

**PENGEMBANGAN BASIS DATA GEN UNTUK VISUALISASI
DAN ANALISIS INTERAKSI GEN KANKER**

SKRIPSI

Diajukan untuk Memenuhi Sebagian dari
Syarat Memperoleh Gelar Sarjana Komputer
Program Studi Ilmu Komputer



Oleh:

Aldian Fallahakbar Nuriza

1700635

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA**

2024

**PENGEMBANGAN BASIS DATA GEN UNTUK VISUALISASI DAN
ANALISIS INTERAKSI GEN KANKER**

Oleh
Aldian Fallahakbar Nuriza
1700635

Sebuah Skripsi yang Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh
Gelar Sarjana Komputer di Fakultas Pendidikan Matematika dan Ilmu
Pengetahuan Alam

© Aldian Fallahakbar Nuriza
Universitas Pendidikan Indonesia
Agustus 2024

Hak Cipta Dilindungi Undang-Undang
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian, dengan dicetak
ulang, difotokopi, atau cara lainnya tanpa izin dari penulis

ALDIAN FALLAHAKBAR NURIZA

1700635

**PENGEMBANGAN BASIS DATA GEN UNTUK VISUALISASI DAN
ANALISIS INTERAKSI GEN KANKER**

DISETUJUI DAN DISAHKAN OLEH PEMBIMBING:

Pembimbing I,



Dr. Rani Megasari, M.T.

NIP. 198705242014042002

Pembimbing II,

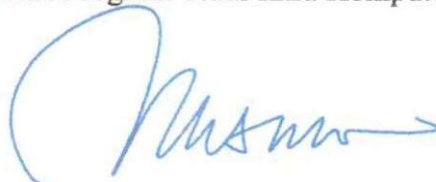


Erna Piantari, S.Kom., M.T.

NIP. 920171219890224201

Mengetahui,

Ketua Program Studi Ilmu Komputer



Dr. Muhammad Nursalman, M.T.

NIP. 197909292006041002

PENGEMBANGAN BASIS DATA GEN UNTUK VISUALISASI DAN ANALISIS INTERAKSI GEN KANKER

Oleh

Aldian Fallahakbar Nuriza – aldian.1998@upi.edu

1700635

ABSTRAK

Abstrak — Kanker merupakan sebuah penyakit ganas yang telah memakan korban dalam jumlah yang sangat besar. Menurut data *World Health Organization* yang dirilis pada tanggal 4 Februari 2024, jumlah kasus kanker di dunia mencapai 20 juta dengan kasus kematian sebesar 9,7 juta kasus. Dalam pengobatan kanker, informasi mengenai mutasi gen sangatlah berharga. Dengan mengetahui interaksi antar gen pengobatan dapat dilakukan dengan lebih terarah dengan menentukan gen sebagai target utama pengobatan. Penelitian ini ditujukan untuk membantu dalam memahami interaksi antar gen untuk penyakit kanker. Data yang digunakan merupakan data yang diunduh melalui situs web *Kyoto Encyclopedia of Genes and Genomes*. Data yang telah dikumpulkan kemudian disimpan dalam basis data MySQL dengan menerapkan tahapan *Database Life Cycle* sebagai tahapan dalam melakukan desain hingga implementasi basis data. Untuk mempermudah analisa interaksi gen, data kanker berupa kumpulan gen direpresentasikan ke dalam bentuk graf berarah yang didasari dengan teori graf. Representasi visualisasi interaksi antar gen pada penyakit kanker ini bertujuan untuk memberikan gambaran yang lebih komprehensif. Hasil penelitian ini membangun sebuah alat visualisasi jaringan kanker beserta interaksi gen yang terlibat untuk membantu proses analisa sebagai informasi pengobatan terapi kanker.

Kata kunci: Kanker, Visualisasi, Teori Graf, Interaksi Gen

PENGEMBANGAN BASIS DATA GEN UNTUK VISUALISASI DAN ANALISIS INTERAKSI GEN KANKER

Arranged by

Aldian Fallahakbar Nuriza – aldian.1998@upi.edu

1700635

ABSTRACT

Abstract — Cancer is a malignant disease that has caused a very large number of victims. According to data from the World Health Organization released on February 4, 2024, the number of cancer cases in the world reached 20 million with 9.7 million deaths. In cancer treatment, information about gene mutations is very valuable. By knowing the interactions between genes, treatment can be carried out more specifically by determining genes as the main target of treatment. This study aims to help understand the interactions between genes for cancer. The data used is data downloaded from the Kyoto Encyclopedia of Genes and Genomes website. The data that has been collected is then stored in a MySQL database by applying the Database Life Cycle stages as stages in designing to implementing the database. To facilitate the analysis of gene interactions, cancer data in the form of a collection of genes is represented in the form of a directed graph based on graph theory. The visualization representation of interactions between genes in cancer aims to provide a more comprehensive picture. The results of this study build a visualization tool for cancer networks along with the interactions of the genes involved to help the analysis process as information for cancer therapy treatment.

Keywords: Cancer, Visualization, Graph Theory, Gene Interaction

DAFTAR ISI

PERNYATAAN.....	iii
ABSTRAK.....	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR	vi
UCAPAN TERIMA KASIH.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Batasan Masalah.....	4
1.5. Manfaat Penelitian	4
1.6. Sistematika Penulisan	4
BAB II KAJIAN PUSTAKA	6
2.1. Kanker.....	6
2.1.1. Kategori Kanker	6
2.1.2. Penyebab Kanker	8
2.2. <i>Oncogene</i>	9
2.3. <i>Tumor Suppressor Gene</i>	11
2.4. <i>Signal Transduction Network</i>	12
2.4.1 Komponen Utama <i>Signal Transduction Network</i>	12
2.5. <i>Cellular Process</i>	13
2.5.1 Komponen Utama Sel	13
2.5.2 Jenis-jenis <i>Cellular Process</i>	14
2.6. <i>Kyoto Encyclopedia of Genes and Genomes</i>	14
2.7. <i>Graph Theory</i>	15
2.7.1. <i>Force-directed graph</i>	16
2.8. Basis Data	17
2.8.1 Basis Data Relasional.....	17

2.8.1.1.	Struktur Basis Data Relasional	18
2.8.1.2.	Normalisasi	18
2.8.1.3.	Bahasa Query SQL.....	19
2.8.1.4.	Siklus Hidup Basis Data	20
2.8.1.5.	Desain Basis Data	22
2.8.2.	Basis Data NoSQL	23
BAB III	METODOLOGI PENELITIAN.....	26
3.1.	Desain Penelitian.....	26
3.1.1.	Analisis.....	26
3.1.2.	Pengumpulan Data	27
3.1.3.	Perancangan Basis Data	29
3.1.3.1.	Analisis Data	29
3.1.3.2.	Desain Konseptual	30
3.1.3.3.	Desain Logika	30
3.1.3.4.	Desain Fisik.....	30
3.1.4.	Pengembangan Situs Web.....	30
3.1.5.	Visualisasi Data.....	32
3.2.	Perangkat yang Digunakan	32
BAB IV	TEMUAN DAN PEMBAHASAN	34
4.1.	Basis Data	34
4.1.1.	Desain Konseptual	34
4.1.2.	Desain Logika	36
4.1.3.	Desain Fisik.....	39
4.1.3.1.	Tabel.....	39
4.1.3.2.	View	43
4.2.	Fitur Tambah Data	44
4.2.1.	Kategori dan Nama Data.....	46
4.2.2.	Alur Mengunggah Berkas ke Basis Data	46
4.3.	Visualisasi Data.....	48
4.3.1.	Alur Visualisasi Jaringan Kanker.....	48
4.3.1.1.	Pilih Data.....	49
4.3.1.2.	Ambil Data	49

4.3.1.3.	Proses Data.....	50
4.3.1.4.	Gambar Graf.....	50
4.3.2.	Alur Visualisasi Detail Interaksi Gen	53
4.3.2.1.	Pilih Data.....	54
4.3.2.2.	Ambil Data	54
4.3.2.3.	Proses Data.....	54
4.3.2.4.	Gambar Graf.....	55
BAB V KESIMPULAN DAN SARAN.....		57
5.1.	Kesimpulan	57
5.2.	Saran.....	57
DAFTAR PUSTAKA		59

DAFTAR GAMBAR

Gambar 2.1 Kasus Jembatan Konigsberg (Ahmed, 2019).....	15
Gambar 2.2 Graf Konigsberg (Hopkins & Wilson, 2004).....	16
Gambar 3.1 Desain Penelitian.....	26
Gambar 3.2 Data Basal Cell Carcinoma	27
Gambar 3.3 Interaksi Gen pada Jaringan Basal Cell Carcinoma	28
Gambar 3.4 Hasil Interaksi Gen dalam Bentuk Graf Berarah	29
Gambar 4.1 Entitas Cancer	34
Gambar 4.2 Entitas Signal Transduction Network	35
Gambar 4.3 Entitas Cellular Process.....	35
Gambar 4.4 Entity Relationship Diagram.....	36
Gambar 4.5 Entitas Gene	36
Gambar 4.6 Relasi Entitas Cancer dengan Entitas Gene	37
Gambar 4.7 Relasi Entitas Signal Transduction Network dengan Entitas Gene ..	38
Gambar 4.8 Relasi Entitas Cellular Process dengan Entitas Gene.....	39
Gambar 4.9 Tabel gene	40
Gambar 4.10 Tabel cancer	40
Gambar 4.11 Tabel cancer_gene.....	40
Gambar 4.12 Tabel stn	41
Gambar 4.13 Tabel stn_gene.....	41
Gambar 4.14 Tabel cellular_process.....	42
Gambar 4.15 Tabel cellular_process_gene	42
Gambar 4.16 View cancer_neighbor.....	43
Gambar 4.17 View stn_neighbor	44
Gambar 4.18 View cellular_process_neighbor	44
Gambar 4.19 Tampilan Menu Tambah Data.....	45
Gambar 4.20 Tahapan Input Data Gen Kanker.....	46
Gambar 4.21 Penggalan Data Prostate Cancer	47
Gambar 4.22 Alur Visualisasi Jaringan Kanker.....	49
Gambar 4.23 Tampilan Visualisasi Tanpa Memilih Jenis Data.....	52
Gambar 4.24 Tampilan Visualisasi Data	52
Gambar 4.25 Alur Visualisasi Detail Interaksi Gen	54

Gambar 4.26 Tampilan Ketika Memilih Node	55
Gambar 4.27 Tampilan Detail Interaksi Gen	56

DAFTAR PUSTAKA

- Ahmed, H. (2019). Graph routing problem using Euler's theorem and its applications.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). From DNA to RNA. In *Molecular Biology of the Cell. 4th edition*: Garland Science.
- Angles, R., & Gutierrez, C. J. A. C. S. (2008). Survey of graph database models. *40*(1), 1-39.
- Barnes, J. A., & Harary, F. J. S. n. (1983). Graph theory in network analysis. *5*(2), 235-244.
- Beckett, D. J. J. o. m. b. (2001). Regulated assembly of transcription factors and control of transcription initiation. *314*(3), 335-352.
- Bertino, E., Sandhu, R. J. I. T. o. D., & computing, s. (2005). Database security-concepts, approaches, and challenges. *2*(1), 2-19.
- Bhowmick, P. R. (2021). Overview of NOSQL databases.
- Card, S. K., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*: Morgan Kaufmann.
- Chaffey, N. (2003). Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. Molecular biology of the cell. 4th edn. In: Oxford University Press.
- Chamberlin, D. D., & Boyce, R. F. (1974). *SEQUEL: A structured English query language*. Paper presented at the Proceedings of the 1974 ACM SIGFIDET (now SIGMOD) workshop on Data description, access and control.
- Chandra, D. G. J. F. G. C. S. (2015). BASE analysis of NoSQL database. *52*, 13-21.
- Collignon, C., Carton, M., Brisse, H. J., Pannier, S., Gauthier, A., Sarnacki, S., . . . Orbach, D. (2020). Soft tissue sarcoma in children, adolescents and young adults: Outcomes according to compliance with international initial care guidelines. *European Journal of Surgical Oncology*, *46*(7), 1277-1286. doi:<https://doi.org/10.1016/j.ejso.2019.11.518>
- Connolly, T. M., & Begg, C. E. (2005). *Database systems: a practical approach to design, implementation, and management*: Pearson Education.
- Coronel, C., & Morris, S. (2019). *Database systems: design, implementation and management*: Cengage learning.
- Croce, C. M. J. N. E. j. o. m. (2008). Oncogenes and cancer. *358*(5), 502-511.
- Damjanov, I. (2009). *Pathology secrets e-book*: Elsevier Health Sciences.
- De Nooijer, J., Lechner, L., De Vries, H. J. P. O. J. o. t. P., Social, & Cancer, B. D. o. (2001). Help-seeking behaviour for cancer symptoms: Perceptions of patients and general practitioners. *10*(6), 469-478.
- Denmeade, S. R., & Isaacs, J. T. J. N. R. C. (2002). A history of prostate cancer treatment. *2*(5), 389-396.
- Druker, B. J., Guilhot, F., O'Brien, S. G., Gathmann, I., Kantarjian, H., Gattermann, N., . . . Stone, R. M. (2006). Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *New England Journal of Medicine*, *355*(23), 2408-2417.
- Elmasri, R. (2008). *Fundamentals of database systems*: Pearson Education India.
- Elmasri, R. (2021). *Fundamentals of database systems seventh edition*.

- Freeman, L. C. (2000). Visualizing social networks. *Journal of social structure*, *1*(1), 4.
- Haerder, T., & Reuter, A. J. A. c. s. (1983). Principles of transaction-oriented database recovery. *15*(4), 287-317.
- Halvorsen, H.-P. (2016). *Structured query language*: University College of Southeast Norway.
- Han, J., Haihong, E., Le, G., & Du, J. (2011). *Survey on NoSQL database*. Paper presented at the 2011 6th international conference on pervasive computing and applications.
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, *144*(5), 646-674. doi:10.1016/j.cell.2011.02.013
- Harary, F. (1969). *Graph theory*. Retrieved from
- Hecht, R., & Jablonski, S. (2011). *NoSQL evaluation: A use case oriented survey*. Paper presented at the 2011 International Conference on Cloud and Service Computing.
- Hoffer, J. A., Ramesh, V., & Topi, H. (2016). *Modern database management*: Pearson.
- Holland, J. F. (2010). *Holland-Frei cancer medicine 8* (Vol. 8): PMPH-USA.
- Holohan, C., Van Schaeybroeck, S., Longley, D. B., & Johnston, P. G. (2013). Cancer drug resistance: An evolving paradigm. *Nature Reviews Cancer*, *13*(10), 714-726.
- Hopkins, B., & Wilson, R. J. J. T. C. M. J. (2004). The truth about Königsberg. *35*(3), 198-207.
- Ideker, T., & Nussinov, R. (2017). Network approaches and applications in biology. In (Vol. 13, pp. e1005771): Public Library of Science San Francisco, CA USA.
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2021). KEGG: Integrating viruses and cellular organisms. *Nucleic acids research*, *49*(D1), D545-D551.
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*, *45*(D1), D353-D361. doi:10.1093/nar/gkw1092
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, *28*(1), 27-30. doi:10.1093/nar/28.1.27
- King, T. (2006). *Elsevier's integrated pathology e-book*: Elsevier Health Sciences.
- Knudson, A. G. J. N. R. C. (2001). Two genetic hits (more or less) to cancer. *1*(2), 157-162.
- Kolch, W., Halasz, M., Granovskaya, M., & Kholodenko, B. N. J. N. R. C. (2015). The dynamic control of signal transduction networks in cancer cells. *15*(9), 515-527.
- Lemmon, M. A., & Schlessinger, J. J. C. (2010). Cell signaling by receptor tyrosine kinases. *141*(7), 1117-1134.
- Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, *88*(3), 323-331.
- Liaudanskaya, V., Sood, D., & Kaplan, D. L. (2019). Mechanical determinants of tissue development. In *Principles of Regenerative Medicine* (pp. 391-404): Elsevier.

- Mausumi, B., Showket, H., Richa, T., Neha, S., & Ravi, M. (2014). Human papillomavirus (HPV): Diagnosis and treatment. 95-120. doi:<https://doi.org/10.1016/B978-0-12-416002-6.00006-7>
- Merico, D., Gfeller, D., & Bader, G. D. (2009). How to visually interpret biological data using networks. *Nature biotechnology*, 27(10), 921-924.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., . . . Ding, W. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266(5182), 66-71.
- Miriam, C., Fabio, C., Diego, F., Elisa, M., Serena, M., Carlo, M., . . . Paolo, V. (2010). HER2 targeting as a two-sided strategy for breast cancer diagnosis and treatment: Outlook and recent implications in nanomedical approaches. *Pharmacological Research*, 62(2), 150-165. doi:<https://doi.org/10.1016/j.phrs.2010.01.013>
- Neves, S. R., Ram, P. T., & Iyengar, R. J. S. (2002). G protein pathways. 296(5573), 1636-1639.
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., & Kanehisa, M. (1999). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 27(1), 29-34.
- Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., . . . Kanehisa, M. (2008). KEGG atlas mapping for global analysis of metabolic pathways. *Nucleic acids research*, 36(suppl_2), W423-W426.
- Painting, F. (2012). IARC monographs on the evaluation of carcinogenic risks to humans. *WORLD HEALTH ORGANIZATION*.
- Paryanti, R., & Thobirin, A. J. U. t., Universitas Ahmad Dahlan, Umbulharjo, Yogyakarta. (2019). Penerapan teori graf untuk mencari lintasan tercepat bus Trans-Jogja.
- Patel, S. R. (2000). Radiation-induced sarcoma. *Curr Treat Options Oncol*, 1(3), 258-261. doi:10.1007/s11864-000-0037-6
- Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T. G., Kossida, S., Aerts, J., . . . Bagos, P. G. (2011). Using graph theory to analyze biological networks. *Biodata mining*, 4, 1-27.
- Pravin, S., & Sudhir, A. (2018). Integration of 3D printing with dosage forms: A new perspective for modern healthcare. *Biomedicine & Pharmacotherapy*, 107, 146-154.
- Rahmadhani, S. D., Sawitri, S., & Sandika, W. J. B. I. K. K. d. K. (2018). Mutasi onkogen dan tumor suppressor gen pada xeroderma pigmentosum. 30(1), 26-33.
- Ralph, H. (2016). Essential DW/BI background and definitions. 59-84. doi:<https://doi.org/10.1016/B978-0-12-396464-9.00004-7>
- Ramakrishnan, R., & Gehrke, J. (2002). *Database management systems*: McGraw-Hill, Inc.
- Rezvan, Y.-R., Atefeh, A., Mohammad, R., Khalil, A., & Seyed Mohammad, T. (2017). Application of aptamers in treatment and diagnosis of leukemia. *International Journal of Pharmaceutics*, 529(1), 44-54. doi:<https://doi.org/10.1016/j.ijpharm.2017.06.058>
- Rous, P. (1979). A transmissible avian neoplasm. (Sarcoma of the common fowl) by Peyton Rous, M.D., *Experimental Medicine for Sept. 1, 1910*, vol. 12, pp.696-705. *J Exp Med*, 150(4), 738-753. doi:10.1084/jem.150.4.729

- Rowinsky, E. K. J. T. O. (2003). Signal events: Cell signal transduction and its inhibition in cancer. *8*, 5-17.
- Samiya, K., Xiufeng, L., Syed Arshad, A., & Mansaf, A. (2023). Bivariate, cluster, and suitability analysis of NoSQL solutions for big graph applications. *Advances in Computers*, *128*, 39-105. doi:<https://doi.org/10.1016/bs.adcom.2021.09.006>
- Scaltriti, M., & Baselga, J. (2006). The epidermal growth factor receptor pathway: A model for targeted therapy. *Clinical cancer research*, *12*(18), 5268-5272.
- Schwaederle, M., Zhao, M., Lee, J. J., Eggermont, A. M., Schilsky, R. L., Mendelsohn, J., . . . Kurzrock, R. (2015). Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. *Journal of clinical oncology*, *33*(32), 3817-3825.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome research*, *13*(11), 2498-2504.
- Silberschatz, A., Korth, H. F., & Sudarshan, S. (2011). Database system concepts.
- Simarmata, A. M., Sinaga, T. P., Rudianto, R., & William, O. J. J. M. I. (2019). Deteksi penyakit limfoma dengan menggunakan algoritma region growing. *4*(2), 25-32.
- Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature*, *458*(7239), 719-724.
- Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature medicine*, *10*(8), 789-799.
- Vousden, K. H., & Prives, C. (2009). Blinded by the light: The growing complexity of p53. *Cell*, *137*(3), 413-431.
- Yusuf, A. F. N., & Rizki, M. J. J. M. H. (2023). Multiple myeloma: Mengenal awitan hingga prognosis. *4*(02 Januari), 3265-3270.
- Zur Hausen, H. (2009). Papillomaviruses in the causation of human cancers—a brief historical account. *Virology*, *384*(2), 260-265.