

# PEMODELAN DIGITALISASI DAN VISUALISASI PENGADUAN KETENAGAKERJAAN MENGGUNAKAN OCR, FEATURE EXTRACTION DAN BUSINESS INTELLIGENCE

Yan Puspitarani, Sukenda, Ari Purno Wahyu, Benny Yustim, Sunjana

Fakultas Teknik

Universitas Widyatama

Jl. Cikutra No. 204A, Bandung

[yan.puspitarani@widyatama.ac.id](mailto:yan.puspitarani@widyatama.ac.id), [kenda@widyatama.ac.id](mailto:kenda@widyatama.ac.id), [ari.purno@widyatama.ac.id](mailto:ari.purno@widyatama.ac.id),

[byustim@widyatama.ac.id](mailto:byustim@widyatama.ac.id), [sunjana@widyatama.ac.id](mailto:sunjana@widyatama.ac.id)

## Abstrak

Penanganan pengaduan ketenagakerjaan merupakan salah satu faktor kinerja Disnakertrans. Kinerja ini berhubungan dengan keputusan yang harus diambil untuk menentukan kebijakan agar pelanggaran norma ketenagakerjaan dapat berkurang. Pengambilan keputusan ini akan lebih mudah jika data pengaduan dapat direkapitulasi secara cepat dan akurat. Rekapitulasi ini akan menjadi alat pantau pimpinan dalam melihat laporan kemajuan status pengaduan yang masuk. Akan tetapi proses administrasinya yang manual membuat rekapitulasinya semakin lambat. Oleh karena itu, penelitian ini akan memodelkan sistem digitalisasi pelaporan pengaduan dengan memanfaatkan OCR dan ekstraksi informasi serta memodelkan visualisasi hasil rekapitulasi menggunakan Business Intelligence. Dengan dibuatnya model ini, diharapkan kinerja Disnakertrans dalam menyelesaikan pengaduan ketenagakerjaan lebih efektif.

Kata kunci :

Pengaduan ketenagakerjaan, visualisasi, business intelligence.

## Abstract

*The handling of labor complaints is one of the performance factors for Disnakertrans. This performance is related to the decisions that must be taken to determine policies so that violations of labor norms can be reduced. Making this decision will be easier if the data on complaints can be recapitulated quickly and accurately. This recapitulation will be a tool to monitor the leadership in seeing the progress report on the status of incoming complaints. However,*

*the manual administrative process makes the recapitulation slower. Therefore, this study will model the complaint reporting digitization system by utilizing OCR and information extraction as well as modeling the visualization of the results of the recapitulation using Business Intelligence. With the creation of this model, it is hoped that the performance of Disnakertrans in resolving labor complaints will be more effective.*

Keywords :

Employment complaints, visualization, business intelligence.

## I. PENDAHULUAN

Terlaksananya norma ketenagakerjaan di tempat kerja menjadi salah satu tugas pokok Dinas Tenaga Kerja. Para pengawas ketenagakerjaan akan memeriksa perusahaan secara berkala atau berdasarkan pengaduan yang diajukan oleh masyarakat. Penyelesaian masalah pengaduan ini menjadi salah satu faktor kinerja Disnakertrans yang berhubungan dengan pengambilan keputusan yang diambil oleh pimpinan dalam menentukan kebijakan. Pimpinan harus mengetahui bagaimana perkembangan status pengaduan yang ditangani oleh Disnakertrans agar dapat menentukan kebijakan dengan tepat. Dengan kebijakan yang tepat, maka pelanggaran terhadap norma ketenagakerjaan dapat berkurang.

Pelanggaran norma ketenagakerjaan seperti Pemutusan Hubungan Kerja(PHK) secara sepihak, gaji yang tidak dibayar, atau hak-hak tenaga kerja yang tidak dipenuhi perusahaan seringkali diajukan

para tenaga kerja. Pengaduan ini dikirimkan kepada Disnakertrans dalam bentuk surat secara resmi. Pesatnya kemajuan teknologi saat ini membuat pengiriman surat pengaduan dilakukan secara digital melalui email atau chat. Hal ini membuat proses rekapitulasi surat yang semakin lambat dikarenakan jumlah surat pengaduan yang masuk semakin cepat sampai.

Kemudahan pengiriman surat ini tentunya mewajibkan para pegawai instansi untuk bekerja lebih cepat dalam menangani pengaduan yang masuk. Sistem rekapitulasi pengaduan yang masih manual akan membuat kinerja pegawai melambat. Oleh karena itu, diperlukan suatu sistem yang dapat merekap surat yang masuk secara otomatis agar dapat segera ditangani.

Hasil penanganan pengaduan ketenagakerjaan yang masuk pun harus diketahui oleh pimpinan agar dapat melihat kinerja para pegawai instansi. Jumlah pengaduan yang berhasil diselesaikan dinas atau progress kemajuan status pengaduan yang masuk harus dapat dipantau oleh pimpinan agar dapat melihat kinerja dinas. Pemantauan ini dapat dilakukan melalui visualisasi data pengaduan yang dapat diperbaharui setiap waktunya.

Oleh karena itu, pada penelitian ini, akan dilakukan pemodelan sistem dengan memanfaatkan OCR dan Ekstraksi Informasi pada penelitian sebelumnya (Puspitarani, Yan; Syukriyah, 2020) dan Business Intelligence untuk menghasilkan visualisasi data pengaduan.

## II. LANDASAN TEORI

Beberapa penelitian telah dilakukan dalam digitalisasi proses persuratan dalam organisasi. Salah satunya adalah Fridhayantie (Fridayanthie & Fauzi, 2019) membuat Sistem Informasi untuk mengelola surat keluar masuk di perusahaan dan Hunaifi (Hunaifi et al., 2019) membuat sistem permohonan surat online. Selain itu, Suprpto (Suprpto & Rahman Prehanto, 2020) juga membuat pengelolaan arsip secara dinamis meliputi klasifikasi arsip, jadwal retensi arsip dan pengusulan tindak lanjut arsip. Ada pula Megawati (Megawaty, 2020) dan Juairiyah (Juairiyah & Hendrixon, 2017) yang merancang sistem e-office pada instansi pemerintahan. Terakhir, telah dilakukan juga penelitian untuk mengekstrak informasi surat pengaduan ketenagakerjaan secara otomatis dengan menggunakan teknologi OCR dan Ekstraksi Informasi seperti nomor surat, tanggal surat,

perihal surat, pengirim surat, dan isi surat (Puspitarani, Yan; Syukriyah, 2020).

### II.1 Optical Character Recognition (OCR)

Optical Character Recognition (OCR) adalah konversi dari hasil pindai atau cetak gambar berisi text dan tulisan tangan menjadi text yang dapat didit untuk pemrosesan selanjutnya. Teknologi ini menggunakan sebuah mesin untuk mengenali text secara otomatis. Jika dianalogikan dengan bagian tubuh manusia, OCR seperti kombinasi dari mata dan pikiran. Mata dapat melihat text pada gambar sedangkan otak memproses interpretasi text yang dibaca oleh mata (Patel et al., 2012).

Untuk dapat mengubah gambar menjadi text, Tahap pertama yang dilakukan adalah menganalisis komponen-komponen dasar dari gambar yang menunjukkan text. Algoritma pencarian baris akan dilakukan untuk menentukan baseline dari text. Baseline merupakan garis batas dimana text pada gambar ditemukan. Setelah posisi baseline yang pas ditemukan, setiap baris yang ditemukan akan dipotong-potong untuk menemukan kata berdasarkan spasi dan kata tersebut dipotong lagi untuk menemukan karakter untuk dilakukan proses pengenalan. Pengenalan karakter dan kata ini memanfaatkan mesin learning berupa klasifikasi berdasarkan ciri-cirinya (Smith, 2007).

### II.2 Feature Extraction

Feature extraction merupakan proses untuk mengurangi dimensionalitas suatu data sehingga data dapat direpresentasikan dalam bentuk bagian-bagian data yang berguna saja untuk efisiensi penyimpanan feature (Agrawal & Samson, 2016). Feature extraction digunakan untuk mengurangi kompleksitas dan menyajikan representasi data secara sederhana. Representasi data yang dihasilkan pada setiap variabel hasil dari feature extraction berupa kombinasi linear dari variabel masukan aslinya (Nasreen & Mary, 2015). Hal ini sangat berguna jika dilakukan terhadap data gambar yang berukuran besar.

Text Feature Extraction melakukan ekstraksi terhadap informasi berupa teks menjadi bagian-bagian kecil sehingga data yang dihasilkan lebih terstruktur. Proses ini menjadi hal yang sangat mendasar dan penting untuk proses text mining dan information retrieval (Liang et al., 2017).

Untuk mendapatkan informasi dari teks, hal pertama yang dilakukan adalah membagi teks menjadi

kalimat dan setiap kalimat akan dipecah kembali menjadi kata-kata melalui proses tokenisasi. Setelah itu, setiap kalimat akan ditandai menggunakan POS tagging (pemberian nama kelas dari kata, contoh : kata benda, kata kerja, kata sifat, dan lain-lain) yang akan digunakan pada langkah berikutnya yaitu named entity detection. Pada tahap ini, dilakukan pencarian entitas-entitas khusus yang berpotensi untuk diberi label pada setiap kalimat. Selanjutnya dilakukan deteksi relasi untuk mencari kemungkinan hubungan antara entitas yang berbeda dalam teks (Bird, Steven; Klein, Ewan; Loper, 2009). Pada penelitian ini, corpus data untuk pengenalan tag menggunakan tagging Bahasa Indonesia (Rashel et al., 2014).

### II.3 Business Intelligence

Business Intelligence menyediakan informasi bisnis dan analisis bisnis yang terdapat konteks proses utama sebuah bisnis yang mengarah pada keputusan dan tindakan. Hal itu menghasilkan peningkatan kinerja bisnis. Secara spesifiknya Business Intelligence berarti meningkatkan informasi dalam proses bisnis utamanya untuk mencapai peningkatan kinerja bisnis (Williams, Steve; Williams, 2007).

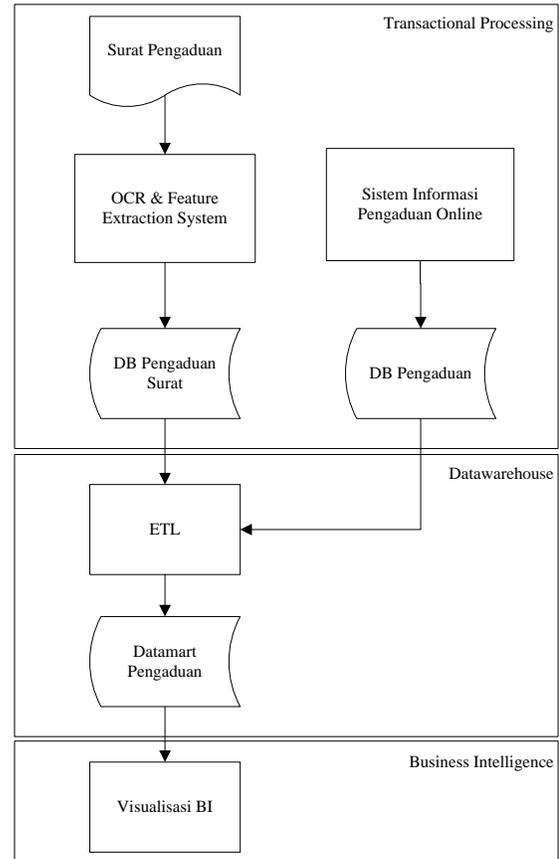
Menurut Vercellis (2009), fungsi utama dari Business intelligence adalah untuk membantu pengguna pengetahuan dengan menggunakan perangkat kerja dan metodologi business intelligence untuk menghasilkan keputusan secara efektif dan tepat pada waktunya (effective decision and timely decision) (Vercellis, 2009).

Business intelligence memberikan kemudahan bagi para pengambil keputusan untuk memilih informasi dan pengetahuan yang ada, dimana informasi dan pengetahuan tersebut dapat diandalkan. Hasilnya, para pengambil keputusan dapat membuat keputusan yang lebih baik dan menemukan langkah selanjutnya yang diperlukan untuk memperoleh hasil yang lebih optimal dan efektif (Vercellis, 2009).

Perusahaan beroperasi pada lingkungan ekonomi yang memiliki tingkat pertumbuhan persaingan dan dinamika yang tinggi. Akibatnya, perusahaan harus memiliki kemampuan bersaing dan beraksi secara cepat terhadap tindakan pesaing dan kondisi pasar yang baru. Kemampuan mengambil keputusan dengan cepat sesuai dengan kondisi pasar merupakan faktor penting yang menentukan keberhasilan dan kelangsungan hidup suatu perusahaan. Dengan adanya business intelligence maka kualitas pengambilan keputusan akan meningkat dengan

bantuan model matematika dan algoritma yang bisa melakukan analisis dalam jumlah besar untuk menghasilkan alternatif pengambilan keputusan (Vercellis, 2009).

### III. PEMODELAN SISTEM

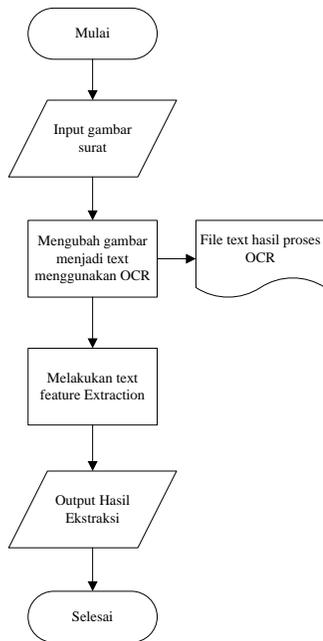


**Gambar 1. Model Digitalisasi dan Visualisasi Pengaduan Ketenagakerjaan**

Gambar 1 menjelaskan pemodelan sistem dalam tiga tahapan, yaitu :

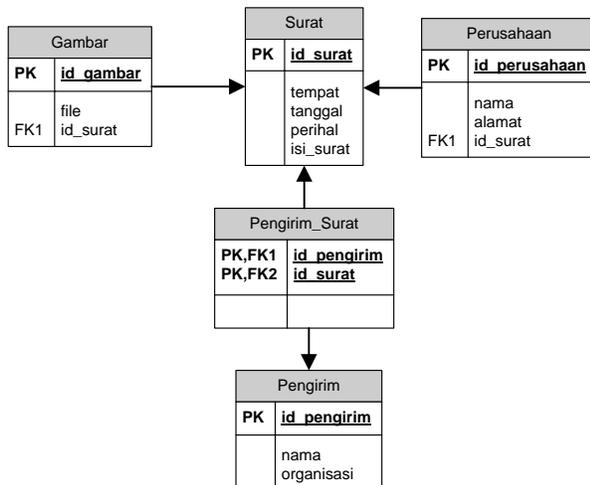
#### 1. Transactional Processing

Tahap ini merupakan tahapan proses digitalisasi dari pengaduan ketenagakerjaan yang diterima Dinas. Proses ini merupakan kegiatan harian Dinas yang dilakukan melalui Sistem Pengaduan Online yang sudah dimiliki Dinas dan Sistem Ekstraksi Surat Pengaduan Otomatis yang pernah dibangun pada penelitian sebelumnya. Gambar 2 menunjukkan alur kerja Sistem Ekstraksi Surat Pengaduan Otomatis.



**Gambar 2. Alur kerja Sistem Ekstraksi Surat Pengaduan Otomatis** (Puspitarani, Yan; Syukriyah, 2020)

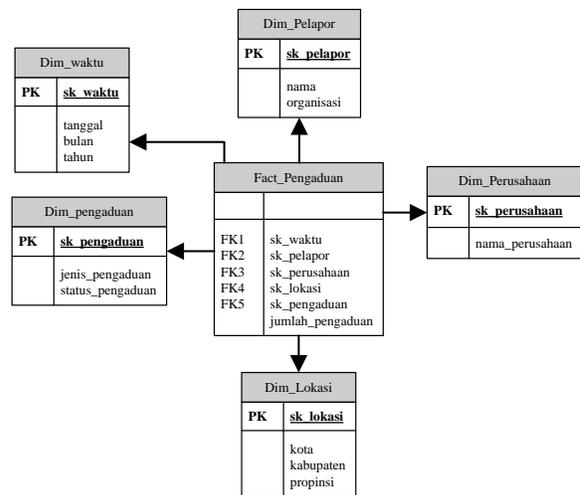
Hasil dari tahap ini adalah dua buah basisdata pengaduan ketenagakerjaan dari masing-masing sistem. Pada Sistem Pengaduan Online dihasilkan data berupa : kode\_pengaduan, nama pengadu, alamat, perusahaan yang diadukan, pengaduan, kategori pengaduan, dan status pengaduan. Sedangkan basisdata Sistem Ekstraksi Surat Pengaduan menghasilkan skema seperti pada gambar 3 berikut :



**Gambar 3. Model Konseptual Basisdata Surat Pengaduan** (Puspitarani, Yan; Syukriyah, 2020)

## 2. Datawarehouse

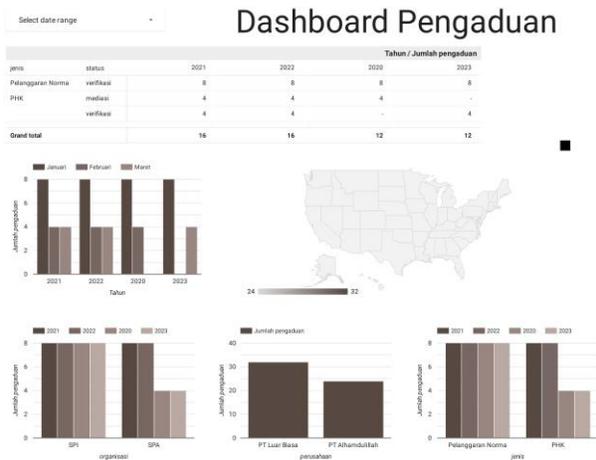
Pada tahap transactional processing menghasilkan dua basisdata yang berbeda struktur. Hal ini membutuhkan penyeragaman struktur dari datawarehouse. Pada tahap ini akan dilakukan proses ETL(Extract, Transform, Load). Pada proses ini data pada kedua basisdata akan diambil (extract) kemudian dilakukan penyesuaian (transform) strukturnya dengan skema datamart pengaduan yang digambarkan pada Gambar 4. Setelah itu, data hasil penyesuaian akan dimasukkan (load) ke dalam datamart Pengaduan. Datamart ini akan menjadi sumber data untuk tahap selanjutnya.



**Gambar 4. Star Schema Datamart Pengaduan**

## 3. Business Intelligence

Tahap ini merupakan tahap pembangunan visualisasi berdasarkan datamart pengaduan. Gambar 5 menunjukkan pemodelan visualisasi yang dapat dihasilkan dari datamart pengaduan.



**Gambar 5. Visualisasi Datamart Pengaduan**

Visualisasi yang dihasilkan datamart pengaduan berdasarkan Gambar 5 di atas terdiri dari :

- Visualisasi Jumlah Pengaduan dalam satuan waktu (hari, bulan, tahun) yang terdiri dari dimensi waktu dan fakta jumlah pengaduan. Pada visualisasi ini, user dapat melakukan drill data secara harian, bulanan, dan tahunan.
- Visualisasi Jumlah Pengaduan setiap lokasi (kota/kabupaten, propinsi) dalam satuan waktu tertentu yang terdiri dari dimensi waktu, dimensi lokasi, dan fakta jumlah pengaduan. Pada visualisasi ini, user dapat melihat jumlah pengaduan melalui peta lokasi di Jawa Barat.
- Visualisasi Jumlah Pengaduan setiap status dan jenis pengaduan dalam satuan waktu tertentu yang terdiri dari dimensi waktu, dimensi pengaduan, dan fakta jumlah pengaduan. Pada visualisasi ini, user dapat melakukan drill data secara harian, bulanan, dan tahunan.
- Visualisasi Jumlah Pengaduan setiap perusahaan dalam satuan waktu tertentu yang terdiri dari dimensi waktu, dimensi perusahaan, dan fakta jumlah pengaduan. Pada visualisasi ini, user dapat melakukan drill data secara harian, bulanan dan tahunan.
- Visualisasi Jumlah Pengaduan setiap organisasi dalam satuan waktu tertentu yang terdiri dari dimensi waktu, dimensi organisasi dan fakta jumlah pengaduan. Pada visualisasi ini, user dapat melihat breakdown jumlah pengaduan yang diajukan organisasi setiap tahunnya.

- Visualisasi Jumlah Pengaduan setiap lokasi dengan status dan jenis pengaduan dalam satuan waktu tertentu yang terdiri dari dimensi waktu, dimensi lokasi, dan dimensi pengaduan. Pada visualisasi ini, user dapat melakukan drill data jenis pengaduan dan status pengaduan setiap tahun.

#### IV. KESIMPULAN DAN SARAN

Penelitian ini menghasilkan pemodelan sistem digitalisasi dan visualisasi pengaduan di lingkungan Dinas Tenaga Kerja dan Transmigrasi Provinsi Jawa Barat. Dengan dibuatnya model ini diharapkan proses pengaduan ketenagakerjaan dapat berjalan secara transparan dan lebih tertib secara administrasi. Selain itu, visualisasi yang dihasilkan dari datamart pengaduan dapat membantu pimpinan dinas dalam melihat capaian kinerja dinas serta dapat membantu dalam penentuan keputusan.

#### UCAPAN TERIMAKASIH

Penelitian ini dibiayai oleh Lembaga Penelitian, Pengabdian kepada Masyarakat, dan Modal Intelektual (LP2M) Universitas Widyatama.

#### REFERENSI

- Agrawal, A., & Samson, S. (2016). A Review on Feature Extraction Techniques and General Approach for Face Recognition. *International Journal of Computer Applications Technology and Research*, 5(3), 156–158. <https://doi.org/10.7753/ijcatr0503.1008>
- Bird, Steven; Klein, Ewan; Loper, E. (2009). *Natural language processing with python* (1st editio). O'Reilly.
- Fridayanthie, E. W., & Fauzi, A. (2019). Rancang Bangun Sistem Informasi Pengelolaan Surat Masuk Dan Surat Keluar Perusahaan. *Paradigma - Jurnal Komputer Dan Informatika*, 21(1), 43–48. <https://doi.org/10.31294/p.v21i1.4942>
- Hunaifi, N., Hikmah, A. B., & Nurhasan, A. (2019). Perancangan Sistem Informasi Pengarsipan Dan Permohonan Surat Online “Sipadu” Di Tingkat Kecamatan Berbasis Web. *JUST IT: Jurnal Sistem Informasi, Teknologi Informasi Dan*

- Komputer*, 10(1), 40–51.
- Juairiyah, O., & Hendrixon. (2017). Penerapan E-Office Dalam Administrasi Perkantoran (Studi Kasus: Balitbangda Prov. Sumsel). *Jurnal Pembangunan Nagari*, 2(1), 75–84.
- Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: a review. *Eurasip Journal on Wireless Communications and Networking*, 2017(1), 1–12. <https://doi.org/10.1186/s13638-017-0993-1>
- Megawaty. (2020). Aplikasi E-Office pada Badan Perencanaan dan Pembangunan Daerah ( BPPD ) Menggunakan Pendekatan Metode Extreme Programming. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 176–181.
- Nasreen, S., & Mary, Q. (2015). *A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning*. July.
- Patel, C., Patel, A., & Patel, D. (2012). Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study. *International Journal of Computer Applications*, 55(10), 50–56. <https://doi.org/10.5120/8794-2784>
- Puspitarani, Yan; Syukriyah, Y. (2020). Pemanfaatan Optical Character Recognition Dan Text Feature Extraction Untuk Membangun Basisdata Pengaduan Tenaga Kerja. *Rekayasa Sistem Dan Teknologi Informasi*, 4, 704–710.
- Rashel, F., Luthfi, A., Dinakaramani, A., & Manurung, R. (2014). Building an Indonesian Rule-Based Part-of-Speech Tagger. *International Conference on Asian Language Processing (IALP 2014)*.
- Smith, R. (2007). An Overview of the Tesseract OCR Engine. *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. <https://doi.org/10.1109/ICDAR.2007.4376991>
- Suprpto, R., & Rahman Prehanto, D. (2020). Pengembangan Sistem Informasi Manajemen Arsip Dinamis Dalam Mendukung Tata Kelola Kearsipan Berbasis Web Menggunakan Metode Sdlc. *Jurnal Manajemen Informatika*, 11(1), 35–42.
- Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. *Business Intelligence: Data Mining and Optimization for Decision Making*, 1–417. <https://doi.org/10.1002/9780470753866>
- Williams, Steve; Williams, N. (2007). *The Profit Impact of Business Intelligence*.