**KONU MODELLEME İLE YAŞAM KALİTESİ LİTERATÜRÜNÜN İNCELENMESİ: TÜRKİYE ÖRNEĞİ**

Özlem Ergüt

**1.GİRİŞ**

Yaşam kalitesi tarih boyunca, ekonomistlerin, politika yapıcıların, sosyal bilimcilerin ilgi konusu olmuştur. Pek çok farklı disiplin tarafından ele alınıp, incelenene bir konu olan yaşam kalitesinin pek çok farklı tanımına rastlamak mümkündür. Dünya Sağlık Örgütü tarafından yaşam kalitesi, bireyin, içinde yaşadığı kültür ve değer sistemleri bağlamında amaçları, beklentileri, standartları ve endişeleriyle ilgili olarak yaşamdaki konumuna ilişkin algısı olarak tanımlanmaktadır. Daha geniş çerçevede ise insanların fiziksel sağlıklarını, psikolojik durumlarını, bağımsızlık düzeylerini, sosyal ilişkilerini, kişisel inançlarını ve çevrenin göze çarpan özellikleriyle olan ilişkilerini birleştiren bir kavram olarak tanımlanmıştır (WHOQOL, 2012).

Sağlık, iş imkanı, eğitim, çevre koşulları, sosyal ilişkiler gibi yaşam kalitesini etkileyen pek çok faktör bulunduğundan, yaşam kalitesi çeşitli disiplinlerde kendine uygulama alanı bulmuştur. Yaşam kalitesinin bireyler ve toplum üzerindeki derin etkisi bu kavramın kamu politikası, sağlık hizmetleri, kentsel planlama gibi çeşitli alanlar üzerinde hedeflerin belirlenmesinde önemli rol oynamıştır.

Literatür incelendiğinde yaşam kalitesinin ölçülmesi ve iyileştirilmesine yönelik kriterler konusunda oldukça fazla sayıda yayına rastlamak mümkündür. Çok boyutlu ve dinamik bir araştırma alanı olan yaşam kalitesinin ölçülmesi ve yaşam kalitesindeki değişikliklerin belirlenmesindeki göstergelerin seçilmesi önemli bir araştırma alanı olarak karşımıza çıkmaktadır. Aynı zamanda yaşam süresinin uzamasıyla birlikte sağlık alanında yapılan çalışmalarda da yaşam kalitesinin ele alınıp sıklıkla incelendiği görülmüştür.

Yaşam kalitesine ilişkin çeşitli bakış açılarını analiz ederken, refahın fiziksel, zihinsel ve sosyal boyutlarının birbirine bağlılığının yanı sıra yapısal ve çevresel faktörlerin bireylerin yaşamları üzerindeki etkisini tanımak önemlidir.

**2.METİN MADENCİLİĞİ VE KONU MODELLEME**

Metin madenciliği, değerli bilgileri keşfetmek ve çıkarmak için yapılandırılmamış metin verilerini analiz etme süreci olarak tanımlanabilir. İnternetin yaşamın her alanında kullanılması ile birlikte üretilen veri miktarı artmış, bu verinin analiz edilip anlamlı bilgiler çıkarılabilmesi için çeşitli teknikler geliştirilmiştir.

Metin madenciliği işleyiş süreci, web siteleri, sosyal medya, e-postalar ve belgeler gibi çeşitli kaynaklardan büyük miktarlarda yapılandırılmamış metin verilerinin toplanmasıyla başlar. Bu veriler çeşitlli tekniklerle işlenerek veriler içindeki kalıplar, eğilimler ve ilişkiler tanımlanabilir hale gelmektedir. Elde edilen bilgiler müşteri tercihleri, pazar eğilimleri, sağlık alanında klinik notlar ve tıbbi kayıtların incelenmesi, hizmet geliştirilmesi için müşteri yorumlarının analizi ve diğer önemli bilgiler hakkında değerli bilgiler sağlayacağından pek çok farklı alanda metin madenciliği tekniklerinden yararlanılmaktadır. Bu tekniklerden birisi de konu modellemedir.

Konu modelleme bir belge koleksiyonunda bulunan soyut konuları keşfetmeyi amaçlayan bir metin madenciliği tekniğidir. Bilgi erişimi, belge sınıflandırma ve öneri sistemleri gibi çeşitli alanlardaki potansiyel etkisi nedeniyle yıllar içinde büyük ilgi görmüştür.

Konu modelleme ile belgelerin anlamlı konulara göre otomatik olarak atanmasına olanak tanınarak büyük metin koleksiyonlarının yönetilmesi ve analiz edilmesi daha kolay hale gelmektedir.

Konu modelleme üç temel üzerine kuruludur. Bunlar: kelimeler (words), belgeler (documents) ve külliyat (corpus)’tır. Kelime, temel veri birimi iken; belge, bir metinsel veri birimi olarak kelime dizisinden oluşmakta; külliyat (:derlem) ise, belgeler topluluğu olarak tanımlanmaktadır. Konu modelleme de kelime dağarcığındaki sıralamanın önemi yoktur. En popüler konu modellemesi teknikleri şunlardır:

- Gizli Anlam Analizi (Latent Semantic Analysis-LSA)

-Olasılıksal Gizli Semantik Analiz (Probabilistic Latent Semantic Analysis -pLSA)

- Gizli Dirichlet Ayrımı (Latent Dirichlet Allocation -LDA)

- Yapısal Konu Modelleme (Structural Topic Modeling-STM)

Bir vektör uzayının boyutunu azaltmak için kullanılan tekniklerden biri olan Gizli Anlam Analizi (GAA), kelimelerin gözlenen birlikte oluşumuna dayalı olarak metnin üstü örtülü bir temsilini türetmek için kullanılan denetimsiz bir öğrenme tekniğidir. Hofmann (1999) tarafından geliştirilen Olasılıksal Gizli Anlam Analizi (OGAA), GAA’dan daha esnek ve daha sağlam bir istatistiksel temele sahip olup, gizli değişkenlerin belirlenmesi için olasılıksal bir yaklaşım sağlamaktadır. Blei, Ng ve Jordan tarafından geliştirilen ve en sık kullanılan olasılıklı konu modelleme tekniklerinden biri Gizli Dirichlet Ayrımı (GDA)’dır. GDA her belgeyi, her konunun kelimeler üzerinde çok terimli bir dağılıma karşılık geldiği, konuların bir karışımı olarak modelleyen üretken bir süreçtir. GDA'nın arkasındaki temel fikir, belgelerin konulardan oluşması ve her konunun o konuyu en iyi şekilde tanımlayan bir dizi kelimeden meydana gelmesidir.

Belgeler içerisindeki konuları bulmayı hedefleyen, konu modelleme yaklaşımlarında sadece o belgedeki metin dikkate alınmakta, diğer bilgiler (yazar, tarih, kaynak, vb.) göz önünde bulundurulmamaktadır. Yapılan kimi çalışmada konuların ortaya çıkartılmasında bu meta veriler önemli olmaktadır. Araştırmacıların geniş bir metin verisi koleksiyonundaki gizli yapıları ortaya çıkarmasına ve analiz etmesine olanak tanıyan istatistiksel bir yöntem olan Yapısal Konu Modelleme (YKM) ile konular ile yazarlık, yayın tarihi ve belge uzunluğu gibi diğer değişkenler arasındaki ilişkilere ilişkin ek bilgiler sağlayabilen çeşitli meta veri kaynaklarını birleştirilerek geleneksel konu modelleme tekniklerinden daha zengin bir veriyle çalışma imkanı sağlanmaktadır. Diğer bir deyişle Yapısal Konu Modelleme’de meta veriler modele dahil edilerek konular belirlenmektedir.

Roberts vd. tarafından geliştirilen YKM, konular arasındaki etkileşimleri ve bunların bir belgenin yazarı veya yazıldığı dönem gibi diğer değişkenlerden nasıl etkilendiğini hesaba katarak konuların ve bunların ilişkilerinin daha kapsamlı bir görünümünü sunarak metin verilerinin altında yatan yapıların daha zengin ve daha ayrıntılı analizini sunarak daha nitelikli bilgi elde edilmesini sağlar.

YKM’de GDA’da olduğu gibi her belge K tane konunun bir karışımı olarak ortaya çıkmaktadır. YKM'de konu oranları (θ) ilişkilendirilebilir ve bu konuların yaygınlığı, standart bir regresyon modeli aracılığıyla bazı ortak değişkenler (X) kümesinden etkilenebilir (θ∼LogisticNormal( )). Yanıttaki her kelime (w) için, yanıta özel dağılımdan bir konu (z) çekilir ve bu konuya bağlı olarak, β tarafından parametrelendirilen kelimeler üzerinden çok terimli bir dağılımdan bir kelime seçilir. Bu dağılım, ikinci bir U ortak değişken kümesini içerebilir.

GDA ile karşılaştırıldığında YKM’nin üç temel farklılığı ön plana çıkmaktadır. Bunlar:

-Konular birbiriyle ilişkilendirilebilir olması

-Her belgenin, genel bir ortalamayı paylaşmak yerine ortak değişken X tarafından tanımlanan, konular üzerinde kendi ön dağılımının olması

-Bir konu içindeki kelime kullanımının ortak değişkene (U) göre değişebilmesidir. (Roberts ve diğ., 2014)

Diğer konu modelleme tekniklerinden farklı olarak ortak değişkenlerin YKM’de mevcut olması belgelerden çıkarım yaparken zengin ve değerli bilgiler ekleyerek konu modelindeki dağılımların yapılandırılmasını sağlar. YKM’nin diğer konu modeleme tekniklerinden farklı olarak, konu yaygınlığının ve içeriğinin tahmin edilmesine de olanak tanımakta; bu da araştırmacıların yalnızca metinde hangi konuların yer aldığını değil aynı zamanda bunların ne ölçüde tartışıldığını ve her konuyu tanımlayan belirli kelimeleri anlamalarına yardımcı olmaktadır.

YKM modeli, araştırmacının yalnızca metin verilerindeki gizli konuların bulunmasını değil, aynı zamanda gizli tematik yapılar ile meta veriler arasındaki mevcut ilişkinin de belirlemesine olanak tanır. Örneğin konuların yaygınlığının zaman içinde ya da bazı bölgelerde nasıl bir değişim gösterdiği saptanabilmektedir.

**3.ÇALIŞMANIN AMACI VE KAPSAMI**

Çalışmanın amacı hem yeni bir teknik olan YKM’nin araştırmacılara tanıtılması hem de yaşam kalitesi literatürüne dair kapsamlı bir inceleme yapmaktır. YKM’nin, yaşam kalitesi gibi güncel bir konunun literatür incelemesinde kullanılması, daha önce bu alanda yapılan çalışmalardan farklılaşmasına neden olmaktadır. Çalışma kapsamında, 2000 yılı ve sonrasında yayınlanan yaşam kalitesi ile ilgili makaleler, YKM tekniği ile analiz edilerek; bu literatürde çalışılan alt başlıklar ve bu başlıkların akademik yayınlardaki yaygınlığı belirlenmiş, yıllara göre değişimleri incelenmiştir.

Bu çalışmada cevap aranan araştırma soruları şunlardır:

1)Yaşam kalitesi ile ilgili son dönemdeki araştırma temaları/konuları nelerdir?

2) Bütün külliyatta yüksek oranda yer alan konular nelerdir?

3) Konular birbirleriyle nasıl ilişkilidir?

4) Yayın yılının konunun önemi üzerindeki etkileri nelerdir?

**4.VERİ SETİ VE TANIMLAYICI İSTATİSTİKLER**

Çalışmanın veri seti, SCI, SCI-Expanded, SSCI ve AHCI’de yayınlanmış Web of Science (WOS) üzerinden “*quality of life*” anahtar kelimesi ile taranan Türk araştırmacılar tarafından kaleme alınan makalelerden oluşmaktadır. Veri seti; makalelerin *başlık, yazarlar, yayın yılı ve özet* bilgilerini içermektedir. Çalışmanın akış sürecine ilişkin akış diyagramı Şekil 1’de gösterilmiştir.

**Şekil 1. Analiz Akış Şeması**

Belirlenen kriterlere göre yapılan tarama işlemi sonucunda toplam 11932 yayına ulaşılmıştır.

Yayınların doküman türüne göre dağılımı Tablo 1’de özetlenmiştir.

**Tablo 1.“Yaşam Kalitesi” Literatürünün Yayın Türüne Göre Dağılımı**

|  |  |
| --- | --- |
| ***Yayın türü*** | ***Yayın Sayısı*** |
| Makale | 10124 |
| Makale değerlendirmesi | 933 |
| Özet bildiri | 747 |
| Kitap Bölümü | 8 |
| Tam metin bildiri | 159 |
| Diğer | 39 |
| *TOPLAM* | 11932 |

Veri setinin analizine yayın türü olarak büyük bir çoğunluğa sahip olan makalelerle devam edilmiştir. Makale sayılarının yıllara, yayın diline göre dağılımını ve WOS kategorilerine göre yayınların kategorizasyonunu içeren performans analizi bulguları aşağıda verilmiştir.

**Şekil 2. Yaşam Kalitesi Konulu Makale Sayısının Yıllara Göre Dağılımı**

Makalelerin yıllara göre dağılımı incelendiğinde 2000 yılı itibariyle hızlı bir artış eğilimi sergilediği görüldüğünden, veri seti 2000 yılından günümüze olacak şekilde derlenmiştir. Şekil 2’de görülebileceği gibi, son 22 yıllık süreçte yaşam kalitesi konusuna akademik ilginin arttığı görülmektedir.

WOS tarafından tanımlanan kategorilere göre makalelerinin en fazla yer aldığı alanlar Tablo 2’de verilmiştir.

**Tablo 2. Yaşam Kalitesi Konulu Makalelerin WOS Kategorilerine Göre Dağılımı**

|  |  |
| --- | --- |
| *Kategori* | *Makale Sayısı* |
| Genel Tıp | 1307 |
| Rehabilitasyon | 769 |
| Klinik nöroloji | 715 |
| Psikiyatri | 689 |
| Ameliyat | 683 |
| Hemşirelik | 629 |
| Onkoloji | 571 |
| Üroloji nefroloji | 563 |
| Romatoloji | 531 |
| Pediatri | 356 |

Tablo 2’de görüldüğü gibi yaşam kalitesi çalışmaları WOS tarafından en fazla tıp alanında, sonra sırasıyla rehabilitasyon ve klinik nöroloji alanlarında kategorize edilmiştir.

Makalelerin yayın diline göre dağılımı Tablo 3’te verilmiştir.

**Tablo 3. Yaşam Kalitesi Konulu Makalelerin Yayın Diline Göre Dağılımı**

|  |  |
| --- | --- |
| *Yayın dili* | *Makale Sayısı* |
| İngilizce | 9172 |
| Türkçe | 937 |
| İspanyolca | 5 |

Makalelerin %90’nın yazım dili İngilizce olup, %9 Türkçe yazılmıştır.

Yayınların üniversitelere göre dağılımı Tablo 4’te verilmiştir.

**Tablo 4. Yaşam Kalitesi Konulu Makalelerin Üniversitelere Göre Dağılımı**

|  |  |
| --- | --- |
| Üniversite | Yayın Sayısı |
| Hacettepe Üniversitesi | 867 |
| İstanbul Üniversitesi | 823 |
| Sağlık Bilimleri Üniversitesi | 546 |
| Dokuz Eylül Üniversitesi | 510 |
| Ege Üniversitesi | 493 |

Yayınların üniversitelere göre dağılımı incelendiğinde ilk sırada Hacettepe Üniversitesi’nin olduğu, ikinci sırada ise bunu yakından takip eden İstanbul Üniversitesi olduğu görülmektedir.

**Tablo 5. Yaşam Kalitesi Konulu Makalelerin Dergilere Göre Dağılımı**

|  |  |
| --- | --- |
| Dergi Adı | *Makale Sayısı* |
| TURKISH JOURNAL OF MEDICAL SCIENCES | 113 |
| RHEUMATOLOGY INTERNATIONAL | 111 |
| CUKUROVA MEDICAL JOURNAL | 89 |
| NOROPSIKIYATRI ARSIVI ARCHIVES OF NEUROPSYCHIATRY | 86 |
| TURKISH JOURNAL OF GERIATRICS TURK GERIATRI DERGISI | 86 |
| JOURNAL OF BACK AND MUSCULOSKELETAL REHABILITATION | 85 |
| ANADOLU PSIKIYATRI DERGISI ANATOLIAN JOURNAL OF PSYCHIATRY | 83 |
| ARCHIVES OF RHEUMATOLOGY | 61 |
| TURKISH JOURNAL OF PHYSICAL MEDICINE AND REHABILITATION | 61 |
| CLINICAL RHEUMATOLOGY | 60 |

Makalelerin dergilere göre dağılımı Tablo 5 yardımıyla incelenecek olursa Turkish Journal Of Medical Sciences ve Rheumatology International dergilerinin ön plana çıktığı görülmektedir.

**5. VERİNİN YAPISAL KONU MODELLEME İLE ANALİZİ VE BULGULAR**

Verilerin toplanması aşamasında sonra analiz edilmesi için R programından yararlanılmıştır. Analizlerin gerçekleştirilmesinde tm, stm, topicmodels, lda paketleri kullanılmıştır.

Verilen analiz edilmeden önce veri ön işleme yapılmıştır. Bu aşamada durak sözcüklerinin (a, an, the, is, vb.), noktalama işaretlerinin, sayıların kaldırılması, kelime kökünün elde edilmesi, tüm harflerin küçük harfe dönüştürülmesi gibi işlemler gerçekleştirilmiştir.

Yapısal Konu Modelleme sürecinde ilk aşamada modellemenin kaç konu için yapılacağına karar verilmesi modelin sonraki aşamalarının uygulanması açısından büyük önem arz etmektedir. Çalışmada konu sayısının belirlenirken farklı metrikler bir arada değerlendirilerek karar verilmiştir.

Verinin içerdiği konu sayısı belirlenirken dikkate alınacak kesin bir sayı veya kural olmadığından ilgili metriklere dayalı grafikler incelenmiş; konu sayısını belirlemek için iki aşamalı bir süreç izlenmiştir. Öncelikle geniş bir aralıkta konu sayısı belirlenmiş daha sonraki aşamada ise aralık daraltılarak uygun konu sayısı saptanmaya çalışılmıştır. Öncelikle konu sayısı 5 ile 20 arasında olmak üzere ve beşer beşer artacak şekilde elde edilen metriklerin grafikleri çizdirilmiştir. İlgili grafikler aşağıda verilmiştir.

**Şekil 3. 5 ile 20 Arasında Konu Sayısı İçin Metrikler**

metin, diyagram, çizgi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

Şekil 3’te verilen grafikler birlikte değerlendirildiğinde; 10’den çok konu sayısı için held-out likelihood değerinin artış eğimi düşerken, artık değerleri azalmakta ve anlamsal tutarlılık değeri de artan ivmeyle düşmektedir. Bundan sonraki aşamada 3 ile 10 arasındaki konu sayıları için metrikler yeniden hesaplanmış ve elde edilen sonuçlar Şekil 4’te sunulmuştur.

**Şekil 4. 3 ile 10 Arasında Konu Sayısı İçin Metrikler**

metin, diyagram, çizgi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

Şekil 4 incelendiğinde, 5’ten sonra özellikle anlamsal tutarlılık değerinde önemli bir kayıp olduğu tespit edilmektedir. Diğer metriklerin seyri de dikkate alınarak konu sayısı 5 olarak belirlenmiştir.

Yapısal Konu Modelleme sürecinde konu sayısı belirlendikten sonra, bu sayı kadar konu üretilip ilgili istatistikler incelenerek konu başlıkları belirlenmesi gerekmektedir. Konu başlıklarına verilecek isimler, Türk araştırmacıların yaşam kalitesi ile alakalı konuyu hangi boyutları ile ele aldıklarını belirlemek açısından önemlidir.

Her bir konu için en yüksek olasılığa sahip 15 kelimeyi gösteren orijinal görüntü, ilk dört konu için Şekil 5’te sunulmuştur.

**Şekil 5. Konu Bazında En Yüksek Olasılığa Sahip 15 Kelime**

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

Külliyatta sıklıkla kullanılan (life, method, quality gibi) kelimeler, birden fazla konu için yüksek olasılıklı kelimeler arasında yer almaktadır. Konu adlarının belirlenmesi aşamasında bu kelimeler ayırt edici katkı sağlayamamaktadır. Konu isimlerinin geçerli olarak belirleyebilmek için Bischof ve Airoldi (2012) tarafından önerilen, bir konuya bağlı kelime sıklığının kelime-konu ayrıcalığına oranı olarak tanımlanan FREX (Frequency-Exclusivity) metriğinden de yararlanılmıştır.

Tablo 6’da her bir konu ile ilişkili en çok kullanılan 10 kelime ve FREX kelimeleri birlikte verilmiştir.

**Tablo 6. Konuyu En İyi Temsil Eden Kelimeler ve FREX Kelimeleri**

|  |  |  |
| --- | --- | --- |
| Konu No | Konuyu En İyi Temsil Eden Kelimeler | FREX Kelimeleri |
| **1** | patient, score, pain, qualiti, diseas, life, signific, group, level | rls, fms, ankylos, spondyl, sleep, diabet, psqi, hemodialysi |
| **2** | life, qualiti, use, care, data, health, result, cancer, particip, social, level, support, caregiv | nurs, citi, urban, student, servic, resid, turkey, profession, countri, care, healthcar, caregiv, covid |
| **3** | group, test, patient, exercis, scale, function, valid, life, qualiti, signific, score, turkish | reliabl, cronbach, exercis, retest, copd, valid, asthma, translat, dyspnea, walk, rhiniti, mwt, strength |
| **4** | patient, treatment, group, month, signific, result, effect, improv, surgeri, follow, life | postop, prostat, resect, preoper, laser, surgeri, inject, graft, surgic, placebo, implant, stent |
| **5** | score, patient, group, depress, anxieti, life, children, qualiti, women, signific, scale, symptom, | psychiatr, sexual, mother, acn, children, anxieti, epilepsi, migrain, adhd, parent, fsfi, stai, trait, |

Analiz sonuçlarının daha rahat değerlendirebilmek için her konu için en yüksek olasılığı olan terimlerini görselleştirip ve özetleyen kelime bulutları oluşturulup, Şekil 6-7'de gösterilmiştir.

**Şekil 6. 1.-2. Konular İçin Kelime Bulutları**

metin, yazı tipi, ekran görüntüsü, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Şekil 7. 3.-5. Konular İçin Kelime Bulutları**

metin, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

YKM sürecinde, konu isimlerinin belirlenmesinde dikkate alınabilecek bir diğer çıktı, her bir konu için bu konuyla yüksek oranda ilişkili belgelerin içeriğidir. Şekil 8’de Konu 1 ve Konu 2 ile en yüksek ilişkiye sahip ikişer belgenin içeriği örnek olarak verilmiştir. Tüm konular için bu tip örnek belge içeriklerini tespit etmek mümkündür.

**Şekil 8. Konu 1 ve Konu 2 ile Yüksek Oranda İlişkili Örnek Belge İçerikleri**

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

Konu isimlerinin belirlenmesinde dikkate alınan tüm istatistiklerin incelenmesi sonucunda, araştırmanın amacı doğrultusunca belirlenen literatüre ilişkin 5 konu başlığı için verilen isimler Tablo 7’de gösterilmiştir.

**Tablo 7. “Yaşam Kalitesi” Literatürü Konu Başlıkları**

|  |  |
| --- | --- |
| Konu No | Konu Adı |
| 1 | Uyku Kalitesi |
| 2 | Covid 19 Dönemi Yaşam Kalitesi |
| 3 | *Solunum Hastalıkları* |
| 4 | *Ameliyat Öncesi ve Sonrası Yaşam Kalitesi* |
| 5 | Sağlıkla İlgili Yaşam Kalitesi |

Konu 1’de, “hasta, hastalık, ağrı, puan, yaş, uyku, rls (huzursuz bacak sendromu), fms, fibromiyalji, uyku kalitesi” vb. kelimeler ön plana çıktığı için konu başlığı *uyku kalitesi olarak* düşünülmüştür. Bu konu başlığında yaşam kalitesi ile huzursuz bacak sendromunun da görülen uyku bozuklukları arasındaki ilişkinin belirlenmesi, fibromiyaljinin yaşam kalitesi üzerindeki etkilerinin incelenmesi, aerobik egzersizlerin sigara içmiş bireylerde uyku ve yaşam kalitesi üzerine etkilerinin saptanması ile ilgili çalışmalar yer almaktadır.

Konu 2’de “yaşam, kullanım, katılımcı, sağlık çalışanları, hemşire, kent, öğrenci, Covid, turizm” vb. kelimeler ön plana çıktığı için konu başlığı olarak *Covid 19 dönemi yaşam kalitesi* düşünülmüştür. Bu konu başlığında, kent ve kırsal kesimde yaşam kalitesinin ölçülmesi, Covid 19 pandemisi döneminde sağlık çalışanlarının yaşam kalitesinin ölçülmesi, Covid 19’dan etkilenen yetişkinlerin Covid 19 korkusu ve yaşam kalitesi arasındaki ilişki, Covid 19 döneminde turizmin yerel halkın yaşam kalitesi üzerindeki etkisi, Covid 19 etkisinin görme engelli üniversite öğrencilerinin akademik yaşam ve yaşam kalitelerindeki etkisinin saptanması ile ilgili çalışmalar yer almaktadır.

Konu 3’te “test, geçerli, sonuç, KOAH, ölçek, rinit, mwt, astım” vb. kelimeler ön plana çıktığı için konu başlığı olarak *solunum hastalıkları* düşünülmüştür. Bu konu başlığında, KOAH hastalarının yaşam kalitesinin ölçülmesi, farklı tedavi yöntemlerinin etkilerinin KOAH ve astımlı hastalardaki etkinliğinin karşılaştırılması ve sonuçlarının yaşam kalitesi ile ilişkilendirilmesi, alerjik rinit tedavilerinin etkinliğinin karşılaştırılmasının sağlıkta yaşam kalitesi ölçeği ve diğer ölçeklerle ölçülmesi ile ilgili çalışmalar yer almaktadır.

Konu 4’te “hasta, tedavi, cerrahi, takip, ameliyat sonrası, ameliyat öncesi, lazer, enjeksiyon, hayatta kalma” vb. kelimeler ön plana çıktığı için konu başlığı olarak *ameliyat öncesi ve sonrası yaşam kalitesi* düşünülmüştür. Bu konu başlığında, çeşitli hastalıklardaki tedavi sonuçlarını değerlendirmek için yaşam kalitesi kriterlerinin ameliyat öncesi ve sonrasında karşılaştırılması, farklı cerrahi yöntemlerin yaşam kalitesi üzerindeki etkisinin değerlendirilmesi ilgili çalışmalar bu başlıkta yer almaktadır.

Konu 5’te “depresyon, kaygı, kadın, semptom, cinsel, menopoz” vb. kelimeler ön plana çıktığı için konu başlığı olarak *sağlıkla ilgili yaşam kalitesi* düşünülmüştür. Bu konu başlığında, kanser hastası çocukların yaşam kalitesi ve rehabilitasyon ihtiyaçlarının belirlenmesi, ses rehabilitasyonunun yaşam kalitesi, depresyon, anksiyete, cinsel işlevler üzerine etkilerinin değerlendirmesi, kadınlara uygulanan ayak refleksolojisinin vazomotor şikayetleri ve yaşam kalitesi üzerine etkilerini belirlemesi, yaygın anksiyete bozukluğu hastalarının yaşam kalitesinin çeşitli parametrelerle araştırılması ile ilgili çalışmalar bu başlıkta yer almaktadır.

Yaşam kalitesi ile ilgili literatürde konu başlıklarının saptanmasından sonra konular arası ilişki durumunun belirlenmesi aşamasına geçilmiştir. Bu aşamada, konu yaygınlığı değerleri ve konular arası korelasyonlar incelenmiştir.

Konu yaygınlığı, belirli bir konu başlığı ile ilintili belgelerin külliyat içindeki oranını vermektedir.

5 konunun her biri için hesaplanan konu yaygınlığı oranları, Şekil 9’da sunulmuştur.

**Şekil 9. Konu Yaygınlık Dağılımı**

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

Şekil 9 incelendiğinde; *ameliyat öncesi ve sonrası* *yaşam kalitesi*nin ana konu başlığı olarak ortaya çıktığı görülmektedir. Bu başlık altında ameliyat öncesi ve sonrasındaki, süreçte kişilerin yaşam kalitesi çeşitli ölçekler yardımıyla ölçülmek istenmektedir.

Konu yaygınlığı en yüksek ikinci ve üçüncü konu başlıkları sırasıyla, *Covid 19 dönemi yaşam kalitesi* ve *uyku kalitesi*dir. En az yaygınlığı olan konu başlıkları ise sırasıyla, *sağlıkla ilgili yaşam kalitesi* ve *solunum hastalıklarıdır*.

Konular arasındaki korelasyon yapısı Şekil 10 yardımıyla incelenmiştir.

**Şekil 10. Konular Arası Ağ Yapısı**

yeşil, renklilik içeren bir resim

Açıklama otomatik olarak oluşturuldu

Şekil 10’da verilen konular arasındaki ağ yapısı incelendiğinde; belirlenen konu başlıklarının bağlantılı olmadıkları görülmektedir.

YKM'nin bir diğer özelliği, belge düzeyindeki ortak değişkenlerin, genelleştirilmiş bir doğrusal modelle konuya ait yaygınlık parametresini nasıl etkilediğini modelleyebilmesidir. Böylece; ortak değişkenler ve konular arasındaki etkileşim araştırılabilmektedir.

Modele dahil edilecek ortak değişkenler, araştırmacı tarafından belirlenmelidir. Bu çalışmada, konu yaygınlığını etkileyen değişken olarak “yayın yılı” seçilmiştir.

Özellikle literatür incelemesi gibi çalışmalarda, ele alınan konu başlıklarının yıllara göre seyri önemli bir husustur. Böylece; konu yaygınlığındaki zamana bağlı değişimler tespit edilebilecek, popülerliği azalan/artan, akademisyenlere sıcak ve soğuk gelen konular tespit edilebilecektir (Ergüt ve Yıldırım, 2022).

“Yayın yılı”, YKM sürecine ortak değişken olarak dahil edilerek, yaşam kalitesi literatürünün her bir konu başlığı için elde edilen trendler Şekil 11’de verilmiştir.

**Şekil 11. Yıllara Göre Konu Yaygınlıklarının Değişimi**

metin, diyagram, çizgi, paralel içeren bir resim

Açıklama otomatik olarak oluşturuldu

Şekil 11’de kırmızı kesikli çizgiler %95 güven aralıklarını göstermektedir. Şekil üzerinde her bir konu için konu yaygınlığının trendi ayrıntılı olarak incelenerek, yükselme eğilimine sahip (sıcak) konular ile düşme eğilimine sahip (soğuk) konuları tespit etmek mümkündür.

Yaşam kalitesi literatürünün alt konu başlıkları ele alındığında, *Covid 19 dönemi yaşam kalitesi* (Konu 2), *solunum hastalıkları* (Konu 3), konu başlıklarının genel olarak artan bir popülariteye sahip olduğu söylenebilir. *Uyku kalitesi* (Konu 1), *ameliyat öncesi ve sonrası yaşam kalitesi* (Konu 4), konu başlıklarında yapılan çalışmaların ise akademik çalışmalar içindeki popülaritesinin azaldığı söylenebilir.

**6. SONUÇ**

Günümüzün veri odaklı dünyasında metin madenciliği işletmeler ve çeşitli alanlardaki araştırmacılar için vazgeçilmez bir araç haline gelmiştir. Yapılandırılmamış metin verilerinden değerli bilgiler elde etme yeteneği, bilgiye erişme ve analiz etmesiyle ön plana çıkan metin madenciliğinin önemi verilerin hacmi büyüdükçe artmaya devam edecektir. Metin madenciliği ile verilerin analizleri sonucunda elde edilen çıktılar ile işletmeler hizmet ve politika değişikliklerine gidebilir rekabet avantajı kazanabilir, bilinçli kararlar alabilir ve bilgi ve teknolojinin ilerlemesine katkıda bulunabilmektedirler. Aynı zamanda akademik araştırma alanında da önemli bir rol oynayan metin madenciliği literatürdeki eğilimleri ve boşlukları belirlemek amacıyla büyük hacimli bilimsel makaleleri, kitapları ve diğer metinsel verileri analiz etmek için de kullanılmaktadır. Bu çalışmada yaşam kalitesi literatürü metin madenciliği tekniklerinden konu modelleme ile incelenmiştir. Metin verilerini analiz etmek ve içindeki gizli yapıları ve dinamikleri ortaya çıkarmak için güçlü ve çok yönlü bir yöntemi temsil eden YKM konu sınıflandırmasının iyileştirilmesine yardımcı olmak için meta verileri birleştirerek ve konu etkileşimlerinin karmaşıklığını hesaba katarak, metin verilerinin daha kapsamlı ve bağlamsal bir görünümünü sağlayarak, altta yatan yapılara ve konular arasındaki ilişkilere dair değerli bilgiler sunmaktadır. Geniş metin koleksiyonlarının içerdikleri zengin bilgiyi keşfetmek ve daha derin bir anlayış elde etmek isteyen araştırmacılar için değerli bir araç olan YKM analizi sonucunda Türk yazarlar tarafından yayınlanan çalışmaların özellikle sağlık alanı temelinde yapıldığı, sağlık alanında yaşam kalitesinin değerlendirilmesine yönelik çalışmaların yoğun olduğu görülmüştür. Son yıllarda sağlık alanında yaşam kalitesine olan ilginin artmasının nedeni, teknolojik gelişmelere paralel olarak sağlık alanında da önemli ilerlemeler kaydedilmesidir. Hastalıkların tedavi edilmesi ile birlikte yaşam süresi uzamış, gerek yeni tedavi gerekse etkili ilaçların piyasaya sürülmesi le hastaların yaşam kalitesi giderek daha önemli hale gelmiştir (Müezzinoğlu, 2005).

Yaşam kalitesinin yapısal konu modelleme ile incelenmesi sonucunda bu konuda yapılan araştırmaların genel bir değerlendirmesi sağlanarak, ilgili alanlarda yapılacak yeni araştırmaların hangi yönlere odaklanabileceği konusunda yön gösterici olma, yaşam kalitesi konusunda politika önerileri geliştirme, araştırma sonuçlarının verimli bir şekilde kullanılması gibi çeşitli alanlarda da faydalı olabileceği düşünülmektedir. Çok çeşitli boyutları ve değerlendirmeleri kapsayan yaşam kalitesinin toplumsal, teknolojik ve çevresel değişikliklere yanıt olarak gelişmeye devam etmesi, yaşam kalitesinin çok yönlü boyutlarının ele alınacağı, bu konunun uzun süreler daha çalışmalarda inceleneceğine bir işarettir. Bu konudaki literatürün incelenmesi mevcut durumun analiz edilmesi, yeni uygulama alanlarının bulunmasına katkıda bulunacaktır.

**KAYNAKÇA**

Bai, X., Zhang, X., Li, K. X., Zhou, Y., & Yuen, K. F. (2021). Research topics and trends in the maritime transport: A structural topic model. *Transport Policy*, 102, 11-24.

Bischof, J., & Airoldi, E. M. (2012). Summarizing topical content with word frequency and exclusivity. In Proceedings of the 29th *International Conference on Machine Learning (ICML-12)* (pp. 201-208)

Blei, D. M., Ng, A. Y., &Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022. https://dl.acm. org/doi/10.5555/944919. 944937

Ergüt, Ö. & Yıldırım, İ. E. (2022). *Yapısal Konu Modelleme ile Kadına Yönelik Şiddet Sorununa Akademik Bakış*. Hiperlink Eğitim İletişim Yayın Gıda Sanayi ve Pazarlama Tic. Ltd. Şti..

Hofmann, T. (1999). Probabilistic latent semantic indexing. In Proceedings of the *22nd annual international ACM SIGIR conference* *on Research and development in information retrieval* (pp. 50-57) New York, USA

Müezzinoğlu, T. (2005). Yaşam kalitesi. *Üroonkoloji Bülteni*, *1*, 25-29.)

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder‐Luis, J., Gadarian, S. K., ... & Rand, D. G. (2014). Structural topic models for open‐ended survey responses. *American Journal Of Political Science*, *58*(4), 1064-1082.

WHO. (2012). WHOQOL: Measuring Quality of Life, https://www.who.int/tools/whoqol (Erişim tarihi: 03.12.2023)

**DOÇ. DR. ÖZLEM ERGÜT**

İstanbul doğumlu olan Özlem Ergüt, 2007 yılında Marmara Üniversitesi İktisadi ve İdari Bilimler Fakültesi Ekonometri Bölümü’nden mezun oldu. 2011 yılında Marmara Üniversitesi Sosyal Bilimler Enstitüsü İstatistik Bilim Dalı’nda yüksek lisansını, aynı Enstitüde 2016 yılında doktorasını tamamladı.  Ergüt, 2009-2017 yılları arasında Marmara Üniversitesi İktisat Fakültesi Ekonometri Bölümünde Araştırma görevlisi olarak görev yapmış olup, 2017-2023 yılları arasında aynı fakülte ve bölümde Doktor Öğretim Üyesi olarak görev yapmıştır. 2023 yılında Doçent unvanını almaya hak kazanan Ergüt’ün akademik ilgi alanları Çok Değişkenli Analiz, Uygulamalı İstatistik ve Metin Madenciliği'dir.