the human deaths in 2007. Cancer, known medically as a malignant neoplasm, is a broad group of various diseases, all involving unregulated cell growth. In cancer, cells divide and grow uncontrollably, forming malignant tumors, and invade nearby parts of the body. The cancer may also spread to more distant parts of the body through the lymphatic system or bloodstream. Not all tumors are cancerous. Benign tumors do not grow uncontrollably, do not invade neighboring tissues, and do not spread throughout the body. There are over 200 different known cancers that afflict humans [2].

This paper presents an ontology based cancer diseases knowledge management that can be used to determine the cancer type and stage the patient has and then recommend treatment options. The paper is organized as follows: section 2 gives an a background on cancer and ontologies, section 3 presents the related work, section 4 discusses the proposed system and finally section 5 contains the conclusions and future work.

## II. MEDICAL BACKGROUND

Cancer is not a disease; rather it is a group of diseases characterized by uncontrolled growth and spread of abnormal cells. If the spread is not controlled, it can result in death. Fig. 1 shows the difference between normal and cancer cells [2]. Cancer is caused by both external factors (tobacco, infectious organisms, chemicals, and radiation) and internal factors (inherited mutations, hormones, immune conditions, and mutations that occur from metabolism). These causal factors may act together or in sequence to initiate or promote the development of cancer. Ten or more years often pass between exposure to external factors and detectable cancer. Cancer is treated with surgery, radiation, chemotherapy, hormone therapy, biological therapy, and targeted therapy.
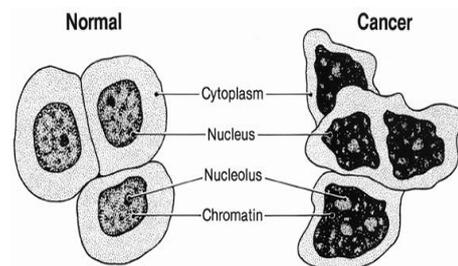


Fig. 1: Normal and Cancer Cells (adapted from [2])

Cancer staging describes the extent or spread of cancer at the time of diagnosis. Proper staging is essential in determining the choice of therapy and in assessing prognosis. A cancer's stage is based on the size or extent of the primary (main) tumor and whether it has spread to other areas of the body. A number of different staging systems are used to classify tumors. A system of summary staging (in situ, local, regional, and distant) is used for descriptive and statistical analysis of tumor registry data. If cancer cells are present only in the layer of cells where they developed and have not spread, the

stage is in situ. If cancer cells have penetrated beyond the original layer of tissue, the cancer is invasive and categorized as local, regional, or distant stage based on the extent of spread. Clinicians typically use the TNM cancer staging system, which assesses tumors in three ways: extent of the primary tumor (T), absence or presence of regional lymph node involvement (N), and absence or presence of distant metastases (M). Once the T, N, and M categories are determined, a stage of 0, I, II, III, or IV is assigned, with stage 0 being in situ, stage I being early, and stage IV being the most advanced disease. Some cancers have alternative staging systems (e.g., leukemia). As the molecular properties of cancer have become better understood, tumor biological markers and genetic features have been incorporated into prognostic models, treatment plans, and/or stage for some cancer sites [2].

## III. RELATED WORK

Ontologies can be used to diagnose cancer diseases. To build an ontology is a similar process than to build an object oriented program, however, classes and objects in a program are about data structures, whereas classes and objects in ontologies are about the domain. A good practice to support ontology building process is to produce an ontology specification document written in natural language with information such as: the purpose of the ontology, its end users, use case scenarios, degree of formality used to codify the ontology and its scope. An ontology specification document should be conciseness, (it must include just relevant terms without duplicates), partial completeness and realism, (meanings of the terms and their relationships making sense in the domain) [3]. Fig. 2 shows a general methodology to build an ontology.
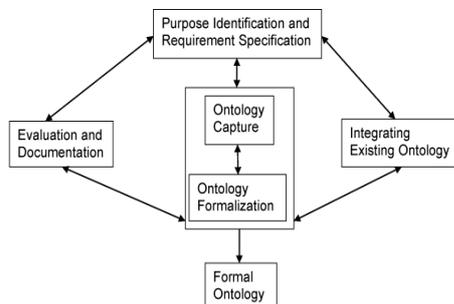


Fig. 2: A General Methodology to build an Ontology.

Many methodologies, languages and tools are used to support ontological engineering process. In [4] the authors provide the different methodologies used to build ontologies, the languages used to codify ontologies and finally the tools that can be used to support ontology building process.

In [5] the authors provide the computerization and execution of a breast cancer follow-up clinical practice guideline. The computerization of the clinical practice guideline led to the development of a breast cancer ontology. They present their breast cancer ontology which models the knowledge inherent within the breast cancer follow-up clinical practice guideline - the breast cancer ontology serves as the knowledge source to determine patient-specific recommendations. They discuss the ontology engineering process that highlights the specification of their breast cancer ontology in terms of clinical concepts and the relationships between the concepts expressed as OWL classes and properties, using the protégé ontology development tool.

In [6] the authors introduce a novel Ontology-based Argumentation Framework (OAF) that links a logic-based argumentation formalism and description logic ontologies. They show how these two formalisms can be tightly coupled by observing a few simple restrictions, and provide features not available in either formalism alone. Their work is evaluated in a large case study on decision-making in treatment choice in breast cancer, where rules are developed from the results of published clinical trials, and they present a small subset of this to demonstrate the use of the system.

In [7] the authors present an Ontology-based Modeling of Clinical Practice Guidelines: A Clinical Decision Support System for Breast Cancer Follow-up Interventions at Primary Care Settings

The ASSIST project aims to facilitate cervical cancer research by integrating medical records containing both phenotypic and genotypic data, and residing in different medical centers or hospitals. The goal of ASSIST is to enable the evaluation of medical hypotheses and the conduct of association studies in an intuitive manner, thereby allowing medical researchers to identify risk factors that can then be used at the point of care to identify women who are at high risk of developing cervical cancer. In [8] the authors present the current status of the ASSIST medical knowledgebase. In particular, they discussed the challenges faced in constructing the ASSIST integrated resource and in enabling query processing through a domain ontology, and the solutions provided using the AutoMed heterogeneous data integration system. They focus on data cleansing issues, on data integration issues related to integrating relational medical data sources into independent domain ontology and also on query processing. Of particular interest is the challenge of providing an easily maintainable integrated resource in a setting where the data sources and the domain ontology are developed independently and are therefore both highly likely to evolve over time.

In [9] a breast cancer ontology has been built and gathers the terms used by lay people to talk about breast cancer. The resulting ontology has been the core of a health consumer query reformulation application. This work proposes the use of spreading activation technique through the ontology to infer new concepts from the ones initially identified in a health consumer question.

In [10] the authors present a benign/malignant breast cancer classification model based on a combination of ontology and case-based reasoning to effectively classify breast cancer tumors as either malignant or benign. This classification system makes use of clinical data. Two

CBR object-oriented frameworks based on ontology are used jCOLIBRI and myCBR. Table 1 illustrates the different ontology based medical systems for cancer diseases currently exist.

Table 1: The Ontology-Based Systems for Cancer Diseases

| System | Approach |
|---|---|
| Breast Cancer Follow-Up Clinical Practice Guideline [5] | Clinical Decision Support System |
| Argument-Based Decision-Making in Breast Cancer [6] | Links a logic-based argumentation formalism and description logic ontologies |
| A Clinical Decision Support System for Breast Cancer Follow-up [7] | Decision Support System |
| The ASSIST project for Cervical Cancer [8] | Data integration approaches |
| Ontology-Based Reformulation of Health Consumer Queries [9] | Spreading Activation Techniques through the ontology |
| A Breast Cancer Classifier [10] | Combination of ontology and case-based reasoning |

## IV. THE PROPOSED SYSTEM

The proposed system contains three basic modules namely; the diagnostic module, the staging module and the treatment recommendation module. In order to detect patient disease, the patient provides his/her signs and symptoms to the diagnostic module, which detects what type of cancer the patient has. Once the type of the cancer is determined, the staging module finds the current stage of the cancer based on the cancer type and the signs and symptoms provided by the patient. Based on the determined cancer type and cancer stage, the treatment recommendation module can recommend a specific treatment for the case at hand.

All the three modules interact with a database of cancer ontologies through the query module, which maps from the query of the asking module to the structure of the vocabulary of the ontologies stored in the database of cancer ontologies. The database of cancer ontologies describes the different types of cancer diseases in detail. Each cancer ontology describes the cancer in terms of its structure, signs and symptoms, staging and treatment. Fig. 3 shows the proposed system.
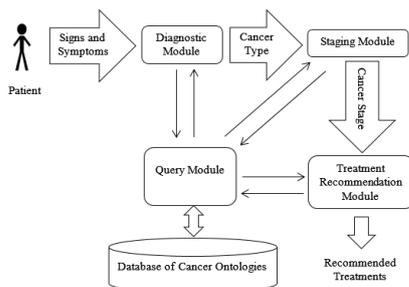


Fig. 3: The Proposed System

Each module (diagnostic, staging or treatment recommendation) can be applied as a stand-alone application according to the user requirements. For example, if the user knows the cancer type and wants to determine the cancer stage it can use the staging module by providing the cancer type and signs and symptoms and the module will determine the current stage and so on.

The system is implemented in Java using the Eclipse IDE for Java developers, the OWL-API and the FaCT++ reasoner. Java is a general-purpose, class-based, object-oriented computer programming language that is specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "Write Once, Run Anywhere" (WORA), meaning that code that runs on one platform does not need to be recompiled to run on another. Java applications are typically compiled to byte code (class file) that can run on any Java virtual machine (JVM) regardless of computer architecture. Java is, as of 2012, one of the most popular programming languages in use with a reported 10 million users. The Eclipse is a multi-language Integrated Development Environment (IDE) comprising a base workspace and an extensible plug-in system for customizing the environment. It can be used to develop applications in Java and, by means of various plug-ins, other programming languages. The OWL API is a Java API and reference implementation for creating, manipulating and serializing OWL Ontologies. FaCT++ is the reasoner for expressive Description Logics (DL). It covers OWL DL based ontology languages. It can be used as a standalone reasoner, or as a back-end reasoner for the OWL API-based application. Now it is used as one of the default reasoners in the Protege4 OWL editor. It is an open-source software.

### A. The Database of Cancer Ontologies

The database of cancer ontologies contains three ontologies, which are the lung cancer ontology [11], the breast cancer ontology [12] and the liver cancer ontology [13]. The authors have built these ontologies. They are encoded in OWL-DL as a representation language because of its reasoning and expressive power using the Protégé-OWL editor which is a free, open source ontology editor developed at Stanford University by Stanford medical informatics team. The three ontologies have been built using Top-Down approach in which classes are constructed from the most abstract to the most concrete. They are described in terms of their risk factors/causes, symptoms, diagnosis, staging and treatment.

### B. The Diagnostic Module

The diagnostic module works as follows: it reads signs and symptoms from user and uses the query module to convert user input to a well formatted query for the ontologies (at the database of ontologies) then it retrieves the signs and symptoms from the existing ontologies to find the type of cancer based on the user input.

The signs and symptoms of the cancers are retrieved from the database of ontologies using FaCT++ reasoner. First the system loads the existing ontologies from the database using the OWL-API then the reasoner gets the signs and symptoms' instances of each ontology.

There are general signs and symptoms that may occur in different cancer types so it is possible that some of the symptoms entered by the user may be found in more than a type of cancer so to allow an accurate diagnosis, we used a similarity measure to find the best-fit cancer type (the cancer type that best match the user input). The similarity measure is defined as:

$$SIMILARITY = MAXIMUM\ (SIM1, SIM2\ \ldots)$$

Where SIM1, SIM2 … etc. represents the degree of match between user' signs and symptoms and the signs and symptoms of the cancers found in the database of ontologies.

The similarity between user' signs and symptoms and the signs and symptoms of the cancers found in the database of ontologies is defined as follows:

$$SIM = N_{ms}\ /\ N_s$$

Where $N_{ms}$ is the number of matched symptoms between user input and cancer symptoms in the ontology and $N_s$ is the total number of symptoms of the cancer in the ontology. The cancer type is determined based on the maximum similarity measure. Once the cancer type is determined, the system moves to the staging module to find the cancer stage.

*C.    The Staging Module*

Once the cancer type is determined, the staging module finds the stage of the cancer defined. The stages of the determined cancer are retrieved from its ontology using FaCT++ reasoner. The reasoner gets the stages' instances from the ontology through the query module. The cancer stages differ based on the cancer type. The stages of the cancer are defined based on a set of rules that differ from a cancer type to another. These rules are retrieved from the cancer ontology through the query module. Tables 2 and 3 show the stages of the breast cancer.

Table 2: TNM Classification of the Breast Cancer

| Tumor Size | |
|---|---|
| Tx | The tumor can't be measured or found. |
| T0 | There isn't any evidence of the primary tumor. |
| Tis | The cancer is "in situ" (the tumor has not started growing into the breast tissue). |
| T1 | The tumor is less than 2 centimeters. |
| T2 | The tumor is from 2 - 5 centimeters. |
| T3 | The tumor is greater than 5 centimeters. |
| T4 | The tumor is of any size that has broken through the skin, or is attached to the chest wall. |
| **Lymph Node Status** | |
| Nx | The nearby lymph nodes can't be measured or found. |
| N0 | Nearby lymph nodes do not contain cancer. |
| N1 | Metastasis to movable, same-side, axillary (armpit) lymph node(s). |
| N2 | Metastasis to same-side lymph node(s) fixed to one another or to other structures |
| N3 | Metastasis to same-side lymph nodes beneath the breastbone (internal mammary nodes) |
| **Distant Metastasis** | |
| Mx | Metastasis can't be measured or found. |
| M0 | There are no distant metastases. |
| M1 | Distant metastases were found. |

Table 3: Stage Groupings Based on the TNM classification of the Breast Cancer

| Overall Stage | T category | N category | M category |
|---|---|---|---|
| Stage 0 | Tis | N0 | M0 |
| Stage I | T1 | N0 | M0 |
| Stage IIA | T0 | N1 | M0 |
| | T1 | N1 | M0 |
| | T2 | N0 | M0 |
| Stage IIB | T2 | N1 | M0 |
| | T3 | N0 | M0 |
| Stage IIIA | T0 | N2 | M0 |
| | T1 | N2 | M0 |
| | T2 | N2 | M0 |
| | T3 | N1 | M0 |
| | T3 | N2 | M0 |
| Stage IIIB | T4 | Any N | M0 |
| Stage IIIC | Any T | N3 | M0 |
| Stage IV | Any T | Any N | M1 |

The lung cancer has two type; small cell lung cancer and non-small cell lung cancer and each type has its own stages. The stages of the small cell lung cancers are:

1. Limited disease: the cancer cells can be seen only in one lung, in nearby lymph nodes, or in fluid around the lung (known as a pleural effusion).
2. Extensive disease: it is clear that the cancer has spread outside the lung, within the chest area or to other parts of the body.

Tables 4 and 5 show the stages of the non-small cell lung cancer.

Table 4: TNM Classification of the Non-Small Cell Lung Cancer

| Primary Tumor | |
|---|---|
| Tx | Primary tumor cannot be evaluated, or tumor proven only by the presence of malignant cells in sputum or bronchial washings but not visualised by imaging or bronchoscopy. |
| T0 | No evidence of primary tumor. |
| Tis | Carcinoma *in situ*. |
| T1 | The tumor is 3cm or less across. |
| T2 | The tumor is more than 3cm across or involves the largest airway (the main bronchus) or the tumor has caused a partial collapse of that lung or the tumour has grown into the inner lining of the chest cavity (the visceral pleura). |
| T3 | T3 is divided into T3a and T3b: T3a means there are several tumours, and at least one is more than 5cm across and T3b means at least one tumour has grown into a major branch of one of the main blood vessels (the portal vein or hepatic vein). |
| T4 | The tumor has grown into the area between the lungs in the middle of the chest (the mediastinum) or there is fluid around the lung that contains cancer cells (a malignant pleural effusion). |
| **Regional Lymph Nodes** | |
| Nx | Lymph node status cannot be evaluated. |
| N0 | There is no cancer in any lymph nodes. |
| N1 | There is cancer in the lymph nodes nearest the affected lung. |
| N2 | There is cancer in lymph nodes in the centre of the chest (mediastinum) but on the same side as the affected lung or there is cancer in lymph nodes that lie just under where the windpipe branches off to each lung. |
| N3 | There is cancer in lymph nodes on the opposite side of the chest from the affected lung or in the lymph nodes above either collar bone. |
| **Distant Metastasis** | |
| Mx | Distant metastasis cannot be evaluated. |
| M0 | There are no signs of cancer spread to another lobe of the lung or any other part of the body. |
| M1 | There is cancer spread to another lobe of the lung or any another part of the body. |

Table 5: Stage Groupings Based on the TNM classification of the Non-Small Cell Lung Cancer

| Overall Stage | T category | N category | M category |
|---|---|---|---|
| occult carcinoma | Tx | N0 | M0 |
| stage 0 | Tis | N0 | M0 |
| stage 1A | T1 | N0 | M0 |
| stage 1B | T2 | N0 | M0 |
| stage 2A | T1 | N1 | M0 |
| stage 2B | T2 | N1 | M0 |
| | T3 | N0 | M0 |
| stage 3A | T1 | N2 | M0 |
| | T2 | N2 | M0 |
| | T3 | N1,N2 | M0 |
| stage 3B | any T | N3 | M0 |
| | T4 | any N | M0 |
| stage 4 | any T | any N | M1 |

Tables 6 and 7 show the stages of liver cancer.

Table 6: TNM Classification of the Liver Cancer

| Primary Tumor | |
|---|---|
| Tx | Primary tumor cannot be assessed. |
| T1 | There is a single small tumour that has not grown into any blood vessels. |
| T2 | Either there is a single tumor that has grown into a blood vessel in the liver OR there are several tumors, all less than 5cm across, but none have grown into blood vessels. |
| T3 | The tumor has grown into the chest wall, the outer lining of the chest cavity (the mediastinal pleura), the muscle at the bottom of the chest cavity (the diaphragm) or the outer covering of the heart (the pericardium) or the tumor has caused the whole lung to collapse. |
| T4 | The cancer has grown into organs close to the liver (not including the gallbladder) OR through the lining that wraps around the internal organs of the abdomen (visceral peritoneum). |
| Regional Lymph Nodes | |
| Nx | Regional lymph nodes cannot be assessed. |
| N0 | There are no cancer cells in nearby lymph nodes. |
| N1 | There are cancer cells in lymph nodes near the liver. |
| Distant Metastasis | |
| Mx | Distant metastasis can't be assessed. |
| M0 | There is no sign that the cancer has spread outside the liver. |
| M1 | There are cancer cells in other body organs such as the lungs or bones. |

Table 7: Stage Groupings Based on the TNM classification of the Liver Cancer

| Overall Stage | T category | N category | M category |
|---|---|---|---|
| Stage 1 | T1 | N0 | M0 |
| Stage 2 | T2 | N0 | M0 |
| Stage 3A | T3a | N0 | M0 |
| Stage 3B | T3b | N0 | M0 |
| Stage 3C | T4 | N0 | M0 |
| Stage 4A | Any T | N1 | M0 |
| Stage 4B | Any T | Any N | M1 |

### D. The Treatment Module

After determining the cancer type and the stage of the cancer, the treatment module is used to recommend the treatment options to the system user. The treatments of the determined cancer are retrieved from its ontology using FaCT++ reasoner. The reasoner gets the treatment instances from the ontology through the query module. The treatment options are determined based on the stage of the cancer at hand.

### Treatment of Breast Cancer by Stage

Stage 0: Treatments differ depending on what kind of stage 0 cancer the patient has. Stage 0 cancer may include:
- Ductal carcinoma in situ (DCIS) or intraductal carcinoma, in which abnormal cells appear in the ducts of the breast.
- Lobular carcinoma in situ (LCIS) develops when abnormal cells appear in the lobes of the breast.
- Paget's disease of the nipple is an uncommon type of cancer that develops in or around the nipple.

Treatments for DCIS include:
- Surgery.
- Radiation therapy.
- Endocrine therapy.
- Biological therapy.

LCIS raises the risk of getting invasive breast cancer. Treatment may include:
- Endocrine therapy.
- Bilateral mastectomy - the removal of both breasts.

Paget's disease of the nipple is usually treated with surgery, without an axillary lymph node dissection, or wide local excision followed by radiation.

Stage I: Treatment options include:
- Surgery. Since the tumor is still small, lumpectomy (wide local excision), is an option. In this procedure, the tumor and some of the normal surrounding tissue are removed. Some patients choose to get a mastectomy. In either case, the surgeon will likely take out one or more of the lymph nodes. After a mastectomy, the patient might choose to get breast reconstruction surgery. Surgical and survival outcomes are equal between lumpectomy and mastectomy. Surgery treats the disease doctors know about. Other treatments, called "adjuvant" treatments, may also be used to treat hidden disease and reduce the risk of breast cancer recurrence.
- Radiation therapy is standard after a lumpectomy. It can kill off any cancer cells that were missed. The patients with stage I cancer who get a mastectomy do not require radiation.
- Chemotherapy is treatment with drugs that attack cancer cells. It's often used after surgery to reduce the risk of the cancer coming back, or recurring.
- Endocrine therapy is sometimes used after surgery in patients who have hormone receptor-positive cancer (tumors whose growth seems dependent on estrogen). In these patients, hormone therapy drugs can prevent the tumor from getting the hormone it needs to grow.

- Biological therapy: In about 25% of women with breast cancer, an excess of a protein known as HER2 makes the cancer spread quickly. Herceptin is a drug that's been approved to treat women with metastatic breast cancer that is HER2-positive. It stops this protein from making the cancer grow and makes chemotherapy more effective. It is most often used in combination with chemotherapy.
- Clinical trials: A clinical trial may allow the patient access to cutting-edge treatments. Many new therapies - new drugs, new treatments, and new combinations - are being studied in ongoing clinical trials.

Stage II: A combination of treatments is used to treat stage II breast cancer.

- Surgery. For smaller tumors, the patient might get a lumpectomy (wide local excision), in which only the tumor and some of the surrounding normal tissue are removed. For larger tumors, the patient might need a mastectomy, in which the entire breast is removed. In either case, the surgeon will remove some of the lymph nodes. After a mastectomy, the patient might get breast reconstruction surgery.
- Radiation therapy is standard for the patients who get a lumpectomy. It can kill cancer cells that were missed during surgery. Some patients with stage II cancer who get a mastectomy will also need radiation, especially if the tumor was large or if there were lymph nodes involved with cancer.
- Chemotherapy is often used after surgery. This treatment can destroy any remaining cancer cells that were missed. Chemotherapy may also be used as neoadjuvant therapy; treatment before surgery to shrink a tumor. If it works, the tumor might then be small enough to remove in a lumpectomy.
- Endocrine therapy is sometimes used after surgery in patients who have hormone receptor-positive cancer. In these patients, endocrine therapy can prevent the tumor from getting the hormone it needs to grow.
- Biological therapy. In about 25% of women with breast cancer, an excess of a protein known as HER2 makes the cancer spread quickly. Herceptin is a drug that's been approved to treat women with metastatic breast cancer that is HER2-positive. It stops this protein from making the cancer grow and makes chemotherapy more effective. It is most often used in combination with chemotherapy.
- Clinical trials: A clinical trial may allow the patient access to cutting-edge treatments. Many new therapies such as new drugs, new treatments, and new combinations, are in current clinical trials.

Stage III: A combination of different treatments often works best and may include:

- Chemotherapy is often used before surgery to shrink the tumor and make it easier to remove. The patient may still need chemotherapy after surgery to help destroy any cancer cells that remain.

- Surgery: The patient might get a lumpectomy, if the neo-adjuvant therapy such as chemotherapy shrinks the tumor to the point in which a surgeon can remove the tumor and surrounding tissue from the breast. The patient may need a mastectomy, in which the whole breast is removed as well lymph nodes. A mastectomy should be performed after neo-adjuvant chemotherapy.
- Endocrine therapy can help patients with hormone receptor-positive cancers. In these patients, medications can prevent the tumor from getting the hormone it needs to grow.
- Biological therapy: In about 25% of women with breast cancer, an excess of a protein known as HER2 makes the cancer spread quickly. Herceptin is a drug that's been approved to treat women with metastatic breast cancer that is HER2-positive. It stops this protein from making the cancer grow and makes chemotherapy more effective. It is most often used in combination with chemotherapy.
- Radiation therapy is standard for patients with stage III cancer who get surgery, including those who have a mastectomy. Radiation can destroy any of the cancer cells that were missed during surgery.
- Clinical trials: A clinical trial may allow the patient access to cutting-edge treatments. Many new therapies, such as new drugs, new treatments, and new combinations are in current clinical trials.

Stage IV: Treatments of stage IV breast cancer include:

- Chemotherapy: It can slow down the growth of the cancer. Chemotherapy is often used in combination with other chemotherapy drugs, biological therapies, or targeted therapies.
- Endocrine therapy for women with hormone receptor-positive cancers. These are cancers that need hormones to grow.
- Biological therapy: In about 25% of women with breast cancer, an excess of a protein known as HER2 makes the cancer spread quickly. Herceptin is a drug that's been approved to treat women with metastatic breast cancer that is HER2-positive. It stops this protein from making the cancer cells grow. It may also boost the immune system, giving it the strength to fight the cancer itself.
- Palliative care: Surgery and radiation are used in some cases to help treat pain and other symptoms. Other medications may be used to help treat some of the side effects of breast cancer treatment, such as nausea, fatigue, and infections.
- Clinical trials: A clinical trial may allow the patient access to cutting-edge treatments. Many new therapies - new drugs, new treatments, and new combinations - are in clinical trials now.

*Treatment of Lung Cancer by Stage*

Treatment of limited-stage small cell lung cancer may include the following:

- Combination chemotherapy and radiation therapy to

the chest, with or without radiation therapy to the brain.
- Combination chemotherapy with or without radiation therapy to the brain in patients with complete response.
- Combination chemotherapy with or without radiation therapy to the chest.
- Surgery followed by chemotherapy or chemotherapy plus radiation therapy to the chest, with or without radiation therapy to the brain.
- Clinical trials of new chemotherapy, surgery, and radiation treatments.

Treatment of extensive-stage small cell lung cancer may include the following:
- Chemotherapy.
- Combination chemotherapy.
- Combination chemotherapy with or without radiation therapy to the brain for patients with complete response.
- Radiation therapy to the brain, spine, bone, or other parts of the body where the cancer has spread, as palliative therapy to relieve symptoms and improve quality of life.
- Clinical trials of new chemotherapy treatments.

Treatment of stage 0 non-small cell lung cancer may include the following:
- Surgery.
- Photodynamic therapy using an endoscope.

Treatment of stage I non-small cell lung cancer may include the following:
- Surgery.
- External radiation therapy (for patients who cannot have surgery or choose not to have surgery).
- Surgery followed by chemotherapy.
- A clinical trial of photodynamic therapy using an endoscope.
- A clinical trial of surgery followed by chemoprevention.

Treatment of stage II non-small cell lung cancer may include the following:
- Surgery.
- External radiation therapy (for patients who cannot have surgery or choose not to have surgery).
- Surgery followed by chemotherapy, with or without other treatments.
- A clinical trial of external radiation therapy following surgery.

Treatment of stage IIIA non-small cell lung cancer may include the following:
- Surgery with or without radiation therapy.
- External radiation therapy alone.
- Chemotherapy combined with other treatments.
- A clinical trial of new ways of giving radiation therapy and chemotherapy.

- A clinical trial of new combinations of treatments.

Treatment of stage IIIB non-small cell lung cancer may include the following:
- External radiation therapy alone.
- Chemotherapy combined with external radiation therapy.
- Chemotherapy combined with external radiation therapy, followed by surgery.
- Chemotherapy alone.
- A clinical trial of new ways of giving radiation therapy.
- A clinical trial of new combinations of treatments.

Treatment of stage IV non-small cell lung cancer may include the following:
- External radiation therapy as palliative therapy, to relieve pain and other symptoms and improve the quality of life.
- Chemotherapy.
- Laser therapy and/or internal radiation therapy.
- A clinical trial of chemotherapy with or without biologic therapy.

*Treatment of Liver Cancer by Stage*

Liver cancers are divided into 3 categories: localized resectable, localized unresectable, and advanced.

Localized resectable liver cancer (some T1 or T2, N0, M0 tumors): If the cancer is at an early stage and the rest of the liver is healthy, surgery (partial hepatectomy) may cure the patient. For some patients with early-stage liver cancer, a liver transplant may be another option.

Localized unresectable liver cancer (some T1 to T4, N0, M0 tumors): Localized, unresectable cancers include cancers that haven't yet spread, but that can't be removed safely by partial hepatectomy. These patients may be treated with a liver transplant if it is possible. Most people must wait at least several months before a liver becomes available for transplant. In many cases a person may get other treatments, such as embolization or ablation, while waiting for a liver transplant. Also ablation of the tumor(s) can be used. Other options may include embolization (with or without chemotherapy or radiation), targeted therapy with sorafenib, chemotherapy (either systemic or by hepatic artery infusion), and/or radiation therapy. In some cases, treatment may shrink the tumor(s) enough so that surgery (partial hepatectomy or transplant) may become possible. Because these cancers can be hard to treat, clinical trials of newer treatments may offer a good option in many cases.

Advanced liver cancer (includes all N1 or M1 tumors): Advanced cancer has spread outside of the liver (either to the lymph nodes or to other organs). Because these cancers are widespread, they cannot be treated with surgery. If the liver is functioning well enough, the targeted therapy sorafenib may help control the growth of the cancer for a time. Also clinical trials of targeted therapies, new approaches to chemotherapy (new drugs and ways to deliver chemotherapy), new forms of radiation therapy, and other new treatments may help the

patient. Treatments such as radiation might also be used to help relieve pain and other symptoms.

## V. Evaluation of the Proposed System

The proposed system is evaluated using 40 test cases of different types of cancers collected from American Cancer Society [14], National Cancer Institute (NCI) [2] and CancerCare [15] .These test cases includes 14 cases for the breast cancer, 12 cases for the lung cancer, 12 cases for the liver cancer and 2 cases for other types of cancers. The system gives 92% correct classification, 3% misclassification and 5% unknown cancer type (since the system does not cover all types of cancer).

## VI. Conclusions and Future Work

This paper discusses the technical aspects of the existing ontology-based medical systems for cancer diseases. These systems focus on a specific type of cancer. These systems use different approaches for the diagnosis of the cancer disease (e.g. decision support system, data integration, ontology …etc.). The paper also proposes a new ontology-based system for cancer diseases knowledge management. This system can be used by patients, students and physicians to decide the type of the cancer the patient has, the stage of the cancer and the treatment options. The main contribution of this system is its extensibility to contain all types of cancers and its capability to provide the treatment recommendation for the case at hand. Currently the database of the cancer ontologies of the system has only three types of cancers, which are lung, liver and breast. In future, it is planned to extend the database with all types of cancer ontologies. The system gives 92% correct classification.

## References

[1] Mar ú Auxilio Medina Nieto. An Overview of Ontologies. Technical report, Center for Research in Information and Automation Technologies, Interactive and Cooperative Technologies Lab, Universidad De Las Am éricas Puebla − México, March 2003.

[2] The National Cancer Institute (NCI): http://www.cancer.gov/cancertopics/cancerlibrary/what-is-cancer.

[3] FIPA, Foundation for Intelligent Physical Agents. FIPA Ontology Service Specification. March 2001. http://www.fipa.org/specs/fipa00086/XC00086D.html.

[4] Abdel-Badeeh M. Salem, Marco Alfonse. "Ontological Engineering in Medicine". Medical Informatics Workshop, Third International Conference on Intelligent Computing and Information Systems, Cairo, Egypt, PP 59-74, 2007.

[5] Abidi, S.R.;, "Ontology-Based Modeling of Breast Cancer Follow-up Clinical Practice Guideline for Providing Clinical Decision Support," Computer-Based Medical Systems, 2007. CBMS '07. Twentieth IEEE International Symposium on, vol., no., pp.542-547, 20-22 June 2007.

[6] Williams, M.; Hunter, A.;, "Harnessing Ontologies for Argument-Based Decision-Making in Breast Cancer," Tools with Artificial Intelligence, 2007. ICTAI 2007. 19th IEEE International Conference on, vol.2, no., pp.254-261, 29-31 Oct. 2007.

[7] Abidi S, Abidi S, Hussain S, Shepherd M. Ontology-based modeling of clinical practice guidelines: a clinical decision support system for breast cancer follow-up interventions at primary care settings. Stud Health Technol Inform, 129(Pt 2):845–9.9, 2007.

[8] Zamboulis, L.; Poulovassilis, A.; Roussos, G.;, "Flexible data integration and ontology-based data access to medical records," BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on, vol., no., pp.1-6, 8-10 Oct. 2008.

[9] Messai, R.; Bricon-Souf, N.; Simonet, M.; Mousseau, M.; , "Ontology-Based Reformulation of Health Consumer Queries Using Spreading Activation Techniques," Intelligent Computing and Cognitive Informatics (ICICCI), 2010 International Conference on , vol., no., pp.98-101, 22-23 June 2010.

[10] Lotfy Abdrabou, E.A.M.; Salem, A.M.; , "A breast cancer classifier based on a combination of case-based reasoning and ontology approach," Computer Science and Information Technology (IMCSIT), Proceedings of the 2010 International Multiconference on , vol., no., pp.3-10, 18-20 Oct. 2010.

[11] Abdel-Badeeh M. Salem, Marco Alfonse. "Building Web-Based Lung Cancer Ontology". The International Journal of Soft Computing Applications, ISSN: 1453-2277 Issue 2, PP 5-14, 2008.

[12] Abdel-Badeeh M. Salem, Marco Alfonse. "Ontological Engineering Approach for Breast Cancer Knowledge Management" Proceedings of Med-e-Tel, the International eHealth, Telemedicine and Health ICT for Education, Networking and Business. PP 320-324, Luxembourg. March 30 – April 5, 2009.

[13] Marco Alfonse, Mostafa M. Aref, Abdel-Badeeh M. Salem. "Ontology-Based Knowledge Representation for Liver Cancer". Proceedings of the International eHealth, Telemedicine and Health ICT Forum for Educational, Networking and Business. Luxembourg, G. D. of Luxembourg, ISSN 1818 – 9334, PP: 821-825, April 18-20, 2012.

[14] American Cancer Society, www.cancer.org.

[15] CancerCare , http://www.cancercare.org/.

## Authors' Profiles

**Marco Alfonse Tawfik** is an Assistant Lecturer at the Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt. His research interests: semantic web, ontological engineering, medical informatics, and artificial intelligence. He has M.Sc. degree in Computer Science. Currently Marco is a PhD candidate of computer science. His PhD thesis title is "Ontological Engineering Approach for Medical Knowledge Sharing". He has published around 13 papers in refereed international journals and conferences.
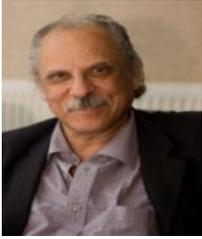
**Prof. Mostafa M. Aref** is Professor of Computer Science and Vice Dean for Graduate studies and Research, Ain Shams University, Cairo Egypt. He got Ph.D. of Engineering Science, June 1988, University of Toledo, Toledo, Ohio, USA. He has more than 50 journal and conference publications. His research areas are Natural Language Processing, Knowledge Representation, Object-oriented Programming, Ontology and Real-Time Strategy Games.

**Prof. Dr. Abdel-Badeeh M Salem** He is a Professor of Computer Science since 1989 at Ain Shams University, Egypt. His research includes intelligent computing, knowledge-based systems, biomedical informatics, and intelligent e-learning. He has published around 250 papers in refereed journals and conferences. He has been involved in more than 400 Confs and workshops as a Keynote Speaker, Scientific Program Committee, Organizer and Session Chair. He is a member of many national and international informatics associations.